

**Virus-Host Interactions during Epstein-Barr virus Latent Infection of B cells**

by

Nicholas Joseph Homa

Department of Molecular Genetics and Microbiology  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Micah Luftig, Supervisor

\_\_\_\_\_  
Dirk Dittmer

\_\_\_\_\_  
Raluca Gordân

\_\_\_\_\_  
Jack Keene

\_\_\_\_\_  
Maria Schumacher

Dissertation submitted in partial fulfillment of  
the requirements for the degree of Doctor  
of Philosophy in the Department of  
Molecular Genetics and Microbiology in the Graduate School  
of Duke University

2017

ABSTRACT

**Virus-Host Interactions during Epstein-Barr virus Latent Infection of B cells**

by

Nicholas Joseph Homa

Department of Molecular Genetics and Microbiology  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Micah Luftig, Supervisor

\_\_\_\_\_  
Dirk Dittmer

\_\_\_\_\_  
Raluca Gordân

\_\_\_\_\_  
Jack Keene

\_\_\_\_\_  
Maria Schumacher

An abstract of a dissertation submitted in partial  
fulfillment of the requirements for the degree  
of Doctor of Philosophy in the Department of  
Molecular Genetics and Microbiology in the Graduate School of  
Duke University

2017

Copyright by  
Nicholas Joseph Homa  
2017

## Abstract

Epstein-Barr virus latently infects over 90% of the adult global population. While it generally results in a completely asymptomatic lifelong infection, EBV is associated with and the causative agent of several human malignancies, particularly in cases of immune suppression. There is extensive interplay between the virus and normal host factors and processes that are critical for maintenance of a successful latent infection. We sought to expand upon the current knowledge of this interplay in three areas: i) appropriation of the Notch signaling pathway via the host factor RBPJ, ii) global regulation of host gene expression and mRNA isoform choice, and iii) host factor regulation of the viral latent/lytic switch. Here we undertook biochemical and structural approaches to investigate the protein-protein interactions between viral latency-associated proteins and RBPJ. This revealed new information differentiating the surfaces of RBPJ bound by viral and Notch signaling proteins. Furthermore we utilized a novel algorithm with microarray data prior to infection and after establishment of latency in B cells to study host mRNA changes associated with latent infection. Here we identified transcription and splicing factors to be uniquely regulated by EBV at the level of mRNA isoform. Finally we identified two host factors that regulate the switch between latent and lytic replication based on their EBV-directed isoform usage.

## **Dedication**

This work is dedicated to Julie. She has been the only ever-present factor throughout this entire process and her support has been unwavering.

# Contents

Abstract .....	iv
List of Tables .....	x
List of Figures .....	xi
Acknowledgements .....	xiii
1. Introduction .....	1
1.1 Epstein-Barr Virus .....	1
1.2 EBV and Human Disease .....	2
1.3 Establishment of Viral Latency and Latent Gene Expression .....	3
1.3.1 EBV Latency Proteins and RNAs .....	3
1.3.2 EBV Latency Gene Expression in LCLs.....	8
1.4 The Role of RBPJ in EBV Latent Infection.....	11
1.4.1 Notch Signaling in Lymphocytes.....	11
1.4.2 Biochemical and Structural Background of RBPJ and NotchIC .....	14
1.4.3 Requirements of EBV Latency Proteins for RBPJ Interaction .....	17
1.5 Maintenance of EBV Latency and Regulation of the Latent/Lytic Switch .....	19
1.5.1 BZLF1 and Lytic Events.....	20
1.5.2 BZLF1 Promoter Regulation.....	21
2. Biophysical Characterization and Preliminary Structural Studies of EBV Latency Proteins and RBPJ .....	23
2.1 Rationale .....	23
2.2 Results .....	26

2.2.1 RBPJ Expression and Purification Strategy .....	26
2.2.2 Bacterially Expressed and Purified RBPJ Protein Retains Normal Biological Properties.....	27
2.2.3 EBNA3 Bioinformatic Characterization and Construct Design for Structural and Biochemical Studies.....	31
2.2.4 E3A-4 Expression and Purification Strategy .....	33
2.2.5 Bacterially Expressed and Purified E3A-4 Protein Disrupts RBPJ-DNA Interactions in EMSA .....	35
2.2.6 RAMANK and E2-ANK Expression and Purification Strategy.....	38
2.2.7 Crystallography Trials of RBPJ Bound to Cp DNA.....	40
2.3 Materials and Methods .....	43
2.4 Discussion.....	47
3. Defining the Molecular Interface on RBPJ for the EBV Latency Proteins and Notch ...	54
3.1 Rationale .....	54
3.2 Results .....	59
3.2.1 Bacterially Expressed RAMANK and E2-ANK Form a Complex with RBPJ and DNA .....	59
3.2.2 Predicted Surfaces of RBPJ Involved in Interactions with Notch and EBV Latency Proteins .....	63
3.2.3 Most, but Not All RBPJ Mutants Retain Normal DNA Binding.....	65
3.2.4 E2-ANK Binding Can Be Differentiated from RAMANK Binding to RBPJ.....	68
3.2.5 Disruption of Binding to RBPJ Directly Impairs Transcriptional Activation at Hes1p .....	71
3.2.6 EBNA2 and EBNA3 Proteins Utilize Both Overlapping and Unique Surfaces of Bind RBPJ.....	74

3.3 Materials and Methods .....	76
3.4 Discussion.....	79
4. Epstein-Barr Virus Induces Global Changes in Cellular mRNA Isoform Usage That Are Important for the Maintenance of Latency .....	87
4.1 Rationale .....	87
4.2 Results .....	89
4.2.1 EBV Infection of Primary B cells Induces Changes in Alternative mRNA Exon Usage.....	89
4.2.2 EBV Infection Induces Host Cell mRNA 3' UTR Shortening and Exon Exclusion.....	92
4.2.3 Diverse Biological Processes and Pathways Are Regulated by EBV Infection at the Level of Alternative mRNA Isoform Usage .....	94
4.2.4 EBV regulates mRNA Isoform Usage of RBPs.....	95
4.2.5 EBV Latency Alters XBP1 mRNA Isoform Usage and Suppresses IRE1-dependent XBP1 Splicing, Thereby Preventing Z Activation .....	99
4.2.6 EBV Latent Infection Inhibits Expression of the Full-length mRNA for TCF4 to Suppress Z Expression.....	104
4.3 Materials and Methods.....	107
4.4 Discussion.....	112
5. Conclusions and Future Directions .....	118
5.1 EBV Regulation of Transcription via RBPJ and Other Host Factors .....	120
5.2 Alternative EBNA Expression Strategies .....	123
5.3 Therapeutic Inhibition of RBPJ-EBNA Interactions .....	123
5.4 Regulation of TCF4 Alternative 5' Initiation .....	125

5.5 Relation of IRE1-Dependent Splicing of XBP1 with Alternative Initiation Isoforms and EBV Infection.....	126
5.6 Lytic Induction Therapeutics .....	127
5.7 Conclusion.....	127
References .....	129
Biography .....	143

## List of Tables

Table 1: EBV Latency Types and Gene Expression .....	5
Table 2: $K_d$ apparent from Tandem FA .....	71
Table 3: Primers Used .....	77
Table 4: Enrichment or depletion of gene ontology groups of alternatively regulated mRNAs. ....	95

## List of Figures

Figure 1: Latency Gene Expression during Transformation.....	10
Figure 2: Notch Protein Organization.....	13
Figure 3: RBPJ Protein Organization and Structure.....	15
Figure 4: Structure of the Notch transcriptional complex.....	17
Figure 5: Alignment of Notch1 and EBNA2 protein sequence centered on the RAM domain of NotchIC. ....	18
Figure 6: Regulation of the BZLF1 promoter. ....	22
Figure 7: RBPJ Expression and Purification. ....	27
Figure 8: Bacterially expressed and purified RBPJ protein retains normal biological properties. ....	30
Figure 9:E3A Bioinformatic Characterization for Structural and Biochemical Construct Design.....	33
Figure 10: E3A-4 Expression and Construct Design.....	35
Figure 11: Bacterially expressed E3A-4 protein disrupts RBPJ-DNA Interactions in EMSA.....	37
Figure 12: RAMANK and E2-ANK purification strategy. ....	40
Figure 13: Crystals of RBPJ bound to Cp DNA. ....	42
Figure 14: RAMANK and E2-ANK can form a complex with RBPJ in <i>in vitro</i> assays.....	62
Figure 15: Surfaces of RBPJ involved in protein-protein interactions.....	65
Figure 16: Most RBPJ mutants bind DNA normally.....	68
Figure 17: Tandem FA of RBPJ mutants on Hes1p with RAMANK and E2-ANK.....	70
Figure 18: RBPJ mutant luciferase transcription.....	74

Figure 19: RBPJ precipitation of EBV latency proteins. ....	76
Figure 20: SplicerEX analysis of mRNA expression and isoform usage. ....	91
Figure 21: Categorization and directionality of SplicerEX-predicted mRNA isoform changes. ....	93
Figure 22: Validation of RNA binding proteins predicted to change at the isoform level by SplicerEX. ....	98
Figure 23: XBP1 alternative isoform usage and lytic reactivation. ....	103
Figure 24: TCF4 alternative isoform usage and lytic reactivation. ....	106

## Acknowledgements

The first thanks goes to my dissertation advisor, Micah. Simply offering me a position in his lab to perform my thesis work was the first step on a long and fruitful road of mentorship, support, and friendship. I look forward to our continuing relationship following this experience. To all of the members of the Luftig Lab, past and present, I offer my thanks. I would like to specifically thank Raul Salinas, Kuan-yi Lu, Ashley Barry, and John Lu for their direct contributions to the work described here. I would also like to thank Pavel Nikitin, Karyn McFadden, Alex Price, Joanne Dai, Amy Hafez, and Josh Messenger for their professional and personal contributions during this time.

Additionally I would like to extend thanks to Arun Shukla, Liyin Huang, and Darrell Capel from the Lefkowitz Lab for their advisement and sharing of laboratory equipment, Porshia Shaw from the Schumacher Lab for her training with new assays, Nathan Nicely from the X-ray Crystallography Shared Resource for his mentorship, as well as the Cullen and Greenleaf labs for their advisement and shared equipment.

I also acknowledge my undergraduate research mentors, Andy VanDemark and Graham Hatfull. Their mentorship contributed greatly to my progression to Duke.

Finally I would like to thank my family, friends, and Julie for their help and support in getting me to a position where I could reach these achievements.

# 1. Introduction

Epstein-Barr virus (EBV) was discovered in 1964 in a case of Burkitt's lymphoma (30). From this original observation, we have now come to learn much about this virus and its role in human disease. Infection by EBV can result in a wide range of health implications from a completely asymptomatic infection to malignancies varying from lymphomas to epithelial carcinomas including Burkitt's lymphoma, Hodgkin's lymphoma, Nasopharyngeal carcinoma (NPC), Gastric carcinoma, Diffuse large B-cell lymphoma (DLBCL) and Post-transplant lymphoproliferative disorder (PTLD). Its intersections with the normal physiological pathways of the cells which it infects are extensive and highly complex. Humans have presumably lived with this virus for many millennia, resulting in extensive coevolution. As with many pathogens, particularly oncogenic ones, EBV does not infect with the intention of causing cancer, but rather its disease states are consequences of various aberrations in the normal host-pathogen interplay that takes place during infection. By studying this virus and its interactions with the host cell during infection we can move closer to a point where the negative health implications of an infection can be minimized, and potentially eventually eliminated.

## 1.1 *Epstein-Barr Virus*

EBV is a large, double-stranded DNA  $\gamma$ -herpesvirus. It is nearly ubiquitous in the human population, with over 90% of adults worldwide infected. Primary infection is

frequently asymptomatic and will result in a lifelong latent infection. These infections generally remain completely asymptomatic throughout our entire lives. Most primary infections occur early in life, prior to reaching adolescence, and frequently are acquired from family and household members. EBV has tropism for human epithelial cells, where primary infection is generally initiated, and for B lymphocytes, where the latent infection is ultimately established and maintained. While the overwhelming majority of infections are asymptomatic, controlled by a strong adaptive immune response, primary infection can occasionally result in infectious mononucleosis (81). Importantly, EBV can also contribute to the development of several different B-cell and epithelial cell malignancies.

## ***1.2 EBV and Human Disease***

Despite the predominance of asymptomatic infections, a primary infection that is initiated later in life, during adolescence or adulthood, can cause infectious mononucleosis. Virus can be shed throughout life in saliva, which is the primary mode of transmission, with oral epithelia being the site of most initial infections. Infectious mononucleosis arises from an expansion of EBV-infected B cells and the large subsequent T-cell response, resulting in the most common symptoms of fever, fatigue, and pharyngitis. This strong cytotoxic T-cell response will eventually contain and control the infection throughout life (28).

Failure to adequately control the initial expansion of B cells following infection, or subsequent disruptions of this immune response can lead to more serious conditions, such as B-cell malignancies arising from the latently infected cells. In most cases, this is a result of a compromised immune state such as in cases with HIV infection or immunosuppressive therapy following organ transplant. Frequently this results in an activation of the virus from a highly restricted latency state to a less restrictive state that would, under normal circumstances, be immunogenic. This less restrictive latency state is capable of driving proliferation of latently infected cells, resulting in B-cell lymphomas. EBV-associated malignancies include Burkitt's lymphoma, Hodgkin's Lymphoma, Diffuse Large B-Cell Lymphoma (DLBCL), and Post-Transplant Lymphoproliferative Disorder (PTLD). In addition to these B-cell disorders, EBV has also been shown to be associated with epithelial cell malignancies including Nasopharyngeal Carcinoma (NPC) and Gastric Carcinoma (114).

### ***1.3 Establishment of Viral Latency and Latent Gene Expression***

#### **1.3.1 EBV Latency Proteins and RNAs**

Despite its genome coding for approximately eighty genes, EBV only expresses nine latency proteins, as well as a handful of noncoding RNAs. The latency proteins include six Epstein-Barr Nuclear Antigens (EBNAs) and three Latent Membrane Proteins (LMPs). There are also two Epstein-Barr Encoded small RNAs (EBERs) that can be expressed during latent infections, as well as additional microRNAs BART and

BHRF1 microRNAs). Found in the nucleus are the EBNA Leader Protein (EBNA-LP), EBNA1, EBNA2, and EBNA3A, -3B, and 3C proteins. LMP1, LMP2A, and LMP2B localize to and span the plasma membrane with their transmembrane helices. The EBERs are localized to the nucleus and can bind to various host proteins (76). The ensemble of proteins expressed dictates the latency type of infected cells. Different EBV-associated malignancies are associated with different latency types.

Latency 0 is the most restricted latency type, where only noncoding RNAs are expressed and no latency proteins are expressed. Latency I is a slightly less restrictive latency type, where EBNA1 is the only latency protein expressed. In Latency II EBNA1 is expressed along with a subset of the remaining latency proteins. Recently, Latency II has been divided into two sub-states in which LMP1 and LMP2A/B are expressed (Latency IIa) or where EBNA2 and the EBNA3s are expressed (Latency IIb) (109). Finally, Latency III, or full latency, includes expression of all nine latency proteins and is the least restrictive—and consequently most immunogenic—latency type. In addition to the latency proteins present, various EBV non-coding RNAs can be present during each of the different latency types (**Table 1**).

**Table 1: EBV Latency Types and Gene Expression**

<b>Latency Type</b>	<b>Viral Proteins</b>	<b>Viral ncRNAs</b>	<b>Associated Diseases</b>
Latency 0	None	EBERs, miR-BARTs	Burkitt's lymphoma
Latency I	EBNA1	EBERs, miR-BARTs	Burkitt's lymphoma
Latency IIa	EBNA1, LMP1, LMP2A/B	EBERs, miR-BARTs	Hodgkin's lymphoma, NPC, Gastric carcinoma
Latency IIb	EBNA1, EBNA2, EBNA-LP, EBNA3s	EBERs, miR-BARTs, miR-BHRF1s	DLBCL, PTLD
Latency III	EBNA1, EBNA2, EBNA-LP, EBNA3s, LMP1, LMP2A/B	EBERs, miR-BARTs, miR-BHRF1s	DLBCL, PTLD

EBNA1 is expressed in most latency types (all but Latency 0) and its primary function is faithful replication and segregation of the EBV genome (153). The genome is maintained extra-chromosomally as an episome. EBNA1 binds the viral origin of replication, *OriP*, and tethers it to host cell chromosomes via a central Gly – Arg-rich region in what is presumed to be a nonspecific manner (20, 103). EBNA1 is essential for faithful segregation of the genome during cell division, but it has no enzymatic properties itself. It is the only EBV latency protein that has the ability to directly bind DNA.

EBNA2 is expressed in less restrictive latency types, including Type IIb and Type III latency. It is a transcriptional activator with two transcriptional activation domains, but requires host DNA binding proteins to access DNA as it has no DNA binding capacity itself. EBNA2 regulates transcription of both host and viral genes. It has the

capacity to activate transcription of all EBV latency genes, including itself. The ability to activate transcription at the viral C promoter (Cp) is critical for EBV to transform B cells and maintain a latent infection *in vitro* (151). In particular, EBNA2 can bind the host factor RBPJ (49), and this interaction is critical for regulating essential viral and host transcription (67). Additionally, variation in EBNA2 alleles is the primary basis for differentiation of the two virus strains infecting humans, Type 1 and Type 2 EBV (140).

EBNA-LP, along with EBNA2, is the first protein expressed upon infection and is maintained in Latency IIb and III (3). It is coded by several repeating exons at the 5' end of the gene, followed by two unique exons forming the C-terminal end of the protein. Depending at which 5' exon transcription is initiated and how the transcript is spliced, the translated protein can be multiple sizes based on how many repeats were incorporated into the N-terminal end of the protein. While not critical for transformation, EBNA-LP has been shown to co-activate transcription with EBNA2 (47). In particular, EBNA-LP has been shown to displace co-repressors from EBNA2-associated transcriptional complexes, specifically those associated with RBPJ (108).

The EBNA3 proteins most likely arose from a gene triplication event in the evolutionary past. They are expressed during Latency IIb and Latency III. While the C-terminal ends of these proteins diverge in sequence and structure, the first approximately one third of the proteins share a high degree of sequence conservation. This region is known as the EBNA3 homology domain (EBNA3hd). The EBNA3 proteins

are traditionally thought of as transcriptional repressors. Like EBNA2, they are unable to directly bind DNA themselves and as such use host proteins to access their transcriptional targets. Also like EBNA2, the EBNA3 proteins, via part of the EBNA3hd, bind RBPJ to regulate transcription (115). EBNA2 and EBNA3 binding to RBPJ is mutually exclusive, and they therefore compete for access to similar transcriptional targets (60, 144). The interaction with RBPJ is critical for *in vitro* transformation and maintenance of latency with regard to EBNA3A and EBNA3C (74, 88). EBNA3B has been shown to be dispensable for transformation *in vitro*, but may play a role as a tumor suppressor *in vivo* (149). EBNA3A and EBNA3C repress transcription through associations with histone methyltransferases (HMTs) and histone deacetylases (HDACs) as well as the polycomb repressive complexes (PRC1 and PRC2) (5, 106, 111, 128).

LMP1 is now known to be the primary viral transforming gene (69). It is associated with Latency IIa and III, and during transformation is not expressed to its full levels until several weeks after infection (110). It is a membrane protein with six transmembrane helices. LMP1 mimics constitutive CD40 receptor activation and signaling through the interactions of its intercellular domains with TNF receptor-associated factor (TRAF) proteins (26, 52, 64). LMP1 activation of the NF $\kappa$ B pathway contributes pro-growth and pro-survival signals in transformed cells.

The LMP2 proteins are made up by the A and B isoforms, where LMP2B lacks a small portion of the N-terminal cytoplasmic domain found in LMP2A. They are

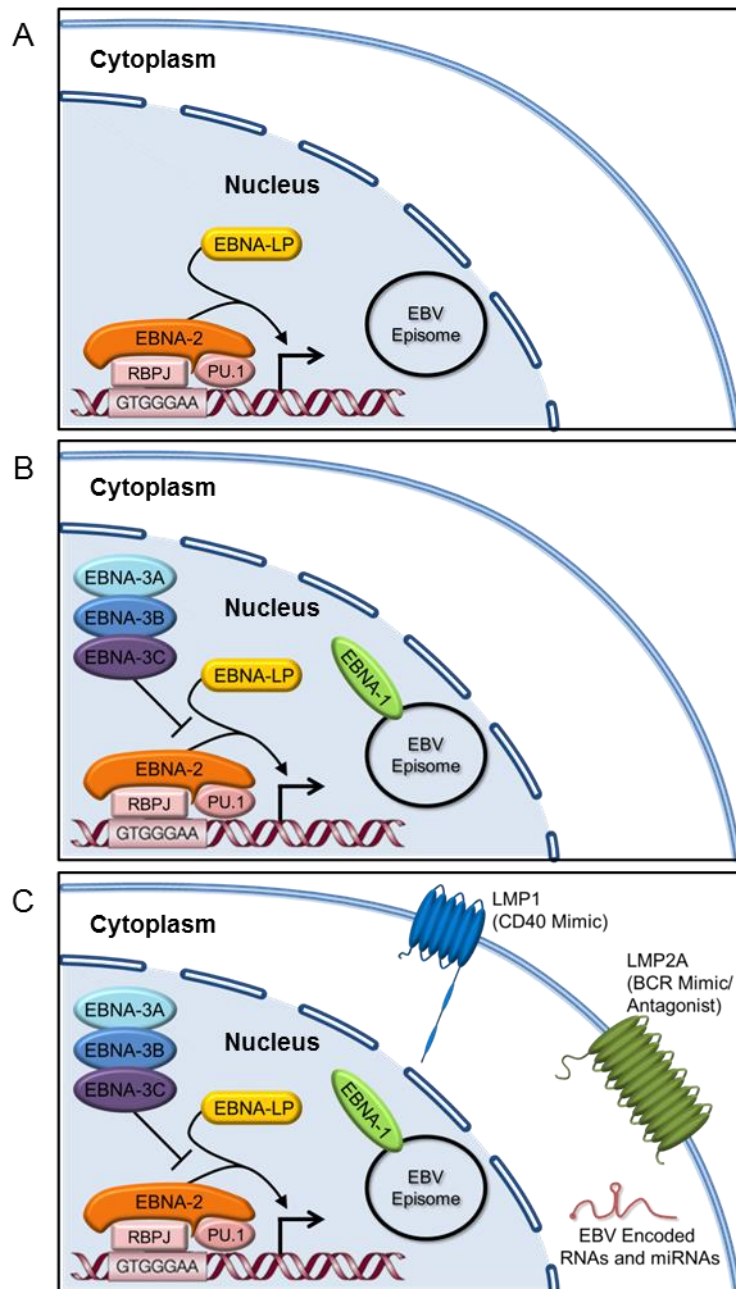
membrane proteins with twelve membrane-spanning helices. The LMP2A isoform, containing the full N-terminal domain, mimics constitutive B-cell receptor activation (2) and signaling and interacts with E3 ubiquitin ligases to coordinate proteasomal degradation of certain cellular proteins (55). LMP2B, which is incapable of activating B-cell receptor signaling, can form heterodimers with LMP2A, inhibiting its signaling ability (83).

### 1.3.2 EBV Latency Gene Expression in LCLs

During normal infections *in vivo* EBV is generally found to exist in Latency 0 in B cells, where no viral proteins are expressed. Most EBV-associated malignancies involve less restrictive latency states, so infections of naïve B cells are performed *in vitro*. These infections ultimately result in transformation of the B cells and establishment of an indefinitely proliferating and karyotypically stable lymphoblastoid cell line (LCL). We use these LCLs to study type III latency where all viral latency proteins are expressed.

Upon infection by EBV, the virion is internalized, the capsid is unpackaged, and the genome is translocated to the nucleus, is circularized, and becomes chromatinized. Shortly thereafter, EBNA-LP, EBNA1, and EBNA2 are expressed (3). EBNA2 and EBNA-LP utilize host transcriptional machinery, including RBPJ, to activate transcription of viral latency genes and host proliferative and pro-survival genes (161) (**Figure 1A**). At approximately 2.5 days post-infection the infected cells begin to proliferate and proceed through a hyperproliferative state where they will divide once every 8-12 hours. At

approximately 5.5 days post-infection the EBNA3 proteins are expressed and attenuate EBNA2-mediated transcriptional activation and proliferation and repress host tumor suppressor transcription (**Figure 1B**). The hyperproliferation slows to a more normal rate of approximately one cell division per 24 hours (102). Over the next several weeks, the LMP1 and LMP2A/B proteins will be expressed to increasing levels and the cells will transition from a Myc-driven proliferative state to an NF $\kappa$ B-driven proliferative state (**Figure 1C**)(110). After approximately five weeks all latency proteins are expressed to their full levels and these cells can be considered an LCL.



**Figure 1: Latency Gene Expression during Transformation.**

Diagram of EBV gene expression following *in vitro* infection of B cells and transformation into an LCL. **(A)** Shortly after infection EBNA2 and –LP are expressed for genome maintenance and viral and host transcriptional activation. **(B)** Later the EBNA3 proteins are expressed and attenuate EBNA2 transcriptional activation as well as host tumor suppressor expression. Additionally, EBNA1 is expressed by this point. **(C)** By

five weeks after infection all latency proteins and non-coding RNAs are expressed and cells can be considered LCLs.

## **1.4 The Role of RBPJ in EBV Latent Infection**

As with any obligate intercellular pathogen, EBV subverts and co-opts host pathways during infection. One of these is the Notch signaling pathway, which communicates extracellular ligand binding—typically from neighboring cells—via signal transduction resulting in modulation of gene expression. It is involved in lymphocyte development and maturation, and its aberrant activation has been implicated in several T-cell malignancies including T-cell Acute Lymphoblastic Leukemia (T-ALL)(96). While this pathway is not normally active in naïve B-cells, some of its components serve a critical role during EBV transformation and latent infection.

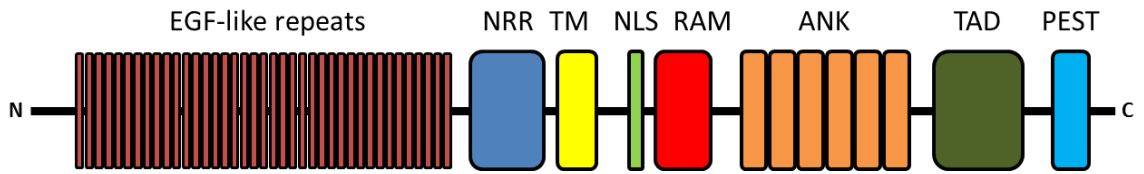
### **1.4.1 Notch Signaling in Lymphocytes**

The Notch signaling pathway is conserved among *C. elegans* and higher eukaryotes. It is an essential mediator of cell-cell interactions during development by responding to external ligands to either facilitate or suppress differentiation steps in lymphocytes through transcriptional activation or repression of target genes. An example of this in B cells is the ability to promote differentiation to marginal zone B cells while suppressing differentiation to follicular B cells in the spleen (72).

There are four Notch proteins (NOTCH1-4) in humans that each have similar structures and possess a single transmembrane domain, localizing them to the plasma membrane (**Figure 2**)(71). A large extracellular domain binds one of the five Notch

pathway ligands (Delta 1, 3, and 4, and Jagged 1 and 2). In mature B cells, Notch2 is the only Notch protein expressed, and its ligand specificity is limited to Delta1 and Jagged 1 and 2 (119). Upon ligand binding, a series of conformational changes and proteolytic events occur, requiring ADAM metalloproteases and the  $\gamma$ -secretase complex. This ultimately releases the intercellular Notch domains (NotchIC) from the plasma membrane. NotchIC is limited to the RBPJ interaction motif (RAM), Ankyrin repeats, a transcriptional activation domain, and PEST domain. It also contains a nuclear localization signal, directing it to the nucleus for transcriptional regulation. Upon translocation to the nucleus NotchIC can bind to RBPJ, the downstream DNA binding component of the Notch signaling pathway, at target gene promoters.

Under non-activating conditions, RBPJ is constitutively bound to DNA at target promoters as a transcriptional repressor (27, 51). It is bound by co-repressors, such as KyoT2 (133), SKIP (160), and the SHARP complex (104), and represses transcription through recruitment of N-CoR2 and HDAC1 (63), among other proteins. Following Notch signaling activation NotchIC is capable of displacing these bound co-repressors. This can result in transcriptional de-repression at target genes. Once formed, the NotchIC-RBPJ complex can bind Mastermind-like protein 1 (MAML-1), which recruits CBP/p300 to activate transcription (100, 143). NotchIC binding does not itself alter the sites on DNA bound by RBPJ (24). Therefore RBPJ is entirely responsible for target site determination in this pathway.

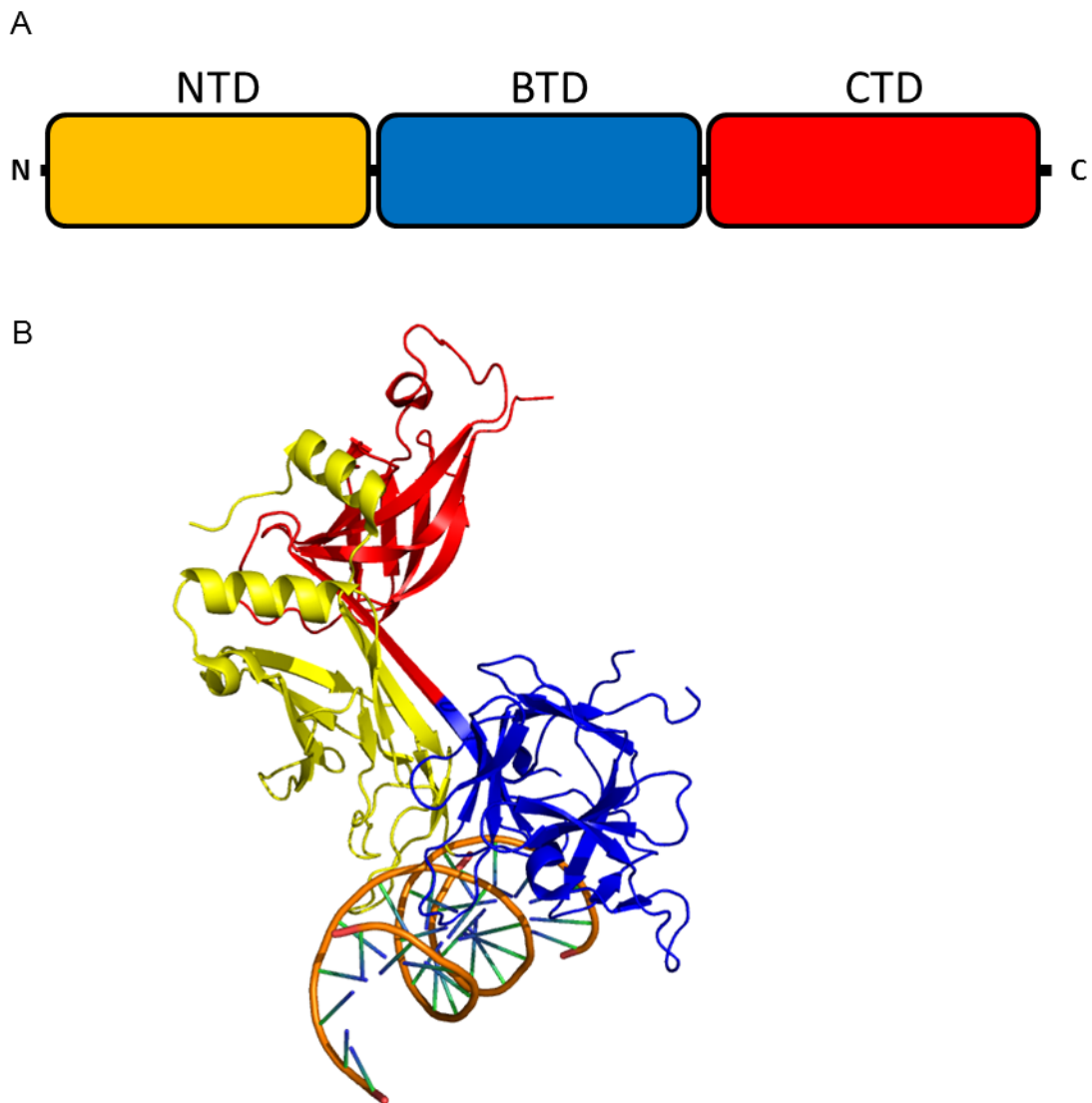


**Figure 2: Notch Protein Organization.**

Epidermal Growth Factor-like motif (EGF-like repeats); Negative Regulator Region (NRR); Transmembrane domain (TM); Nuclear Localization Signal (NLS); RBPJ Association Motif (RAM); Ankyrin repeats (ANK); Transcriptional Activation Domain (TAD); PEST domain.

### 1.4.2 Biochemical and Structural Background of RBPJ and NotchIC

RBPJ has three domains: an N-terminal domain (NTD), a  $\beta$ -Trefoil domain (BTD), and C-terminal domain (CTD) (**Figure 3A**). The NTD and CTD show structural similarity to Rel proteins. Unlike Rel proteins, however, RBPJ binds DNA as a monomer (17, 101) through its NTD and BTD (**Figure 3B**), and there is no evidence for RBPJ dimerization. RBPJ recognizes the consensus DNA sequence 5'-GTGGGAA-3' and binds it with an affinity in the low-to-mid nanomolar range (38). There is a hydrophobic pocket on the surface of the BTD that serves as a binding site for KyoT2 (19), an RBPJ co-repressor. Additionally this pocket serves as a binding site for the RAM peptide of NotchIC, which displaces the co-repressors and allows for binding of ANK to the CTD of RBPJ. This hydrophobic pocket has been shown to be critical for NotchIC binding and function (16, 39).

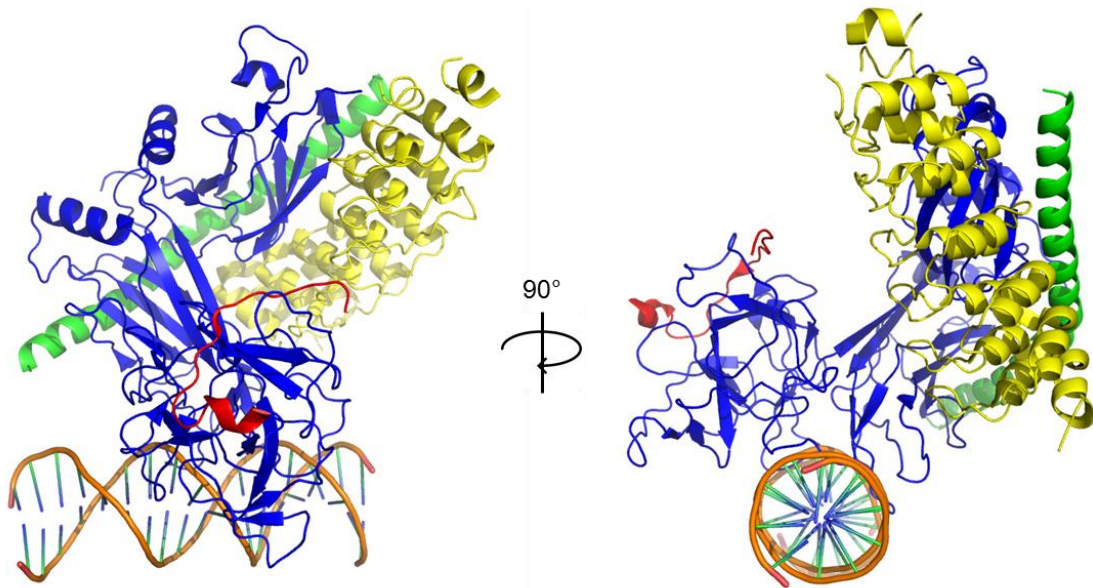


**Figure 3: RBPJ Protein Organization and Structure.**

(A) RBPJ protein organization. N-terminal and C-terminal domains (NTD, CTD) are Rel homology domains flanking the  $\beta$ -trefoil domain (BTM). (B) Crystal structure of RBPJ bound to Hes1p DNA (PDB ID: 3BRG (39)). NTD is shown in yellow; BTM is shown in blue; CTD is shown in red.

Several crystal structures have been solved of NotchIC bound to RBPJ at DNA (6, 16, 100, 150). These structures have revealed much about the formation of this complex, including with MAML-1 (Figure 4). First, the RAM peptide of NotchIC binds to RBPJ

through interactions of a conserved  $\Phi W \Phi P$  motif (where  $\Phi$  represents a hydrophobic amino acid) with the hydrophobic pocket on the RBPJ BTB. This mechanism of binding is conserved with KyoT2 (19), which has also been shown to bind the hydrophobic pocket with a similar motif. Second, the Ankyrin repeats of NotchIC bind the CTD of RBPJ, with an unstructured loop connecting it to RAM. Third, binding of NotchIC (and specifically RAM) induces an “opening” of a small loop on the RBPJ NTD which allows for binding of MAML-1 (39). This represents the core Notch transcriptional complex. Additionally, NotchIC has been shown capable of forming homodimers via its Ankyrin repeats (6). This allows for formation of tandem transcriptional complexes in a head-to-head manner resulting in the ability to coordinate transcription at composite DNA sites.



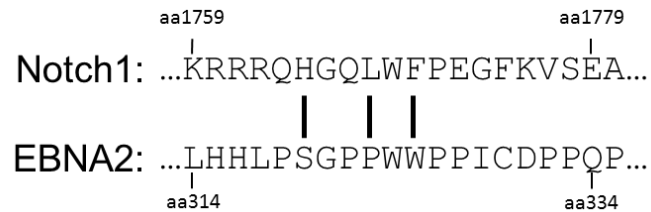
**Figure 4: Structure of the Notch transcriptional complex.**

Crystal structure of RBPJ bound to RAMANK and MAML-1 at Hes1p DNA, rotated 90°. RBPJ is shown in blue; the Ankyrin repeats of RAMANK are shown in yellow; MAML-1 is shown in green; the RAM peptide of RAMANK is shown in red. RAM and ANK are covalently joined as a single polypeptide, however the linker region is unstructured and unable to be visualized (PDB ID: 3V79 (16)).

### **1.4.3 Requirements of EBV Latency Proteins for RBPJ Interaction**

None of the EBV latency proteins have the ability to bind DNA with the exception of EBNA1. Therefore the EBNA2 and EBNA3 proteins, which are transcriptional regulators, require intermediary host proteins for access to DNA. RBPJ is one mechanism through which these proteins accomplish this. EBNA2 utilizes RBPJ in a manner that appears to mimic activated Notch signaling (161). Indeed, Notch and EBNA2 expression are capable of partially rescuing each other with regard to transcriptional activation at certain loci (43). EBNA2 also contains a  $\Phi W \Phi P$  motif that, like RAM of NotchIC, is the primary mediator of the protein-protein interaction (**Figure**

5) (61). Disruption of the ability of EBNA2 to bind to RBPJ results in arrested proliferation of LCLs and would presumably prevent transformation from *de novo* infections as well (31). Unlike Notch1C, EBNA2 does not require recruitment of MAML-1 for transcriptional activation as it is capable of recruiting CBP/p300 directly (146). Outside of the RAM-like motif, EBNA2 shares no homology to the rest of Notch1C, although it does contain a dimerization domain. Aside from binding the hydrophobic pocket of RBPJ the extent of mimicry of the Notch signaling pathway by EBNA2 from a structural perspective is unclear.



**Figure 5: Alignment of Notch1 and EBNA2 protein sequence centered on the RAM domain of Notch1C.**

EBNA3A and EBNA3C similarly require interaction with RBPJ for transformation and maintenance of proliferation in LCLs (74, 88). Previous truncation and deletion experiments have identified regions within the EBNA3hd that are required for binding RBPJ. These have identified a 71 residue stretch in EBNA3A and an overlapping 48 residue stretch in EBNA3C that are critical for binding. Contained within this region is a conserved TxLC motif that has been shown to disrupt interactions with RBPJ. EBNA3C also contains a 90 residue stretch sixty amino acids downstream of the

aforementioned regions that also disrupts RBPJ interaction when deleted, but is not critical for EBNA3A. Whether this is truly a difference between EBNA3A and EBNA3C interactions or is simply an artifact of the experimental design is unclear. A further discrepancy between EBNA3C and EBNA3A or EBNA3B is a putative  $\Phi W\Phi P$  motif that is only found in EBNA3C (VWTP) (13). The corresponding EBNA3A sequence is TFKL and the EBNA3B sequence is IVQI. Based on what is known about RAM/EBNA2 binding to RBPJ in the hydrophobic pocket, the W and P residues are absolutely critical. Mutation of the W in this motif of EBNA3C has a mild effect on RBPJ binding; however it does block the ability of EBNA3C to bind an isolated RBPJ BTD. Like EBNA3A, EBNA3C can bind both the NTD and BTD independently. This suggests the EBNA3 proteins may bind RBPJ at multiple sites, or at a site shared by the NTD and BTD, and the putative  $\Phi W\Phi P$  motif in EBNA3C is binding the BTD specifically.

### ***1.5 Maintenance of EBV Latency and Regulation of the Latent/Lytic Switch***

The default infection pattern for EBV is to establish a latent infection. This ensures a persistent infection and replication of the viral genome as a consequence of cellular proliferation, as well as persistence in a non-immunogenic state. EBV also can undergo lytic replication where new virions are produced and shed from host cells in order to infect new cells and new individuals. Regulation of the balance between latent and lytic replication is a crucial and complex aspect of EBV biology. The key to regulation of this switch is the viral BZLF1 gene.

### 1.5.1 BZLF1 and Lytic Events

BZLF1 encodes the main lytic transactivator protein for EBV, Zta/ZEBRA/Z. Consistent with herpesvirus nomenclature it is an immediate-early gene. Latent versus lytic replication can be controlled by repression or activation respectively of transcription at the BZLF1 promoter (Z promoter) and Z translation (95). Z is a member of the bZIP family of transcription factors and has four functional domains: a transactivation domain, a regulatory domain, a basic DNA binding domain, and a coiled-coil dimerization domain. It is very similar to AP-1 proteins such as Jun and Fos in structure and in the DNA elements in which it targets (32). Z is differentiated however from AP-proteins in that it binds specifically to methylated DNA (10). Indeed, alanine-to-serine mutants of Jun and Fos homologous to Z allow them to bind methylated DNA and in fact disrupt EBV latent infection by inducing lytic replication (156). Z-mediated transcriptional activation is a result of recruitment of histone acetyltransferases (HATs) to target gene promoters as well as removal of repressive elements from the promoter (1).

Lytic replication can arise spontaneously at low frequency in latently infected cells, or it can be induced by HDAC inhibitors (82), phorbol ester agonists of protein kinase C (PKC) (36), induction of the unfolded protein response (UPR) (127), or activation of the B-cell receptor (BCR) (132). Acetylation of the Z promoter allows for active transcription of Z mRNA and subsequently translation of Z protein (58). Z is

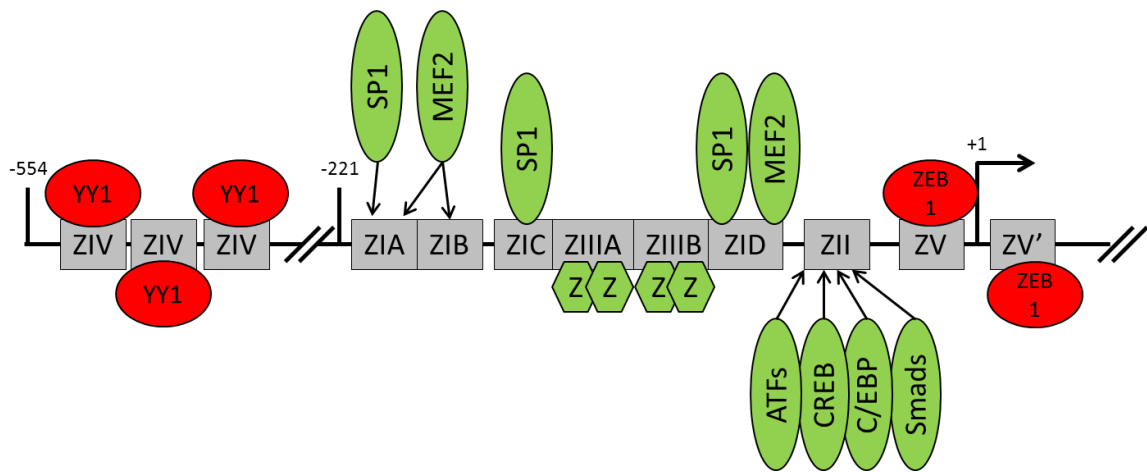
capable of inducing its own transcription in a feed-forward loop (154). Z can also activate transcription of viral lytic genes by binding specifically to the methylated EBV genome. Its primary role is to activate transcription of the BRLF1 gene which codes for the Rta/R protein (9). In addition to another feed-forward loop where R can activate transcription of Z, R and Z initiate a transcriptional cascade of separate classes of lytic genes that ultimately results in the replication of the viral genome and production of new infectious virions. As Z expression is the critical factor, ectopic expression of Z is sufficient to induce entry into lytic replication (21).

### **1.5.2 BZLF1 Promoter Regulation**

Activation of transcription at the Z promoter is the trigger for the switch to lytic replication. As such, the Z promoter is tightly regulated. There are several *cis*-DNA elements that regulate the promoter, numbered ZI-ZV (**Figure 6**).

Four ZI elements are found throughout the promoter and are recognized by SP1 and MEF2 family transcription factors (79, 80). Activated SP1 and MEF2D recruit HATS to the ZI elements. The ZII element is recognized by several bZIP family transcription factors and is essential for Z transcription (147). The two ZIII elements, ZIIIA and ZIIIB, are Z response elements (ZREs) and are bound by Z protein as part of the feed-forward expression loop (35). ZIV is a silencing element, but has not been extensively studied (98). ZV and ZV' can be bound by ZEB1 and ZEB2 respectively, zinc-finger E-box binding proteins that repress transcription (29).

It has been shown that histone acetylation at the Z promoter is a primary determinant of activation, and as such HDAC inhibitors are capable of inducing lytic replication (58). PKC agonism causes an accumulation of  $\beta$ -catenin and results in induction of the UPR (23). This has the ultimate consequence of inducing activated BLIMP1 and XBP1 which are capable of activating the Z promoter in epithelial cells and lymphocytes respectively (8, 113). Treatment of latently infected cells with anti-immunoglobulins to cross-link the BCR also can induce Z expression through a signal transduction pathway that, like phorbol esters, activates PKC as well as PI3K and MAPK (23, 56, 122).



**Figure 6: Regulation of the BZLF1 promoter.**

BZLF1 *cis*-regulatory elements are shown with the *trans*-acting factors which bind them. ZI-ZIII elements are positive regulatory elements and ZIV and ZV are negative regulatory elements. The positively acting *trans* factors are shown in green and the negatively acting *trans* factors are shown in red.

## **2. Biophysical Characterization and Preliminary Structural Studies of EBV Latency Proteins and RBPJ**

### **2.1 Rationale**

It has been well established that the EBNA2 and EBNA3 proteins play critical roles during B cell infection and latency in LCLs. Infection with a virus strain lacking the EBNA2 coding region fails to transform B cells (18). Additionally, infection with a recombinant virus that contains premature termination codons in the EBNA3A or EBNA3C open reading frames also fails to transform B cells (138). Conversely, EBNA3B is not required for B cell transformation (137, 138). In addition to their requirement for initial transformation and establishment of latency, EBNA2, EBNA3A, and EBNA3C are each required to maintain cell growth during latent infections (73, 87, 89). Furthermore, the EBNA2 and EBNA3 proteins are key contributors to the pathogenesis of a number of EBV-associated malignancies, including post-transplant lymphoproliferative disorder (PTLD), diffuse large B-cell lymphoma (DLBCL), and Hodgkin's lymphoma, among others.

Activation of Notch signaling in lymphocytes induces the intercellular portion of the Notch protein (NotchIC) to bind to the pathway's downstream DNA binding component RBPJ. This interaction, which leads to the formation of the active Notch transcriptional complex, is mediated primarily by the RBPJ-association motif (RAM) of

Notch. A  $\Phi W\Phi P$  motif within RAM is the primary contributor of the thermodynamics of the protein-protein interaction by binding within a hydrophobic pocket on the surface of RBPJ (61). This  $\Phi W\Phi P$  motif is conserved among all members of the mammalian Notch family, *Drosophila* and *C. elegans* Notch homologues, as well as the cellular RBPJ corepressor KyoT2, all of which use this motif to bind the hydrophobic pocket on the beta-trefoil domain (BTD) of RBPJ (19). Interestingly, EBNA2 also contains a  $\Phi W\Phi P$  motif, and this has also been shown to drive its interaction with RBPJ (31, 61, 77). Furthermore, EBNA2 and NotchIC can functionally replace each other with regard to transcriptional activation to a degree (70, 120, 129, 161). Atomic resolution structures of Notch and KyoT2 bound to RBPJ present a similar theme with the  $\Phi W\Phi P$  motif binding the hydrophobic pocket of RBPJ, suggesting this would be common to EBNA2 as well. However, differences between Notch and KyoT2 binding outside of the  $\Phi W\Phi P$  motif, coupled with the fact that EBNA2 shares no homology to either protein other than the  $\Phi W\Phi P$  motif allow for the possibility of some unique aspects to the interaction.

The EBNA3A, -3B, and -3C proteins were the presumed result of a gene triplication event, and each contain an approximately 300 amino acid homology domain (EBNA3hd) near their N-termini. The EBNA3 proteins also bind RBPJ and it has been demonstrated that the interactions of EBNA3A and EBNA3C with RBPJ are in part responsible for these proteins' critical roles in transformation and LCL growth (74, 88). A 2011 study (13) described that both EBNA3A and EBNA3C utilize both the N-terminal

domain (NTD) and BTD of RBPJ. The study also mapped an approximately 50 amino acid region of EBNA3A and a 23 amino acid region of EBNA3C that are critical for binding RBPJ. Interestingly this region of EBNA3C contains a VWTP sequence that is similar to the  $\Phi W\Phi P$  found in the RAM peptide of NotchIC and EBNA2, but is not conserved in EBNA3A or EBNA3B. Mutation of this sequence resulted in loss of binding of EBNA3C to the BTD of RBPJ (13). However discrepancies between the abilities of various mutant EBNA3 proteins to bind RBPJ based on bacterial versus mammalian expression systems leaves the mechanisms of interaction with RBPJ unclear still.

An atomic resolution view of the interactions between the EBNA proteins and RBPJ has the potential to reveal the extent of EBNA2 mimicry of Notch signaling and a clear understanding of the mechanisms of binding to RBPJ utilized by the EBNA3 proteins. For these reasons, we have undertaken a structural approach to study these interactions with X-ray crystallography. In this chapter we provide an extensive biophysical characterization of the production and purification of the viral and cellular proteins for these studies. The structural approaches include co-crystallization of a portion of EBNA3A bound to RBPJ, a peptide of EBNA2 bound to RBPJ, and a chimeric protein composed of NotchIC with the RAM peptide substituted for the corresponding 20 amino acids from EBNA2. Due to evidence that the affinity of EBNA2 for RBPJ is significantly lower than that of Notch for RBPJ (61) we devised a strategy to incorporate

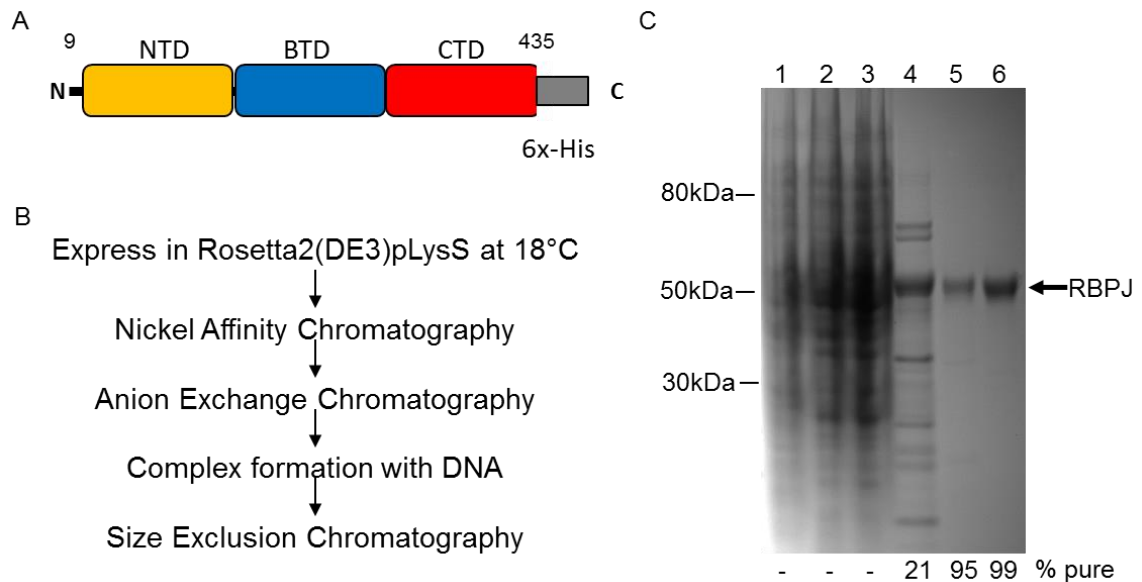
the Ankyrin repeat domain of Notch in order to increase the affinity of the EBNA2  $\Phi$ W $\Phi$ P motif-containing construct for Notch.

## **2.2 Results**

### **2.2.1 RBPJ Expression and Purification Strategy**

Previous structural studies have determined the atomic resolution structures of RBPJ bound to Hes1p DNA—a canonical promoter targeted by Notch signaling—in various states as part of the Notch transcriptional complex (16, 39, 100, 150). However there is no high resolution structures that have been solved of RBPJ bound to any EBV viral proteins. We used a construct of human RBPJ, courtesy of Stephen Blacklow, designed for bacterial expression, purification, and X-ray crystallography (**Figure 7A**) (100). The construct includes amino acids 9-435, which represents nearly the entire protein, with a C-terminal 6xHis-tag. The protein was expressed in *E. coli* bacteria at 18°C and purified in three steps (**Figure 7B**). First, affinity chromatography with Ni-NTA agarose exploited the 6xHis-tag. Second, the RBPJ protein construct has an isoelectric point of 7.96, and so possesses a net negative charge in most alkaline buffers. As such, we performed anion exchange chromatography for the RBPJ protein recovered from affinity chromatography. Third, we performed size exclusion chromatography as a final purification step to remove any remaining contaminants. This purification strategy

produced RBPJ protein to a purity level of >99% (**Figure 7C**). This purity level is sufficient for X-ray crystallography and other sensitive biochemical assays.



**Figure 7: RBPJ Expression and Purification.**

**(A)** RBPJ expression construct. Includes NTD, BTM, and majority of CTD with a C-terminal 6x-His tag. **(B)** RBPJ expression and purification strategy. **(C)** Coomassie stained gel of RBPJ purification steps. 1=uninduced bacterial lysate; 2=induced bacterial lysate; 3=soluble lysate; 4=nickel affinity elution; 5=anion exchange elution; 6=size exclusion peak fraction following DNA binding.

### 2.2.2 Bacterially Expressed and Purified RBPJ Protein Retains Normal Biological Properties

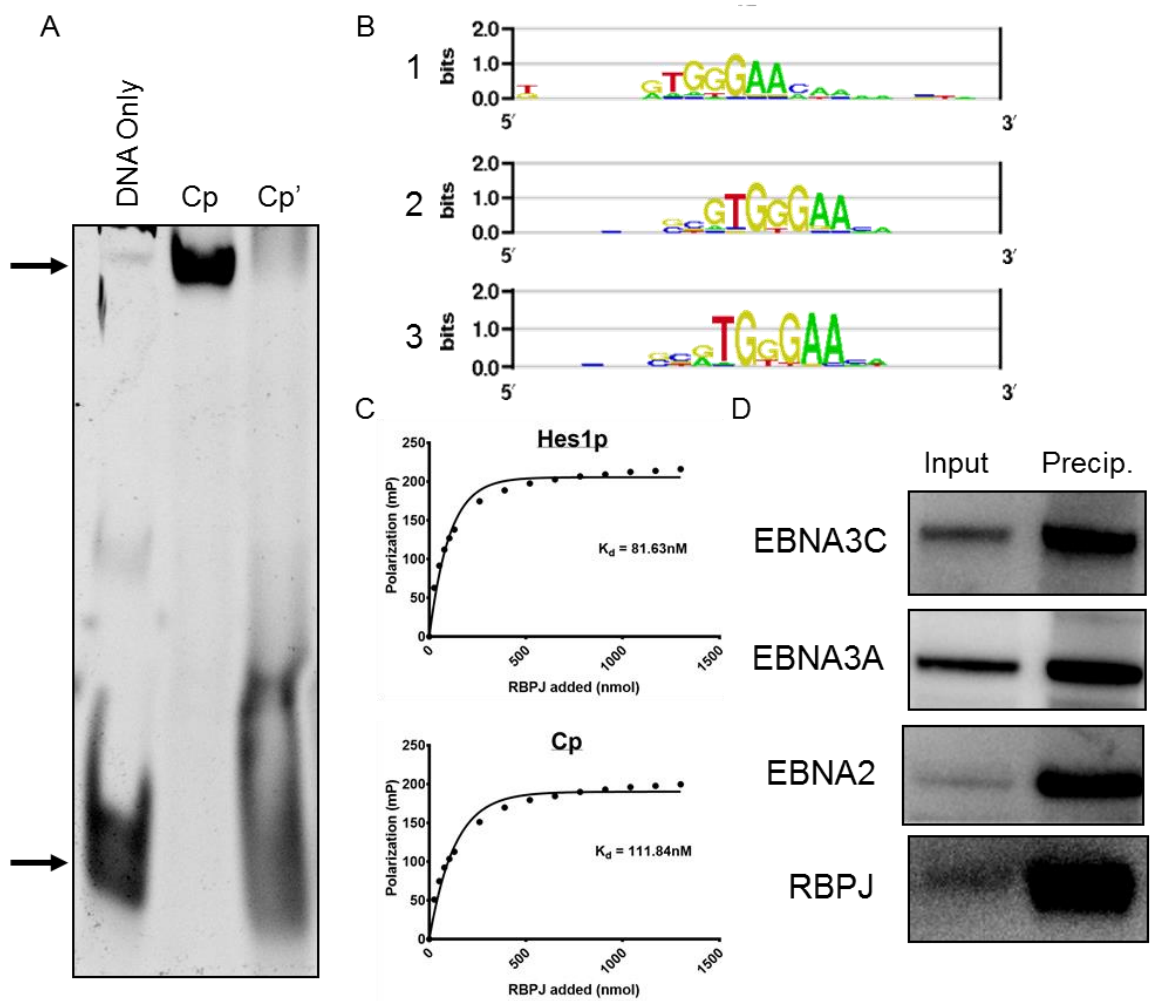
As mammalian protein expression in bacterial systems frequently presents problems with solubility, proper folding, post-translational modifications, and activity, it is critical to interrogate the biological properties of any protein expressed and purified in this manner. As such, we performed several experiments to confirm the fidelity of our purified RBPJ protein, focusing on its ability to bind DNA with specificity and to form

canonical protein-protein interactions (**Figure 8**). RBPJ binds to a consensus DNA sequence of 5'-GTGGGAA-3' (139). Bacterially expressed and purified RBPJ is capable of binding the viral C promoter (Cp), which contains this sequence, in an electrophoretic mobility shift assay (EMSA) (**Figure 8A**). However when the recognition motif is mutated, RBPJ is no longer capable of binding, demonstrating sequence-specific DNA affinity. Additionally, we performed protein binding microarray (PBM) experiments in which bacterially expressed and purified RBPJ protein was incubated with all possible 8-mers of DNA on a microarray (**Figure 8B**). This revealed that the top DNA sequence motifs to which RBPJ binds all contain the consensus 5'-GTGGGAA-3' sequence, with little contribution outside of this sequence to specificity.

In order to address the affinity of the bacterially expressed and purified RBPJ protein for its DNA recognition sequence we performed fluorescence anisotropy (FA) with the cellular Hes1p DNA and viral Cp DNA (**Figure 8C**). In these experiments, we found RBPJ to bind Hes1p with a  $K_d$  of 82nM and bind Cp with a  $K_d$  of 112nM. These  $K_d$ 's were similar to those obtained in isothermal calorimetry (ITC) experiments that were performed with RBPJ binding to Hes1p under similar conditions (38). These data confirm that bacterially expressed and purified RBPJ protein exhibits normal DNA binding activity.

During EBV latent infection of B cells RBPJ is capable of binding the EBNA2 and EBNA3 proteins. We investigated the ability of bacterially expressed and purified RBPJ

to bind EBNA2, EBNA3A, and EBNA3C from LCLs. RBPJ protein was incubated with soluble protein lysate generated from LCLs and Ni-NTA agarose beads. RBPJ protein was then precipitated by the Ni-NTA beads via its 6xHis-tag and analyzed by Western Blot to identify EBV latency proteins that were co-precipitated (**Figure 8D**). This revealed that RBPJ was capable of forming canonical protein-protein interactions. All of these data taken together confirm that RBPJ that is expressed and purified from bacteria retains normal biological properties.



**Figure 8: Bacterially expressed and purified RBPJ protein retains normal biological properties.**

**(A)** EMSA of RBPJ with Cp DNA. Cp' denotes Cp DNA in which the RBPJ binding site has been mutated. The bottom arrow indicates unbound DNA. The top arrow indicates RBPJ-bound DNA. **(B)** Top three TRANSFAC motifs from PBM with RBPJ. **(C)** FA of RBPJ on Hes1p (top) and Cp (bottom) DNA. **(D)** Co-precipitation of EBV latency proteins from LCL lysate with RBPJ. Input lane represents 3% the total reaction. Precipitated lane represents 1/3 total reaction.

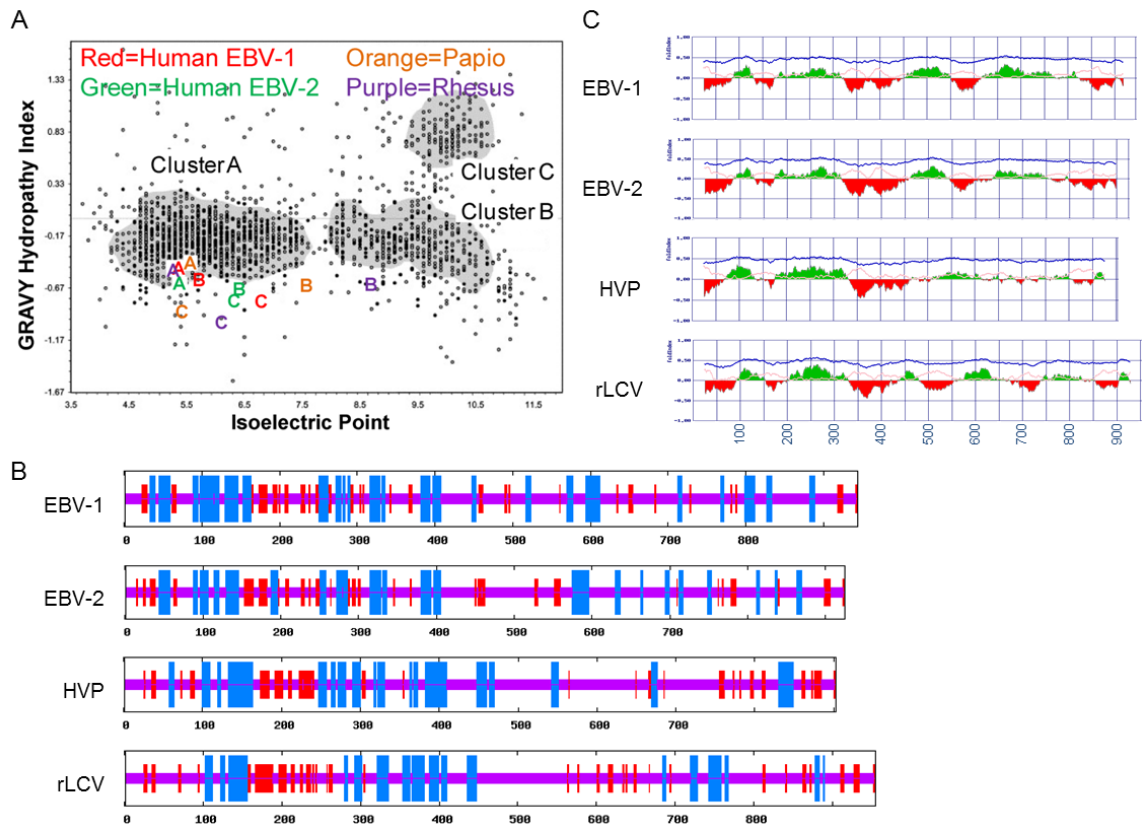
### 2.2.3 EBNA3 Bioinformatic Characterization and Construct Design for Structural and Biochemical Studies

While RBPJ has been characterized structurally in the context of the Notch transcriptional complex, no successful structural studies on the viral EBNA3 proteins have been done. We sought to solve a crystal structure of an EBNA3 protein bound to RBPJ. In order to increase the likelihood of success we analyzed each of the EBNA3A, -3B, and -3C proteins from the two types of human virus (EBV-1, EBV-2), the strain infecting rhesus macaques (rhesus lymphocryptovirus (rLCV)), and the strain infecting baboons (*herpesvirus papio* (HVP)) computationally for likelihood of solubility and crystallization success.

Plotting isoelectric point (pI) of a protein versus its grand average of hydropathicity (GRAVY) has been shown to predict the likelihood of a protein's ability to be solubilized and ultimately crystallized if it falls within one of three clusters on the plot (14). When plotting the twelve EBNA3 proteins in this manner, the EBNA3A proteins are located within a region that has a higher likelihood of crystallization success, and cluster together more so than the EBNA3B and -3C proteins (**Figure 9A**). Using EBNA3A as a starting point, we next looked at secondary structure and disorder predictions based on primary amino acid sequence (**Figure 9B, C**). While EBNA3A proteins across all four virus strains exhibited some predicted disorder within the region hypothesized to bind to RBPJ, this was to a lesser degree than the EBNA3B or -3C proteins. Based on these predictions, we selected EBNA3A from the baboon strain, HVP,

to pursue further. It is worth noting that the amino acid sequence of baboon RBPJ is identical to that of human. This also serves as evidence that the mechanisms of binding between various strains of virus are conserved, and hence we expect the biochemical interactions of HVP proteins to replicate those of EBV-1 or EBV-2 in humans.

We designed several constructs incorporating the EBNA3 homology domain and RBPJ binding region and tested for solubility with bacterial expression. The only construct that produced any soluble protein when expressed contained amino acids 72-365 from HVP. This construct, designated E3A-4, was used in all downstream experiments for EBNA3 protein expression, purification, and biochemistry. This construct includes the entire EBNA3hd, which contains the RBPJ binding regions, a nuclear localization signal, and a charged domain on the C-terminal end. Furthermore, this region is indispensable for RBPJ binding, transcriptional regulation at Cp, and maintenance of a latent infection and LCL growth.



**Figure 9: E3A Bioinformatic Characterization for Structural and Biochemical Construct Design.**

(A) pI vs. GRAVY plot of EBNA3 proteins across viral strains. EBNA annotations plotted with data from (14). (B) Gor IV secondary structure predictions for EBNA3A proteins across viral strains. Blue bars indicate predicted helical regions. Red bars indicate predicted  $\beta$ -strand regions. Purple line in absence of blue or red bars indicates predicted unstructured loops. (C) FoldIndex disorder predictions for EBNA3A proteins across viral strains. A positive (green) value predicts the region exists in an ordered fold. A negative (red) value predicts the region is disordered. Blue line indicates relative hydrophobic potential. Pink line indicates the potential for an ionic charge in the region.

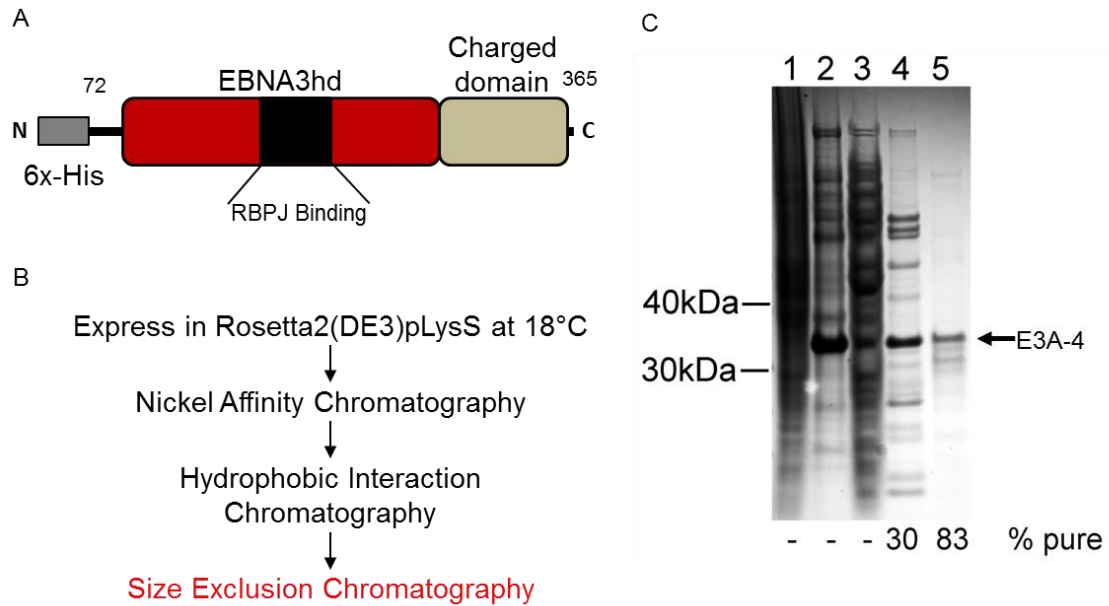
### 2.2.4 E3A-4 Expression and Purification Strategy

To date there have been no successful attempts to purify and crystallize any EBNA3 protein for structural studies. As such, we were required to empirically

determine a strategy for expression and purification for the E3A-4 construct (**Figure 10A, B**). Low temperature expression (18°C) was required in order to generate any soluble protein upon cell lysis, despite high total protein expression levels. Hence, the RBPJ protein expression conditions were utilized for E3A-4 as well. Several lysis conditions, including variations in pH, salt concentrations, and inclusion of detergents, glycerol, or chaotropic agents were tested, but none were able to increase the yield of soluble protein greater than ~15% of total protein expressed. While relative yield was low, this could be overcome by simply increasing the scale the expression, providing sufficient protein for downstream applications. Despite the variety of lysis conditions tested, the same minimal buffer system used for RBPJ lysis and purification performed the best with regard to soluble protein yield. This provided an optimal scenario as we could expect no complications or incompatibilities resulting from buffering conditions when performing biochemical experiments with E3A-4 and RBPJ.

Following cell lysis, soluble lysate was collected and, similar to RBPJ, was purified first with Ni-NTA agarose via its N-terminal 6x-His tag. After this first round of purification, rather than ion affinity, we sought to take advantage of the potential inherent hydrophobicity and “stickiness” of the E3A-4 protein that results in low levels of soluble protein during expression and lysis. For this reason we employed hydrophobic interaction chromatography (HIC). Binding to phenyl sepharose agarose beads in the presence of ammonium sulfate allowed us to further purify E3A-4 protein

to greater than 80% purity (**Fig. 10C**). This level of purity is insufficient for crystallography; however it is suitable for other useful applications such as EMSA, FA, and co-precipitation.



**Figure 10: E3A-4 Expression and Construct Design.**

(A) E3A-4 expression construct encompassing amino acids 72-365 from HVP. (B) E3A-4 expression and purification strategy. Size exclusion chromatography has not yet been successful. (C) Coomassie stained gel of E3A-4 purification steps. 1=uninduced lysate; 2=induced lysate; 3=soluble lysate; 4=nickel affinity elution; 5=HIC elution.

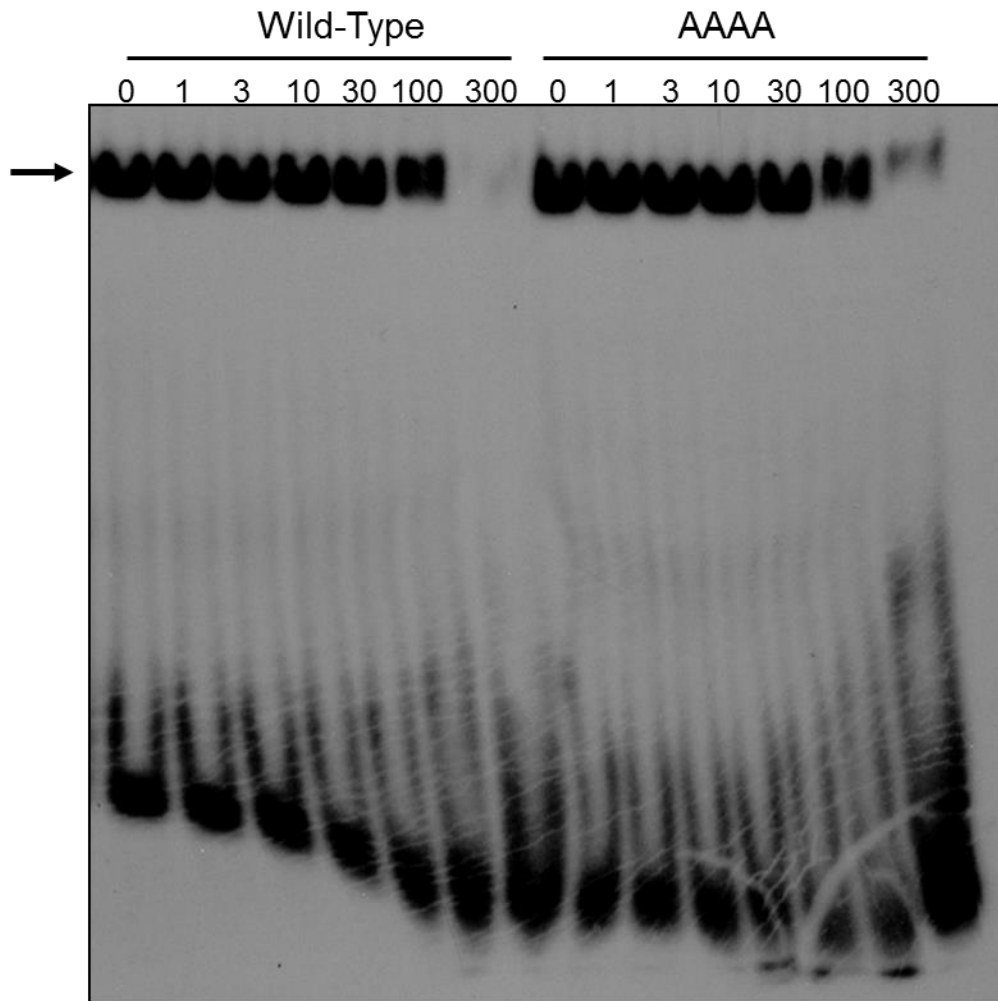
### 2.2.5 Bacterially Expressed and Purified E3A-4 Protein Disrupts RBPJ-DNA Interactions in EMSA

In the absence of active Notch signaling and EBV latency protein expression RBPJ is constitutively bound to DNA containing the RBPJ recognition motif (27, 51). It has been clearly demonstrated that active Notch signaling and NotchIC binding does not alter the DNA binding preferences of RBPJ (24). The role of RBPJ's interactions with

EBV latency proteins in transcriptional regulation is indisputable; however several *in vitro* experiments including EMSA have demonstrated RBPJ loses DNA binding upon interactions with the EBNA3 proteins (144). This has given rise to a theory that the EBNA3 proteins function with RBPJ to disrupt interactions with DNA and potentially repress transcription by this mechanism. Regardless whether this is an artifact of *in vitro* assays or the EBNA3 proteins do in fact disassociate RBPJ from DNA during latent infection, this has become an accepted readout of EBNA3-RBPJ binding in EMSA.

We expressed and purified E3A-4 protein as describe in Section 2.2.4. In addition we similarly expressed and purified a mutant of E3A-4 in which the amino acids TLAC at positions 201-204 (199-202 in EBV-1 EBNA3A) were each mutated to alanine (called “AAA mutant”). This motif is highly conserved among all EBNA3 proteins across all virus strains, with the cysteine absolutely conserved in all proteins. This mutation has previously been shown to disrupt interactions with RBPJ for EBNA3A and EBNA3C from EBV-1 (74, 88). We used EMSA to interrogate the ability of our bacterially expressed E3A-4 proteins to bind RBPJ with loss of DNA binding as a readout (**Figure 11**). Wild-type E3A-4 was able to disrupt RBPJ interactions with Cp DNA, albeit at relatively high molar ratios (>30:1 E3A-4:RBPJ). The AAA mutant was only able to partially disrupt DNA binding at the highest molar concentrations tested (300:1). At this molar ratio wild-type E3A-4 could completely disrupt RBPJ-DNA interactions,

indicating the ability of the AAAA mutant to bind RBP is reduced relative to wild-type between 3- and 10-fold.



**Figure 11: Bacterially expressed E3A-4 protein disrupts RBPJ-DNA Interactions in EMSA.**

EMSA of RBPJ bound to Cp DNA with wild-type or AAAA E3A-4 protein. Arrow indicates RBPJ-Cp complex. Values at top indicate E3A-4:RBPJ molar ratios.

Here we observed yet reproducible reduction in affinity for RBPJ by the mutant E3A-4. While the difference is not large, we can clearly identify it using loss of DNA

binding as a proxy. This assay requires large amounts of E3A-4 protein, yet it is sufficient to detect differences in binding affinities between constructs. This assay would be suitable to qualitatively interrogate the ability of EBV latency proteins to bind RBPJ, particularly from the perspective of RBPJ mutants deficient for particular interactions due to the requirement of high quantities of E3A-4 protein. The high protein requirement and difficulty of E3A-4 purification would most likely be better suited for screening wild-type E3A-4 against multiple RBPJ mutants rather than the inverse.

### **2.2.6 RAMANK and E2-ANK Expression and Purification Strategy**

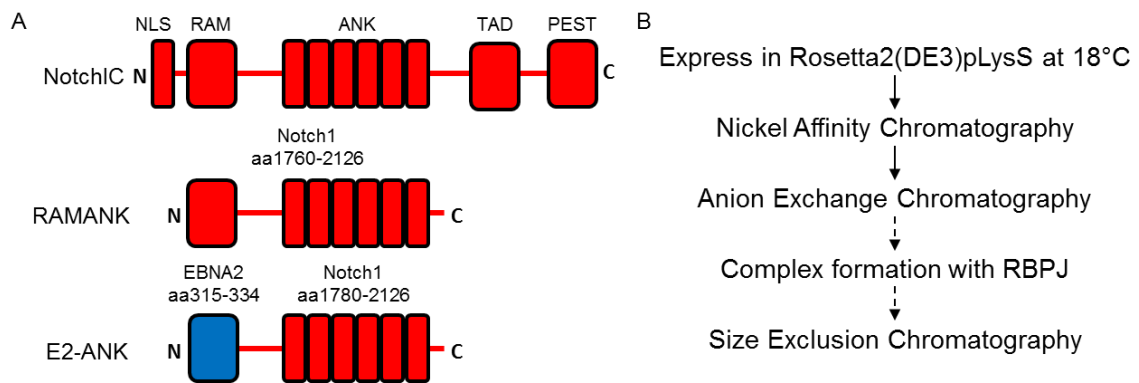
Bacterial expression of EBNA2 protein has proven to be inviable due to the inability to recover soluble protein, regardless of construct design or expression and lysis conditions. This may be due to required post-translational modifications unavailable during bacterial expression. Additionally, the EBNA2 protein may be inherently unstable due to unique features, such as a stretch where forty-three out of forty-six amino acids are proline. In light of this, an alternative mechanism for EBNA2 expression and purification for crystallography is needed.

The Blacklow Laboratory has used a construct of NotchIC in their crystallography studies incorporating the RAM peptide, the Ankyrin repeats, and the linker connecting these two regions called "RAMANK" (**Figure 12A**) (100). Their studies, among others, have demonstrated that the RAM peptide of this construct is critical for binding RBPJ, while the Ankyrin repeats play an accessory role. It has also

been shown previously that the region of EBNA2 responsible for RBPJ binding has a substantially reduced affinity for RBPJ than that of the RAM peptide (61). For these reasons, we designed a chimeric protein where the RAM peptide of the RAMANK construct is replaced with the corresponding region of EBNA2 (**Figure 12A**). By including the Ankyrin repeats of NotchIC, which bind to the CTD of RBPJ with low affinity, E2-ANK can compensate for the inherently lower affinity of EBNA2 for RBPJ. Even small amounts of ANK binding to RBPJ will effectively increase the local concentration of the EBNA2 portion of the chimera, increasing the likelihood of full complex formation. As EBNA2 and RBPJ share no homology outside of their  $\Phi W\Phi P$  motifs, differences in the EBNA2 portion of the complex could reveal binding mechanisms unique from RAMANK. A similar mechanism has been demonstrated for SMRT/MINT, which can bind RBPJ BTD or CTD individually with low affinity, but together form a high affinity interaction (141).

RAMANK can be expressed and purified under similar conditions to RBPJ from *E. coli* bacteria (**Figure 12B**). Expression was carried out at 18°C and lysis was performed under identical conditions to RBPJ. Lysate was then cleared and soluble RAMANK protein was purified first by Ni-NTA affinity chromatography via its 6x-His tag. Recovered protein was purified further by anion exchange chromatography, and finally by size exclusion chromatography. This yielded highly pure protein. E2-ANK protein was expressed and purified in an identical manner, and yielded similar results.

Interestingly, size exclusion chromatography of RAMANK or E2-ANK in solution revealed a tendency for the proteins to elute from the column at a size representative of a homodimer. Incidentally, NotchIC has shown the capacity to dimerize via its Ankyrin repeats, lending support to the possibility of a dimer in our purifications (6). Here we are able to generate RAMANK and E2-ANK protein to sufficient quantity and purity for x-ray crystallography studies.



**Figure 12: RAMANK and E2-ANK purification strategy.**

**(A)** Schematic of ICN1 with RAMANK and E2-ANK chimeric protein for expression and purification. RAMANK includes only the RAM and ANK domains of NotchIC and excludes the NLS, TAD, and PEST domains. The 20 amino acids of RAM are replaced with the corresponding 20 amino acids of EBNA2 in the E2-ANK construct while the remainder of the RAMANK protein is unchanged. **(B)** Expression and purification strategy of RAMANK and E2-ANK constructs. Protein can be used for biochemical assays following anion exchange chromatography, or it may be used to form a complex with RBPJ and undergo final purification by size exclusion chromatography for crystallography.

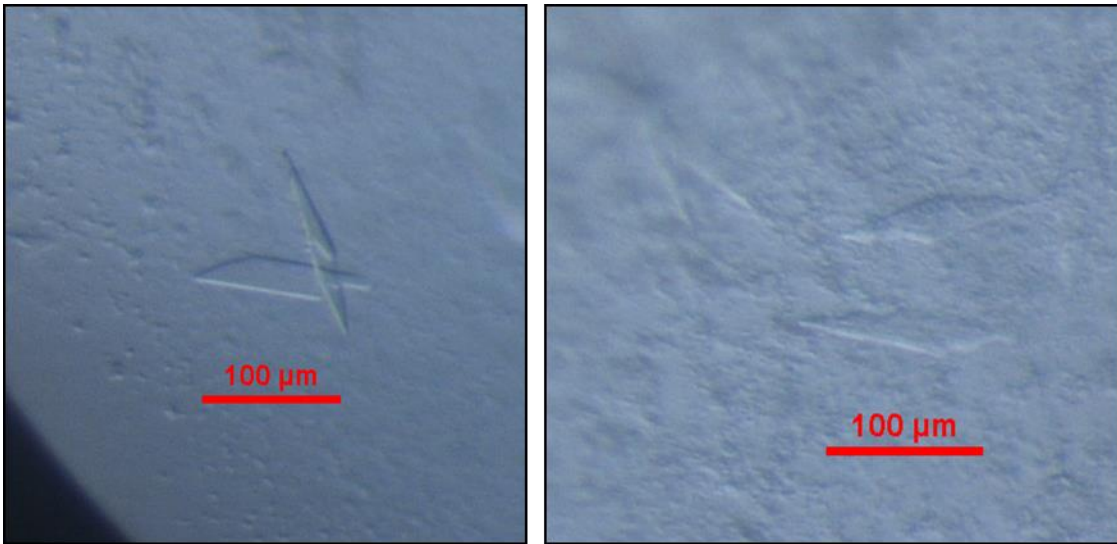
### 2.2.7 Crystallography Trials of RBPJ Bound to Cp DNA

One approach to solving a crystal structure of EBV latency proteins bound to RBPJ is to form the protein complex prior to crystallization trials and to co-crystallize the

complex. While this remains a goal, we have not yet had sufficient quantities of protein at high enough purity to perform these experiments. We have a second approach, based on pre-forming crystals of RBPJ bound to DNA and subsequently soaking peptides of the EBV proteins into these pre-formed crystals and crystallant solution. By including only peptides of the relevant portions of the proteins, we may identify important aspects of these interactions. Following purification of RBPJ we bound it to Cp DNA in a 1:1 molar ratio and isolated the complex by size exclusion chromatography.

RBPJ has been crystallized while bound to Hes1p DNA, a canonical RBPJ target, in the absence of NotchIC (PDB ID: 3BRG (39)). The lattice of this crystal structure reveals sufficient solvent channels to allow access for EBV latency protein peptides to their hypothesized binding sites on RBPJ. We first attempted to replicate the crystallization conditions from this structure with two key differences. First, we used the human version of RBPJ, whereas the 3BRG structure was solved with RBPJ from mouse. There are three amino acid variants between human and mouse RBPJ used here, although they are either buried in the core of the protein or are subtle changes so we did not anticipate their involvement in crystal packing. Second, we used RBPJ bound to the viral Cp rather than the canonical Hes1p. There are no published structures of RBPJ bound to any DNA other than sites from the HES1 promoter. Unfortunately, these targeted crystallographic trials yielded no crystals.

We additionally performed traditional sparse matrix screening to identify new potential crystallization conditions for our RBPJ-Cp complex. This yielded a single condition that resulted in crystals. Several small (~100 $\mu$ m long) rhomboid crystals were found in the same well (**Figure 13**). No other conditions generated any crystals. This condition was not similar at all to that of the previously published RBPJ crystal structure. Crystals were harvested and screened for diffraction; however no crystals generated reflections that would be indicative of a well-ordered protein crystal. Further optimization of these conditions may yield diffraction quality crystals, or further screening may be required.



**Figure 13: Crystals of RBPJ bound to Cp DNA.**

Crystals were generated in 50% (v/v) MPD, 0.1M Tris pH8.5, and 0.2M  $\text{NH}_4\text{H}_2\text{PO}_4$  in 1:1 protein:crystallant ratio at 20°C by sitting drop vapor diffusion.

## **2.3 Materials and Methods**

**Plasmids and Protein Expression.** RBPJ constructs were originally provided by Stephen Blacklow. The construct used in expression incorporated amino acids 9-435 of RBPJ and included a C-terminal 6x-His tag. RBPJ was expressed from pSK-Duet vector. The RAMANK construct encompassed amino acids 1761-2127 of Notch1 and corresponded to the RAMANK constructs used in crystal structures solved by the Blacklow Laboratory. E3A-4 encompassed amino acids 72-365 from EBNA3A of *Herpesvirus papio*. It was cloned from cDNA generated from S594 cells (Fred Wang). These constructs were cloned into the pDEST17 vector by Gateway cloning (Thermo Fisher), using pDONR221 as a donor vector. The E2-ANK chimera was synthesized based on the coordinates of the RAMANK construct and was synthesized by Thermo Fisher.

In all cases protein expression was performed in Rosetta2(DE3)pLysS *E. coli* cells (71400; EMD Millipore). Cultures were grown at 37°C to  $OD_{600} \approx 0.6A$ , at which point they were moved to 18°C. Once cultures reached  $OD_{600} \approx 0.8A$ , they were induced for expression with 0.5mM IPTG (Bioline). Cells grew with induction for approximately 18h, at which point they were harvested by centrifugation.

**Bioinformatics.** EBNA3 expression construct design was done using various bioinformatic tools available online. General protein characteristics were calculated with the ExPASy ProtParam tool. Secondary structure predictions were generated by the

PRABI Gor IV secondary structure prediction method. Disorder predictions were generated by the FoldIndex Disorder Prediction Server.

**Protein Lysis and Purification.** All proteins were resuspended in the same lysis buffer containing 50mM Tris pH 8.0, 500mM NaCl, and 5mM imidazole in a 1:20 lysis buffer:culture volume ratio and frozen at -80°C. After thawing resuspended cells,  $\beta$ -mercaptoethanol to 5mM and the protease inhibitors aprotinin, leupeptin, pepstatin A (Sigma-Aldrich) and phenylmethylsulfonyl fluoride (Amresco) to concentrations of 1 $\mu$ g/ml, 1 $\mu$ g/ml, 5 $\mu$ g/ml and 1 $\mu$ M respectively were added. Cells were lysed on ice by sonication. Lysates were cleared by centrifugation.

Soluble lysate was incubated by batch-binding with HisPur Ni-NTA beads (Thermo Fisher) and purified by gravity flow chromatography. Beads were washed with lysis buffer, then wash buffers containing 50mM Tris pH8.8, 500mM NaCl, and 10 and 15mM imidazole. Protein was eluted with wash buffer containing 250mM imidazole.

RBPJ, RAMANK, and E2-ANK were further purified by anion exchange on a 1ml ResourceQ column (17117710; GE Healthcare Life Sciences). Protein was diluted with Buffer A (50mM Tris pH 8.0, 0M NaCl, 5mM DTT) to 100mM NaCl and passed through a 0.22 $\mu$ m filter. Protein was loaded onto the column, washed with loading buffer and eluted with a gradient of increasing concentrations of Buffer B (50mM Tris pH 8.0, 1M NaCl, 5mM DTT) on an ÄKTA FPLC system (Amersham Biosciences).

E3A-4 protein recovered from Ni-NTA chromatography was further purified by hydrophobic interactions chromatography (HIC). Protein was dialyzed into start buffer containing 20mM sodium phosphate and 1.2M ammonium sulfate. It was then passed through a 0.22µm filter and loaded onto a 1ml HiTrap Phenyl HP column (17-1351-01; GE Healthcare Life Sciences). Column was washed with decreasing concentrations of ammonium sulfate.

Some RBPJ protein was further purified for crystallography by size exclusion chromatography (SEC). Protein was first incubated for 30 min. at room temperature with annealed Cp DNA (5'-ACGCCGTGGGAAAA-3', 5'-ATTTTCCCACGGC-3') with two-base 5' overhangs on each end in 50mM Tris pH8.8, 50mM NaCl, and 1mM DTT. RBPJ-DNA complex was concentrated with Amicon spin concentrators (UFC803096, EMD Millipore) and Vivaspin 500 spin concentrators (28-9322-25, GE Healthcare Life Sciences). Protein was fractionated over a Superdex200 10/30 column (17-1088-01, Amersham Biosciences).

**Electrophoretic mobility shift assays.** Double stranded Cp DNA

(5'-GGAAACACGCCGTGGGAAAAATTTGGC-3',

5'-GCCAAATTTTTCCCACGGCGTGTTC-3') and mutated Cp DNA

(5'-GGAAACACGCCGTGGCTAAAAATTTGGG-3',

5'-CCCAAATTTTTAGCCACGGCGTGTTC-3') (Thermo Scientific) was incubated

with RBPJ at 30°C for 30 minutes. Reaction buffer included 20mM HEPES pH7.9, 60mM

potassium chloride, 5mM magnesium chloride, 10mM DTT, 0.2mg/ml BSA, 10% (v/v) glycerol, and 13µg/ml poly-dGdC. Samples were run on 10% TGE gels in TGE running buffer at 140V. Gels were stained and visualized with SYBR Gold (S11494; Life Technologies). Reactions including E3A-4 protein were done similarly using <sup>32</sup>P-labeled oligos.

**Protein binding microarrays.** Array for PBM contained all possible 8-mers (Agilent).

Array was blocked in 2% milk for 1 hour followed by washing in PBS with 0.1% Tween-20 and PBS with 0.01% TX-100. Protein was incubated with array at 100µM in binding buffer containing 2% milk in PBS, 0.2µg/µl BSA, 0.05µg/µl salmon sperm DNA, and 0.02% TX-100 for 1 hour. Following protein incubation, array was washed with 0.5% Tween-20 in PBS and 0.01% TX-100 in PBS. Array was incubated with Penta-His A488-conjugated antibody (35310; Qiagen) at a 1:20 dilution in 2% milk in PBS for 1 hour and then washed in 0.05% Tween-20 in PBS, and then in PBS. Finally array was scanned with a GenePix 4400A Scanner (Molecular Devices).

**Fluorescence anisotropy.** Cp (5'-ACGCCGTGGGAAAAA-3', 5'-TTTTTCCCACGGCGT-3') and Hes1p (5'-TTACTGTGGGAAAGAAA-3', 5'-TTTCTTCCCACAGTAA-3') DNAs were labeled on a single strand with 5' 6-FAM and annealed (IDT). Titrations were performed in 50mM NaCl and 50mM Tris pH8.8 at 25°C. Measurements were taken with a Panvera Beacon 2000.  $K_d$  was calculated with the following equation:

$$y = \frac{m_1 \cdot [\text{protein}]}{m_2 + [\text{protein}]} + m$$

where  $m_1$ =maximum mP;  $m_2$ = $K_d$ ;  $m_3$ =y-intercept.

**Crystallographic screening.** Traditional sparse matrix screening was used to screen for RBPJ-Cp crystallization conditions. Qiagen screens were dispensed with a Phoenix RE liquid handling robot (Art Robbins Instruments) and individual wells were set up with an Oryx4 (Douglas Instruments Ltd) in a 1:1 protein:crystallant ratio. Screens were incubated at 20°C and tracked with Minstrel HT-UV device and CrystalTrak software (Rigaku). Crystals were screened for diffraction with a Rigaku Micro-Max-007 generator and RAXIS IV++ image plate detector with copper  $K\alpha$  radiation at a wavelength of 1.54Å. Crystals were generated from condition A4 from NeXtal JCSG Core Suite 1 screen (Qiagen).

## **2.4 Discussion**

Here we have bacterially expressed and purified RBPJ, EBNA3A, and RAMANK and its chimeric analog E2-ANK protein. These proteins can serve as substrates for biochemical experiments to characterize and interrogate protein-protein interactions and for X-ray crystallography for structural determination. In each case the proteins were initially purified by Ni-NTA affinity chromatography via their 6x-His tag, and subsequently by either anion exchange chromatography or hydrophobic interaction

chromatography. In the case of RBPJ, it has been further purified to homogeneity by size exclusion chromatography.

As RBPJ forms the intersection of latent viral infection and a normal host signaling pathway in these studies, we confirmed the functionality of the bacterially expressed and purified RBPJ protein. We found that it can bind host and viral canonical promoter DNA elements. Furthermore, as measured by FA it binds Hes1p DNA with an affinity comparable to that determined previously by ITC. Additionally, in an unbiased protein binding microarray, RBPJ bound preferentially to DNA 8-mers which contained a consensus sequence, 5'-GTGGGAA-3', a canonical RBPJ recognition motif. Finally, it was capable of precipitating the EBNA2, EBNA3A, and EBNA3C proteins from LCL lysate. The sum of all of these data indicates that our bacterially expressed and purified protein is completely functional with regard to DNA binding and interactions with viral latency proteins, even in the absence of eukaryotic chaperones or post-translational modifications. Furthermore, it confirms that post-translational modifications—at least from the perspective of RBPJ—are not required for the interaction with EBNA proteins and therefore bacterial expression is sufficient for downstream biochemical experiments.

We additionally sought to express and purify EBNA3 protein from bacteria as well. There is an absence of clear success in this area in the literature, so we sought a completely objective procedure for designing, expressing, and purifying EBNA3 constructs. Using bioinformatic techniques and the literature to guide selection and

construct design of a stable, soluble, RBPJ-binding domain, we ultimately employed a portion of EBNA3A from the baboon virus HVP, called E3A-4. While only a small fraction of the total expressed protein is soluble, E3A-4 was unmatched in this regard by any other construct or lysis strategy we tested. Conveniently, our best results were found using comparable expression and lysis conditions as RBPJ, allowing us the benefit of ensuring compatible buffering systems for future experiments incorporating the two proteins.

Furthermore, we demonstrated that despite only being able to recover a small percentage of the expressed E3A-4 protein, this protein was capable of binding RBPJ in EMSA experiments, while a mutant shown to disrupt RBPJ binding demonstrated a mildly weaker affinity. The readout in these assays is loss of binding to DNA by RBPJ. Several previous publications have demonstrated that *in vitro* binding of EBNA3 proteins to RBPJ induces loss of DNA binding, despite evidence that RBPJ and the EBNA3 proteins co-localize at DNA during latent infection. CHIP data demonstrates that RBPJ and EBNA3 proteins can be found at DNA together (59, 62), and the simplest mechanism is that this can occur through RBPJ binding DNA and an EBNA3 protein simultaneously. Additionally, a primary function of EBNA3C and EBNA3A is to suppress expression of the tumor suppressor gene p16 by means of polycomb repressive complex recruitment to the promoter. This transcriptional repression is dependent on

RBPJ binding (90), suggesting that RBPJ is most likely playing some role associated with DNA binding.

While DNA binding by RBPJ is most likely involved in at least some fraction of EBNA3-mediated transcriptional regulation, this does not necessarily have to occur directly at the promoter of the gene being regulated. As transcriptional activators can act at enhancer elements to promote DNA looping, EBNA3 proteins may do something similar. The EBNA3 proteins are capable of forming complexes with a number of other DNA-binding proteins (4). These complexes can form simultaneously with RBPJ binding, allowing the EBNA3 proteins to act as a bridge between two distinct DNA binding proteins. Indeed, there is evidence that RBPJ can be found more distal to particular DNA sites relative to EBNA3C (59), suggesting that it is bridging at least one more DNA binding protein with RBPJ. Understanding the mechanism of interaction between EBNA3 proteins and RBPJ could inform on the potential other interactions involved in complexes such as these. An interesting experiment would be to disrupt EBNA3-RBPJ binding and observe how this affects, first transcriptional regulation at these bridged sites and, second, the composition of the larger complex in the absence of RBPJ. This would provide valuable insight into the larger role of the EBNA3 proteins in transcriptional regulation in latent infections.

Bacterial expression of EBNA2 faced similar challenges to EBNA3 in that no soluble protein was able to be recovered. As such, we adopted an alternative strategy to

interrogate its interaction with RBPJ. Additionally, the affinity of EBNA2 for RBPJ is substantially lower than that of NotchIC. Hence we sought to take advantage of the low affinity of the ANK domain of NotchIC for RBPJ to increase the overall affinity of an EBNA2 construct to RBPJ. RAMANK has been successfully expressed, purified, and crystallized in complex with RBPJ, so we replaced the NotchIC RAM domain with the corresponding EBNA2 amino acids to create the E2-ANK chimera, the rationale being that we could observe genuine EBNA2 interactions with RBPJ assisted by ANK functioning to increase the local concentration of EBNA2 at RBPJ. In the RAMANK-RBPJ crystal structures, RAM is connected to ANK via an unstructured loop, suggesting that covalent attachment to ANK should not substantially constrain the EBNA2 portion. This would allow us, in a crystal structure, to observe a mostly unconstrained EBNA2 peptide binding RBPJ.

Furthermore, as RAMANK-RBPJ structures have previously been solved, we could take advantage of those crystallization conditions to attempt to replicate a similar crystal, substantially reducing the time, reagents, and effort involved in this rate limiting step. Additionally, as the twenty amino acids difference in our potential structure would make up a very minor proportion of the overall structure, molecular replacement for phasing may be possible. This would further reduce time, reagent, and effort requirements associated with anomalous dispersion techniques. While *de novo* crystallization condition determination and anomalous dispersion are viable

alternatives, the E2-ANK construct provides us an attractive starting point for crystallographic studies.

In addition to the E2-ANK approach, we are also pursuing a small peptide approach. As EBNA2 binding is mediated by its  $\Phi W\Phi P$  motif interacting with the hydrophobic pocket of RBPJ, EBNA2 peptides centered around this motif can still provide much of the same useful information that a larger protein might. Outside of the ANK domain, the interaction of NotchIC with RBPJ is limited to the one dozen or so amino acids of RAM. It would follow that the equivalent EBNA2 interactions might utilize a similar pattern where only a handful of residues could inform us of the binding mechanism. As the conservation between EBNA2 and Notch is essentially limited to the  $\Phi W\Phi P$  motif, even a small peptide offers the opportunity to learn much about the consequences of this lack of conservation.

A small peptide approach can include two techniques. First, we attempted pre-forming RBPJ crystals and subsequently soaking in EBNA2 peptide at high concentrations. Based on previous structures of RBPJ(39), packing of the crystal lattice provides sufficiently large solvent channels to allow penetration of the exogenous peptide to the hydrophobic pocket. To this point we have been able to generate crystals of RBPJ bound to the viral Cp, however none of the crystals screened were of acceptable diffraction quality. Repeating these trials with the previously crystallized Hes1p DNA would be prudent. Second, we can attempt co-crystallization of RBPJ with EBNA2

peptides. This technique has been successfully employed previously with RAM peptide (38). This could provide an opportunity for replication of crystallization conditions, although the different peptide and slight differences in the RBPJ constructs used may render this approach moot. Regardless, traditional sparse matrix screening is an acceptable alternative to this.

Overall the expression and purification of EBNA proteins has proven troublesome, although not completely futile. E3A-4 protein can be purified to a reasonably high purity for biochemical assays, although the relative high quantity required for these assays predicated a low-throughput approach. The E2-ANK chimera, due to its high expression and relative ease of purification, may prove to be a valuable alternative to EBNA2 structural studies. While it is an artificial system, alternative approaches to expression may allow us to more easily produce purified protein. For example, bacterial expression would prevent any necessary post-translational modifications to proteins. Whether these modifications are required for proper folding or for interaction with RBPJ, a eukaryotic expression system could overcome this hurdle. A baculovirus expression system may be well-suited for this approach. Additionally, mammalian expression may be required for proper folding and function. In this case, we could employ a RBPJ-knockout cell line (84) to express viral proteins. The lack of RBPJ in this expression system would allow for non-interference during biochemical experiments and for expression of an EBNA construct suitable for crystallography.

### **3. Defining the Molecular Interface on RBPJ for the EBV Latency Proteins and Notch**

Transformation of B cells into LCLs and maintenance of proliferation and latent infection is critically dependent upon EBV latency proteins to regulate transcription through the host factor RBPJ(49, 60, 74, 88). As such, the protein-protein interactions between the latency proteins and RBPJ are also absolutely critical. The viral EBNA2 and EBNA3 proteins are all capable of binding RBPJ. With the exception of EBNA3B, which is dispensable during *in vitro* infection, these interactions with RBPJ are each essential for *in vitro* transformation of B cells and maintenance of latency(49, 74, 88). Despite Notch, EBNA2, and the EBNA3s all binding RBPJ in order to access DNA and to regulate transcription, they can vary widely in both their target genes as well as transcriptional effects.

#### **3.1 Rationale**

Available ChIP-Seq data reveals that EBNA2, EBNA3A, EBNA3B, and EBNA3C each co-localizes with RBPJ at hundreds of sites within the human genome (59, 92, 123, 159), as well as at sites within the viral genome. In particular, EBNA2 is predominantly found at RBPJ co-occupied sites, with over 80% of all EBNA2 sites also containing RBPJ(159). Conversely, a minority of EBNA3A, -3B, and -3C sites also contain RBPJ. In addition, a substantial portion of the RBPJ co-occupied sites among the EBNA3 proteins are not also shared with EBNA2, suggesting that despite utilizing the same host factor to

access DNA, EBNA2 and EBNA3 proteins demonstrate certain differences in the complexes they form with RBPJ. There are no obvious discernable similarities between EBNA2 and EBNA3 proteins with regard to their mechanism of binding to RBPJ.

This absence of a shared mechanism of binding is based on lack of sequence homology in the RBPJ-binding domains of the EBNA2 and EBNA3 proteins. Furthermore, it has been shown that binding of EBNA2 to RBPJ shares a conserved mechanism with Notch signaling, in which binding is predicated by a  $\Phi W\Phi P$  motif (where  $\Phi$  represents a hydrophobic amino acid) that directly interacts with a hydrophobic pocket on RBPJ. This motif is necessary for binding, as well as transcriptional activation (44, 61). Despite this shared motif, EBNA2 and Notch complexes with RBPJ exhibit several differences.

First, while some genes can be co-activated by either EBNA2 or Notch signaling with RBPJ, the lists of activated genes are not completely overlapping (70). Along these same lines, expression of intercellular Notch (NotchIC) protein can only partially replace and rescue EBNA2 expression in latently infected B cells (43). While some genes are similarly expressed during latent infection with NotchIC, expression of the viral LMP1 protein, one of the major EBV transforming factors, is repressed when B cells are infected with EBV and co-cultured with stromal cells expressing the Notch2 ligand DLL1. Furthermore, activation of Notch signaling via DLL1 ligand in LCLs also represses LMP1 expression (118)

Second, the additional factors associated with Notch-mediated transcription differ from those associated with EBNA2-mediated transcription. One major difference is the requirement for MAML-1 in the Notch-mediated pathway. Under circumstances of activated Notch signaling, the RAM peptide of NotchIC drives the thermodynamics of the protein-protein interaction, and additionally induces an allosteric conformational change in RBPJ (39). Whether EBNA2 similarly induces this allosteric change upon binding RBPJ is unclear, however the absence of this requirement suggests that the biochemical interactions with RBPJ are not identical to those of Notch.

Third, it has been shown previously in yeast two-hybrid experiments that there are several mutations that can be made on the surface of RBPJ that preferentially disrupt binding—and hence transcriptional activation—by either EBNA2 or Notch (40). This directly supports the hypothesis that EBNA2 does not perfectly mimic Notch interaction with RBPJ. This evidence, taken with the abovementioned, strongly suggests that the EBNA2 interaction with RBPJ differs in ways significant enough to potentially explain the disparities in Notch- versus EBNA2-mediated transcriptional activation.

While NotchIC and EBNA2 clearly share many similarities in their binding to RBPJ and transcriptional regulation, the same cannot be said for the EBNA3 proteins. The  $\Phi W\Phi P$  motif shared between EBNA2 and the RAM peptide of Notch is not common to the EBNA3 proteins. EBNA3A, -3B, and -3C share a homology domain (EBNA3hd) among each other which is responsible for, among other things, binding to RBPJ(115).

While a putative  $\Phi W \Phi P$  motif has been identified within the homology domain of EBNA3C, it is not conserved among EBNA3A or EBNA3B. Furthermore, it is unclear based on the current literature whether this motif follows a similar mechanism as that of EBNA2 and Notch, or if it is even involved in the interaction at all (13). While this could potentially be a mechanism unique to EBNA3C, its absence from EBNA3A and EBNA3B along with a lack of convincing evidence in favor suggest that it is not the critical component that is observed with EBNA2 and Notch.

Additionally, genes regulated by, as well as bound in the genome by EBNA2 and the EBNA3 proteins only partially overlap. This overlap becomes even less pronounced when considering only sites also bound by RBPJ in ChIP-Seq experiments (123). Some of these disparities may be explained by differences in additional factors recruited to the transcriptional complex altering DNA affinity/sequence specificity or forming composite recognition sequences. However, the lack of a conserved mechanism of protein-protein interaction with RBPJ suggests a more fundamental difference at work. Furthermore, both EBNA2 and the EBNA3 proteins utilize RBPJ to access DNA at the viral C promoter (Cp) during latent infection. Interaction with RBPJ at Cp by the EBNA2 and EBNA3 proteins (with the exception of EBNA3B) is absolutely critical for establishing and maintaining a Type III Latency state *in vitro*. For this reason, the ability to more comprehensively decipher the biochemical interactions with RBPJ is essential toward

developing a complete understanding of the roles of the EBNA2 and EBNA3 proteins during latent infection of B cells.

In this light, the literature gives us a clear understanding of the Notch-mediated transcriptional complex with several atomic resolution structures of portions of NotchIC bound to RBPJ, both with and without MAML-1, as well as in monomeric and dimeric complexes. Furthermore, there is structural data describing the interactions of other RBPJ-binding proteins. However, there is essentially no structural information regarding the EBNA2 and EBNA3 proteins, particularly concerning their interactions with RBPJ. While there is biochemical evidence to support a mechanism for EBNA2 binding that is, in part, conserved with Notch, there is very little data describing with much resolution the mechanism for EBNA3 binding. This is limited to experiments utilizing relatively large truncation and deletion mutants of the EBNA3 proteins, which only offers the resolution of a couple dozen amino acids involved in binding (74, 88). Conversely, data concerning the EBNA2 binding have the resolution of individual amino acids, albeit not to the extent to be able to map the interactions completely (61).

Even more neglected is the mapping of EBNA-RBPJ interactions from the perspective of RBPJ. The literature on the subject is dominated by addressing the contributions to binding from the perspective of the EBNA2 and EBNA3 proteins. Experiments focused on the contributions of RBPJ to the interactions are limited to the aforementioned yeast two-hybrid assay identifying a couple mutations that uniquely

disrupt EBNA2 versus Notch binding and a series of truncations of RBPJ with which *in vitro* binding assays demonstrated that the EBNA3 proteins are capable of binding the N-terminal and  $\beta$ -trefoil domains (NTD and BTD) of RBPJ, but not the C-terminal domain (CTD)(13).

The current state of knowledge regarding the biochemical basis of RBPJ interactions makes this area ripe for further study, particular from the perspective of RBPJ. Here I describe several sets of experiments to begin to address this gap in knowledge.

## **3.2 Results**

### **3.2.1 Bacterially Expressed RAMANK and E2-ANK Form a Complex with RBPJ and DNA**

For *in vitro* biochemical interaction experiments with RBPJ, RAMANK and E2-ANK protein constructs were expressed in bacteria to serve as surrogates for activated Notch signaling and EBNA2 expression respectively (see **Figures 7** and **12**). The expression and purification of these constructs are described in detail in section 2.2.6. The RAMANK construct is identical to that which was crystallized by the Blacklow laboratory and convincingly shown to functionally replicate NotchIC *in vitro*. The E2-ANK construct was originally designed for crystallographic purposes, however due to the inability to successfully express and purify soluble EBNA2 protein from bacteria, it is utilized here in *in vitro* interaction experiments as a stand-in for EBNA2 protein. As the interactions with RBPJ are critically dependent on the  $\Phi W\Phi P$  motifs in EBNA2 and

RAMANK plus the handful of amino acids flanking this sequence, replacement of the Notch RAM domain with the corresponding EBNA2 motif provides a suitable model for addressing questions concerning differences in their binding to RBPJ.

We used an electrophoretic mobility shift assay (EMSA) to interrogate the ability of the bacterially expressed RAMANK and E2-ANK constructs to interact with RBPJ while bound to a relevant DNA sequence, the viral Cp. RBPJ binding to Cp DNA induces an upward shift during gel electrophoresis. Indeed, addition of RAMANK or E2-ANK protein induces a supershift of this complex (**Figure 14A**), indicating formation of a RBPJ-RAMANK or RBPJ-E2-ANK complex on DNA. E2-ANK is required at higher concentrations to generate the same supershift as RAMANK. However EBNA2 has been shown to have a significantly lower affinity for RBPJ than Notch (61), and therefore this observation is not of great concern. This data supports the notion that bacterially expressed and purified RAMANK and E2-ANK proteins are functional in the ability to bind RBPJ and are suitable for *in vitro* biochemical interaction experiments.

In order to more precisely define and characterize the complexes of RAMANK and E2-ANK with RBPJ on DNA, we used fluorescence anisotropy (FA) to quantify the affinities of these interactions. We employed a modified FA approach called “Tandem FA” (**Figure 14B**). RBPJ binding to fluorescently labeled DNA is detected by an increase in polarization of the emitted light, just as in traditional FA experiments, and as shown in section 2.2.2 (**Figure 8C**). In Tandem FA experiments, an additional factor—in this

case RAMANK or E2-ANK—is added to the mixture with near-saturating amounts of RBPJ relative to DNA. Binding of the additional factor results in a further change in mass of the fluorophore-containing complex, and hence an additional change in polarization that can be detected. From this we can quantify relative affinities of RAMANK and E2-ANK for RBPJ.

We used a fluorescently labeled DNA corresponding to a canonical RBPJ binding site in the human genome, the HES1 promoter (Hes1p), in Tandem FA with RAMANK and E2-ANK binding to RBPJ (**Figure 14C**). Here both RAMANK and E2-ANK proteins were able to induce a further change in the polarization of the emitted fluorescent signal from the RBPJ-DNA complex. Based on these data, we have calculated a “ $K_d$  apparent” of ~79nM and 280nM for RAMANK and E2-ANK respectively for binding RBPJ in these conditions.

Due to the design of the experiment, we are only able to calculate a  $K_d$  apparent for relative comparisons rather than absolute affinities. Nevertheless, the  $K_d$  apparent of RAMANK is very similar to the  $K_d$  of the RAM peptide of Notch for RBPJ determined by the Barrick group by isothermal calorimetry titration (ITC) under comparable conditions (61), suggesting that our Tandem FA experiments serve to address affinities for RBPJ with reasonable accuracy.

Additionally, the relative affinity of E2-ANK for RBPJ is approximately 3.5-fold lower than that of RAMANK. This replicates the observations made from the EMSA

experiments where a higher molar ratio of E2-ANK was required to produce the same supershift observed with RAMANK. Taken together, this data indicates that bacterially expressed RAMANK and E2-ANK are capable of binding RBPJ on DNA and are suitable to further interrogate the biochemical interactions between these proteins.

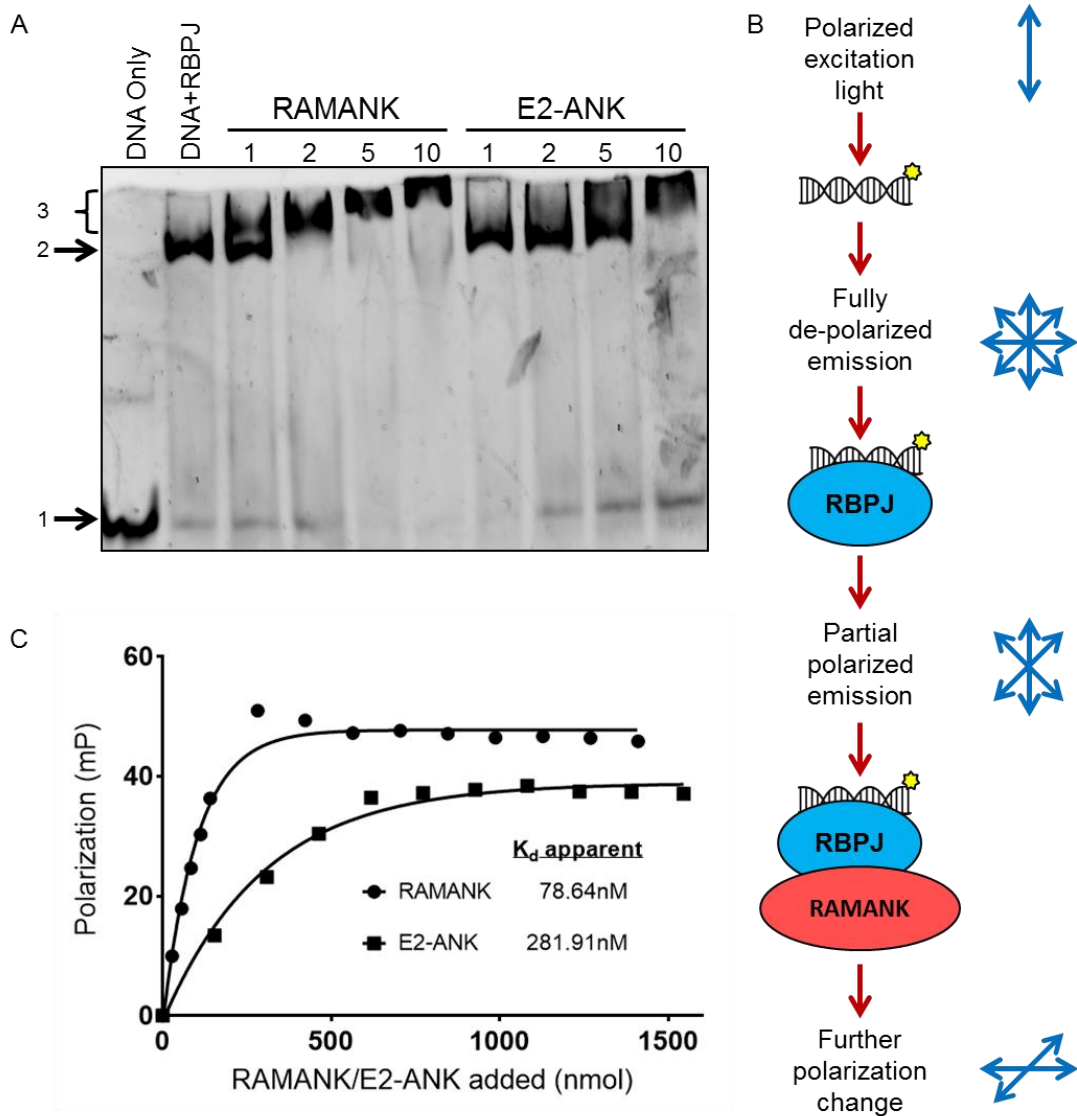


Figure 14: RAMANK and E2-ANK can form a complex with RBPJ in *in vitro* assays.

(A) EMSA of bacterially expressed and purified RAMANK and E2-ANK with RBPJ on Cp DNA. Values above lanes represent the molar ratio of RAMANK/E2-ANK-RBPJ protein. (B) Schematic of Tandem FA experimental procedure. Upon binding of RBPJ to a fluorescently labeled dsDNA oligonucleotide polarizes emitted fluorescent light (see **Figure 8C**). Addition of a subsequent molecule, in this case RAMANK (or E2-ANK), to RBPJ-DNA complex induces a subsequent, measurable change in the polarization of the emitted light. (C) Tandem FA of RAMANK and E2-ANK with RBPJ bound to Hes1p DNA.

### 3.2.2 Predicted Surfaces of RBPJ Involved in Interactions with Notch and EBV Latency Proteins

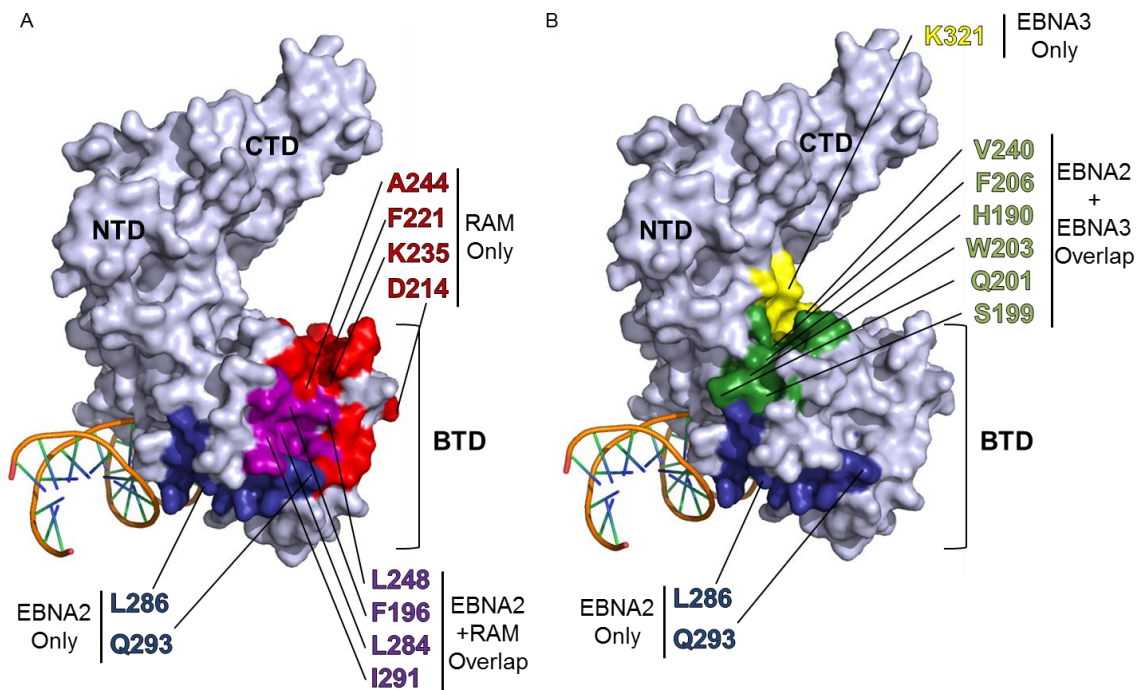
In order to make logical predictions about the surfaces on RBPJ which NotchIC and the EBNA2 and EBNA3 proteins utilize to bind, we operated under four primary assumptions based on biochemical, bioinformatic, genetic, and structural data in the literature:

First, EBNA2 and NotchIC, via their  $\Phi W\Phi P$  motifs, both bind within the hydrophobic pocket on the BTD of RBPJ, but each also utilize distinct, unique binding surfaces (40, 61). Second, due to high conservation within the EBNA3hd, which contains the region responsible for RBPJ binding, EBNA3A, -3B, and -3C all bind similar surfaces of RBPJ (115). Third, as EBNA3C has been shown to bind the BTD and NTD of RBPJ independently, the EBNA3 proteins bind both the BTD and NTD simultaneously (13). Fourth, since binding of any of the EBNA2 or EBNA3 proteins to RBPJ is mutually exclusive, EBNA2 and the EBNA3 proteins utilize surfaces on RBPJ that overlap at least partially.

Based on the literature and these assumptions we have constructed a model for the surfaces utilized by the NotchIC, EBNA2, and EBNA3 proteins (**Figure 15**). The

colored surfaces correspond to amino acids predicted to be involved in the various interactions mapped onto the crystal structure of RBPJ bound to RAMANK at Hes1p DNA (PDB ID 3V79 (16)). Unique NotchIC surfaces are shown in red and unique EBNA2 surfaces are shown in blue. The purple surface denotes where NotchIC and EBNA2 binding overlaps and is essentially limited to the hydrophobic pocket. Unique EBNA3 surfaces are shown in yellow. The green surface identifies that which is shared between EBNA2 and EBNA3 binding.

EBNA2 and NotchIC share the hydrophobic pocket of RBPJ. NotchIC-specific surfaces extend distal to the DNA binding surface of RBPJ based on the crystal structures that include the RAM peptide of Notch. The EBNA2-specific surfaces extend from the hydrophobic surface along a groove on the BTD toward DNA and turning towards the hinge connecting the BTD and NTD. Yeast two-hybrid experiments indicate mutations in this region disrupting EBNA2 binding, but not NotchIC binding (40). The EBNA3 binding surface includes the same hinge connecting the BTD and NTD as EBNA2, but extends further along it than the EBNA2 surface does, preserving the mutual exclusivity of EBNA2 and EBNA3 binding while still providing an EBNA3-specific surface. Based on this model, we have designed, constructed and expressed several single amino acid mutants of RBPJ in order to interrogate each of these surfaces, which are indicated on the proposed model.



**Figure 15: Surfaces of RBPJ involved in protein-protein interactions.**

Surface rendering of a crystal structure of RBPJ bound to Hes1p DNA (based on PDB ID: 3V79). Colored areas represent surfaces known (in the case of RAMANK) or predicted (in the cases of EBNA2 and EBNA3) to be involved in protein-protein interactions. **(A)** EBNA2 vs Notch RAM predicted binding surfaces. Red indicates the known unique Notch RAM binding surface. Blue indicates the predicted EBNA2 unique binding surface. Purple indicates the predicted shared hydrophobic pocket between RAM and EBNA2 binding. **(B)** EBNA2 vs EBNA3 predicted binding surfaces. Blue indicates the predicted EBNA2 unique binding surface. Yellow indicates the predicted unique EBNA3 binding surface. Green indicates the predicted shared binding surface between EBNA2 and EBNA3.

### 3.2.3 Most, but Not All RBPJ Mutants Retain Normal DNA Binding

The RBPJ single amino acid mutants were designed to disrupt binding of Notch or EBV latency protein binding. As such, all mutations were made on the surface of RBPJ and correspond only to the predicted binding sites of those proteins. While these mutations were not predicted *a priori* to disrupt normal protein folding or function of

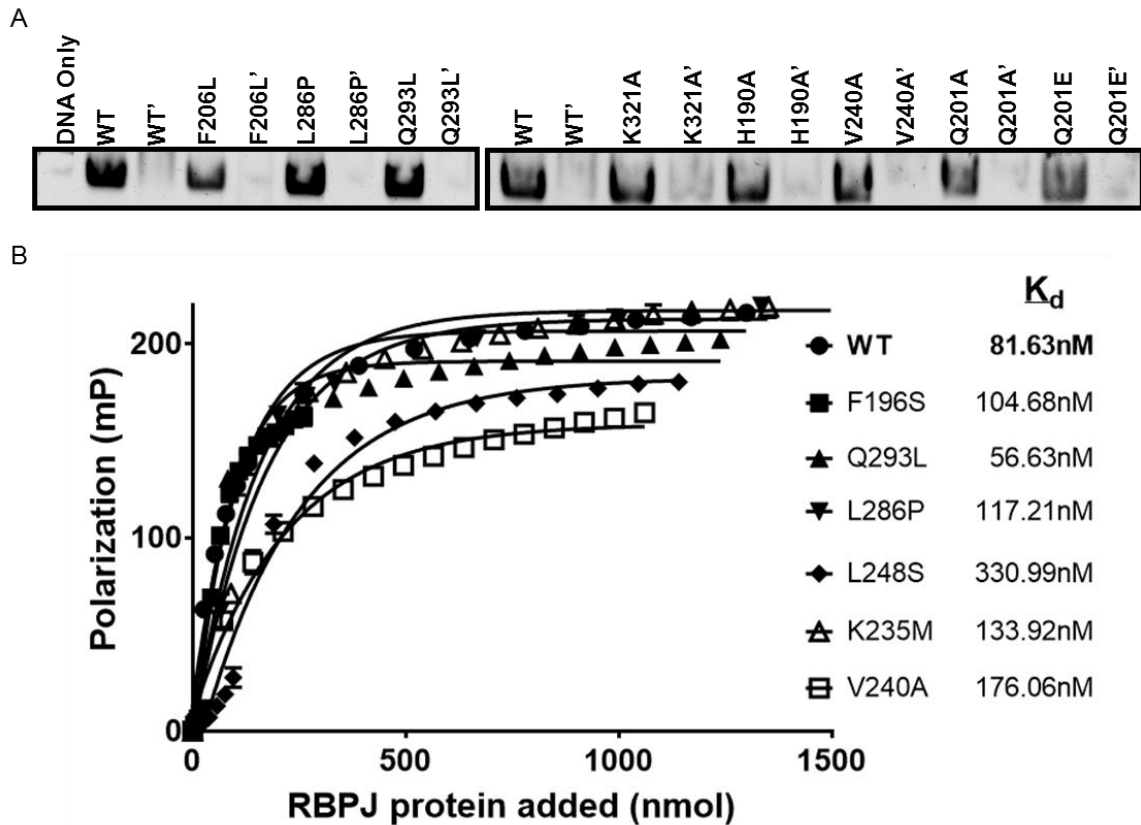
RBPJ (outside of the targeted protein-protein interactions), it is still necessary to determine this empirically as the Tandem FA experiments require DNA binding by RBPJ.

The RBPJ mutants were expressed and purified in the same manner as wild-type RBPJ protein (see section 2.2.1). Despite some mutants displaying variability in total expression levels or solubility in bacteria, the majority of mutants tested were able to be purified comparably to wild-type RBPJ. Following initial purification by Ni-NTA affinity, mutants were tested for their ability to bind Cp DNA in EMSA experiments (**Figure 16A**). Protein was incubated with DNA that contained either wild-type or a mutated RBPJ binding site in Cp. Wild-type and all mutants tested were capable of binding Cp DNA similarly by EMSA. Additionally, the sequence specificity of this binding was maintained as no constructs bound Cp DNA containing a mutated RBPJ binding sequence.

Several RBPJ mutants were purified further by anion exchange chromatography and concentrated before testing for DNA binding to Hes1p by FA. Of the constructs tested, wild-type, Q293L, K235M, L286P, L248S, F196S, and V240A were all capable of binding Cp DNA with comparable affinities (**Figure 16B**). All mutants had a  $K_d$  within ~2-fold of wild-type RBPJ and similar maximum polarization levels with the exception of L248S. This mutant bound Hes1p DNA with a  $K_d$  of ~4.0-fold weaker affinity than wild-type, although this is not a substantial enough deviation to discard from future

assays. Mild to moderate reductions in DNA affinity were not wholly unexpected among the RBPJ mutants and should not interfere with downstream applications.

Conversely, the H190A and W203A mutants exhibited signs of very weak and/or non-specific binding to DNA. Changes in polarization upon protein titration were significantly lower than wild-type and other mutants and advanced in a linear fashion. This linear progression of polarization is indicative of non-specific binding. H190A had previously been shown capable of sequence-specific DNA binding in EMSA, however the binding conditions between EMSA and FA have certain different requirements that could potentially result in this discrepancy. For example glycerol and non-specific DNA competitors required in EMSA but absent in FA could provide additional stability and sequence specificity to a mutant that is inherently less stable or properly folded. Regardless, these mutants could not be employed in further protein interaction experiments, but a majority of all mutants tested were still suitable for downstream applications.



**Figure 16: Most RBPJ mutants bind DNA normally.**

(A) EMSA of RBPJ mutants binding to Cp in a sequence dependent manner. ['] indicates Cp with the RBPJ binding site mutated. (B) FA of RBPJ mutants on Hes1p DNA.

### 3.2.4 E2-ANK Binding Can Be Differentiated from RAMANK Binding to RBPJ

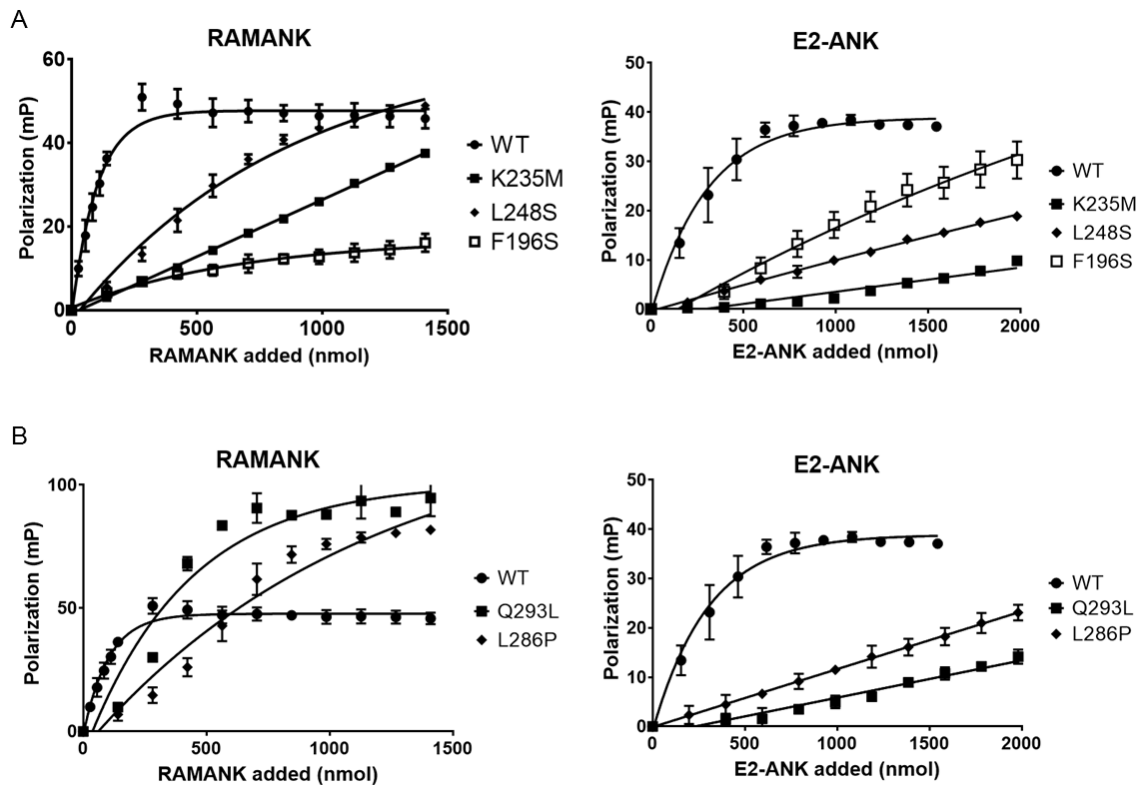
Because most RBPJ mutants we tested are capable of replicating DNA binding in FA, and RAMANK and E2-ANK constructs are adequate substrates for Tandem FA, we interrogated the ability of several of these mutants to bind either RAMANK or E2-ANK by Tandem FA (summarized in **Table 2**). In doing so we found that the F196S, K235M and L248S mutants were unable to bind RAMANK or E2-ANK with specificity (**Figure 17A**). Here any change in polarization upon RAMANK or E2-ANK titration appears to

be attributable to non-specific interactions due to the linearity of the change in polarization curves. This data supports the hypotheses present in **Figure 15** that residues F196 and L248, which are contained within the hydrophobic pocket of RBPJ, are directly involved in the interactions with RAMANK and E2-ANK, and subsequently EBNA2. Conversely the K235M mutant was originally hypothesized to specifically disrupt RAMANK and not E2-ANK/EBNA2 binding; however it disrupted binding of both constructs. While not contained within it, the proximity of K235 to the hydrophobic pocket may be relevant to its binding of the conserved  $\Phi W\Phi P$  motif.

Additionally we found that the L286P and Q293L mutants are both capable of binding RAMANK with specificity, but are unable to bind E2-ANK similarly (**Figure 17B**). In the case of each mutant, binding of RAMANK was reduced, however specific binding of E2-ANK was essentially undetectable. In these experiments, any change in polarization upon E2-ANK titration appears to be attributable to non-specific interactions. The Q293L and L286P data with RAMANK also displays a higher maximum polarization than observed with wild-type RBPJ, however a “leveling off” of the polarization effect can begin to be seen at higher concentrations of RAMANK. This suggests that there are still specific interactions occurring, but there may also be other non-specific interactions that are contributing to the readings.

This data indicates that though that while RAMANK binding is weakened by these mutations, residues L286 and Q293 are directly involved in E2-ANK binding, and

subsequently EBNA2 binding. All together this data begins to separate NotchIC binding from EBNA2 binding and yet clearly demonstrates the shared requirement for the hydrophobic pocket on RBPJ.



**Figure 17: Tandem FA of RBPJ mutants on Hes1p with RAMANK and E2-ANK.**

**(A)** RBPJ mutants around the hydrophobic pocket do not bind RAMANK (left) or E2-ANK (right) with specificity in Tandem FA. **(B)** RBPJ mutants that are capable of binding RAMANK (left) with specificity but not E2-ANK (right).

Table 2:  $K_d$  apparent from Tandem FA

<b>Mutation</b>	<b>RAMANK <math>K_d</math> apparent (nM)</b>	<b>E2-ANK <math>K_d</math> apparent (nM)</b>
WT	78.64 (4.30)	281.81 (76.02)
F196S	780.49*	N/A
K235M	N/A	N/A
L248S	1221.64*	8053.66*
L286P	2849.72 (795.18)	N/A
Q293L	597.79 (34.69)	N/A

\*Calculated  $K_d$  apparent resulting from predominantly non-specific interactions  
Standard error of the mean shown in parentheses

### 3.2.5 Disruption of Binding to RBPJ Directly Impairs Transcriptional Activation at Hes1p

The ability to bind RBPJ gives the EBNA2 and EBNA3 proteins access to DNA as they do not possess any direct DNA binding activity. This access to DNA allows EBNA2 to recruit CBP/p300 and the EBNA3s to recruit histone deacetylases (HDACs) and histone methyl transferases (HMTs) to both host and viral promoter and enhancer elements. Disruption of RBPJ binding has transcriptional consequences that can dictate the ability of EBV to transform B cells and maintain a latent infection. Therefore we sought to interrogate the physiological and transcriptional consequences of disruption of EBNA2 or NotchIC binding to RBPJ with our mutants.

We performed luciferase assays with the luciferase gene being driven by Hes1p with transfection of DNAs expressing wild-type or mutant RBPJ and NotchIC or EBNA2 (**Figure 18**). In order to eliminate effects from endogenous RBPJ we utilized the B-cell line DG75 which had been knocked out for RBPJ expression (SM224.9) (84). Wild-type RBPJ is not sufficient on its own to induce transcription, but addition of NotchIC or EBNA2 each induce luciferase expression. L286P and Q293L, two mutants shown by Tandem FA to disrupt E2-ANK binding more so than RAMANK, reduced luciferase expression by greater than 50% with EBNA2 transfection. NotchIC-mediated expression was reduced also, however to a lesser extent than that of EBNA2. This corroborates the Tandem FA data and supports the prediction that these amino acids preferentially disrupt EBNA2 function on RBPJ over NotchIC. It should be noted that these experiments were done with a single amount of EBNA2 or NotchIC DNA transfected into cells, but the Tandem FA experiments involved titration of at least ten different concentrations of E2-ANK or RAMANK protein. While the luciferase assays support the interpretations of the Tandem FA assays, it would be prudent to explore a wider range of E2-ANK and RAMANK protein expression with the luciferase readouts. This would serve to confirm that the data we are collecting is not being influenced by this particular protein concentration.

When transfected with the K235M or L248S mutants, both NotchIC- and EBNA2-mediated transcription were significantly reduced (>50%). The reduction in luciferase

expression was to similar extents between NotchIC and EBNA2 suggesting that the mutations are equally disruptive to their functions. This data corroborates the Tandem FA data and confirms the hypothesis that L248 is critical to NotchIC and EBNA2 binding and transcription. Further, the similar ability of K235M to disrupt transcriptional activation supports the observation from the Tandem FA experiments that this residue is not in fact unique to the interaction with NotchIC, but also contributes to the interaction with EBNA2. The K235M mutation is near the perimeter of the hydrophobic pocket, and therefore is not surprising that it could contribute to the binding of both NotchIC and EBNA2.

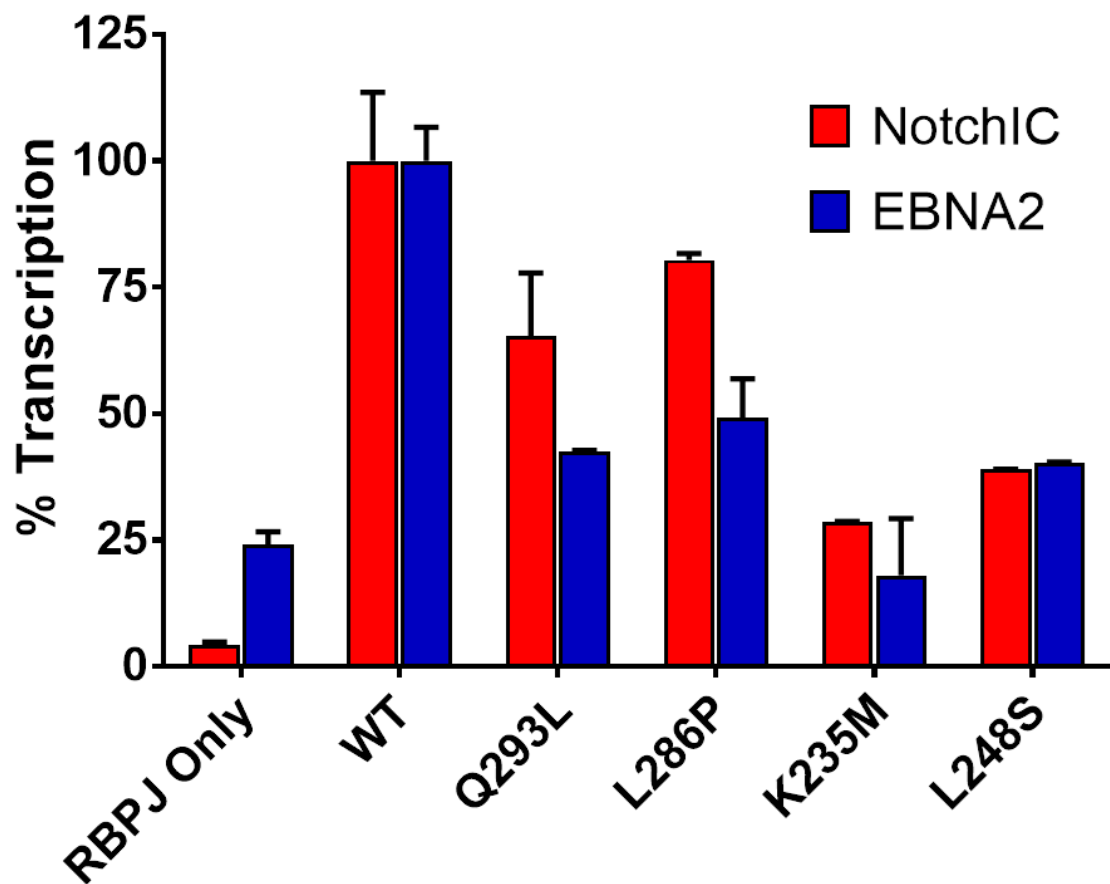


Figure 18: RBPJ mutant luciferase transcription.

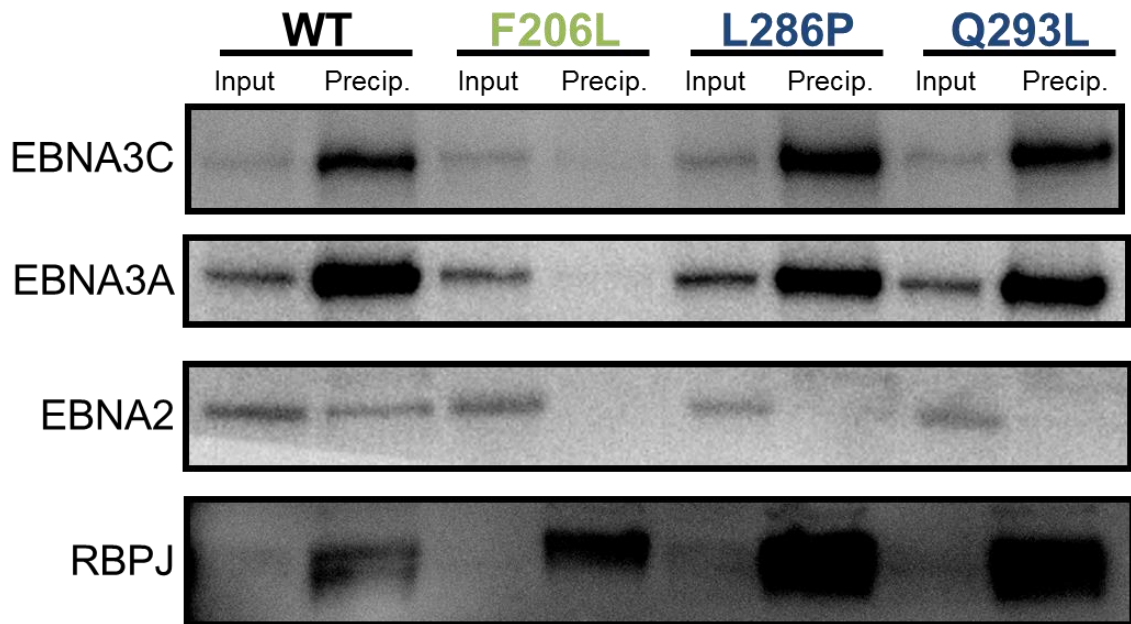
Luciferase assays of wild-type and mutant RBPJ with NotchIC or EBNA2 on Hes1p in RBPJ-deficient DG75 cells. Wild-type luciferase transcription is set at 100%.

### 3.2.6 EBNA2 and EBNA3 Proteins Utilize Both Overlapping and Unique Surfaces of Bind RBPJ

Due to the inability to produce sufficient quantities at high enough purity of EBNA3 protein to perform the Tandem FA experiments, we were required to pursue additional methods to query the ability of RBPJ mutants to bind the EBNA3 proteins. Hence we utilized an affinity-based co-precipitation technique by incubating 6x-His-

tagged RBPJ protein with LCL lysate and precipitating RBPJ with beads conjugated to Ni-NTA (**Figure 19**). Wild-type RBPJ is capable of precipitating EBNA2, EBNA3A, and EBNA3C from LCL lysate in this assay. Incubation with the F206L, L286P, and Q293L mutants however did not precipitate any detectable EBNA2 protein in each case. This supports the hypothesis that each of these residues is critical for EBNA2 interaction with RBPJ. In addition to its inability to precipitate EBNA2, F206L was similarly unable to precipitate both EBNA3A and EBNA3C. This data fits the model where this particular residue is involved in the interaction of RBPJ with both EBNA2 and the EBNA3 proteins.

Interestingly L286P and Q293L were equally capable of precipitating EBNA3A and EBNA3C compared to wild-type RBPJ while still lacking the ability to bind EBNA2. This supports the model that these residues are specific to interactions with EBNA2 only. Importantly, this represents the first instance of differentiating EBNA2 and EBNA3 binding from each other.



**Figure 19: RBPJ precipitation of EBV latency proteins.**

Precipitation of bacterially expressed and purified wild-type or mutant RBPJ protein by Ni-NTA with EBV latency proteins from LCL lysate. Input lanes represent 5% the total reaction. Precipitation lanes represent 1/3 the total reaction. Colors of mutants correspond to Fig. 15, where green represents an EBNA2/EBNA3 shared surface and blue represents an EBNA2 unique surface.

### **3.3 Materials and Methods**

**Mutagenesis.** Site-directed mutagenesis was performed according to QuikChange Site-Directed Mutagenesis Kit (Agilent). Successful mutants were screened by sequencing (Genewiz).

**Table 3: Primers Used**

<b>Mutant</b>	<b>Primer 1 (5'→3')</b>	<b>Primer 2 (5'→3')</b>
H190A	ACAGTTAGTACCAGATACTTGGCTGTAGAAGGA GGTAATTTTCA	TGAAAATTACCTCCTTCTACAGCCAAGTATCTGGT ACTAACTGT
F196S	GATACTTGCATGTAGAAGGAGGTAATTCTCATG CCAGTTCACAGCAGTGGGGA	CTCCCCACTGCTGTGAACTGGCATGAGAATTACCT CCTTCTACATGCAAGTATC
Q201A	AATTTTCATGCCAGTTCAGCGCAGTGGGGAGCCT TTTT	AAAAAGGCTCCCCACTGCGCTGAACTGGCATGAA AATT
Q201E	TTTCATGCCAGTTCAGAGCAGTGGGGAGCCT	AGGCTCCCCACTGCTCTGAACTGGCATGAAA
W203A	ATGCCAGTTCACAGCAGGCGGGAGCCTTTTTTAT TC	GAATAAAAAAGGCTCCCGCCTGCTGTGAACTGGC AT
K235M	GGCTACATCCATTATGGACAAACAGTCATGCTT GTGTGCTCAGTTACTGGCATGGCA	GTGCCATGCCAGTAACTGAGCACACAAGCATGAC TGTTTGTCCATAATGGATGTAGCC
V240A	AACTTGTGTGCTCAGCTACTGGCATGGCACT	AGTGCCATGCCAGTAGCTGAGCACACAAGTT
L248S	GCTCAGTTACTGGCATGGCACTCCAAGATCGA TAATTAGGAAAGTTGATAAGCAGACCCG	GCGGTCTGCTTATCAACTTTCCTAATTATCGATCTT GGGAGTGCCATGCCAGTAACTGAGC
L286P	ACAGAAAGAATGTATTTGTGCCCTTCTCAAGAA AGAATAATTCAA	TTGAATTATTCTTTCTTGAGAAGGGCACAAATACA TTCTTTCTGT
Q293L	TTCTCAAGAAAGAATAATTCTATTTTCAGGCCACT CCATGTC	GACATGGAGTGGCCTGAAATAGAATTATTCTTTCT TGAGAA
K321A	GGACAATCATTAGCACAGATGCGGCAGAGTATA CATTTTATG	CATAAAATGTATACTCTGCCGCATCTGTGTAATGA TTGTCC

**Electrophoretic Mobility Shift Assays.** EMSA experiments were performed as described in Section 2.3 above.

**Fluorescence Anisotropy.** FA experiments were performed as described in Section 2.3 above. For Tandem FA, RBPJ protein was added to near saturation, at 90% the calculated maximum mP.  $K_d$  apparent was calculated with the equation above, with the starting mP with RBPJ set to 0.

**Luciferase Assays.** SM224.9 cells (84) were transfected using a Digital Bio MicroPorator MP-100 system with one pulse at 1,400V for 30ms in a 100 $\mu$ l tip at a cell density of 10 million cells/ml using a Neon transfection system (Life Technologies). 1 $\mu$ g of SV-40 Renilla luciferase DNA was transfected with 3 $\mu$ g of Hes1p-Firefly luciferase DNA. 3 $\mu$ g each of RBPJ, NotchIC, and EBNA2 DNA were included in appropriate reactions. Transfected cells were grown in 1ml RPMI-1640 supplemented with 10% fetal bovine serum for 48 hours before harvesting. Luciferase signals were read on a Lumat LB 9507 luminometer (Berthold Technologies). Renilla luciferase was used to normalize for transfection efficiency.

**Co-precipitation and Western Blotting.** LCLs were washed in PBS and lysed with rotation at 4°C for 30 minutes in Pulldown Lysis Buffer (PLB), which contained 1% (v/v) NP-40, 40mM Tris-HCl pH7.5, 150mM sodium chloride, 10mM magnesium chloride and cComplete Protease Inhibitor Cocktail (Roche). Lysate was cleared by centrifugation at

4°C. LCL lysate was pre-cleared with rotation at 4°C for 1 hour in Ni-NTA agarose beads (Thermo Scientific) that were equilibrated into PLB. Pre-cleared lysate was combined with RBPJ protein and 50µl equilibrated Ni-NTA beads and incubated with rotation at 4°C for 2 hours. Beads were washed extensively in PLB and protein was eluted with PLB supplemented with 500mM imidazole.

Western Blot analysis was performed on eluted samples. All blotting was done in blocking buffer with 5% milk in TBST. Primary Antibodies:  $\alpha$ -His (RBPJ) (GE Healthcare) at 1:3000; EBNA2 (PE-2 hybridoma supernatant) at 1:100; EBNA3A (ExAlpha) at 1:500; EBNA3C (ExAlpha) at 1:500. Secondary antibodies:  $\alpha$ -mouse and  $\alpha$ -sheep (Sigma) at 1:3000.

### **3.4 Discussion**

The biochemical interactions of EBV latency proteins are critically important to the establishment and maintenance of a latent infection in LCLs. Currently the interaction of Notch with RBPJ is well established through an extensive literature encompassing genetic, biochemical and structural data. However this extensive knowledge does not also include the interactions of EBV latency proteins with RBPJ. Despite similarities between EBNA2 and NotchIC binding there are clearly differences, and the quantity of knowledge with regard to EBNA2 does not match that of Notch. While EBNA-specific disruption of RBPJ interactions can certainly be a therapeutic target, a more fundamental understanding of these interactions is a pressing need.

Here we have used bacterially expressed and purified protein to interrogate these differences between NotchIC and EBNA2 interactions with RBPJ. A surrogate EBNA2 construct, the E2-ANK chimera, served to functionally replace EBNA2 in these experiments; however it was sufficient to provide the desired insight. Despite only containing a small portion of EBNA2 protein sequence it was able to test our own hypotheses, as well as confirm some assumptions in the field, such as the requirement of the RBPJ hydrophobic pocket for binding. While several of the mutants tested here confirmed parts of the hypotheses presented in **Figure 15**, the K235M mutant was an exception. Originally predicted to disrupt specifically NotchIC binding, we found it disrupts both EBNA2 and NotchIC interactions. The transcriptional data was consistent with the binding data, lending confidence to these observations. The proximity of K235 to the shared hydrophobic pocket substantiates the plausibility of these results further.

While EMSA and Tandem FA experiments both indicated a lower affinity of E2-ANK for RBPJ than that of RAMANK, it has been demonstrated in the past that corresponding peptide from EBNA2 does indeed have a lower affinity for RBPJ than the RAM peptide. Additionally, the fact that only a small peptide differentiates E2-ANK from RAMANK in terms of amino acid sequence lends credibility to the theory that the critical similarities and differences between NotchIC and EBNA2 binding can be found here.

The E2-ANK construct does have limitations however. While the  $\Phi W\Phi P$  motif shared between EBNA2 and NotchIC clearly is the primary thermodynamic driver of the interaction, the absence of the rest of the EBNA2 protein prohibits us from identifying other potentially interesting interfaces but which may not be absolutely essential. Furthermore, the chimeric nature of the construct may introduce certain constraints on the EBNA2 peptide that would not allow it to faithfully replicate folding, movement, or more extensive interactions. Nevertheless, we were able to generate useful data to map the interacting surfaces on RBPJ.

With regard to the Tandem FA experiments, the curves generated from the data collected in some cases do not closely match what one would expect for a standard one-phase association profile for two molecules binding each other at a single site. The parameters of the experiment only allow for a bulk change in signal, and this must occur in a complex with the fluorescently labeled molecule—in this case the DNA substrate for RBPJ. As such, any non-specific interactions or interactions that do not conform to the standard one-phase association paradigm can contribute aberrantly to the observed change in polarization signal. There are several possibilities in the experiments presented here. First, non-specific interactions from the minor impurities of the protein mixes may become detectable in the presence of a mutant that has even mildly reduced specific affinity. Second, RAMANK and E2-ANK interacts with RBPJ at two distinct surfaces on RBPJ—at the hydrophobic pocket via the RAM/E2 peptide and at the CTD

via the Ankyrin repeats. The Ankyrin repeats have a substantially lower affinity for RBPJ than the RAM peptide, and the kinetics of this interaction can usually not be discerned independently in the context of RAMANK, however — as noted above — the weaker affinity for a mutant RBPJ or by the E2 peptide may allow this binding to be observed as part of the data. Third, RAMANK has displayed the ability to form homodimers mediated by its Ankyrin repeats. This could result in several different species of complexes, each potentially with a unique polarization property, making a quantitative interpretation of the data very complicated. RAMANK has been shown to exist as a monomer in solution, but studies have not been done at the high concentrations used in these experiments. Indeed, we have evidence from size exclusion chromatography experiments that RAMANK or E2-ANK, when applied to the column at high concentrations, elutes in at a volume indicative of a ~90kD protein as opposed to its predicted ~44kD size, suggesting homodimerization. Nevertheless, despite this Tandem FA not being a particularly quantitative assay, we can still clearly differentiate between some specific binding (even weakened binding) and complete absence of specific binding in the example of the Q293L and L286P mutants with RAMANK versus E2-ANK.

Furthermore, luciferase assays lend physiological support for the biochemical observation, connecting protein-protein interactions to transcriptional regulation. Despite the experimental design not allowing for calculation of absolute affinities, the

Tandem FA results reveal relative changes in the affinities of various RBPJ mutants for RAMANK and E2-ANK. Taken together with the luciferase assay data, we have begun to decipher the contribution of individual RBPJ residues to the formation of transcriptional complexes and transcriptional regulation.

With regard to the EBNA3 interactions with RBPJ we were limited in our ability to perform biochemical experiments as a result of the protein production and purification process. Additionally, the functional luciferase assay would be insufficient to completely map the EBNA3 interacting surfaces if these proteins do in fact overlap with the EBNA2 interacting surfaces. While the EBNA3 proteins are traditionally thought of as transcriptional repressors, the normal readout for their activity is repression of EBNA2-mediated transcriptional activation. For any RBPJ mutants that were deficient for EBNA2 binding we would be unable to query the ability of the EBNA3 proteins to repress this activation. Therefore, we have employed the co-precipitation assay to begin to address EBNA3 binding. While this is a qualitative rather than quantitative readout, we have already generated convincing data that the EBNA2 and EBNA3 proteins utilized both similar and unique surfaces on RBPJ.

It has been demonstrated that the EBNA2 and EBNA3 proteins bind in a mutually exclusive fashion to RBPJ, and hence a theory that arose is that they share a common mechanism of binding—a mechanism which, in light of the similarities of EBNA2 to Notch, would suggest a conserved mechanism among all of these proteins.

The ability of the L286P and Q293L mutants to separate EBNA2 from EBNA3 binding represents the first clear evidence that these mechanisms are not identical. Further elucidation of these mechanisms can reveal the extent of their similarities and differences, and may identify a mechanism of RBPJ binding not previously observed with any protein.

Despite the differences, the F206L mutant still disrupted binding to both EBNA2 and EBNA3 proteins. This residue is not involved at all in NotchIC binding, and forms the bottom of a narrow but deep hydrophobic pocket located in the groove between the NTD and BTD, distal from DNA. The literature has identified asymmetrically dimethylated arginine on EBNA2 as a mediator of the interaction with RBPJ *in vitro* and *in vivo* (45). EBNA2 contains an Arg-Gly repeat sequence that begins approximately fifteen amino acids past the  $\Phi W\Phi P$  motif, putting it in a position to access this F206 pocket if our model is correct. Asymmetrically dimethylated arginine is the only amino acid or modified amino acid that would be able to stably fit within this pocket and be long enough to contact the F206 residue at its bottom. Further investigation of this mechanism is warranted, particularly if this modification could be utilized as a potential therapeutic target (57).

One aspect of these interactions that was not addressed by these experiments is the potential for RBPJ binding partners to define transcriptional targets. As the lists of EBNA2-, EBNA3-, and Notch-mediated transcriptional targets only share partial

overlap, there must be other mechanisms at play. One possibility is the ability of the EBNA2 and/or EBNA3 proteins to directly alter the DNA binding specificity of RBPJ. This could be accomplished through conformational change upon binding. While NotchIC induces an allosteric change when bound, there are no changes in the DNA interacting residues and therefore no change in DNA sequence specificity or affinity.

This may not be the case for the EBNA2 or EBNA3 proteins however. Without structural data for these proteins, we cannot make truly educated hypotheses. It has been shown in *in vitro* experiments though that EBNA2 and EBNA3 binding induces RBPJ to be released from DNA. This may even be the case *in vivo*, but ChIP data still puts the EBV latency proteins and RBPJ in proximity to DNA, suggesting that this may be an artifact of these experiments, or that there are other factors contributing to the association with DNA. It is plausible, based on our proposed model, that EBNA2 could alter the conformation of the DNA binding surface of RBPJ as we hypothesize that EBNA2 binds adjacent to this surface. Further, the proximity of EBNA2 to the DNA binding surface may in fact sterically alter the ability of RBPJ to access DNA. Finally, EBNA2 has been shown to dimerize, and in fact contributes to its transcriptional activation potential (37). An EBNA2 dimer while bound to RBPJ has the potential to target two RBPJ molecules to DNA, in *cis* or *trans*, that could create composite recognition sequences. NotchIC has been shown to dimerize via its Ankyrin repeats to

coordinate two RBPJ molecules on DNA, so this could represent another conserved mechanism between EBNA2 and Notch signaling.

Finally, *in vivo* there are certainly other factors, both host and viral, that can participate in a larger complex with RBPJ and the EBNA2 and EBNA3 proteins. For example, EBNA3C has been shown to associate with RUNX, BATF, IRF4 and SPI1 in addition to RBPJ at DNA (4, 59). The participation of each of these proteins in complexes can certainly dictate alternative sequence specificity that creates a composite recognition site. Furthermore, this could in fact remove the necessity for RBPJ association with DNA, satisfying the hypothesis that EBV latency proteins disrupt RBPJ interactions with DNA. Likely some combination of viral latency protein binding to RBPJ and composite recognition sites contribute to the diversity of binding sites and targets of transcriptional regulation among Notch and the EBV latency proteins.

## **4. Epstein-Barr Virus Induces Global Changes in Cellular mRNA Isoform Usage That Are Important for the Maintenance of Latency**

*This chapter is based on a research article published by Nicholas J. Homa et al. in the Journal of Virology in 2013 ((50)).*

### **4.1 Rationale**

The levels of mRNAs in cells are regulated by transcription as well as co- and post-transcriptional events that can generate diverse mRNA isoforms. Alternative isoform usage provides a means to greatly increase the diversity of protein species within a cell as well as to regulate the localization, stability, and translation of mRNAs through altering recognition by RNA binding proteins (RBPs) and microRNAs (miRNAs) (65, 121). Recent genome-wide studies have demonstrated that 94% of all protein-coding genes generate multiple mRNA transcripts (145). Indeed, a large number of human genetic diseases are thought to be caused by aberrant mRNA splicing (33). Furthermore, alternative mRNA processing that functionally alters protein output has been implicated in virus infections and cancer (15, 97, 142). Therefore, it is likely that an oncogenic virus such as EBV will dramatically alter not only cellular mRNA abundance but also cellular mRNA isoform usage during B-cell infection, and this may have important consequences for disease progression.

Latent infection established in LCLs strongly represses lytic virus replication. The regulation of the lytic cycle is primarily enacted through the promoter of the major

lytic transactivator BZLF1, or Z (68). The Z promoter receives inputs, including stressors, such as DNA damage and phorbol ester treatment, signaling through the B-cell receptor (BCR), and cues downstream of plasma cell differentiation (8, 46, 130). The critical *cis*-acting elements in the Z promoter have been well characterized, as have the *trans*-acting factors that promote Z transcription (68). Histone acetylation also plays an important role in the regulation of Z, and the balance between recruitment of histone acetyltransferases versus histone deacetylases by these *trans*-acting factors promotes the switch between latency and lytic reactivation (58). The downstream cascade of viral gene products activated by Z, including the transcription factor BRLF1, or R, and the initiation of viral DNA replication, capsid assembly, and virion maturation follow the rubric established in other lytic herpesvirus infections (95). The distinction in the setting of EBV is that B-cell infection is typically tightly latent, and therefore, Z-promoting factors in these cells must be kept in check to prevent triggering of the lytic infection cycle.

Here we used the SplicerEX algorithm to investigate changes in host mRNAs regulated by EBV infection of primary B cells upon transformation into LCLs at the level of both total mRNA abundance and mRNA isoform usage. The role of alternative isoform choice in the regulation of lytic reactivation has yet to be explored. Ultimately we identified two genes, XBP1 and TCF4, whose alternative isoform usage plays and

important role in the regulation of the switch between latent and lytic replication of the virus through transcriptional regulation of the Z promoter.

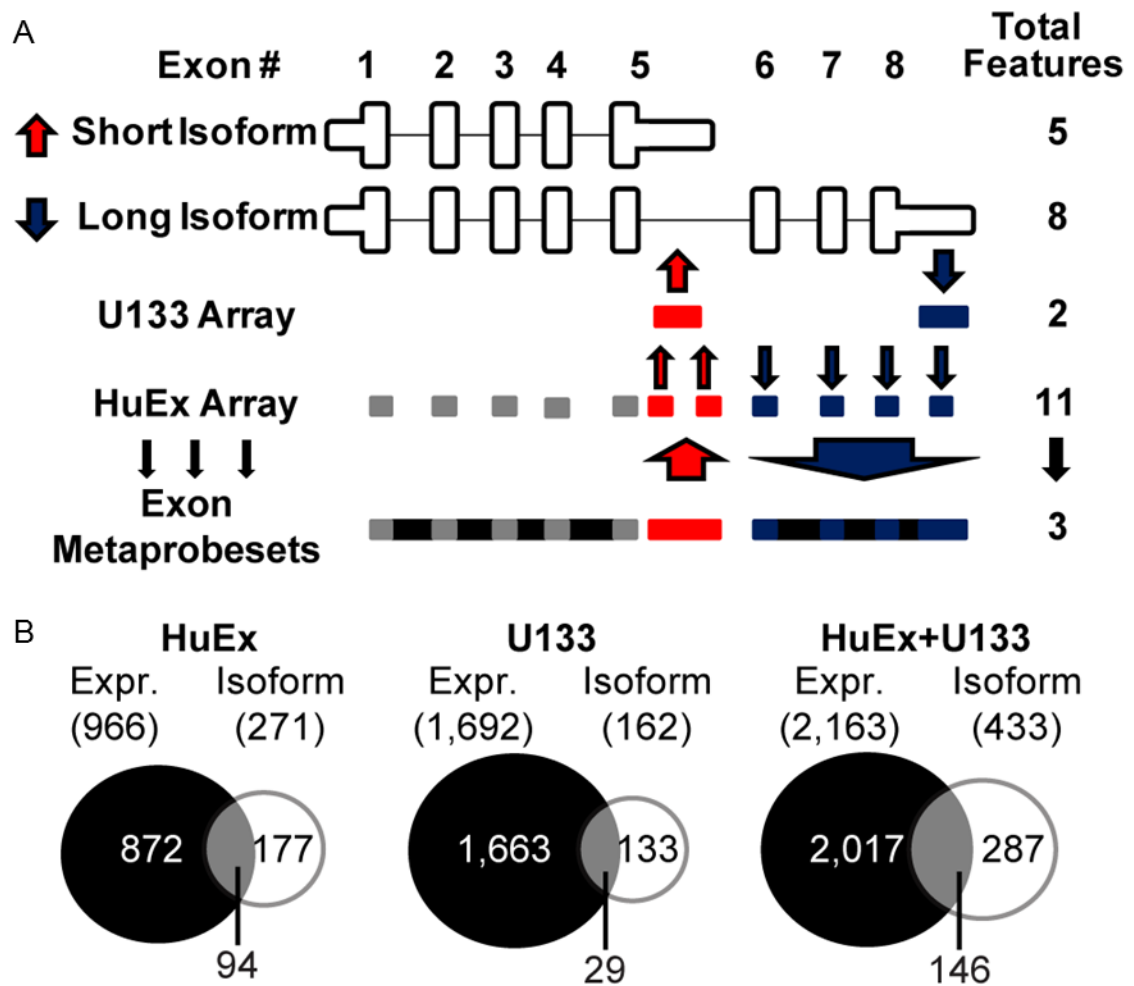
## **4.2 Results**

### **4.2.1 EBV Infection of Primary B cells Induces Changes in Alternative mRNA Exon Usage**

Infection of primary B cells by EBV dramatically reorganizes the cellular transcriptome via multiple mechanisms, ranging from transcriptional activation, effects on RNA stability, and alternative splicing (12, 66, 110, 117). While most studies have focused on EBV-induced mRNA abundance changes, we used the primary B-cell infection system to assay the effects of EBV on mRNA isoform usage. Resting Primary B cells from four healthy donors were infected with the B95-8 strain of EBV at limiting dilution to generate monoclonal LCLs. Transcript isoform changes were analyzed using two microarray platforms: the U133 Plus 2.0 array and the Human Exon 1.0 ST array. U133 arrays predominantly detect the 3' untranslated regions (UTRs) of mRNAs, while HuEx arrays detect exons throughout the entirety of mRNAs, and the combined use of these two platforms provides better coverage of all events occurring across entire transcripts (7, 117). HuEx arrays detect nearly every exon in the genome, with 56,598 probe sets querying 8,006 expressed genes, with an average of 7 exons being interrogated per gene. Importantly, 2,613 genes (of 8,409 expressed) are detected on the U133 array targeting multiple exons, such that differential expression of mRNA isoforms can be studied on this platform (116). The SplicerEX algorithm was used to assess

alternative exon usage and overall mRNA abundance changes mediated by EBV infection from data obtained using both microarray platforms (**Figure 20A**).

HuEx arrays detected 966 mRNAs whose abundance changed >2-fold between B cells and LCLs ( $P < 0.01$ ), with 723 increasing and 243 decreasing, while U133 arrays detected 1,692 mRNAs whose abundance changed, with 1,484 increasing and 208 decreasing. Importantly, there was substantial overlap in mRNA abundance changes among commonly detected genes on these two platforms (117). EBV infection changed 433 cellular mRNAs at the level of isoform usage, with 271 of these events detected on the HuEx array and 162 detected on the U133 array (**Figure 20B**). Only 35% of the mRNAs that displayed altered isoform usage were also changed in overall abundance (**Figure 20B**) (HuEx array, 94/271; U133 array 29/162; total, 146/433). Similarly, only ~7% of total differentially expressed mRNAs changed at the level of isoform usage (146/2,163 were common between both platforms). As has been noted in other analyses (126), our data indicate that the set of mRNA isoforms regulated by EBV is largely orthogonal with the set of mRNAs changed at the level of overall abundance.



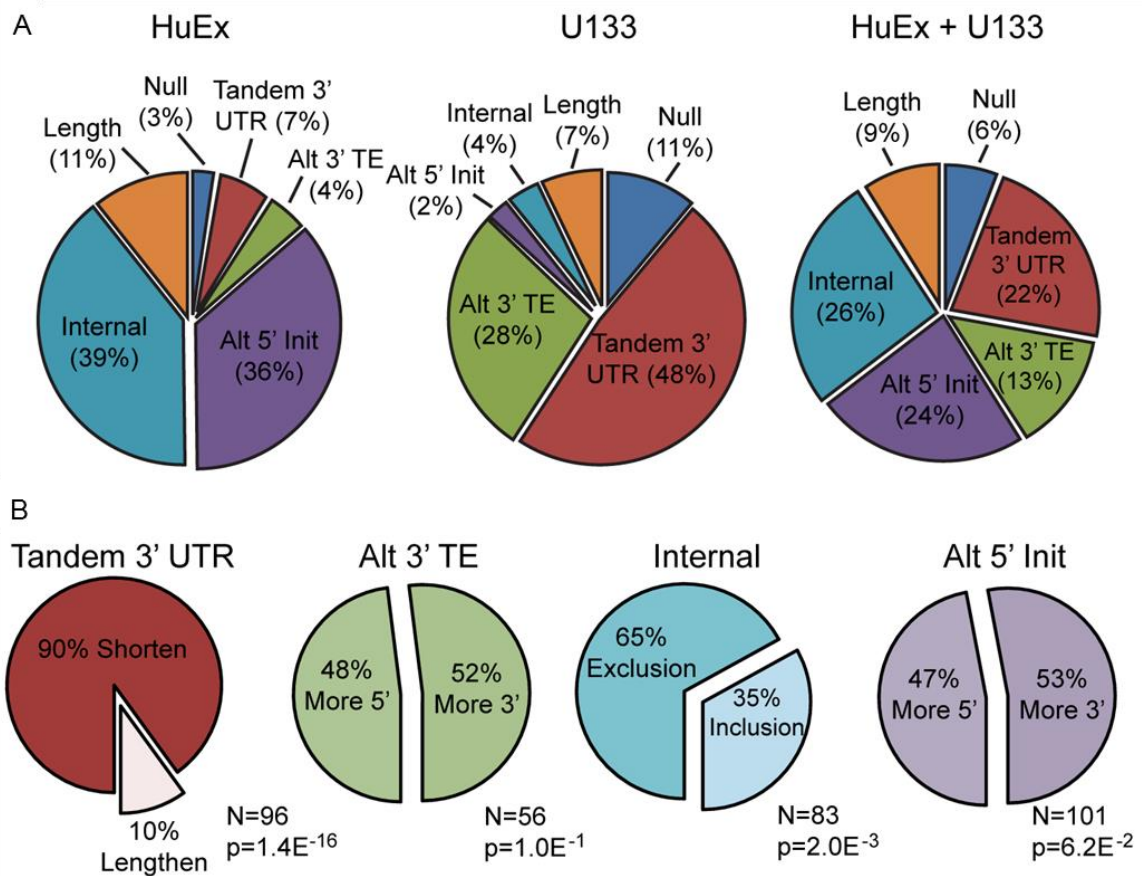
**Figure 20: SplicerEX analysis of mRNA expression and isoform usage.**

(A) Schematic of SplicerEX method used to identify alternative isoforms in a generic mRNA. SplicerEX combines probes that are changing similarly into larger meta-probe sets. Probes detecting upregulated exons (UP probes) are represented in red and increase in abundance relative to the other probes for sequences within the gene upon transformation. Probes detecting downregulated exons (DOWN probes) are represented in blue and decrease in abundance relative to the other probes for sequences within the gene. Gray blocks are probe sets that are not changing significantly upon transformation. SplicerEX uses the UP and DOWN features to predict the alternative isoform usage. (B) Venn diagrams of genes whose mRNAs are changing either in abundance or at the level of alternative isoforms, as predicted by SplicerEX between the two arrays.

## 4.2.2 EBV Infection Induces Host Cell mRNA 3' UTR Shortening and Exon Exclusion

The SplicerEX algorithm categorizes mRNA alternative isoform usage into distinct classes of transcriptional or processing events, including i) alternative 5' initiation, ii) internal cassette exon inclusion or exclusion, iii) alternative 3' terminal exon choice, iv) tandem 3' UTR choice, and v) changes in transcript length that do not fall into any of the aforementioned categories. Consistent with their focused array design, the EBV-induced mRNA isoform changes detected on the HuEX platform were predominantly internal and 5' initiation changes, and those detected on the U133 platform were predominantly tandem 3' UTR length or 3' terminal exon choice changes (**Figure 21A**). When combining the data from both platforms, however, we failed to identify a preferential mRNA isoform event class regulated by EBV (**Figure 21A**).

In contrast, EBV infection preferentially altered the directionality of cellular mRNA isoform choice. Specifically, EBV infection strongly favored the shortening of tandem 3' UTRs, with 86/96 of mRNAs whose 3' UTR changed being shortened rather than lengthened (**Figure 21B**). These data are consistent with recent observations indicating that cell proliferation triggers mRNA 3' UTR shortening (91, 121). EBV infection also favored exclusion of internal cassette exons (54 excluded/83 total cassette events). No preference of EBV-regulated mRNA isoforms with regard to the directionality (upstream versus downstream sites) of alternative transcriptional initiation or alternative 3' terminal exon choice was observed.



**Figure 21: Categorization and directionality of SplicerEX-predicted mRNA isoform changes.**

**(A)** Categories of isoform changes, as described in the text and (117) detected on each array individually or on both arrays combined. Categories include tandem 3' UTRs, an alternative 3' terminal exon (TE) choice, alternative 5' initiation (Init) sites, inclusion/exclusion of an internal cassette exon, changes in the length of the transcript that could not be categorized in any of the previous groups, or a null hypothesis if the two primary distinguishing features (most significant UP or DOWN meta-probe sets) do not overlap a known UCSC transcript. **(B)** Directionality of specific categories of isoform changes detected on both arrays combined. Relative to the isoforms present in B cells, tandem 3' UTRs are classified as lengthened or shortened 3'-terminal exon choice, 5' initiation sites are classified as more 5' or more 3', and internal cassette exons are classified as inclusion or exclusion. *P* values were calculated using a binomial probability distribution.

### 4.2.3 Diverse Biological Processes and Pathways Are Regulated by EBV Infection at the Level of Alternative mRNA Isoform Usage

In order to assess the biological significance of the EBV-regulated mRNA isoform changes, the PANTHER algorithm was used to identify pathways and biological processes that were enriched in the set of 433 genes whose mRNA isoforms changed following EBV infection. A parallel PANTHER analysis was performed using the 2,163 genes whose overall mRNA abundance changed due to EBV infection. Genes associated with the biological processes of cell cycle and mitosis were strongly enriched in both of these gene sets (**Table 4**); however, genes associated with exocytosis and protein transport were exclusively enriched among those regulated at the mRNA isoform level.

Molecular functions, cellular components, and protein classes were even more divergent between mRNAs regulated at the level of abundance versus isoform. The proteins encoded by the set of genes corresponding to those mRNAs changed at the level of isoform were enriched for cysteine proteases and nucleic acid binding proteins (**Table 4**). More specifically, RNA binding proteins, including splicing factors, were strongly depleted in the set of mRNAs regulated by EBV at the level of overall transcript abundance (**Table 4**). These data further strengthen the observation that the mRNAs regulated by isoform changes are distinct from those regulated at the level of abundance, both in absolute terms and in broad functional terms.

**Table 4: Enrichment or depletion of gene ontology groups of alternatively regulated mRNAs.**

Green, enrichment of genes in a given group; red, depletion of genes in a given group

Biological process or pathway	Alternative mRNA isoform			Total mRNA abundance		
	Observed abundance	Expected abundance	<i>P</i> value	Observed abundance	Expected abundance	<i>P</i> value
Cellular processes, cell cycle	66	49.89	1.12E <sup>-02</sup>	308	257.44	5.64E <sup>-04</sup>
Molecular function						
Nucleic acid binding	117	101.37	4.31E <sup>-02</sup>	193	230.73	1.96E <sup>-03</sup>
DNA binding	70	60.29	1.01E <sup>-01</sup>	130	134.95	3.42E <sup>-01</sup>
RNA binding	24	15.44	2.37E <sup>-02</sup>	12	35.81	2.78E <sup>-06</sup>
Splicing factors	12	8.14	1.20E <sup>-01</sup>	4	18.72	4.19E <sup>-05</sup>

#### 4.2.4 EBV regulates mRNA Isoform Usage of RBPs

Strong enrichment of RNA binding proteins (RBPs) was observed in EBV-regulated mRNAs changing at the level of isoform usage, while less than the expected number of RBPs changed at the level of overall mRNA abundance (**Table 4**). EBV infection altered the mRNA isoform of 24 RBPs, 12 of which were splicing factors. To validate our microarray-based observations, we assessed a subset of the predicted isoform changes in primary B cells and EBV-transformed LCLs using RT-PCR and qRT-PCR. We focused on two significant mRNA isoform changes: those in the splicing factor SRSF7, also known as 9G8, and those in the mRNA stability factor ELAVL1, also known as HuR.

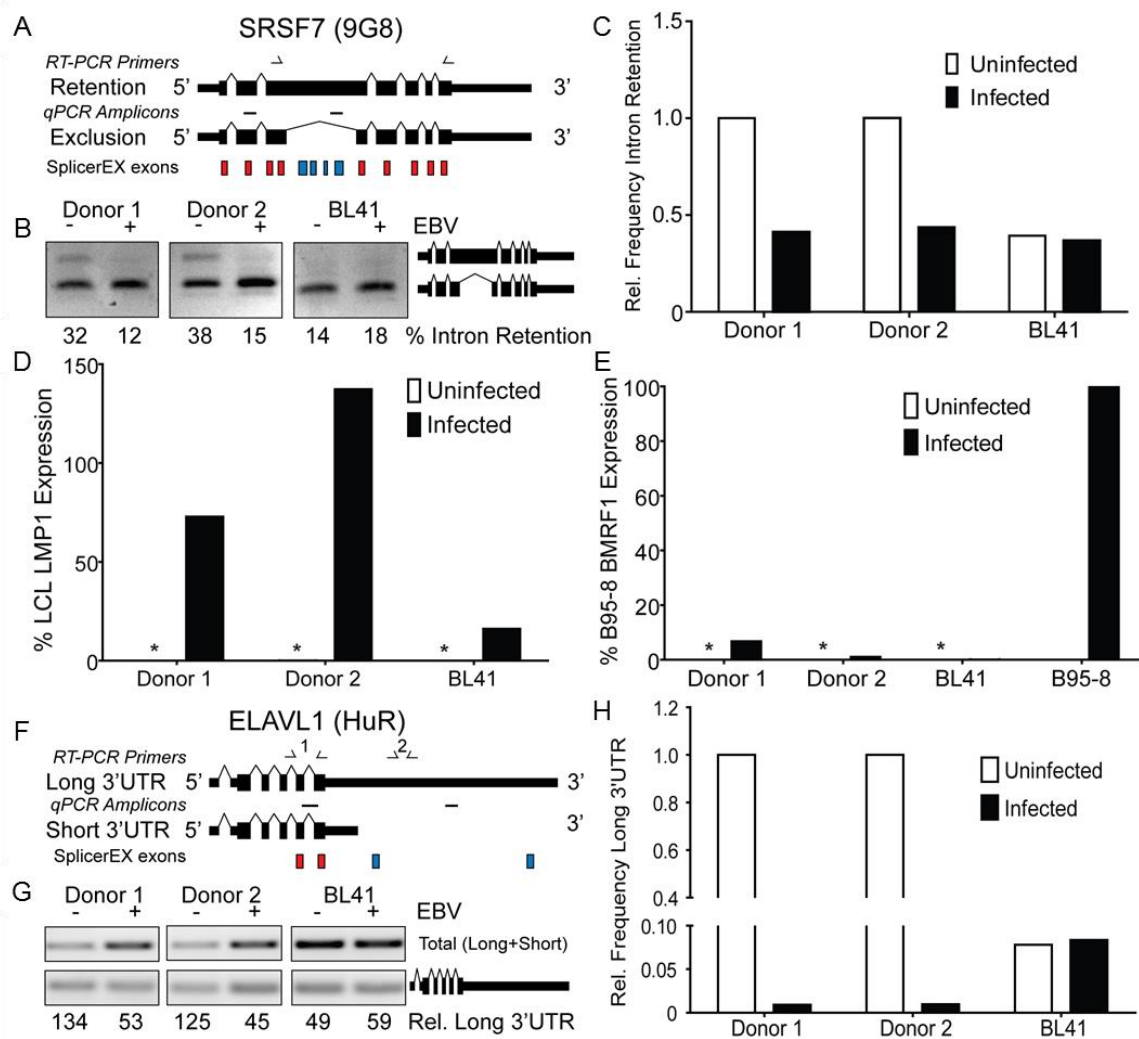
SplicerEX predicted that EBV infection inhibited the retention of an intron between exons 3 and 4 of SRSF7 (**Figure 22A**) (75, 107). To validate this prediction, we assayed the retention of intron 3 in B-cell (with B cells from two healthy human donors) and LCL pairs by RT-PCR. EBV infection decreased the relative expression level of SRSF7 mRNAs containing intron 3 (**Figure 22B**). To determine whether latent EBV

infection contributed to the loss of SRSF7 intron 3 or if EBV-driven proliferation of resting B cells was responsible, we assayed this event in the EBV-negative Burkitt's lymphoma cell line BL41 and its latent EBV-infected counterpart, BL41/B95-8. In this case, BL41 cells already displayed minimal expression of the intron 3-retained SRSF7 isoform, and latent EBV infection of BL41 cells did not change this (**Figure 22B**). Moreover, qRT-PCR confirmed that the exclusion of intron 3 from SRSF7 was favored upon EBV infection of primary B cells, but the level of the intron 3-retained isoform was already low in BL41 cells and unchanged in BL41/B95-8 cells (**Figure 22C**). Latent infection was verified by LMP1 expression in LCLs and BL41/B95-8 cells, but not in uninfected cells, while little lytic BMRF1 transcript was detected compared to the amount detected in lytic B95-8 cells (**Figure 22D, E**). Therefore, EBV-driven B-cell proliferation rather than latent infection *per se* led to the loss of an SRSF7 mRNA isoform retaining intron 3 in LCLs.

Another RNA binding protein, ELAVL1, which has a role in mRNA stability (11), was also changed at the mRNA isoform level upon EBV infection of resting B cells. Specifically, in LCLs the ELAVL1 mRNA was predicted to have a shortened 3' UTR relative to the length of the 3' UTR in the isoform in resting B cells (**Figure 22F**). RT-PCR experiments confirmed this predicted change, as the total ELAVL1 mRNA level increased from B cells to LCLs, while the level of ELAVL1 with the long 3' UTR did not increase (**Figure 22G**). These experiments were corroborated by quantitative RT-PCR

specific for the long isoform (**Figure 22H**). These data suggest that the ELAVL1 mRNA isoform containing the long 3' UTR represents a significantly smaller proportion of the total ELAVL1 mRNA species in LCLs than in resting B cells.

Similar to what was observed in the case of SRSF7, EBV infection *per se* did not affect the change in the ELAVL1 mRNA isoform. BL41 cells and EBV-infected BL41/B95-8 cells displayed similar levels of the ELAVL1 mRNA species with a long 3'UTR (**Figure 22G, H**). This level, however, was already quite low as a percentage of the total relative to that observed in resting B cells, suggesting that B-cell proliferation is consistent with the previously recognized widespread 3' UTR shortening observed in proliferating cells and tissues (121).



**Figure 22: Validation of RNA binding proteins predicted to change at the isoform level by SplicerEX.**

**(A)** Exon schematics of SRSF7 gene. Red blocks, SplicerEX UP probe sets; blue blocks, DOWN probe sets; single-headed arrows above schematics, the primers used in RT-PCRs; horizontal hashes between the isoform schematics, amplicons generated in qRT-PCR. **(B)** Gel of RT-PCR using the primers shown in panel A flanking the retained/excluded intron in SRSF7 from two donor-matched B-cell and LCL pairs and BL41 versus BL41/B95-8 cells. The upper band represents intron retention, and the lower band represents intron exclusion. The quantification below the gel represents the frequency of intron retention. **(C)** Isoform-specific qRT-PCR depicting the frequency on intron retention in SRSF7 relative to the B-cell average using the amplicons shown in panel A. Rel., relative. **(D)** qRT-PCR depicting the latency-associated LMP1 mRNA

expression in the B-cell and LCL pairs and BL41 versus BL41/B95-8 cells. A value of 100% represents the average LMP1 expression in the LCLs. \*, a signal that was not detectable over the background or where the standard deviation between technical replicates did not meet our threshold of 0.5 cycle. **(E)** qRT-PCR depicting lysis-associated BMRF1 mRNA expression in the B-cell and LCL pairs and BL41 versus BL41/B95-8. A value of 100% represents the BMRF1 expression in the B95-8 cell line. \*, a signal that was not detectable over the background. **(F)** Exon schematics of ELAVL1 mRNA in the cells (primer set 1); lower band, amount of the long 3' UTR isoform specifically (primer set 2). The quantification below the gel represents the relative frequency of the long 3' UTR isoform. **(H)** Isoform-specific qRT-PCR of the frequency of ELAVL1 mRNAs containing the long 3' UTR relative to the total amount using the amplicons shown in panel F and normalized to the B-cell average.

#### **4.2.5 EBV Latency Alters XBP1 mRNA Isoform Usage and Suppresses IRE1-dependent XBP1 Splicing, Thereby Preventing Z Activation**

In addition to proliferation-associated changes that may be important for EBV latency, we were interested in identifying EBV-specific changes in mRNA isoform usage. We therefore investigated genes that are known to be associated with the EBV-specific process of reactivation from latency. A number of stimuli, including treatment with anti-immunoglobulins (131), histone deacetylase inhibitors (HDACi) (82), and phorbol esters (36), have been shown to induce reactivation in EBV-infected cell lines. Each of these stimuli ultimately induces activation of transcription of the major EBV lytic transactivator BZLF1, or simply Z, which is capable of upregulating its own transcription and beginning the cascade of lytic transcriptional events that ultimately leads to the production of infectious virions (35, 41). During latency, Z expression is tightly repressed by both host transcription factors (29, 78, 124, 157) and chromatin structure (22, 58, 99, 112).

XBP1 and TCF4 were two of the transcription factors regulated by EBV at the level of mRNA isoform change. Both of these proteins have been implicated in regulation of Z expression (8, 93, 130, 135). SplicerEX predicted an alternative 5' initiation of the plasma cell differentiation factor XBP1, yielding a shorter mRNA (**Figure 23A**). We confirmed the short XBP1 mRNA isoform, thereby substantially decreasing the frequency of long XBP1 (**Figure 23B**). EBV latent infection of BL41 cells modestly increased short XBP1 mRNA levels but did not substantially decrease the frequency of long XBP1 (**Figure 23B**). We corroborated these findings by qRT-PCR, where the B-cell-to-LCL transition led to a strong reduction in the frequency of long XBP1 relative to the total amount of XBP1 (**Figure 23C**). BL41 cells already contained a reduced frequency of long XBP1 mRNA relative to the total amount, and this was unaffected by EBV infection. As suggested above for the cases of the SRSF7 and ELAVL1 transcripts, B-cell proliferation is likely responsible for this isoform change in XBP1.

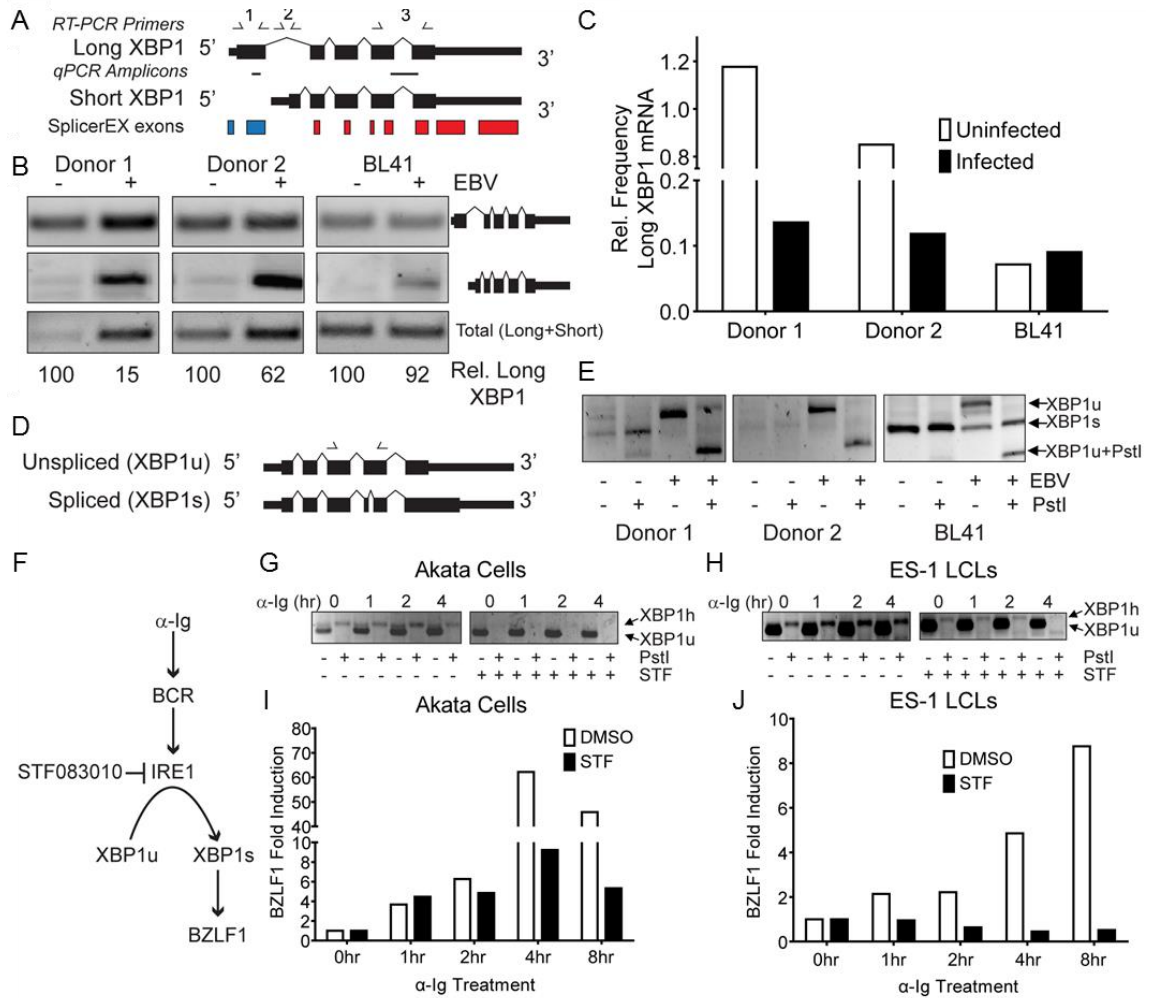
While the HuEx array identified changes in the 5' end of the mRNA, another splicing event that regulates XBP1 function is the noncanonical processing of XBP1 mRNA in the cytoplasm by the endonuclease IRE1 downstream of endoplasmic reticulum (ER) stress (155). IRE1-dependent splicing of the inactive XBP1u mRNA to the active XBP1s mRNA in response to ER stress is due to an endonucleolytic cleavage event that removes 26 nucleotides from the coding sequence and generates a frame-shifted protein product (155) (**Figure 23D**). The exon in which this event takes place was not

queried by the microarrays, but given the importance of this gene in EBV biology, we sought to assess the status of this IRE1-dependent splicing event in XBP1, as this event ultimately controls its activity as a transcription factor. We therefore assayed the extent of IRE1-dependent splicing in resting B cells and LCLs by RT-PCR at the IRE1-spliced site. XBP1u contains a PstI restriction site, and this restriction site is lost in XBP1s, making it resistant to PstI digestion. We were therefore able to assess the presence of XBP1u and XBP1s by susceptibility to PstI digestion. We found that EBV infection strongly increased overall XBP1u levels but potently suppressed XBP1s, consistent with inhibition of lytic gene expression in the tightly latent LCL system (**Figure 23E**). Similarly, latent EBV infection of BL41 cells reduced the spontaneous splicing of XBP1u to XBP1s.

In order to discern the functional significance of IRE1-dependent splicing of XBP1 in lytic reactivation, we took advantage of a new small-molecular inhibitor of the IRE1 endonuclease, STF083010 (STF) (105). Cross-linking of the B-cell receptor (BCR) with anti-Ig antibody has been shown to induce IRE1-dependent cleavage of XBP1u transcripts to generate XBP1s (93), which activates the Z promoter (8, 130). We hypothesized that inhibition of IRE1-dependent splicing of XBP1u would suppress Z-promoter activation (**Figure 23F**). We first demonstrated that STF is capable of inhibiting IRE1 endonuclease activity by preventing XBP1s formation after anti-Ig cross-linking in Akata BL cells (**Figure 23G**). In addition, we queried the role of XBP1 splicing in

LMP2A-deficient LCLs (ES-1), which are primed to trigger Z expression in response to anti-Ig due to loss of B-cell receptor antagonism by LMP2A (94) (**Figure 23H**).

Importantly, IRE1 inhibition markedly suppressed anti-Ig-induced Z activation in both Akata cells (**Figure 23I**) and ES-1 LCLs (**Figure 23J**). These data indicate that strong suppression of the IRE1-dependent splicing of XBP1 in LCLs likely prevents the aberrant activation of the promoter of the key lytic transactivator Z, thereby maintaining latency in immortalized B cells. Interestingly, the strong increase in XBP1u levels also suggests that these latently infected cells are primed to transition to lytic cycle upon sensing IRE1-activating stimuli.



**Figure 23: XBP1 alternative isoform usage and lytic reactivation.**

(A) Exon schematics of the XBP1 gene. Red blocks, SplicerEX UP probe sets; blue blocks, DOWN probe sets; arrows above the schematics, the primers used in RT-PCRs; horizontal hashes between the isoform schematics, amplicons generated in qRT-PCRs. (B) Gels of RT-PCR specifically targeting the long (top; primer set 1) and short (middle; primer set 2) XBP1 isoforms as well as a downstream region common to both isoforms (bottom; primer set 3) from two donor-matched B-cell and LCL pairs and BL41 versus BL41/B95-8 cells using the primers shown in panel A. The quantification below the gel represents the frequency of the long XBP1 isoform relative to the total amount of XBP1 mRNA present. (C) Isoform-specific qRT-PCR depicting the frequency of long XBP1 relative to the total amount of XBP1 using the amplicons shown in panel A. (D) Exon schematics depicting the IRE1-dependent alternative splicing of XBP1. Arrows above the schematics, the primers used in RT-PCRs. (E) Gel of RT-PCR querying IRE1-mediated

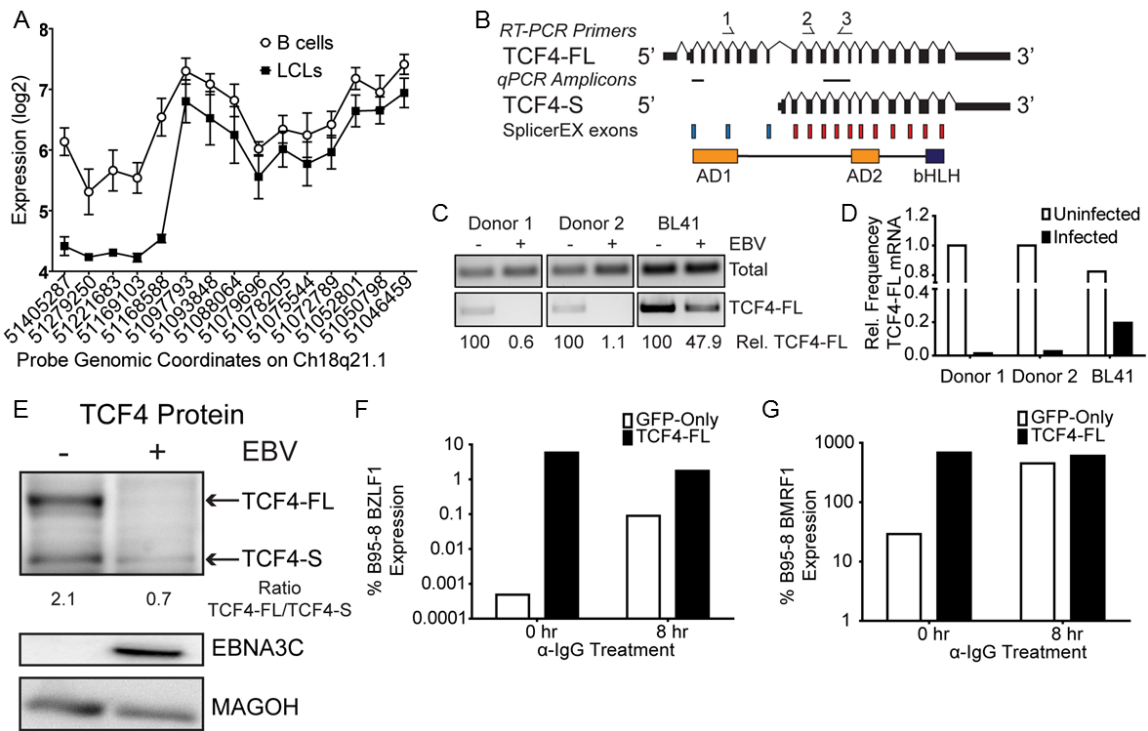
splicing. XBP1u contains a PstI restriction site, and this site is lost in XBP1s. Digestion of the PCR product with PstI results in selective cleavage of the XBP1u product. **(F)** Schematic of Ig antibody-induced transcriptional activation of Z via IRE1-dependent splicing of XBP1. **(G)** Gel of RT-PCR of Akata cells treated with 5 $\mu$ g/ml anti-IgG antibody for the indicated times in the presence or absence of the IRE1 inhibitor STF. Arrows, XBP1u, the unspliced transcript, and XBP1h, the hybrid product. **(H)** Gel of RT-PCR of ES-1 cells treated with anti-Ig antibody in the presence or absence of STF, as described for panel G. **(I)** qRT-PCR of the induction of endogenous Z by Ig antibody and inhibition of STF in EBV-positive Akata cells. **(J)** qRT-PCR of the induction of endogenous Z by anti-Ig antibody and inhibition by STF in ES-1 LCLs.

#### **4.2.6 EBV Latent Infection Inhibits Expression of the Full-length mRNA for TCF4 to Suppress Z Expression**

We identified an mRNA isoform change of the TCF4 gene product upon EBV infection of primary B cells. TCF4, also known as E2-2, is an E-box binding protein that plays a major role in B-cell development and has previously been shown to bind elements in the EBV lytic Z promoter (135). We were therefore interested in characterizing this isoform change and its significance. As shown in **Figure 24A**, we detected similar expression of TCF4 across all exons in resting B cells. However, in LCLs, the expression of the first five meta-probe sets detecting the 5' end of the mRNA was significantly lower than that in B cells (**Figure 24A**). These data suggest that the TCF4 gene is transcribed from an alternative 5' initiation site, which would give rise to a protein product lacking the first transcriptional activation domain, AD1, but retaining AD2 and the basic helix-loop-helix dimerization and DNA binding domain (**Figure 24B**). To confirm this hypothesis, we used exon-specific RT-PCR primers and observed downregulation of the full-length TCF4 isoform (TCF4-FL) in EBV-infected LCLs relative to its regulation in EBV-negative B cells (**Figure 24C**). These data were

corroborated by a quantitative RT-PCR specific for TCF4-FL mRNA relative to total TCF4, where LCLs contained much less TCF4-FL (**Figure 24D**). We also assessed the relative levels of TCF4-FL in BL41 cells that were either uninfected or latently infected with EBV. In this setting, TCF4 also displayed a preference for lower expression of the full-length isoform in the infected cells (**Figure 24C, D**). Finally, Western analysis of the TCF4 protein product indicated a reduction of the full-length isoform expressed in EBV-negative BL41 cells relative to that in the latently infected BL41/B95-8 cells, where the full-length protein was essentially undetectable (**Figure 24E**).

TCF4 has previously been shown to bind E boxes present in the Z promoter (135), and EBV preferentially downregulates the TCF4-FL isoform, which contains both transactivation domains. Therefore, we hypothesized that TCF4-FL would be capable of inducing transcription at the endogenous Z promoter. Indeed, transfection of TCF4-FL into EBV-positive Akata BL cells was sufficient to induce the endogenous Z mRNA at levels comparable to those induced by cross-linking the BCR (**Figure 24F**). Furthermore, TCF4-FL expression also induced expression of a downstream Z target, the EBV polymerase processivity factor BMRF1, to the levels observed upon BCR cross-linking (**Figure 24G**). These results support the hypothesis that EBV latent infection suppresses TCF4-FL expression to prevent the spontaneous induction of Z expression and downstream targeting of lytic genes.



**Figure 24: TCF4 alternative isoform usage and lytic reactivation.**

(A) Plot of expression levels of TCF4 meta-probe sets in the 5' → 3' direction from left to right in B cells and LCLs. (B) Exon schematics of the TCF4 gene. Red blocks, SplicerEX UP probe sets; blue blocks, DOWN probe sets; single-headed arrows above the schematics, the primers used in RT-PCRs; horizontal hashes between isoform schematics, amplicons generated in qRT-PCRs. Below the isoform schematics is a representation of the TCF4 protein organization. bHLH, basic helix-loop-helix. (C) Gel of isoform-specific RT-PCR of two donor-matched B-cell and LCL pairs and BL41 versus BL41/B95-8 cells using the primers depicted in panel B. Upper band (primers 2 and 3), total amount of the TCF4 transcript in the cells; lower band (primers 1 and 3), amount of TCF4-FL specifically. (D) Isoform-specific qRT-PCR depicting the frequency of TCF4-FL relative to the B-cell average using the amplicons shown in panel B. (E) Western blot of TCF4-FL in BL41 and BL41/B95-8 cells. Densitometry was performed, and the ratio of TCF4-FL/TCF4-S detected is shown. The results for EBNA3C as a control for EBV latent infection are shown. The results for MAGOH as a loading control are shown. (F) Endogenous BZLF1 expression in Akata cells in response to anti-IgG antibody treatment and/or TCF4-FL overexpression. GFP-only, samples transfected with GFP only; TCF4-FL, sample were transfected with GFP and a plasmid expressing TCF4-FL. Values are relative to the Z mRNA levels in B95-8 cells. (G) BMRF1 expression in Akata cells in

response to anti-IgG antibody treatment and/or TCF4-FL overexpression. Values are relative to the BMRF1 levels in B95-8 cells.

### **4.3 Materials and Methods**

**Cell lines and culture conditions.** B cells were purified from healthy donor buffy coats through the Carolina Red Cross and Gulf Coast Regional Blood Center and transformed to monoclonal LCLs as described previously (117). BL41 is an EBV-negative Burkitt's lymphoma cell line. BL41/B95-8 is BL41 latently infected with the B95-8 strain of EBV (BL41 lines were gifts of Elliott Kieff, Harvard Medical School). Akata is an EBV-positive Burkitt's lymphoma cell line (kind gift of Kenzo Takada). ES-1 is an LCL that is deficient for LMP2A and was a gift from Richard Longnecker. Cells were grown in RPMI 1640 supplemented with 10% fetal bovine serum (FBS), 20mM L-glutamine, 1% streptomycin.

**RNA and cDNA preparation and PCR.** RNA was isolated from cells using a Qiagen RNeasy Plus kit or Qiagen RNeasy microkit. cDNA synthesis was performed using an Applied Biosystems high-capacity cDNA reverse transcription kit. Quantitative RT-PCR was performed using Qanta Biosciences SYBR green on an Applied Biosystems StepOne Plus quantitative PCR instrument in technical duplicate or triplicate with a standard deviation between technical replicates of <0.5. The threshold cycle ( $\Delta\Delta C_T$ ) method of relative abundance changes was used to determine the fold changes in the abundance of targets relative to the average abundance of that target in the B cells tested. Reactions were normalized to those for exons from the genes SETDB1 and SYNE1, which were determined by SplicerEX to remain stable between resting B cells and LCLs. All

experiments reported were performed in biological triplicates (e.g., with cells from at least three healthy human donors), where the results for two representative donors are shown in each figure. Fold change data were used to perform the statistical analyses whose findings are present in Results. RT-PCRs were performed using Invitrogen Platinum *Taq* DNA polymerase high fidelity in an Eppendorf Mastercycler apparatus, and the results were visualized on 1% or 2% Tris-acetate-EDTA (TAE) agarose gels. Quantification of gel bands was performed using the GeneTools software from Syngene.

IRE1-dependent splicing of XBP1 assays were carried out by pretreating cells for 1h with either 0.1% dimethyl sulfoxide or 100 $\mu$ M STF083010 (STF; Sigma-Aldrich). Cells were then washed in phosphate-buffered saline before being returned to RPMI 1640 and treated with 100 $\mu$ g/ml anti-Ig antibody (Jackson ImmunoResearch) for 0, 1, 2, 4, or 8h. RNA was extracted and cDNA was synthesized as described above. PCR was performed using primers flanking the inositol-requiring protein 1 (IRE1) splice site. Following PCR, one half of the reaction mixture was digested with PstI restriction enzyme (NEB) for 2h at 37°C, while the other half was left undigested. The reactions were visualized on a 2% TAE agarose gel. In the Akata and ES-1 cells, XBP1h is a hybrid product that is generated as a result of the annealing of one strand of the unspliced XBP1 (XBP1u) PCR product, which is resistant to PstI digestion, as described in references (93) and (127). Here, XBP1h indicates the presence of the spliced transcript.

**Plasmids and cloning.** The pCEP4-EGFP plasmid was a generous gift from Seiji Maruo. Full-length TCF4 (TCF4-FL) was cloned from cDNA purchased from Open Biosystems (material no. MHS4426-99625743) by Gateway recombination cloning technology (Life Technologies). pDONR221 (a gift from Bryan Cullen) was used as the donor vector, and pSG5 engineered for Gateway cloning with an N-terminal 6x-His tag and a hemagglutinin-tagged protein expression cassette (a gift from Eric Johannsen) was used as the destination vector.

**Microarray analysis.** The analysis of U133 and human exon (HuEx) arrays from resting human B cells and LCLs from four independent donors (GEO accession number GSE29301) was performed as described in reference (117). Briefly, SplicerEX uses a maximum likelihood ratio (MLR) to compare the relative likelihood that changes in probe set expression levels are described by alternative processing versus overall transcription level changes. This base MLR is modified by a multiple-probe-set correction to adjust for multiple paired groups of probe sets, an expression cutoff modifier for the minimum required change between isoforms, and a centering modifier that preferentially ranks genes with probe sets whose levels are changing in opposite directions. These parameters are used to generate a final splice score. For SplicerEX to categorize an mRNA to be undergoing alternative processing, it must have a positive splice score and analysis of variance and SplicerEX *P* values of <0.01.

SplicerEX computed the overall mRNA abundance from the U133 arrays using the probe set that reported the most statistically significant change in expression by a simple *t* test log<sub>2</sub> expression values between B cells and LCLs. SplicerEX computed overall mRNA abundance for the HuEx arrays by identifying the largest meta-probe set within each gene (composed of the most individual probe sets) and then performed a two-tailed, homoscedastic *t* test of meta-probe set expression in values in B cells and LCLs. For an mRNA to be called increasing or decreasing in overall abundance it had to have a >2-fold change. Correction for multiple-hypothesis testing was not performed to minimize the likelihood of falsely rejecting an overlap between overall and isoform-specific mRNA expression changes.

**Gene ontology analysis.** Gene ontology analyses were performed using the PANTHER classification system (136). SplicerEX-predicted mRNA isoform changes (splice score >0; *P* <0.01) were compared to the values for a reference set of all genes whose U133 probe sets had log<sub>2</sub> expression of >6 or whose HuEx probe sets had log<sub>2</sub> expression of >8.

Analysis was carried out for the SplicerEX data generated from each microarray platform independently, as well as combined.

**Transfection and anti-Ig antibody treatment.** Akata cells were transfected using a Digital Bio MicroPorator MP-100 system with one pulse at 1,470V for 30ms in a 100µl tip at a cell density of 10 million cells/ml using a Neon transfection system (Life Technologies). 5µg pCEP4-EGFP plasmid or 1µg pCEP4-EGFP and 4µg pSG5-TCF4-FL

plasmids were used per 1 million cells. Cells were allowed to grow in RPMI 1640 supplemented with 15% FBS (R15) for 48h post-transfection. Green cells were sorted using a BD DiVa cell sorter on the basis of green fluorescent protein (GFP) positivity with a 488-nm argon laser and 530/30-nm filter. GFP-positive cells were placed back in the R15 medium and treated for 0 or 8h with 100 $\mu$ /ml anti-IgG antibody from Jackson ImmunoResearch before harvesting for RNA.

**Western blotting and antibodies.** Western blots were run on NuPAGE 4-12% bis-Tris polyacrylamide gels (Life Technologies) and were transferred to a polyvinyl difluoride membrane in the presence of 20% methanol in Tris-glycine at 50V for 75min. TCF4 membranes were blocked with 3% milk in Tris-buffered saline-Tween 20 for 1h at room temperature, blotted with primary antibody (sc-48947; Santa Cruz Biotechnology) in 1% bovine serum albumin (BSA) overnight at 4°C, and blotted with horseradish peroxidase (HRP)-conjugated secondary antibody (705-035-003; Jackson ImmunoResearch) in 3% milk for 1h at room temperature. The exon junction complex component MAGOH was used as a control for protein expression between uninfected and infected cells. MAGOH membranes were blocked with 5% BSA for 1h at room temperature, blotted with primary antibody (sc-56724; Santa Cruz Biotechnology) in 5% BSA overnight at 4°C, and blotted with HRP-conjugated secondary antibody (715-035-150; Jackson ImmunoResearch) in 5% BSA for 1h at room temperature. EBNA3C membranes were blocked with 5% milk for 1h at room temperature, blotted with primary antibody

(F125P; Exalpa Biologicals) in 5% milk overnight at 4°C, and blotted with HRP-conjugated secondary antibody (713-035-003; Jackson ImmunoResearch) in 5% milk for 1h at room temperature. Blots were developed using the SuperSignal West Femto maximum-sensitivity substrate (Thermo Scientific). Quantification of bands was performed using GeneTools software from Syngene.

#### **4.4 Discussion**

In this study, we aimed to understand how virus infection alters the cellular mRNA landscape of the infected cell. We studied changes in both mRNA abundance and isoform usage altered by Epstein-Barr virus infection of primary human B cells using SplicerEX, a novel algorithm to analyze human exon and conventional microarray data. We identified 433 mRNAs changing at the level of isoform usage and 2,163 host mRNAs changing in overall abundance after EBV infection. These two sets of mRNAs were largely orthogonal, where less than one-third of those regulated by isoform usage were also changed at the level of abundance. This result suggests that mRNAs regulated at the level of the isoform may have functions distinct from those regulated by abundance, consistent with the findings of other global transcriptomic studies (126). Indeed, mRNAs changing at the level of the isoforms were enriched for nucleic acid binding proteins, including splicing factors and transcription factors, while these protein classes were depleted among the mRNAs regulated at the level of abundance. We validated several of the mRNA isoform changes and found that EBV altered the isoform-

level expression of two transcription factors, XBP1 and TCF4, which prevented accumulation of active forms capable of inducing lytic EBV reactivation. Therefore, EBV latent infection is maintained, in part, through the regulation of specific cellular mRNA isoforms.

The SplicerEX algorithm generates hypotheses regarding the mechanism and directionality of mRNA isoform changes. While analysis of combined data from both the HuEx and U133 array platforms failed to identify any preferential mechanistic class of isoform change driven by EBV infection, EBV infection and proliferation of primary resting B cells promoted a strong preference for shortening rather than lengthening of the 3' UTRs of transcripts. This result is consistent with the findings of other studies indicating that proliferation correlates with expression of mRNAs containing shorter 3' UTRs (91, 121). The mRNAs with shortened 3' UTRs are often translated more efficiently due to the loss of negative regulatory elements, such as miRNA binding sites (91, 121). In addition to shorter 3' UTRs, we also observed a preference toward exclusion rather than inclusion of internal cassette exons during EBV transformation. While most of what is known about miRNA-mediated mRNA repression comes from the studies of regulatory effects of miRNA binding sites within 3' UTRs, miRNAs also may bind to and regulate translation through sites located in the coding sequence of mRNAs (48). Therefore, the 3'UTR-shortened and cassette exon-excluded transcripts regulated by EBV likely derive from changes induced by EBV-mediated B-cell proliferation and may

consequently alter the infected cell proteome through changes in mRNA regulatory elements. Examples of transcripts with these proliferation-associated mRNA isoform changes include ELAVL1 and SRSF7, which may play an important role in B-cell transformation through promoting translation and alternative splicing of viral and cellular mRNAs. In addition to proliferation-associated changes, we hypothesized that other mRNAs may be regulated by EBV latent infection even in the background of a proliferating cell. Such changes were illuminated in the characterization of gene ontology groups enriched in the set of EBV-regulated mRNA isoform changes.

Strong enrichment of nucleic acid binding proteins, including splicing factors and transcription factors, was detected within the set of mRNAs regulated by EBV at the level of isoform usage. Interestingly, genes in these protein classes were depleted in the set of mRNAs regulated by overall abundance, suggesting distinct modes of regulating gene expression. It is not uncommon for RBPs and splicing factors to be regulated themselves by alternative splicing (25, 152). In fact, a recent study identified enrichment in splicing factors among mRNAs changing at the isoform level in Kaposi's sarcoma-associated herpesvirus-infected versus uninfected endothelial cells (15). In our study, we identified 24 RNA binding proteins to be regulated at the isoform level in EBV-infected cells, with 12 of these being splicing factors, including SF1, SRSF7, SRSF15, CUGBP2, MBNL1, MBNL2, and RAVR1. We would propose that these factors may be important

in EBV-regulated processing and the stability of viral latent mRNAs as well as cellular mRNAs during infection.

Another important category of mRNAs regulated by EBV at the isoform level is the group of transcripts encoding DNA binding proteins. Previous reports have identified IRF5 to be a gene whose mRNA transcripts undergo alternative isoform changes as a result of latent EBV infection (86). While our arrays did not differentially detect those specific IRF5 isoforms, we did identify several transcription factors that were regulated at the isoform level, including STAT5B, MEF2B, EGR1, XBP1, and TCF4. The last two proteins have been shown to bind the promoter of the key EBV lytic switch gene, Z (130, 135). Interestingly, EBV latent infection suppressed isoforms of XBP1 and TCF4 capable of activating the Z promoter.

XBP1 is a basic leucine zipper protein and plasma cell differentiation factor that responds to the unfolded protein response (UPR) pathway and ER stress. IRE1 is embedded in the ER membrane, and upon ER stress, IRE1 endonucleolytically cleaves XBP1u transcripts, leading to the generation of an mRNA, termed XBP1s, that encodes the active XBP1 transcription factor (155). XBP1s has been shown to activate transcription at the Z promoter (8, 93, 130), and recently, EBV lytic reactivation has been shown to be activated by ER stress, suggesting a role for XBP1s in this process (134).

SplicerEX detected an alternative 5' initiation of XBP1 mRNA in LCLs relative to resting B cells, and we additionally queried the IRE1-dependent splicing of XBP1 in

LCLs. While XBP1u was dramatically upregulated, the levels of XBP1s were suppressed, suggesting inhibition of the IRE1-mediated splicing during latent infection. Importantly, we demonstrated that inhibition of IRE1 endonuclease activity prevented XBP1 splicing and Z induction downstream of BCR activation. It remains unclear whether the alternative 5' initiation of XBP1 directly contributes to the high XBP1u/XBP1s ratio in LCLs. It has been suggested that alternative splicing may in some cases be dependent on promoter usage (85). In the case of XBP1, the alternative 5' initiation could potentially alter the secondary structure of the mRNA in such a way as to suppress the ability of IRE1 to bind and/or cleave the XBP1u transcript. Alternatively, EBV latent infection may broadly antagonize the IRE1-mediated splicing pathway, perhaps through EBNA3C (42), resulting in suppression of XBP1s and, therefore, suppression of Z transcription and lytic reactivation.

A complementary EBV-regulated isoform change is in an mRNA encoding the E-box binding protein TCF4 (also known as E2-2). EBV latent infection suppressed the expression of full-length TCF4 (TCF4-FL), while it retained expression of a shorter isoform (TCF4-S). TCF4-FL contains a basic helix-loop-helix DNA binding and dimerization domain along with two activation domains (AD1 and AD2), while TCF4-S lacks AD1. TCF4 shares high amino acid sequence homology in its AD1 with AD1 from TCF3 (also called E2A), particularly in a conserved motif that recruits the transcriptional coactivators CBP and p300 (53, 54). We would therefore expect that TCF4-FL, which

contains AD1 and is downregulated in LCLs, promotes transcription at the Z promoter through recruitment of CBP and p300. The absence of AD1 in TCF4-S is likely to reduce its ability to activate transcription at target promoters (125), including Z, thereby preventing lytic reactivation. Furthermore, TCF4-S retains DNA binding activity and may potentially bind to the E-boxes in the Z promoter (135) and act as a dominant negative inhibitor of Z transcription by blocking the activity of residual TCF4-FL in LCLs.

The data presented here demonstrate an important role for mRNA isoform regulation in the maintenance of EBV latency. We have shown that SplicerEX is capable of predicting isoform changes that are functionally relevant to EBV infection of primary B cells, specifically, in the regulation of lytic reactivation. We anticipate that in the future, analysis of the RNA binding proteins that are regulated by EBV may provide clues into how viral transcripts are spliced and processed during latency as well as how global isoform changes in cellular mRNAs may be regulated. Our studies provide a new angle in understanding the virus-host interactions with regard to mRNA processing. This important area of study is now significantly more tractable with new technologies and computational approaches, and our study has provided us with a solid foundation for the detailed analysis of mRNA isoform regulation in EBV transformation.

## 5. Conclusions and Future Directions

This dissertation has explored several nodes of host-pathogen interactions during EBV latent infection of B cells. It addressed aspects with the interactions of viral latency proteins with the host factor RBPJ. Specifically the strategies for structural biological approaches and an interrogation of the protein-protein interactions with RBPJ at a molecular level were delineated. Additionally it detailed the effects of viral latent infection on a broader scale by examining modulation of host mRNA isoform choice. Further, these broad changes informed a novel mechanism for the regulation of a critical viral process of the switch between latent and lytic replication.

First, we used a structural biology approach to investigate the nuances of EBNA2 and EBNA3 protein-protein interactions with RBPJ. While structural studies of RBPJ have been successful in the context of canonical Notch signaling, there is a dearth of structural information pertaining to its interactions with EBV latency proteins. This is in part due to the difficulties with expression and purification of these proteins. While we experienced some of the same difficulties in our work, we have regardless taken steps to improve the methods. We have identified an EBNA3 construct and lysis protocol that, while it still yields mostly insoluble protein, produces enough useful material to perform informative biochemical assays. Furthermore, we have devised a novel strategy to overcome the hurdles of EBNA2 expression and purification. The chimeric E2-ANK

construct provides an easily purified and biologically relevant surrogate for EBNA2. Finally, we have undertaken crystallography trials for peptide soaking and co-crystallization of RBPJ with viral protein. While these trials have yet to yield diffraction quality crystals, these approaches offer little *a priori* guidance for experimental design. However, further experimentation may yield viable results.

Second, we began to delineate the nuances of the interactions between RBPJ and the EBNA2 and EBNA3 proteins. Using the E2-ANK chimera we have begun to more distinctly map the surfaces on RBPJ involved in EBNA2 and NotchIC binding. This has confirmed the hypothesis of overlap at the hydrophobic pocket on RBPJ. It has also revealed binding surfaces unique to each of EBNA2 and NotchIC. This data can begin to illuminate the biochemical discrepancies between Notch signaling and EBNA2 expression during EBV infection and potentially explain physiological consequences of these differences. In addition we have identified both common and unique requirements for the interactions of EBNA2 and the EBNA3 proteins with RBPJ. This has reinforced the theory of mutual exclusivity in RBPJ binding and also further suggests a unique mechanism of binding between the proteins. This is the first instance where EBNA2 and EBNA3 binding have been separated from each other.

Third we examined the changes in host gene expression upon EBV latent infection, specifically at the level of mRNA isoform choice. This revealed changes in isoform choice independent from overall expression changes. Further, changes in

isoform choice were enriched for nucleic acid binding proteins, including transcription and splicing factors which have the ability to regulate gene expression on a global scale. Finally, we identified two host genes, XBP1 and TCF4, whose isoform choice is modulated by EBV infection and also reinforces repression of the viral lytic Z promoter in order to maintain a latent infection. Regulation of isoform choice represents a novel mechanism of modulation of the switch between latent and lytic replication.

This dissertation builds on what is known about EBV latent infection in B cells. Both interactions with RBPJ and regulation of the latent/lytic switch are critical aspects of the EBV lifecycle. The data presented here begins to reveal some of the hidden intricacies of these processes. Disruption of either of these can be critically deleterious for latent infection. As such, they may provide targets for directed therapies against certain EBV-associated malignancies.

## ***5.1 EBV Regulation of Transcription via RBPJ and Other Host Factors***

The critical role of EBNA2 and EBNA3 interactions with RBPJ during EBV latent infection of B cells has been well documented (31, 74, 88, 115, 144). While the importance of these protein-protein interactions have been demonstrated both *in vivo* and *in vitro*, one cannot discount the contributions of other host factors that may be involved in transcriptional regulation alongside the EBNA2 and EBNA3 proteins and RBPJ. While EBNA2 ChIP sites are nearly exclusively found at RBPJ co-occupied sites, the majority of EBNA3A or EBNA3C ChIP sites do not coincide with loci also occupied by RBPJ (59,

123). This indicates that other DNA binding factors must be involved in these EBNA3-containing transcriptional complexes. Indeed, EBNA3 ChIP-seq sites show enrichment for a PU.1 recognition motif at these loci (92), as well as co-occupation by factors such as PU.1, IRF4, BATF and RUNX3 (59, 123). Furthermore, co-occupancy of multiple co-factors with the EBNA3 proteins is associated with strong versus weak promoter elements, based on the histone modification profile at these loci, and the particular combination of co-factors is associated with the strength of the enhancer element at which they are found (59, 123). This represents strong evidence toward combinatorial regulation of transcription via the EBNA3 proteins at composite DNA elements which could result in different transcriptional effects based on the particular target.

There are two potential mechanisms for incorporation of additional DNA binding proteins in EBNA3-containing complexes when considering the role of RBPJ across the EBNA3 proteins. First, RBPJ binding precludes binding of other factors to the EBNA3 proteins presenting an “either/or” scenario where the EBNA3 proteins are targeted to DNA by *either* RBPJ *or* an additional DNA binding protein. For example, RBPJ is depleted in EBNA3C ChIP sites which are co-occupied by IRF4 and RUNX3 or BATF and RUNX3 (59), suggesting that binding of either of these pairs of factors to EBNA3C could be excluding RBPJ from the complex. Second, the EBNA3 proteins are capable of associating with DNA binding factors in addition to RBPJ simultaneously leading to combinatorial effects. The latter mechanism is more intriguing as it allows for

greater diversity in EBNA3-mediated transcriptional regulation. In addition to incorporating composite recognition elements in DNA binding, it can also allow the possibility of DNA looping, with the EBNA3 proteins acting as part of the bridge spanning enhancer and promoter elements. Indeed, this model has been proposed previously (59, 92). Furthermore, it has also been shown previously that RBPJ and EBNA3C can be found by ChIP to occupy the same loci, but with RBPJ more distal to DNA than EBNA3C (59). This suggests the presence of an additional DNA binding protein acting as the tether between EBNA3C and the DNA.

By mapping the interactions of viral latency proteins with RBPJ we have begun to unravel the knot that is viral and host co-regulation of transcription during infection. We have done this with RBPJ due to its known critical role in EBV latent infection. A similar biochemical description and mapping of the protein-protein interactions of the EBNA2 and EBNA3 proteins with additional DNA binding factors will begin to reveal some of the additional layers of EBV-mediated transcriptional regulation. This can inform on which of the abovementioned mechanisms are at work at specific genomic loci during infection. Additionally, this dissertation has not explored the potential for unique sets of binding partners among the EBNA3A, -3B, and -3C proteins. Even if the mechanism of binding to RBPJ is shared between the three proteins, this may not hold true for other factors, further adding to the variety and complexity of EBNA3-mediated transcriptional regulation.

## **5.2 Alternative EBNA Expression Strategies**

Solubility is commonly a concern when expressing and purifying protein for crystallographic studies. This can be particularly problematic in the case of mammalian proteins with bacterial expression systems. While bacterial expression is generally simple, efficient, and cost-effective it lacks the chaperones and post-translational modifications that mammalian proteins may require for proper folding and function. With this in mind, other expression systems can be considered. Baculovirus infections of insect cells have the potential to circumvent these potential problems for the EBNA2 and EBNA3 proteins. Expression in eukaryotic cells can provide the necessary factors to promote proper protein folding. Furthermore, if a mammalian expression system—or even a B-cell environment—is required, we can take advantage of a DG75 cell line created by the Kempkes laboratory that has been knocked out for RBPJ expression. There are a number of proteins in these cells that could interact with EBNA2 or EBNA3hd *in vivo*, predicating more extensive and stringent purification procedures. However the absence of RBPJ removes one of the major purification hurdles as well as ensures there will be no interference in other biochemical interaction experiments in which the purified protein may be used.

## **5.3 Therapeutic Inhibition of RBPJ-EBNA Interactions**

As described here, EBNA2 and EBNA3 interaction with RBPJ is critical to continued proliferation and maintenance of a latent infection in LCLs. This could be the

case for several EBV malignancies, particularly those associated with Latency IIb and III where those proteins are expressed. Targeting and disrupting these interactions serve as the basis for a potential therapeutic against certain EBV-associated disease. Most therapeutic small molecules inhibit enzymatic activity of target molecules. There are however examples of small molecules which disrupt protein-protein interactions, most notably the nutlins which block the interaction of MDM2 with p53 (158). The hydrophobic pocket of RBPJ is a potential target for one such molecule. We know that this pocket is critical with regard to EBNA2 binding, and may even play some role in EBNA3 binding as well. Additionally, as we know that the  $\Phi W\Phi P$  motif mediates the interactions here, efforts could be focused on molecules which resemble this moiety.

Targeting the hydrophobic pocket of RBPJ would also disrupt normal Notch signaling as well as other endogenous transcriptional regulators that utilize it. Indeed, an EBNA2 peptide has already shown potential as a therapeutic in this regard (31). By gaining a more complete map of the surfaces on RBPJ involved in interactions with viral latency proteins we may be able to apply targeted drug design to those regions uniquely utilized by the virus and not by other Notch components. Those residues which block E2-ANK binding but not RAMANK binding could serve to initially direct these efforts. A structural biology-based approach for highly targeted drug design in this context has not been done.

## **5.4 Regulation of TCF4 Alternative 5' Initiation**

The studies performed here clearly indicate that TCF4 preferentially undergoes alternative 5' initiation in B cells upon EBV latent infection. Rescue of the full-length protein by ectopic expression removes repression of the Z promoter and lytic replication. This makes the regulation of TCF4 transcription of particular interest. The Allday group published microarray data of infection of a Burkitt's Lymphoma cell line, BL31, in 2010 which included isoform-specific probes, including for TCF4 isoforms (148). Their infections included wild-type virus as well as various EBNA3 knockouts. Their data showed that TCF4-FL was reduced upon infection, similar to what we observed in Chapter 4. When infecting with a virus knocked out for the entire EBNA3 locus or specifically EBNA3C that reduction in TCF4-FL was partially ablated while EBNA3A and EBNA3B knockouts had TCF4-FL expression levels comparable to wild-type virus. This suggests that EBNA3C specifically is at least partially responsible for the preference away from TCF4-FL isoform choice during latent infection.

As EBNA3C is capable of recruiting HDACs to promoters to repress transcription it would be interesting to examine the histone modification status of the TCF4-FL promoter in the presence and absence of EBNA3C. The EBNA3 proteins are also capable of repressing transcription through the polycomb repressive complexes (PRC1 and PRC2) via recruitment of C-terminal binding protein (CtBP). However from the Allday data we see that viruses containing EBNA3 proteins knocked out for CtBP

interactions specifically are still capable of repressing TCF4-FL. This suggests that EBNA3C-mediated repression is occurring in a polycomb-independent manner.

### ***5.5 Relation of IRE1-Dependent Splicing of XBP1 with Alternative Initiation Isoforms and EBV Infection***

The alternative splicing event that determines the ability of XBP1 to regulate the Z promoter was not actually queried by the SplicerEX algorithm on our microarrays. The IRE1-dependent splicing event was queried independently by us following identification of XBP1 as utilizing an alternative 5' initiation by SplicerEX. EBV clearly suppresses the long isoform and promotes the short isoform of XBP1 as well as suppresses the IRE1-spliced isoform, however the connection between these two processes is unclear. Identifying whether the long and short isoforms are spliced with different efficiencies in an EBV-independent system would inform on whether this is a virally regulated process or is an inherent property of the different mRNA molecules. If the suppression of XBP1 splicing by IRE1 is indeed modulated by the virus a prudent course of action would be to identify the viral genes responsible for the suppression of this pathway, and the mechanisms by which it is accomplished. XBP1 and IRE1 form part of the UPR as well as play roles in plasma cell differentiation, a process which EBV co-opts during infection. If EBV were to suppress these pathways in an established latent infection this would be very relevant to how the virus works to suppress reactivation of the Z promoter and lytic replication.

## **5.6 Lytic Induction Therapeutics**

Lytic induction has already been an area targeted by researchers as a potential therapeutic for EBV-associated malignancies (34). A deeper understanding of initiation of lytic replication and regulation of the Z promoter is important to our ability to exploit this process in the clinic. The novel mechanisms of Z promoter repression described here could lead to additional knowledge that may be able to be applied to therapeutic development. The directions proposed in Sections 5.3 and 5.4 here could provide this information. For example, if it were determined that TCF4-FL transcription is suppressed by EBNA3C recruitment of HDAC1 to its promoter, certain HDAC inhibitors may be applicable, especially considering investigations into clinical applications have already been initiated for several of these. Such therapies could result in substantial improvements to the current standards of care for certain EBV-associated malignancies.

## **5.7 Conclusion**

Despite its generally benign character, EBV remains a public health concern throughout the world. The extreme prevalence of this virus among the global population represents the potential for disease development in nearly every individual. Decades of research has brought this virus from complete anonymity to a very extensively studied pathogen. Yet much remains to be discovered concerning some of the basic underlying aspects of infection. This dissertation represents an extension of our

knowledge of this virus into the biochemistry of protein-protein interactions essential for a latent proliferative infection as well as the regulation of host gene expression and its role in maintenance of latency. Host-pathogen interactions are always critical to understanding the role of an infectious entity in disease. New insights into these interactions will continue to expand our understanding of human disease and will get us closer to addressing the public health concerns associated with EBV.

## References

1. **Adamson, A. L., and S. Kenney.** 1999. The Epstein-Barr virus BZLF1 protein interacts physically and functionally with the histone acetylase CREB-binding protein. *J Virol* **73**:6551-6558.
2. **Alber, G., K. M. Kim, P. Weiser, C. Riesterer, R. Carsetti, and M. Reth.** 1993. Molecular mimicry of the antigen receptor signalling motif by transmembrane proteins of the Epstein-Barr virus and the bovine leukaemia virus. *Current biology : CB* **3**:333-339.
3. **Alfieri, C., M. Birkenbach, and E. Kieff.** 1991. Early events in Epstein-Barr virus infection of human B lymphocytes. *Virology* **181**:595-608.
4. **Allday, M. J., Q. Bazot, and R. E. White.** 2015. The EBNA3 Family: Two Oncoproteins and a Tumour Suppressor that Are Central to the Biology of EBV in B Cells. *Current topics in microbiology and immunology* **391**:61-117.
5. **Anderton, E., J. Yee, P. Smith, T. Crook, R. E. White, and M. J. Allday.** 2008. Two Epstein-Barr virus (EBV) oncoproteins cooperate to repress expression of the proapoptotic tumour-suppressor Bim: clues to the pathogenesis of Burkitt's lymphoma. *Oncogene* **27**:421-433.
6. **Arnett, K. L., M. Hass, D. G. McArthur, M. X. Ilagan, J. C. Aster, R. Kopan, and S. C. Blacklow.** 2010. Structural and mechanistic insights into cooperative assembly of dimeric Notch transcription complexes. *Nature structural & molecular biology* **17**:1312-1317.
7. **Bemmo, A., D. Benovoy, T. Kwan, D. J. Gaffney, R. V. Jensen, and J. Majewski.** 2008. Gene expression and isoform variation analysis using Affymetrix Exon Arrays. *BMC genomics* **9**:529.
8. **Bhende, P. M., S. J. Dickerson, X. Sun, W. H. Feng, and S. C. Kenney.** 2007. X-box-binding protein 1 activates lytic Epstein-Barr virus gene expression in combination with protein kinase D. *J Virol* **81**:7363-7370.
9. **Bhende, P. M., W. T. Seaman, H. J. Delecluse, and S. C. Kenney.** 2005. BZLF1 activation of the methylated form of the BRLF1 immediate-early promoter is regulated by BZLF1 residue 186. *J Virol* **79**:7338-7348.
10. **Bhende, P. M., W. T. Seaman, H. J. Delecluse, and S. C. Kenney.** 2004. The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome. *Nature genetics* **36**:1099-1104.
11. **Brennan, C. M., and J. A. Steitz.** 2001. HuR and mRNA stability. *Cellular and molecular life sciences : CMLS* **58**:266-277.

12. **Cahir-McFarland, E. D., K. Carter, A. Rosenwald, J. M. Giltane, S. E. Henrickson, L. M. Staudt, and E. Kieff.** 2004. Role of NF-kappa B in cell survival and transcription of latent membrane protein 1-expressing or Epstein-Barr virus latency III-infected cells. *J Virol* **78**:4108-4119.
13. **Calderwood, M. A., S. Lee, A. M. Holthaus, S. C. Blacklow, E. Kieff, and E. Johannsen.** 2011. Epstein-Barr virus nuclear protein 3C binds to the N-terminal (NTD) and beta trefoil domains (BTD) of RBP/CSL; only the NTD interaction is essential for lymphoblastoid cell growth. *Virology* **414**:19-25.
14. **Canaves, J. M., R. Page, I. A. Wilson, and R. C. Stevens.** 2004. Protein biophysical properties that correlate with crystallization success in *Thermotoga maritima*: maximum clustering strategy for structural genomics. *Journal of molecular biology* **344**:977-991.
15. **Chang, T. Y., Y. H. Wu, C. C. Cheng, and H. W. Wang.** 2011. Differentially regulated splice variants and systems biology analysis of Kaposi's sarcoma-associated herpesvirus-infected lymphatic endothelial cells. *Nucleic Acids Res* **39**:6970-6985.
16. **Choi, S. H., T. E. Wales, Y. Nam, D. J. O'Donovan, P. Sliz, J. R. Engen, and S. C. Blacklow.** 2012. Conformational locking upon cooperative assembly of notch transcription complexes. *Structure* **20**:340-349.
17. **Chung, C. N., Y. Hamaguchi, T. Honjo, and M. Kawaichi.** 1994. Site-directed mutagenesis study on DNA binding regions of the mouse homologue of Suppressor of Hairless, RBP-J kappa. *Nucleic Acids Res* **22**:2938-2944.
18. **Cohen, J. I., F. Wang, J. Mannick, and E. Kieff.** 1989. Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. *Proc Natl Acad Sci U S A* **86**:9558-9562.
19. **Collins, K. J., Z. Yuan, and R. A. Kovall.** 2014. Structure and function of the CSL-KyoT2 corepressor complex: a negative regulator of Notch signaling. *Structure* **22**:70-81.
20. **Coppotelli, G., N. Mughal, S. Callegari, R. Sompallae, L. Caja, M. S. Luijsterburg, N. P. Dantuma, A. Moustakas, and M. G. Masucci.** 2013. The Epstein-Barr virus nuclear antigen-1 reprograms transcription by mimicry of high mobility group A proteins. *Nucleic Acids Res* **41**:2950-2962.
21. **Countryman, J., H. Jenson, R. Seibl, H. Wolf, and G. Miller.** 1987. Polymorphic proteins encoded within BZLF1 of defective and standard Epstein-Barr viruses disrupt latency. *J Virol* **61**:3672-3679.
22. **Daigle, D., L. Gradoville, D. Tuck, V. Schulz, R. Wang'ondou, J. Ye, K. Gorres, and G. Miller.** 2011. Valproic acid antagonizes the capacity of other histone deacetylase inhibitors to activate the Epstein-barr virus lytic cycle. *J Virol* **85**:5628-5643.

23. **Davies, A. H., R. J. Grand, F. J. Evans, and A. B. Rickinson.** 1991. Induction of Epstein-Barr virus lytic cycle by tumor-promoting and non-tumor-promoting phorbol esters requires active protein kinase C. *J Virol* **65**:6838-6844.
24. **Del Bianco, C., A. Vedenko, S. H. Choi, M. F. Berger, L. Shokri, M. L. Bulyk, and S. C. Blacklow.** 2010. Notch and MAML-1 complexation do not detectably alter the DNA binding specificity of the transcription factor CSL. *PLoS One* **5**:e15034.
25. **Dembowski, J. A., and P. J. Grabowski.** 2009. The CUGBP2 splicing factor regulates an ensemble of branchpoints from perimeter binding sites with implications for autoregulation. *PLoS Genet* **5**:e1000595.
26. **Devergne, O., E. Hatzivassiliou, K. M. Izumi, K. M. Kaye, M. F. Kleijnen, E. Kieff, and G. Mosialos.** 1996. Association of TRAF1, TRAF2, and TRAF3 with an Epstein-Barr virus LMP1 domain important for B-lymphocyte transformation: role in NF-kappaB activation. *Mol Cell Biol* **16**:7098-7108.
27. **Dou, S., X. Zeng, P. Cortes, H. Erdjument-Bromage, P. Tempst, T. Honjo, and L. D. Vales.** 1994. The recombination signal sequence-binding protein RBP-2N functions as a transcriptional repressor. *Mol Cell Biol* **14**:3310-3319.
28. **Dunmire, S. K., K. A. Hogquist, and H. H. Balfour.** 2015. Infectious Mononucleosis. *Current topics in microbiology and immunology* **390**:211-240.
29. **Ellis, A. L., Z. Wang, X. Yu, and J. E. Mertz.** 2010. Either ZEB1 or ZEB2/SIP1 can play a central role in regulating the Epstein-Barr virus latent-lytic switch in a cell-type-specific manner. *J Virol* **84**:6139-6152.
30. **Epstein, M., B. Achong, and Y. Barr.** 1964. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet* **1**.
31. **Farrell, C. J., J. M. Lee, E. C. Shin, M. Cebrat, P. A. Cole, and S. D. Hayward.** 2004. Inhibition of Epstein-Barr virus-induced growth proliferation by a nuclear antigen EBNA2-TAT peptide. *Proc Natl Acad Sci U S A* **101**:4625-4630.
32. **Farrell, P. J., D. T. Rowe, C. M. Rooney, and T. Kouzarides.** 1989. Epstein-Barr virus BZLF1 trans-activator specifically binds to a consensus AP-1 site and is related to c-fos. *The EMBO journal* **8**:127-132.
33. **Faustino, N. A., and T. A. Cooper.** 2003. Pre-mRNA splicing and human disease. *Genes Dev* **17**:419-437.
34. **Feng, W. H., G. Hong, H. J. Delecluse, and S. C. Kenney.** 2004. Lytic induction therapy for Epstein-Barr virus-positive B-cell lymphomas. *J Virol* **78**:1893-1902.
35. **Flemington, E., and S. H. Speck.** 1990. Autoregulation of Epstein-Barr virus putative lytic switch gene BZLF1. *J Virol* **64**:1227-1232.
36. **Flemington, E., and S. H. Speck.** 1990. Identification of phorbol ester response elements in the promoter of Epstein-Barr virus putative lytic switch gene BZLF1. *J Virol* **64**:1217-1226.

37. **Friberg, A., S. Thumann, J. Hennig, P. Zou, E. Nossner, P. D. Ling, M. Sattler, and B. Kempkes.** 2015. The EBNA-2 N-Terminal Transactivation Domain Folds into a Dimeric Structure Required for Target Gene Activation. *PLoS Pathog* **11**:e1004910.
38. **Friedmann, D. R., and R. A. Kovall.** 2010. Thermodynamic and structural insights into CSL-DNA complexes. *Protein science : a publication of the Protein Society* **19**:34-46.
39. **Friedmann, D. R., J. J. Wilson, and R. A. Kovall.** 2008. RAM-induced allostery facilitates assembly of a notch pathway active transcription complex. *J Biol Chem* **283**:14781-14791.
40. **Fuchs, K. P., G. Bommer, E. Dumont, B. Christoph, M. Vidal, E. Kremmer, and B. Kempkes.** 2001. Mutational analysis of the J recombination signal sequence binding protein (RBP-J)/Epstein-Barr virus nuclear antigen 2 (EBNA2) and RBP-J/Notch interaction. *European journal of biochemistry* **268**:4639-4646.
41. **Garcia-Blanco, M. A., and B. R. Cullen.** 1991. Molecular basis of latency in pathogenic human viruses. *Science* **254**:815-820.
42. **Garrido, J. L., S. Maruo, K. Takada, and A. Rosendorff.** 2009. EBNA3C interacts with Gadd34 and counteracts the unfolded protein response. *Virology* **6**:231.
43. **Gordadze, A. V., R. Peng, J. Tan, G. Liu, R. Sutton, B. Kempkes, G. W. Bornkamm, and P. D. Ling.** 2001. Notch1IC partially replaces EBNA2 function in B cells immortalized by Epstein-Barr virus. *J Virol* **75**:5899-5912.
44. **Grabusic, K., S. Maier, A. Hartmann, A. Mantik, W. Hammerschmidt, and B. Kempkes.** 2006. The CR4 region of EBNA2 confers viability of Epstein-Barr virus-transformed B cells by CBF1-independent signalling. *J Gen Virol* **87**:3169-3176.
45. **Gross, H., S. Barth, R. D. Palermo, A. Mamiani, C. Hennard, U. Zimmer-Strobl, M. J. West, E. Kremmer, and F. A. Grasser.** 2010. Asymmetric Arginine dimethylation of Epstein-Barr virus nuclear antigen 2 promotes DNA targeting. *Virology* **397**:299-310.
46. **Hagemeier, S. R., E. A. Barlow, Q. Meng, and S. C. Kenney.** 2012. The cellular ataxia telangiectasia-mutated kinase promotes Epstein-Barr virus lytic reactivation in response to multiple different types of lytic reactivation-inducing stimuli. *J Virol* **86**:13360-13370.
47. **Harada, S., and E. Kieff.** 1997. Epstein-Barr virus nuclear protein LP stimulates EBNA-2 acidic domain-mediated transcriptional activation. *J Virol* **71**:6611-6618.
48. **Hausser, J., A. P. Syed, B. Bilen, and M. Zavolan.** 2013. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome research* **23**:604-615.

49. **Henkel, T., P. D. Ling, S. D. Hayward, and M. G. Peterson.** 1994. Mediation of Epstein-Barr virus EBNA2 transactivation by recombination signal-binding protein J kappa. *Science* **265**:92-95.
50. **Homa, N. J., R. Salinas, E. Forte, T. J. Robinson, M. A. Garcia-Blanco, and M. A. Luftig.** 2013. Epstein-Barr virus induces global changes in cellular mRNA isoform usage that are important for the maintenance of latency. *J Virol* **87**:12291-12301.
51. **Hsieh, J. J., and S. D. Hayward.** 1995. Masking of the CBF1/RBPJ kappa transcriptional repression domain by Epstein-Barr virus EBNA2. *Science* **268**:560-563.
52. **Huen, D. S., S. A. Henderson, D. Croom-Carter, and M. Rowe.** 1995. The Epstein-Barr virus latent membrane protein-1 (LMP1) mediates activation of NF-kappa B and cell surface phenotype via two effector regions in its carboxy-terminal cytoplasmic domain. *Oncogene* **10**:549-560.
53. **Hyndman, B. D., P. Thompson, R. Bayly, G. P. Cote, and D. P. LeBrun.** 2012. E2A proteins enhance the histone acetyltransferase activity of the transcriptional co-activators CBP and p300. *Biochim Biophys Acta* **1819**:446-453.
54. **Hyndman, B. D., P. Thompson, C. M. Denis, S. Chitayat, R. Bayly, S. P. Smith, and D. P. LeBrun.** 2012. Mapping acetylation sites in E2A identifies a conserved lysine residue in activation domain 1 that promotes CBP/p300 recruitment and transcriptional activation. *Biochim Biophys Acta* **1819**:375-381.
55. **Ikeda, M., A. Ikeda, L. C. Longan, and R. Longnecker.** 2000. The Epstein-Barr virus latent membrane protein 2A PY motif recruits WW domain-containing ubiquitin-protein ligases. *Virology* **268**:178-191.
56. **Iwakiri, D., and K. Takada.** 2004. Phosphatidylinositol 3-kinase is a determinant of responsiveness to B cell antigen receptor-mediated Epstein-Barr virus activation. *J Immunol* **172**:1561-1566.
57. **James, L. I., J. E. Beaver, N. W. Rice, and M. L. Waters.** 2013. A synthetic receptor for asymmetric dimethyl arginine. *Journal of the American Chemical Society* **135**:6450-6455.
58. **Jenkins, P. J., U. K. Binne, and P. J. Farrell.** 2000. Histone acetylation and reactivation of Epstein-Barr virus from latency. *J Virol* **74**:710-720.
59. **Jiang, S., B. Willox, H. Zhou, A. M. Holthaus, A. Wang, T. T. Shi, S. Maruo, P. V. Kharchenko, E. C. Johannsen, E. Kieff, and B. Zhao.** 2014. Epstein-Barr virus nuclear antigen 3C binds to BATF/IRF4 or SPI1/IRF4 composite sites and recruits Sin3A to repress CDKN2A. *Proc Natl Acad Sci U S A* **111**:421-426.
60. **Johannsen, E., C. L. Miller, S. R. Grossman, and E. Kieff.** 1996. EBNA-2 and EBNA-3C extensively and mutually exclusively associate with RBPJkappa in Epstein-Barr virus-transformed B lymphocytes. *J Virol* **70**:4179-4183.

61. **Johnson, S. E., M. X. Ilagan, R. Kopan, and D. Barrick.** 2010. Thermodynamic analysis of the CSL x Notch interaction: distribution of binding energy of the Notch RAM region to the CSL beta-trefoil domain and the mode of competition with the viral transactivator EBNA2. *J Biol Chem* **285**:6681-6692.
62. **Kalchschmidt, J. S., A. C. Gillman, K. Paschos, Q. Bazot, B. Kempkes, and M. J. Allday.** 2016. EBNA3C Directs Recruitment of RBPJ (CBF1) to Chromatin during the Process of Gene Repression in EBV Infected B Cells. *PLoS Pathog* **12**:e1005383.
63. **Kao, H. Y., P. Ordentlich, N. Koyano-Nakagawa, Z. Tang, M. Downes, C. R. Kintner, R. M. Evans, and T. Kadesch.** 1998. A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. *Genes Dev* **12**:2269-2277.
64. **Kaye, K. M., O. Devergne, J. N. Harada, K. M. Izumi, R. Yalamanchili, E. Kieff, and G. Mosialos.** 1996. Tumor necrosis factor receptor associated factor 2 is a mediator of NF-kappa B activation by latent infection membrane protein 1, the Epstein-Barr virus transforming protein. *Proc Natl Acad Sci U S A* **93**:11085-11090.
65. **Keene, J. D.** 2007. RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* **8**:533-543.
66. **Keller, S. A., E. J. Schattner, and E. Cesarman.** 2000. Inhibition of NF-kappaB induces apoptosis of KSHV-infected primary effusion lymphoma cells. *Blood* **96**:2537-2542.
67. **Kempkes, B., M. Pawlita, U. Zimmer-Strobl, G. Eissner, G. Laux, and G. W. Bornkamm.** 1995. Epstein-Barr virus nuclear antigen 2-estrogen receptor fusion proteins transactivate viral and cellular genes and interact with RBP-J kappa in a conditional fashion. *Virology* **214**:675-679.
68. **Kieff, E., and A. Rickinson.** 2007. Epstein-Barr virus and its replication, p. 2603-2654. *In* D. M. Knipe and P. M. Howley (ed.), *Fields Virology*, 5 ed, vol. 2.
69. **Kieser, A., and K. R. Sterz.** 2015. The Latent Membrane Protein 1 (LMP1). *Current topics in microbiology and immunology* **391**:119-149.
70. **Kohlhof, H., F. Hampel, R. Hoffmann, H. Burtscher, U. H. Weidle, M. Holz, D. Eick, U. Zimmer-Strobl, and L. J. Strobl.** 2009. Notch1, Notch2, and Epstein-Barr virus-encoded nuclear antigen 2 signaling differentially affects proliferation and survival of Epstein-Barr virus-infected B cells. *Blood* **113**:5506-5515.
71. **Kovall, R. A., and S. C. Blacklow.** 2010. Mechanistic insights into Notch receptor signaling from structural and biochemical studies. *Current topics in developmental biology* **92**:31-71.
72. **Kuroda, K., H. Han, S. Tani, K. Tanigaki, T. Tun, T. Furukawa, Y. Taniguchi, H. Kurooka, Y. Hamada, S. Toyokuni, and T. Honjo.** 2003. Regulation of

- marginal zone B cell development by MINT, a suppressor of Notch/RBP-J signaling pathway. *Immunity* **18**:301-312.
73. **Lee, J. M., K. H. Lee, C. J. Farrell, P. D. Ling, B. Kempkes, J. H. Park, and S. D. Hayward.** 2004. EBNA2 is required for protection of latently Epstein-Barr virus-infected B cells against specific apoptotic stimuli. *J Virol* **78**:12694-12697.
  74. **Lee, S., S. Sakakibara, S. Maruo, B. Zhao, M. A. Calderwood, A. M. Holthaus, C. Y. Lai, K. Takada, E. Kieff, and E. Johannsen.** 2009. Epstein-Barr virus nuclear protein 3C domains necessary for lymphoblastoid cell growth: interaction with RBP-Jkappa regulates TCL1. *J Virol* **83**:12368-12377.
  75. **Lejeune, F., Y. Cavaloc, and J. Stevenin.** 2001. Alternative splicing of intron 3 of the serine/arginine-rich protein 9G8 gene. Identification of flanking exonic splicing enhancers and involvement of 9G8 as a trans-acting factor. *J Biol Chem* **276**:7850-7858.
  76. **Lerner, M. R., N. C. Andrews, G. Miller, and J. A. Steitz.** 1981. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc Natl Acad Sci U S A* **78**:805-809.
  77. **Ling, P. D., D. R. Rawlins, and S. D. Hayward.** 1993. The Epstein-Barr virus immortalizing protein EBNA-2 is targeted to DNA by a cellular enhancer-binding protein. *Proc Natl Acad Sci U S A* **90**:9237-9241.
  78. **Liu, P., S. Liu, and S. H. Speck.** 1998. Identification of a negative cis element within the ZII domain of the Epstein-Barr virus lytic switch BZLF1 gene promoter. *J Virol* **72**:8230-8239.
  79. **Liu, S., A. M. Borrás, P. Liu, G. Suske, and S. H. Speck.** 1997. Binding of the ubiquitous cellular transcription factors Sp1 and Sp3 to the ZI domains in the Epstein-Barr virus lytic switch BZLF1 gene promoter. *Virology* **228**:11-18.
  80. **Liu, S., P. Liu, A. Borrás, T. Chatila, and S. H. Speck.** 1997. Cyclosporin A-sensitive induction of the Epstein-Barr virus lytic switch is mediated via a novel pathway involving a MEF2 family member. *The EMBO journal* **16**:143-153.
  81. **Longnecker, R. M., E. Kieff, and J. I. Cohen.** 2013. Epstein-Barr virus p. 1898-1959. *In* D. M. Knipe and P. M. Howley (ed.), *Fields Virology*, 6th ed, vol. 2. Lippincott, Williams, and Wilkins, Philadelphia.
  82. **Luka, J., B. Kallin, and G. Klein.** 1979. Induction of the Epstein-Barr virus (EBV) cycle in latently infected cells by n-butyrate. *Virology* **94**:228-231.
  83. **Lynch, D. T., J. S. Zimmerman, and D. T. Rowe.** 2002. Epstein-Barr virus latent membrane protein 2B (LMP2B) co-localizes with LMP2A in perinuclear regions in transiently transfected cells. *J Gen Virol* **83**:1025-1035.
  84. **Maier, S., M. Santak, A. Mantik, K. Grabusic, E. Kremmer, W. Hammerschmidt, and B. Kempkes.** 2005. A somatic knockout of CBF1 in a human B-cell line reveals that induction of CD21 and CCR7 by EBNA-2 is strictly

- CBF1 dependent and that downregulation of immunoglobulin M is partially CBF1 independent. *J Virol* **79**:8784-8792.
85. **Mancl, M. E., G. Hu, N. Sangster-Guity, S. L. Olshalsky, K. Hoops, P. Fitzgerald-Bocarsly, P. M. Pitha, K. Pinder, and B. J. Barnes.** 2005. Two discrete promoters regulate the alternatively spliced human interferon regulatory factor-5 isoforms. Multiple isoforms with distinct cell type-specific expression, localization, regulation, and function. *J Biol Chem* **280**:21078-21090.
  86. **Martin, H. J., J. M. Lee, D. Walls, and S. D. Hayward.** 2007. Manipulation of the toll-like receptor 7 signaling pathway by Epstein-Barr virus. *J Virol* **81**:9748-9758.
  87. **Maruo, S., E. Johannsen, D. Illanes, A. Cooper, and E. Kieff.** 2003. Epstein-Barr Virus nuclear protein EBNA3A is critical for maintaining lymphoblastoid cell line growth. *J Virol* **77**:10437-10447.
  88. **Maruo, S., E. Johannsen, D. Illanes, A. Cooper, B. Zhao, and E. Kieff.** 2005. Epstein-Barr virus nuclear protein 3A domains essential for growth of lymphoblasts: transcriptional regulation through RBP-Jkappa/CBF1 is critical. *J Virol* **79**:10171-10179.
  89. **Maruo, S., Y. Wu, S. Ishikawa, T. Kanda, D. Iwakiri, and K. Takada.** 2006. Epstein-Barr virus nuclear protein EBNA3C is required for cell cycle progression and growth maintenance of lymphoblastoid cells. *Proc Natl Acad Sci U S A* **103**:19500-19505.
  90. **Maruo, S., B. Zhao, E. Johannsen, E. Kieff, J. Zou, and K. Takada.** 2011. Epstein-Barr virus nuclear antigens 3C and 3A maintain lymphoblastoid cell growth by repressing p16INK4A and p14ARF expression. *Proc Natl Acad Sci U S A* **108**:1919-1924.
  91. **Mayr, C., and D. P. Bartel.** 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**:673-684.
  92. **McClellan, M. J., C. D. Wood, O. Ojeniyi, T. J. Cooper, A. Kanhere, A. Arvey, H. M. Webb, R. D. Palermo, M. L. Harth-Hertle, B. Kempkes, R. G. Jenner, and M. J. West.** 2013. Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. *PLoS Pathog* **9**:e1003636.
  93. **McDonald, C., C. E. Karstegl, P. Kellam, and P. J. Farrell.** 2010. Regulation of the Epstein-Barr virus Z<sub>p</sub> promoter in B lymphocytes during reactivation from latency. *J Gen Virol* **91**:622-629.
  94. **Miller, C. L., J. H. Lee, E. Kieff, and R. Longnecker.** 1994. An integral membrane protein (LMP2) blocks reactivation of Epstein-Barr virus from latency following surface immunoglobulin crosslinking. *Proc Natl Acad Sci U S A* **91**:772-776.

95. **Miller, G., A. El-Guindy, J. Countryman, J. Ye, and L. Gradoville.** 2007. Lytic cycle switches of oncogenic human gammaherpesviruses. *Adv Cancer Res* **97**:81-109.
96. **Mirandola, L., P. Comi, E. Cobos, W. M. Kast, M. Chiriva-Internati, and R. Chiaramonte.** 2011. Notch-ing from T-cell to B-cell lymphoid malignancies. *Cancer letters* **308**:1-13.
97. **Misquitta-Ali, C. M., E. Cheng, D. O'Hanlon, N. Liu, C. J. McGlade, M. S. Tsao, and B. J. Blencowe.** 2011. Global Profiling and Molecular Characterization of Alternative Splicing Events Misregulated in Lung Cancer. *Mol Cell Biol* **31**:138-150.
98. **Montalvo, E. A., M. Cottam, S. Hill, and Y. J. Wang.** 1995. YY1 binds to and regulates cis-acting negative elements in the Epstein-Barr virus BZLF1 promoter. *J Virol* **69**:4158-4165.
99. **Murata, T., Y. Kondo, A. Sugimoto, D. Kawashima, S. Saito, H. Isomura, T. Kanda, and T. Tsurumi.** 2012. Epigenetic histone modification of Epstein-Barr virus BZLF1 promoter during latency and reactivation in Raji cells. *J Virol* **86**:4752-4761.
100. **Nam, Y., P. Sliz, L. Song, J. C. Aster, and S. C. Blacklow.** 2006. Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes. *Cell* **124**:973-983.
101. **Nam, Y., A. P. Weng, J. C. Aster, and S. C. Blacklow.** 2003. Structural requirements for assembly of the CSL.intracellular Notch1.Mastermind-like 1 transcriptional activation complex. *J Biol Chem* **278**:21232-21239.
102. **Nikitin, P. A., C. M. Yan, E. Forte, A. Bocedi, J. P. Tourigny, R. E. White, M. J. Allday, A. Patel, S. S. Dave, W. Kim, K. Hu, J. Guo, D. Tainter, E. Rusyn, and M. A. Luftig.** 2010. An ATM/Chk2-mediated DNA damage-responsive signaling pathway suppresses Epstein-Barr virus transformation of primary human B cells. *Cell Host Microbe* **8**:510-522.
103. **Norseen, J., A. Thomaе, V. Sridharan, A. Aiyar, A. Schepers, and P. M. Lieberman.** 2008. RNA-dependent recruitment of the origin recognition complex. *The EMBO journal* **27**:3024-3035.
104. **Oswald, F., U. Kostezka, K. Astrahantseff, S. Bourteele, K. Dillinger, U. Zechner, L. Ludwig, M. Wilda, H. Hameister, W. Knochel, S. Liptay, and R. M. Schmid.** 2002. SHARP is a novel component of the Notch/RBP-Jkappa signalling pathway. *The EMBO journal* **21**:5417-5426.
105. **Papandreou, I., N. C. Denko, M. Olson, H. Van Melckebeke, S. Lust, A. Tam, D. E. Solow-Cordero, D. M. Bouley, F. Offner, M. Niwa, and A. C. Koong.** 2011. Identification of an Ire1alpha endonuclease specific inhibitor with cytotoxic activity against human multiple myeloma. *Blood* **117**:1311-1314.

106. **Paschos, K., G. A. Parker, E. Watanatanasup, R. E. White, and M. J. Allday.** 2012. BIM promoter directly targeted by EBNA3C in polycomb-mediated repression by EBV. *Nucleic Acids Res.*
107. **Popielarz, M., Y. Cavaloc, M. G. Mattei, R. Gattoni, and J. Stevenin.** 1995. The gene encoding human splicing factor 9G8. Structure, chromosomal localization, and expression of alternatively processed transcripts. *J Biol Chem* **270**:17830-17835.
108. **Portal, D., B. Zhao, M. A. Calderwood, T. Sommermann, E. Johannsen, and E. Kieff.** 2011. EBV nuclear antigen EBNA1P dismisses transcription repressors NCoR and RBPJ from enhancers and EBNA2 increases NCoR-deficient RBPJ DNA binding. *Proc Natl Acad Sci U S A* **108**:7808-7813.
109. **Price, A. M., and M. A. Luftig.** 2015. To be or not IIb: a multi-step process for Epstein-Barr virus latency establishment and consequences for B cell tumorigenesis. *PLoS Pathog* **11**:e1004656.
110. **Price, A. M., J. P. Tourigny, E. Forte, R. E. Salinas, S. S. Dave, and M. A. Luftig.** 2012. Analysis of Epstein-Barr Virus-Regulated Host Gene Expression Changes through Primary B-Cell Outgrowth Reveals Delayed Kinetics of Latent Membrane Protein 1-Mediated NF-kappaB Activation. *J Virol* **86**:11096-11106.
111. **Radkov, S. A., R. Touitou, A. Brehm, M. Rowe, M. West, T. Kouzarides, and M. J. Allday.** 1999. Epstein-Barr virus nuclear antigen 3C interacts with histone deacetylase to repress transcription. *J Virol* **73**:5688-5697.
112. **Ramasubramanian, S., K. Osborn, K. Flower, and A. J. Sinclair.** 2012. Dynamic chromatin environment of key lytic cycle regulatory regions of the Epstein-Barr virus genome. *J Virol* **86**:1809-1819.
113. **Reusch, J. A., D. M. Nawandar, K. L. Wright, S. C. Kenney, and J. E. Mertz.** 2015. Cellular differentiation regulator BLIMP1 induces Epstein-Barr virus lytic reactivation in epithelial and B cells by activating transcription from both the R and Z promoters. *J Virol* **89**:1731-1743.
114. **Rickinson, A., and E. Kieff.** 2007. Epstein-Barr virus p. 2603-2654. *In* D. M. Knipe and P. M. Howley (ed.), *Fields Virology*, 5th ed. Lippincott, Williams, and Wilkins, Philadelphia.
115. **Robertson, E. S., J. Lin, and E. Kieff.** 1996. The amino-terminal domains of Epstein-Barr virus nuclear proteins 3A, 3B, and 3C interact with RBPJ(kappa). *J Virol* **70**:3068-3074.
116. **Robinson, T. J., M. A. Dinan, M. Dewhirst, M. A. Garcia-Blanco, and J. L. Pearson.** 2010. SplicerAV: a tool for mining microarray expression data for changes in RNA processing. *BMC Bioinformatics* **11**:108.
117. **Robinson, T. J., E. Forte, R. E. Salinas, S. Puri, M. Marengo, M. A. Garcia-Blanco, and M. A. Luftig.** 2012. SplicerEX: a tool for the automated detection and

- classification of mRNA changes from conventional and splice-sensitive microarray expression data. *RNA* **18**:1435-1445.
118. **Rowe, M., S. Raithatha, and C. Shannon-Lowe.** 2014. Counteracting effects of cellular Notch and Epstein-Barr virus EBNA2: implications for stromal effects on virus-host interactions. *J Virol* **88**:12065-12076.
  119. **Saito, T., S. Chiba, M. Ichikawa, A. Kunisato, T. Asai, K. Shimizu, T. Yamaguchi, G. Yamamoto, S. Seo, K. Kumano, E. Nakagami-Yamaguchi, Y. Hamada, S. Aizawa, and H. Hirai.** 2003. Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. *Immunity* **18**:675-685.
  120. **Sakai, T., Y. Taniguchi, K. Tamura, S. Minoguchi, T. Fukuhara, L. J. Strobl, U. Zimmer-Strobl, G. W. Bornkamm, and T. Honjo.** 1998. Functional replacement of the intracellular region of the Notch1 receptor by Epstein-Barr virus nuclear antigen 2. *J Virol* **72**:6034-6039.
  121. **Sandberg, R., J. R. Neilson, A. Sarma, P. A. Sharp, and C. B. Burge.** 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**:1643-1647.
  122. **Satoh, T., Y. Hoshikawa, Y. Satoh, T. Kurata, and T. Sairenji.** 1999. The interaction of mitogen-activated protein kinases to Epstein-Barr virus activation in Akata cells. *Virus genes* **18**:57-64.
  123. **Schmidt, S. C., S. Jiang, H. Zhou, B. Willox, A. M. Holthaus, P. V. Kharchenko, E. C. Johannsen, E. Kieff, and B. Zhao.** 2015. Epstein-Barr virus nuclear antigen 3A partially coincides with EBNA3C genome-wide and is tethered to DNA through BATF complexes. *Proc Natl Acad Sci U S A* **112**:554-559.
  124. **Schwarzmann, F., N. Prang, B. Reichelt, B. Rinkes, S. Haist, M. Marschall, and H. Wolf.** 1994. Negatively cis-acting elements in the distal part of the promoter of Epstein-Barr virus trans-activator gene BZLF1. *J Gen Virol* **75** ( Pt 8):1999-2006.
  125. **Sepp, M., K. Kannike, A. Eesmaa, M. Urb, and T. Timmusk.** 2011. Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PLoS One* **6**:e22138.
  126. **Shapiro, I. M., A. W. Cheng, N. C. Flytzanis, M. Balsamo, J. S. Condeelis, M. H. Oktay, C. B. Burge, and F. B. Gertler.** 2011. An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet* **7**:e1002218.
  127. **Shirley, C. M., J. Chen, M. Shamay, H. Li, C. A. Zahnow, S. D. Hayward, and R. F. Ambinder.** 2011. Bortezomib induction of C/EBPbeta mediates Epstein-Barr virus lytic activation in Burkitt lymphoma. *Blood* **117**:6297-6303.
  128. **Skalska, L., R. E. White, M. Franz, M. Ruhmann, and M. J. Allday.** 2010. Epigenetic repression of p16(INK4A) by latent Epstein-Barr virus requires the interaction of EBNA3A and EBNA3C with CtBP. *PLoS Pathog* **6**:e1000951.

129. **Strobl, L. J., H. Hofelmayr, G. Marschall, M. Brielmeier, G. W. Bornkamm, and U. Zimmer-Strobl.** 2000. Activated Notch1 modulates gene expression in B cells similarly to Epstein-Barr viral nuclear antigen 2. *J Virol* **74**:1727-1735.
130. **Sun, C. C., and D. A. Thorley-Lawson.** 2007. Plasma cell-specific transcription factor XBP-1s binds to and transactivates the Epstein-Barr virus BZLF1 promoter. *J Virol* **81**:13566-13577.
131. **Takada, K.** 1984. Cross-linking of cell surface immunoglobulins induces Epstein-Barr virus in Burkitt lymphoma lines. *Int J Cancer* **33**:27-32.
132. **Takada, K., and Y. Ono.** 1989. Synchronous and sequential activation of latently infected Epstein-Barr virus genomes. *J Virol* **63**:445-449.
133. **Taniguchi, Y., T. Furukawa, T. Tun, H. Han, and T. Honjo.** 1998. LIM protein KyoT2 negatively regulates transcription by association with the RBP-J DNA-binding protein. *Mol Cell Biol* **18**:644-654.
134. **Taylor, G. M., S. K. Raghuvanshi, D. T. Rowe, R. M. Wadowsky, and A. Rosendorff.** 2011. Endoplasmic reticulum stress causes EBV lytic replication. *Blood* **118**:5528-5539.
135. **Thomas, C., A. Dankesreiter, H. Wolf, and F. Schwarzmann.** 2003. The BZLF1 promoter of Epstein-Barr virus is controlled by E box-/HI-motif-binding factors during virus latency. *J Gen Virol* **84**:959-964.
136. **Thomas, P. D., M. J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, and A. Narechania.** 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome research* **13**:2129-2141.
137. **Tomkinson, B., and E. Kieff.** 1992. Use of second-site homologous recombination to demonstrate that Epstein-Barr virus nuclear protein 3B is not important for lymphocyte infection or growth transformation in vitro. *J Virol* **66**:2893-2903.
138. **Tomkinson, B., E. Robertson, and E. Kieff.** 1993. Epstein-Barr virus nuclear proteins EBNA-3A and EBNA-3C are essential for B-lymphocyte growth transformation. *J Virol* **67**:2014-2025.
139. **Tun, T., Y. Hamaguchi, N. Matsunami, T. Furukawa, T. Honjo, and M. Kawaichi.** 1994. Recognition sequence of a highly conserved DNA binding protein RBP-J kappa. *Nucleic Acids Res* **22**:965-971.
140. **Tzellos, S., and P. J. Farrell.** 2012. Epstein-barr virus sequence variation-biology and disease. *Pathogens* **1**:156-174.
141. **VanderWielen, B. D., Z. Yuan, D. R. Friedmann, and R. A. Kovall.** 2011. Transcriptional repression in the Notch pathway: thermodynamic characterization of CSL-MINT (Mx2-interacting nuclear target protein) complexes. *J Biol Chem* **286**:14892-14902.

142. **Venables, J. P.** 2006. Unbalanced alternative splicing and its significance in cancer. *BioEssays : news and reviews in molecular, cellular and developmental biology* **28**:378-386.
143. **Wallberg, A. E., K. Pedersen, U. Lendahl, and R. G. Roeder.** 2002. p300 and PCAF act cooperatively to mediate transcriptional activation from chromatin templates by notch intracellular domains in vitro. *Mol Cell Biol* **22**:7812-7819.
144. **Waltzer, L., M. Perricaudet, A. Sergeant, and E. Manet.** 1996. Epstein-Barr virus EBNA3A and EBNA3C proteins both repress RBP-J kappa- EBNA2-activated transcription by inhibiting the binding of RBP-J kappa to DNA. *J Virol* **70**:5909-5915.
145. **Wang, E. T., R. Sandberg, S. Luo, I. Khrebtkova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, and C. B. Burge.** 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**:470-476.
146. **Wang, L., S. R. Grossman, and E. Kieff.** 2000. Epstein-Barr virus nuclear protein 2 interacts with p300, CBP, and PCAF histone acetyltransferases in activation of the LMP1 promoter. *Proc Natl Acad Sci U S A* **97**:430-435.
147. **Wang, Y. C., J. M. Huang, and E. A. Montalvo.** 1997. Characterization of proteins binding to the ZII element in the Epstein-Barr virus BZLF1 promoter: transactivation by ATF1. *Virology* **227**:323-330.
148. **White, R. E., I. J. Groves, E. Turro, J. Yee, E. Kremmer, and M. J. Allday.** 2010. Extensive co-operation between the Epstein-Barr virus EBNA3 proteins in the manipulation of host gene expression and epigenetic chromatin modification. *PLoS One* **5**:e13979.
149. **White, R. E., P. C. Ramer, K. N. Naresh, S. Meixlsperger, L. Pinaud, C. Rooney, B. Savoldo, R. Coutinho, C. Bodor, J. Gribben, H. A. Ibrahim, M. Bower, J. P. Nourse, M. K. Gandhi, J. Middeldorp, F. Z. Cader, P. Murray, C. Munz, and M. J. Allday.** 2012. EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors. *J Clin Invest* **122**:1487-1502.
150. **Wilson, J. J., and R. A. Kovall.** 2006. Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA. *Cell* **124**:985-996.
151. **Woisetschlaeger, M., X. W. Jin, C. N. Yandava, L. A. Furmanski, J. L. Strominger, and S. H. Speck.** 1991. Role for the Epstein-Barr virus nuclear antigen 2 in viral promoter switching during initial stages of infection. *Proc Natl Acad Sci U S A* **88**:3942-3946.
152. **Wollerton, M. C., C. Gooding, E. J. Wagner, M. A. Garcia-Blanco, and C. W. Smith.** 2004. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol Cell* **13**:91-100.
153. **Yates, J. L., N. Warren, and B. Sugden.** 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature* **313**:812-815.

154. **Yin, Q., K. Jupiter, and E. K. Flemington.** 2004. The Epstein-Barr virus transactivator Zta binds to its own promoter and is required for full promoter activity during anti-Ig and TGF-beta1 mediated reactivation. *Virology* **327**:134-143.
155. **Yoshida, H., T. Matsui, A. Yamamoto, T. Okada, and K. Mori.** 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* **107**:881-891.
156. **Yu, K. P., L. Heston, R. Park, Z. Ding, R. Wang'ondy, H. J. Delecluse, and G. Miller.** 2013. Latency of Epstein-Barr virus is disrupted by gain-of-function mutant cellular AP-1 proteins that preferentially bind methylated DNA. *Proc Natl Acad Sci U S A* **110**:8176-8181.
157. **Yu, X., P. J. McCarthy, H. J. Lim, T. Iempridee, R. J. Kraus, D. A. Gorlen, and J. E. Mertz.** 2011. The ZIIR element of the Epstein-Barr virus BZLF1 promoter plays a central role in establishment and maintenance of viral latency. *J Virol* **85**:5081-5090.
158. **Zhang, B., B. T. Golding, and I. R. Hardcastle.** 2015. Small-molecule MDM2-p53 inhibitors: recent advances. *Future medicinal chemistry* **7**:631-645.
159. **Zhao, B., J. Zou, H. Wang, E. Johannsen, C. W. Peng, J. Quackenbush, J. C. Mar, C. C. Morton, M. L. Freedman, S. C. Blacklow, J. C. Aster, B. E. Bernstein, and E. Kieff.** 2011. Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. *Proc Natl Acad Sci U S A* **108**:14902-14907.
160. **Zhou, S., M. Fujimuro, J. J. Hsieh, L. Chen, A. Miyamoto, G. Weinmaster, and S. D. Hayward.** 2000. SKIP, a CBF1-associated protein, interacts with the ankyrin repeat domain of NotchIC To facilitate NotchIC function. *Mol Cell Biol* **20**:2400-2410.
161. **Zimber-Strobl, U., and L. J. Strobl.** 2001. EBNA2 and Notch signalling in Epstein-Barr virus mediated immortalization of B lymphocytes. *Semin Cancer Biol* **11**:423-434.

## Biography

Nicholas Joseph Homa was born in Derry, New Hampshire in October 1987. He attended the University of Pittsburgh where he graduated with a BS in Microbiology in 2010. He also completed minor areas of study in chemistry and the Latin Language, as well as a certificate in the conceptual foundations of medicine. During his tenure in Pittsburgh, he worked in the Laboratory of Dr. Andrew VanDemark and alongside Dr. Graham Hatfull where he was a Howard Hughes Undergraduate Research Fellow from 2008-2010. There he published his work, "The structure of Xis reveals the basis for filament formation and insight into DNA bending within a mycobacteriophage intasome". In 2010 he began his graduate career at Duke University in the Department of Molecular Genetics and Microbiology under the mentorship of Dr. Micah Luftig. Here he published his work "Epstein-Barr virus induces global changes in cellular mRNA isoform usage that are important for the maintenance of latency". He was also a member of the Viral Oncology Training Grant from 2011-2012 and was a member of the Duke Scholars in Molecular Medicine for infectious disease from 2015-2016.