BASIS OF GENE SPECIFIC ACTIVATION BY EBV SM PROTEIN

By

ZHAO HAN

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2008

To Liping and Yichun; my parents, my anchor and my sea and to Ben; my love, my life

ACKNOWLEDGMENTS

I would like to express my deepest appreciation and gratitude to all who helped and supported me emotionally, academically, and structurally. My family and friends who always believed in me and stood behind the choices I made. I thank my mentor, Dr. Sankar Swaminathan, for giving me this exciting project, providing me outstanding intellectual guidance, and presenting me with endless opportunities to learn and explore science. I would also like to thank the past and present members of the Swaminathan Lab, particularly Dinesh Verma, Dorit Muller, Bindhu Monica, Nagaraja Tirumuru, Melusine Gaillard and Mike Nekorchuk, for all the stimulating discussions, captivating memories and technical advice. I also extend a special thanks to Dr. Eric Johannsen for the kind gift of p3HR1-ZHT and B958-ZHT cell lines and Dr. Dirk Dittmer and the Dittmer Lab for processing and performing the RT-PCR microarray described in this thesis. In addition, I want to thank Joyce Conners and Sony Kuruppacherry for their administrative support and relieving much of my concerns regarding class registrations, semester deadlines, and graduation requirements among countless other things. Finally, I would like to thank my committee members, Dr. David Bloom and Rolf Renne, both Principal Investigators with their own productive labs, for their generous guidance and invaluable assistance for which I owe a debt of gratitude.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

4-HT	4-hydroxytamoxifen
α-BFRF3	anti-BFRF3
aa	amino acid
ARM	arginine rich region
CAT	chloramphenicol acetyl transferase
cDNA	complementary deoxyribonucleic acid
CR2	complement receptor 2
CT	cycle number
EBV	Epstein-Barr virus
GST	glutathione-S-transferase
hCMV	Human cytomegalovirus
HHV-4	Human herpesvirus 4
HHV-8	Human herpesvirus 8
HSV	herpes simplex virus
IE	immediate early
IM	infectious mononucleosis
IP	immunoprecipitation
IR	Internal direct repeats
kb	kilo-base
kbp	kilo-basepairs
kDa	kilo Dalton
KSHV	Kaposi’s sarcoma associated herpesvirus
mM	millimolar
ORF	open reading frame

PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PI	pre-immune
QPCR	quantitative polymerase chain reaction
R	BRLF1
RNA	ribonucleic acid
RT	reverse transcription
TE	Tris EDTA
TR	terminal direct repeats
UL	unique long
US	unique short
UTR	untranslated region
VCA	viral capsid antigen
VZV	varicella-zoster virus
Z	BZLF1

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

BASIS OF GENE SPECIFIC ACTIVATION BY EBV SM PROTEIN

By

Zhao Han

December 2008

Chair: Sankar Swaminathan
Major: Medical Sciences

Epstein-Barr virus (EBV) is a human herpes virus that infects and persists in approximately 90% of the adult human population and is associated with a number of human malignancies such as Burkitt's lymphoma, Hodgkins's lymphoma and nasopharyngeal carcinoma. The EBV SM protein is a posttranscriptional regulatory protein expressed early during lytic replication and is essential for virus production. SM is an RNA binding protein which enhances accumulation of its target mRNAs but its exact mechanism of action remains to be determined. We had previously shown that SM enhances accumulation of some EBV transcripts over others. However the basis of such specificity has not been investigated. Understanding the basis of gene specific activation by SM should provide insights into the regulation of lytic EBV replication and possibly opportunities for specific therapeutic interventions.

This study is aimed at determining the basis of specific RNA recognition by SM. To ask whether SM associates more efficiently with specific EBV transcripts, we employed an RNA immunoprecipitation/RT-QPCR assay. We used cell lines derived from lymphoma infected with EBV that have been modified to permit high level lytic EBV replication in an inducible manner. Induced cells were lysed, and SM/RNA complexes were immunoprecipitated with SM antibody.

RNA was isolated from each immunoprecipitation and analyzed by RT-QPCR microarray for all EBV open reading frames (ORF). We found that there is general enrichment of EBV RNA in SM-immunoprecipitates, suggesting that SM has some non-specific RNA binding capability. However, there were several RNAs which were highly enriched by SM, suggesting that SM does bind preferentially to specific RNAs. In order to map high affinity SM-binding sites on BFRF3 RNA, protein-RNA crosslinking assays were employed to compare the affinity of SM for various portions of the BFRF3 RNA. SM bound to the full-length BFRF3 transcript but not to the anti-sense sequence of BFRF3. Furthermore, SM bound preferentially to the first 189 bases of the BFRF3 transcript suggesting that there may be some specific sequence or structural motif in this region of the BFRF3 RNA which allows for preferential binding by SM.

CHAPTER 1
INTRODUCTION

Basic Overview

Epstein - Barr virus (EBV) is a member of the virus family *Herpesviridae* and the γ-herpesvirus subfamily. It is a human herpesvirus (HHV-4) that infects and persists in approximately 90% of the adult population and is associated with a number of human malignancies such as Burkitt's lymphoma, Hodgkin's lymphoma and nasopharyngeal carcinoma. Primary infection of EBV can occur early during childhood and is usually asymptomatic or later during adolescence or adulthood, when it can cause infectious mononucleosis (IM). Transmission of EBV generally occurs through the exchange of saliva in which the virions infect the epithelial cells of the oropharynx. EBV can also infect primary B-lymphocytes where it establishes latent infection and maintains a persistent infection. Although EBV is usually latent, it can become permissive for lytic virus replication and virus production.

Two types of EBV infect humans; EBV type 1 and 2 (also called type A and type B). Differences between the two types are primarily in the sequences that code for latent proteins. While type 1 is prevalent in the United States, Europe, and Southeast Asia, both type 1 and 2 are widespread throughout equatorial Africa and New Guinea (9).

Epstein - Barr virus

Virus Structure and Genome

Like other herpesviruses, EBV contains a toroidal DNA core surrounded by an icosahedral nucleocapsid and further surrounded by a lipid bilayer envelope with a tegument in between the nucleocapsid and the envelope (9). The EBV virion contains a single copy of the linear double-stranded DNA (dsDNA) genome that is 184 kilobase pairs (kbp) and encodes over 90 proteins. Many of the genes are named after the BamHI restriction fragment on which they are encoded

from (2). The genome is comprised of two unique segments; the unique short (US) and the unique long (UL), which are flanked by repeated sequences; terminal direct repeats (TR) and internal direct repeats (IR) (9).

EBV Life Cycle

Entry of EBV into B lymphocytes is mediated by binding of the most abundantly expressed viral glycoprotein, gp350/220, to the host cell receptor, CD21 (also know as CR2) (27). Upon attachment, the virion enters through the endocytic pathway and fuses with the cellular vesicle membrane allowing the release of the viral capsid into the cytoplasm (16). EBV capsids are then transported to the nuclear pore along microtubules (9). Disassembly of the capsid is required for the release of the genomic DNA into the nucleus, where it circularizes and persists as an episome (9).

Infection of a B lymphocyte usually results in latent (quiescent) infection in which the viral episome is replicated once per cell cycle by the host DNA polymerase, a few latent genes are expressed but no viral proteins or new virus progeny are produced (1,9). Latently infected cells can also be induced into lytic (productive) infection in which all viral genes are expressed, viral DNA is replicated through a rolling circle replication by a virally encoded polymerase, and new virions are produced (28, 29).

EBV Gene Expression during Lytic Infection

Upon reactivation or induction of EBV, gene expression shifts from latent genes to expression of lytic genes. EBV lytic gene expression occurs in a temporally regulated manner such that the transcription and translation of lytic genes are classified into three main periods; immediate early (IE), early (E) and late (L). IE genes are involved in transcriptional activation of early genes (9). These genes do not require viral proteins for their transcription and translation. The principal IE genes are BZLF1 (Z) and BRLF1 (R), which are both

transactivators of early EBV genes (7, 18, 19, 31). Early genes are mostly genes that code for viral proteins involved in DNA replication and late genes are mostly those that code for structural proteins and proteins required for new virion formation (9).

The EBV SM Protein

The EBV SM protein, also known as EB2, Mta, and BMLF1 is one of the first early genes to be expressed during lytic EBV replication (26). SM is essential for viral DNA replication and the production of infectious virions (12, 24). It has homologs in other herpesviruses including ORF57 of Kaposi's sarcoma associated herpesvirus (KSHV, HHV-8) (3), ICP27 of herpes simplex virus (HSV) (23), ORF4 of varicella-zoster virus (VZV) (8), and UL69 of human cytomegalovirus (hCMV) (30). Although these genes are functionally and structurally related, there are significant differences including when during lytic replication they are expressed, their sequences, their abilities to rescue mutant recombinant viruses and other modes of action (3, 8, 11, 23, 25, 30).

Gene and Protein Structure

SM is a 55 kilo Dalton (kDa) nuclear protein encoded by an 1.7 kb mRNA transcript spliced from the Bam S and Bam M regions of the EBV genome (BSLF2 and BMLF1 exons) giving it a total size of 479 amino acids (aa) in length (15). The functional regions of SM have been under much investigation. SM amino acids 60-140 and 218-237 are nuclear export signals which may mediate nuclear to cytoplasmic shuttling of RNAs (4, 6). Studies have shown that the RXP (Arg-X-Pro) RNA binding domain located at amino acids 152-172 binds RNA *in vitro* however, its role in SM function have been shown to be non-essential (5, 20). Amino acids 470-474 contain a highly conserved motif, GLFF motif (gly-leu-phe-phe), which is required for SM function and proper folding (21).

Regulation of Gene Expression

SM is a post-transcriptional regulator of cellular and viral gene expression. Earlier studies showed that when co-transfected with certain reporter constructs, such as chloramphenicol acetyl transferase (CAT), SM greatly increases the expression of CAT without increasing the transcription initiation rate (22). Furthermore, the activation of CAT by SM is independent of the promoter used further indicating that SM's main mechanism of action is post-transcriptional (22).

The properties of SM-mediated activation were further studied using various reporter genes in co-transfection assays of BJAB cells. Activation by SM of CAT was greatly enhanced whereas there were no effects by SM on firefly luciferase and the human growth hormone gene expression (20,22). These data suggests that effects of SM on target genes are gene specific. This conclusion is supported by cDNA microarray studies using a SM-deleted recombinant EBV strain which demonstrate that only half of the EBV lytic genes are SM-dependent and that expression of some EBV genes were more dependent on SM than others (12).

RNA Binding

In initial studies looking at the RNA binding capability of SM, bacterially derived SM-glutathione-S-transferase (GST) fusion proteins were made and shown to bind radioactively labeled RNA *in vitro* indicating that SM is a RNA binding protein (22). Later studies suggested that the RXP domain of SM was essential for RNA binding *in vitro*. When made with a mutated or deleted RXP domain, SM loses its ability to bind RNA *in vitro* as indicated by Northwestern assays (5, 20). However, deletion of the RXP region does not eliminate or alter normal localization of SM and function of gene activation which prompted further research to explore the necessity of the RXP motifs for RNA binding *in vivo*. Co-transfection assays of RXP-deleted SM with target CAT gene were performed and co-immunoprecipiation of CAT mRNAs were

demonstrated with SM immunoprecipitates confirming that the RXP region is not required for RNA binding *in vivo* (20).

In attempt to identify the RNA binding region of SM, one study found that when various bacterially synthesized GST-SM fusion peptides were incubated with radioactively labled RNAs, that an arginine-rich region (ARM) located at aa 190-223 of SM bound RNA nonspecifically (13). However, this lack of binding specificity does not explain the selective effects of SM on mRNA accumulation. Furthermore, the basic properties of any arginine-rich region might allow binding of negatively charged groups such as RNA and not necessarily be reflective of the SM domains actually binding RNA targets *in vivo*. Although the RNA binding domain of SM is still poorly defined, the specificity of SM's effects on target mRNAs suggests that a sequence or structure-specific interaction between SM and its target RNAs exists.

Purpose and Significance

Although much has been reported concerning the functions and mechanisms of the EBV SM protein, there are several important aspects yet to be explained such as the actions of SM on specific mRNA transcripts and whether there are specific sequence or structural elements in RNA molecules that allows for binding of SM. Identification of specific SM-RNA interactions and specific SM binding sites on its target RNAs will help determine the basis of specific RNA recognition by SM and perhaps provide a better understanding of the nature of protein-RNA interactions. Understanding the nature of SM specificity will allow to predict which genes, both cellular and viral, SM might affect and which cellular RNA-binding proteins SM might compete with, thus providing new insights into the mechanisms by which SM post-transcriptionally regulates cellular and viral gene expression and ultimately a better understanding of how infection of EBV alters host cell functions. Finally, because SM is required for DNA replication of EBV and there are no cellular homologs of SM makes SM an attractive target for antiviral

compounds. By increasing our knowledge on the molecular mechanisms by which the EBV SM protein functions inside the cell, more opportunities and better strategies for therapeutic developments can be created.

CHAPTER 2
MATERIALS AND METHODS

BFRF3, BFRF3 subsets, and α-BFRF3 Plasmids

BFRF3 (the gene encoding a component of the small viral capsid antigen of EBV's VCA) containing its 5'UTR and cleavage and polyadenylation signal and the α-BFRF3 (antisense of the entire BFRF3 sequence) were constructed by high fidelity PCR amplification using AccuPrime Pfx DNA Polymerase (Invitrogen). The BFRF3 clone was generated from B95-8 EBV genome positions 49056 to 49806 (accession number NC00705) using primers (Table 2-1) flanking the entire BFRF3 gene including 5' and 3' UTRs. PCR products were directionally cloned into the HindIII to EcoRV sites of pcDNA3 (Invitrogen) in opposite orientations. The orientation of the inserts was determined by restriction enzyme digestion and confirmed by DNA sequencing.

Subclones of BFRF3 encompassing approximately one-fourth of the gene extending from the 5'UTR to 20bps downstream of the cleavage and polyadenylation signal of BFRF3 were also constructed by PCR amplification. PCR was performed using different 5' and 3' primers (Table 2-1) and PCR products representing the different subsets of BFRF3 ranged between 186 to 196 bps in length. Each cloned PCR product was screened by restriction digestion and confirmed by DNA sequencing.

Cell Lines

P3HR1-ZHT is a Burkitt's lymphoma cell line infected with the type 2 EBV and B958-ZHT is a marmoset B cell line transformed by type 1 EBV (14), kind gift of Eric Johannsen, Harvard Medical School. Both cell lines stably express a fusion protein containing the BZLF1 transactivator of early lytic cycle replication fused to the hormone domain of the estrogen receptor protein in which during the presence of 4-hydroxytamoxifen allows the release of

BZLF1, inducing lytic EBV replication. P3HR1-ZHT and B958-ZHT cell lines were propagated in RPMI supplemented with 10% fetal bovine serum (HyCLone), 0.8mg/ml G418 (AG Scientific), and L-glutamine (Invitrogen). Cos7 cells are African green monkey kidney fibroblast cells transformed with a mutant simian virus 40 (SV40) (Gluzman). These cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum and L-glutamine. All cells were grown at 37^{o}C in 5% CO_2.

Transfection Assays

Cos7 cells were transfected with empty vector plasmid (pcDNA3, Invitrogen) or SM expression vector (EW63, 22) DNA using Lipofectamine Plus (Invitrogen) in 100 mm plates with 6 ug of DNA per transfection plate, according to manufacturer's protocol. Cells were harvested 48 hours after transfection.

Immunoprecipitation and RNA Isolation

Lytic replication was induced in 180 x 10^6 P3HR1-ZHT or B958-ZHT cells at 5 x 10^5 cells/ml by adding 100nM 4-hydroxytamoxifen to the cell growth medium. Cells were harvested 48 hours after treatment by 4-HT and lysed in ice-cold immunoprecipitation (IP) lysis buffer (Tris-buffered saline [pH 7.4], 1% Triton X 100 and protease inhibitor cocktail [Sigma, P2714]). Cells were incubated in lysis buffer for 10 minutes on ice with frequent mixing and sonicated to ensure maximum lysis. The lysed cell suspension was centrifuged at 4^{o}C for 10 min at 10^5 x g. Supernatant was transferred to fresh tubes and cleared with normal rabbit IgG (Bethyl) and protein A-conjugated agarose beads (Sigma) followed by immunoprecipitation with either preimmune serum or anti-SM antibody and protein A agarose beads and washed four times in immunoprecipitation wash buffer (500 mM NaCl, 25mM Tris, 27 mM KCl, 1% NP40 [pH 7.4]). Co-immunoprecipitated RNA was isolated from the immunoprecipitates using RNA-bee

(Teltest) and RNeasy columns (Qiagen) containing an on-column DNAse treatment (Qiagen) and eluted with RNAse-free Tris-EDTA (TE) buffer.

Complementary DNA Microarray and Analysis

Real-time quantitative PCR arrays containing PCR primers targeting all ORF of the EBV genome were designed and performed by our collaborators at the University of North Carolina, Lineberger Comprehensive Cancer Center. RNA used in the microarray analysis was co-immunoprecipitated and purified as described above. Reverse transcription of RNA was performed with reverse transcriptase (Invitrogen), 2 mM deoxyribonucleoside triphophates, 2.5 mM $MgCl_2$, RNAsin (Applied Biosystems Inc.) and random hexamers. Cycling conditions for RT are 42^{o}C for 45 min, 52^{o}C for 30 min, and 70^{o}C for 10 min. Following RT, the removal of excess RNA was done by incubation of each RT reaction with 1 U of RNAse H at 37^{o}C in order to prepare the samples for PCR amplification. Real-time PCR was performed in triplicates for each sample with SYBR Green PCR mix (Applied Biosystems) using universal cycling conditions (17). Raw cycle number (CT) values were determined and used directly to compare fold-differences.

Northern Blotting

P3HR1-ZHT and B958-ZHT cells were induced into lytic replication by incubating the cells at 5 x 10^5 cells/ml in 100 nM 4-HT. 10^7 cells were harvested for total RNA at 0 and 48 hours after induction and co-immunoprecipitated RNA was harvested as described above. RNA samples were loaded and electrophoresed in a 1% denaturing formaldehyde agarose gel, transferred to Zeta probe membrane, and UV-crosslinked to membrane. Gene specific probes were generated by PCR amplification, gel purification and [α-^{32}P]dCTP labeling. The probes were hybridized to blots overnight at 65^{o}C, washed and exposed to film and a phosphorimager screen for quantification by ImageQuant software.

In vitro Photocrosslinking

Cos7 cells were transfected with empty vector or SM as described above. Cells were washed with fresh warm complete DME and gently scrapped off and pelleted by centrifugation at 20^{o}C for 5 min at 900 x g. Cells were lysed by re-suspending the pellets in twice the pellet volume of ice-cold lysis buffer (20mM HEPES [pH 7.9], 10 mM NaCl, 10% glycerol, 3 mM $MgCl_2$, 0.2 mM EDTA, 1 mM dithiothreitol, and protease inhibitor cocktail [Sigma]) and incubating at 4^{o}C for 15 min with frequent gentle mixing. The lysed cell suspensions were centrifuged at 4^{o}C for 5 min at 700 x g. Cleared supernatants were transferred to a fresh tube and high salt buffer (20 mM HEPES [pH 7.9], 400 mM KCl, 20% glycerol, 0.2 mM EDTA, 0.4 mM phenylmethlysulfonyl fluoride, 1 mM dithiothreitol, and protease inhibitor cocktail [Sigma]) was added at one-third the volume of the supernatant. Aliquots of protein extracts were snap-frozen and stored at -80^{o}C.

RNAs were synthesized using BFRF3, subsets of BFRF3 and α-BFRF3 plasmid DNA, previously linearized with EcoRV, in the presence of [α^{-32}P]rUTP (Perkin Elmer) and T7 RNA Polymerase (NEB). Radioactively labeled RNA transcripts were ran on denaturing urea-polyacrylamide gels and full length transcripts were excised out and purified. Crosslinking was performed by incubating 2 x 10^6 cpm of purified radiolabeled RNA with 8 ul of whole cell extract, 2 ul of 20 mM magnesium acetate, 2 ul of 10 mM ATP, 2 ul of 2000 mM potassium glutamate, 2 ul of 50 mM creatine phosphate, 1 ul of tRNA (1 ug/ul), and 1 ul of RNAsin (Promega) in a total volume of 20 ul. Reactions were incubated at 30^{o}C for 30 min and RNA-protein complexes were UV-crosslinked on ice in a Stratalinker (Stratagene) for a total of 0.6 J/cm^2. Samples were digested with RNAse A at 100 ug/ul for 1 hour at 37^{o}C and immunopurified using anti-SM antibody and protein A agarose beads. Purified proteins were

separated on 10% SDS-PAGE gels followed by autoradiography exposure on a phosphorimager screen and film.

Table 2-1. List of Primers

Primer Name	Primer Sequence
BFRF3 5'#7HindIII	5'ctgaaagctttatttaactttgcggacagagg3'
BFRF3 3'#7EcoRV	5'tcaggatatcgttgggcttgggcagccggcgtg3'
BFRF3 5'#8HindIII	5'ctgaaagcttcaagcccaccctccag3'
BFRF3 3'#8EcoRV	5'tcaggatatcggcaccccaaaagtcctctgcac3'
BFRF3 5'#9HindIII	5'ctgaaagctttcggcgccaacgcgccatagacaag3'
BFRF3 3'#9EcoRV	5'tcaggatatcgatgaagaaacagaggggtc3'
BFRF3 5'#10HindIII	5'ctgaaagctttcatctattagcagcctc3'
BFRF3 3'#10EcoRV	5'tcaggatatcagtttttgtatctgtaattg3'
BFRF3 5'#11EcoRV	5'tcagggatatctatttaactttgcggacagag3'
BFRF3 3'#11HindIII	5'ctgaaagcttagtttttgtatctgtaattg3'

CHAPTER 3
RESULTS

Isolation of SM Target RNA

The EBV SM protein is known to bind RNA, however many aspects of its interaction with RNA remain to be elucidated. In this study, we wished to determine whether SM preferentially binds EBV RNA. In order to isolate SM/RNA complexes that form *in vivo*, we treated P3HR1-ZHT and B958-ZHT cells with 4-hydroxytamoxifen to permit lytic EBV replication. By employing an immunoprecipitaion assay with anti-SM antibody under conditions that preserve protein/RNA complexes, we were able to isolate specific EBV RNAs bound by SM.

Characterization of In Vivo Transcript Specificity of SM

To identify specific EBV RNA targets bound by SM, we took induced P3HR1-ZHT and B958-ZHT cell lysates and immunoprecipitated with either anti-SM antibody (αSM) or preimmune serum (PI, control). By immunoprecipitating SM, we were able to specifically isolate SM and any RNA molecules bound by SM. To determine which RNAs were present in each immunoprecipitate, we isolated SM or PI co-immunoprecipiated RNAs and prepared cDNA for qPCR microarray analysis using primers for all EBV ORFs as previously described (17). Raw CT values representing the relative abundance of transcripts bound by SM versus PI were determined. The amount of RNA in the SM IP was compared to that in the pre-immune IP for each ORF, and the relative enrichment by SM for each RNA target was determined. Our data indicated that there was a general enrichment in the SM-IP versus control (PI-IP) suggesting that SM bound more RNA than the control. In P3HR1-ZHT cells, approximately seven times more RNA was immunoprecipitated by SM than the control. In B958-ZHT, around fifty-five times more RNA was associated with SM than in the control IP. This overall enrichment of RNA (the majority of EBV RNAs) in the SM IPs in both induced cell lines suggested that SM has some

non-specific RNA binding capability. From the raw CT values generated by the PI-IP and SM-IP microarray data, we calculated the mean enrichment values for each target using a linear regression model. Interestingly, we found that in addition to the general enrichment by SM, there were several RNAs that were particularly enriched by SM in both the P3HR1-ZHT and B958-ZHT cells (Table 3-1 and Table 3-2), with CT values 2-5 times greater than the mean enrichment. These highly enriched RNAs included the BFRF3, BGLF5, BDLF3, BTRF1, BCRF1, BBRF3, BBRF2, BLRF1, and BLRF3 from the P3HR1-ZHT microarray. In B958-ZHT, the RNAs most highly enriched by SM included BDLF3, BNLF2b, BBRF3, BFRF3, BGLF2, BTRF1, BBRF2, BLRF2, and BXLF2. Among the highly enriched RNAs, BFRF3, BDLF3, BBRF3, and BTRF1 were highly bound by SM in both P3HR1-ZHT and B958-ZHT cells. This disproportionate binding of RNAs by SM indicated that SM selectively bound certain RNAs more avidly than it bound other RNAs.

To rule out the possibility that the greater enrichment by SM was due to increased fold induction of certain RNAs (Figure 3-1) or the greater abundance of particular RNAs (Figure 3-2), we isolated total RNA from uninduced and induced P3HR1-ZHT and B958-ZHT cells and compared the relative amounts of each RNA and the fold induction in the input RNA to the amounts present of each RNA in the SM immunoprecipitate. By taking the difference in CT values from total uninduced and induced RNA samples, we determined the fold induction as well as the relative abundance of each RNA target. We saw no correlation between the enrichment of each RNA target by SM to its relative fold induction and relative input abundance. There were highly induced as well as less highly induced RNA targets associated with SM (Figure 3-1). Similarly, RNA targets were enriched by SM independent of their relative abundance in the input sample (Figure 3-2). These data suggested that the binding of SM to its targets are RNA-

specific and not merely a consequence of the greater abundance of certain target RNA molecules.

Analysis of SM's RNA Binding Partners

In order to confirm preferential binding by SM as indicated by the array data, we analyzed the RNAs used in QPCR microarray analyses from both P3HR1-ZHT (Figure 3-3) and B958-ZHT (Figure 3-4 and Figure 3-5) by Northern blotting. By analyzing equal amounts of total RNA from induced and uninduced cells, we verified that induction into lytic EBV replication by 4-HT enhances the expression of the genes of interest. The data from our Northern blots were also consistent with the microarray data which showed that SM co-immunoprecipitated more RNA than control (PI or IgG) in both P3HR1-ZHT and B958-ZHT cells and that SM has some non-specific binding to RNA in general. To compare the binding of SM for specific RNA targets, we ran equal percentages of the total SM co-immunoprecipitated and control (PI or IgG) co-immunoprecipitated RNAs and hybridized blots with gene-specific ^{32}P-labeled probes. These gene-specific probes represented the RNAs in both P3HR1-ZHT and B958-ZHT cells that were highly enriched by SM, such as BFRF3 and BBRF3, and RNAs that were less enriched in SM IPs, such as BALF2. Although induction of BFRF3, BBRF3 and BALF2 in P3HR1-ZHT cells was comparable, SM co-immunoprecipitated 68.6 times more BFRF3 and 27.4 time more BBRF3 but only 3.6 times more BALF2 than control (Figure 3-3). Similarly in B958-ZHT cell, SM co-immunoprecipitated BFRF3 38 times more and BBRF3 58.9 times more than control while BALF2 was co-immunoprecipitated only 3.5 times more than control (Figure 3-4 and Figure 3-5).

To further verify that BFRF3, BBRF3 but not BALF2 were highly enriched by SM, we calculated the percentage of total input RNAs that were co-immunoprecipitated by SM (Table 3-4). We showed that 10.2 % of input BFRF3 and 8.4 % of input BBRF3 transcripts from P3HR1-

ZHT cells were brought down by SM. As expected only 0.54% of BALF2 were co-immunoprecipitated by SM. Similarly in B958-ZHT cells, 38.8% of BFRF3 and 23.7% of BBRF3 were co-immunoprecipitated while only 1% of BALF2 were brought down in the SM-IP. These data confirm our hypothesis that SM does exhibit some preferential binding to its target RNA molecules. Interestingly, we also noticed that the SM co-immunoprecipitated RNAs contained a disproportionate amount of 18S ribosomal RNA compared to the 28S suggesting that SM may have higher preference for the 18S ribosomal RNA (data not shown).

Mapping of SM-binding on BFRF3 RNA Transcripts

To ask whether SM exhibited sequence or structure specific binding, we performed an *in vitro* cross-linking assay by incubating SM-containing cell extracts with radioactively labeled transcripts (Figure 3-6A) consisting of the entire sense or anti-sense BFRF3 sequences (Figure 3.6B). SM/RNA complexes were cross-linked by UV irradiation. Unbound RNAs were hydrolyzed by RNase treatment and SM was immunoprecipitated using SM-specific antibody. This procedure renders those proteins that were in physical contact with the target transcript radiolabeled by virtue of covalent crosslinking to radiolabled uridine residues. The immuno-purified RNA-labeled SM samples were visualized by SDS-PAGE and autoradiography. If binding of SM to RNA were sequence or structure-specific, then the binding affinity of SM for BFRF3 and α-BFRF3 should be different since the α-BFRF3 RNA would be predicted to process a different sequence and secondary structure than BFRF3 despite having an identical GC content. As shown in Figure 3-6B, a 55 kDa band corresponding to the size of SM was detected only in the sample containing SM and the sense strand of BFRF3, and not in the α-BFRF3/SM sample indicating that the SM protein was labeled by covalent-linkage of radiolabeled uridine from the sense strand and not by α-BFRF3 (Figure 3-6B). A 55-kDa band was not detected in

the lanes containing empty vector (pcDNA3) and BFRF3 RNA immunoprecipitated with anti-SM antibody or SM and BFRF3 RNA immunoprecipitated with PI serum.

To map potential high affinity SM-binding site on BFRF3 RNA, different subclones of BFRF3 were constructed. Each subclone was designed to encompass one-fourth of the entire BFRF3 which included the 5' and 3' UTRs. The subclones were similar in length and were all directionally cloned into the HindIII and EcoRV sites of pcDNA3 (Invitrogen). The RNA transcripts representing each region of BFRF3 were synthesized and labeled with ^{32}P-UTP *in vitro* (Figure 3-6A). Protein-RNA cross-linking assays were employed to compare the affinity of SM for the various portions of the BFRF3. The more a specific RNA transcript binds SM, the more SM will be labeled with that radioactive RNA transcript and therefore a stronger signal will be generated. The cross-linking results obtained for the different regions of BFRF3 indicated that SM showed much greater affinity for the first 189 nucleotide sequence (the 5'UTR region) of the BFRF3 RNA (Figure 3-6C). Interestingly, the 2nd quarter of BFRF3 also demonstrated some cross-linking activity and almost no binding by SM was shown in the 3rd and 4th quarter regions of BFRF3. Overall, these studies indicate that SM specifically binds to BFRF3 around the 5'UTR.

Table 3-1. The RNA enrichment by SM in P3HR1-ZHT cells

	Sample CT Pre-immune	Sample CT SM	Difference in CT	Enrichment over mean (CT)	Fold enrichment over mean
BFRF3	29.79	24.15	5.64	3.4	10.6
BGLF5	32.05	26.78	5.27	3.3	9.8
BDLF3	34.51	29.67	4.84	3.1	8.6
BTRF1	32.84	28.20	4.64	2.7	6.5
BCRF1	36.88	32.87	4.01	2.6	6.1
BBRF3	27.79	23.04	4.75	2.3	4.9
BBRF2	36.58	33.05	3.53	2.1	4.3
BLRF1	27.63	23.32	4.31	1.8	3.5
BLRF3	30.19	26.42	3.77	1.6	3.0
BALF2	29.10	27.69	1.41	-0.8	0.6

Table 3-2. The RNA enrichment by SM in B958-ZHT cells

	Sample CT Pre-immune	Sample CT SM	Difference in CT	Enrichment over mean (CT)	Fold enrichment over mean
BDLF3	40	30.04	>9.96	5.23	37.5
BNLF2b	40	30.13	>9.87	5.14	35.3
BBRF3	34.44	27.51	6.93	1.95	3.9
BFRF3	32.66	25.2	7.46	1.88	3.7
BGLF2	33.98	26.69	7.29	1.86	3.6
BTRF1	35.76	28.93	6.83	1.61	3.1
BBRF2	34.44	27.51	6.93	1.55	2.9
BLRF2	31.13	23.86	7.27	1.51	2.8
BXLF2	33.59	26.8	6.79	1.31	2.5
BALF2	31.66	30.26	1.4	-4.30	0.05

Table 3-3. The RNA recovery by immunoprecipitation

	P3HR1-ZHT		B958-ZHT	
	IgG (%)	SM (%)	PI (%)	SM (%)
BFRF3	0.15	10.2	1.0	38.8
BBRF3	0.30	8.4	0.39	23.7
BALF2	0.15	0.54	0.92	3.2

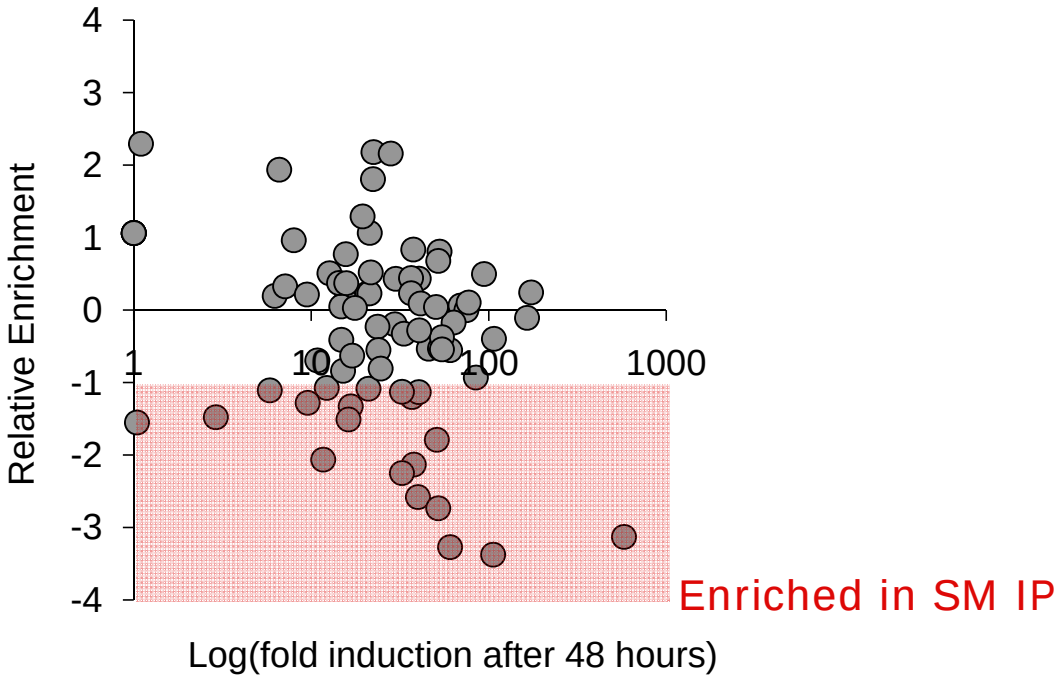

Figure 3-1. Semi-quantitative inter-gene comparison using fold induction. The fold enrichment of each EBV ORF was plotted against the fold induction. Each circle represents an EBV gene. The region of SM enrichment are highlighted in red.

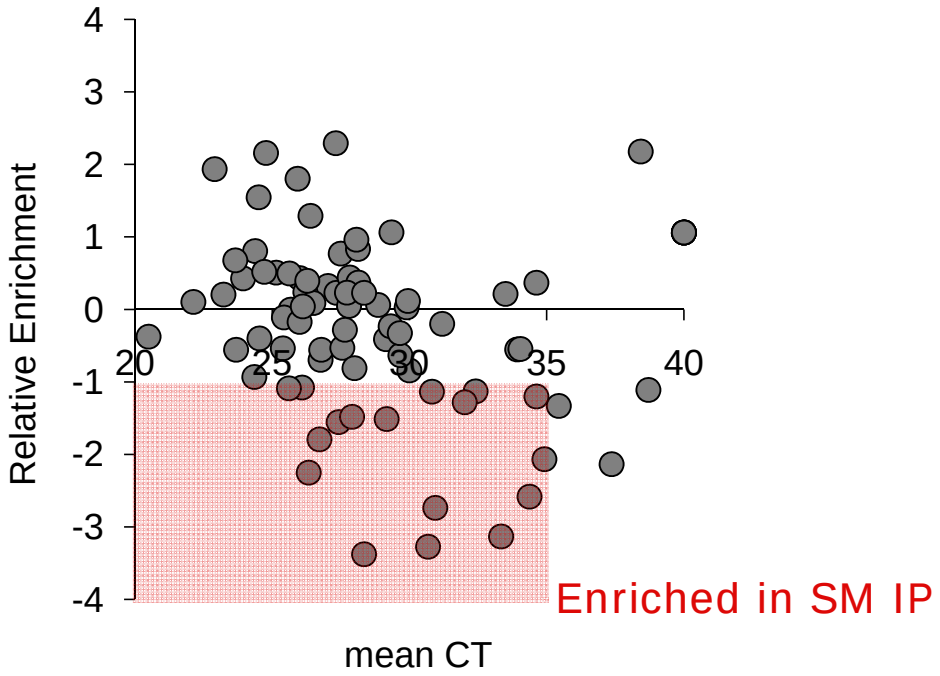

Figure 3-2. Inter-gene comparison of fold enrichment based on mean CT. The fold enrichment of each EBV ORF was plotted against the mean CT. Each circle represents an EBV gene. The region of SM enrichment are highlighted in red.

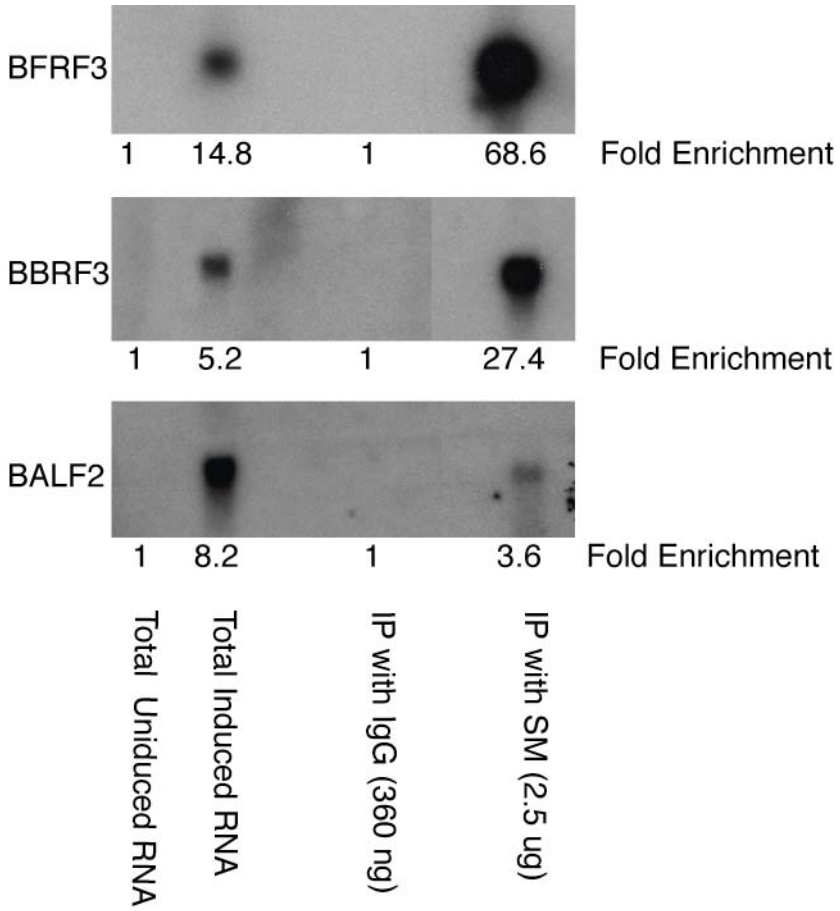

Figure 3-3. Enrichment of RNAs by SM in P3HR1-ZHT cells. Total RNA from induced and uninduced P3HR1-ZHT cells were prepared. IgG (control) or SM co-immunoprecipitated RNAs were isolated. Enrichment of BFRF3, BBRF3, and BALF2 RNAs were measured by Northern blotting. Fold Induction and SM-enrichment of RNAs were determined by phosphorimaging detection.

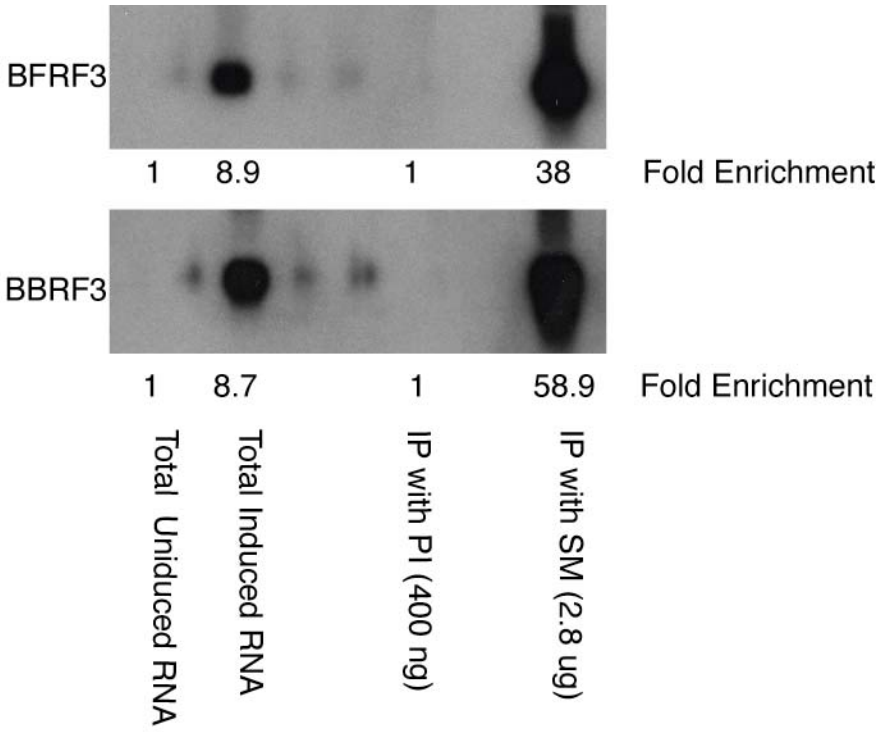

Figure 3-4. The BFRF3 and BBRF3 enrichment by SM in B958-ZHT cells. Total RNA from B958-ZHT cells were harvested at 0 and 48 hours after induction. Co-immunoprecipitated RNA from PI-IP (preimmune serum, control) and SM-IP were isolated. BFRF3 and BBRF3 were detected by Northern blotting using gene-specific probes. Fold induction and SM-enrichment of BFRF3 and BBRF2 were determined by phosphorimaging detection.

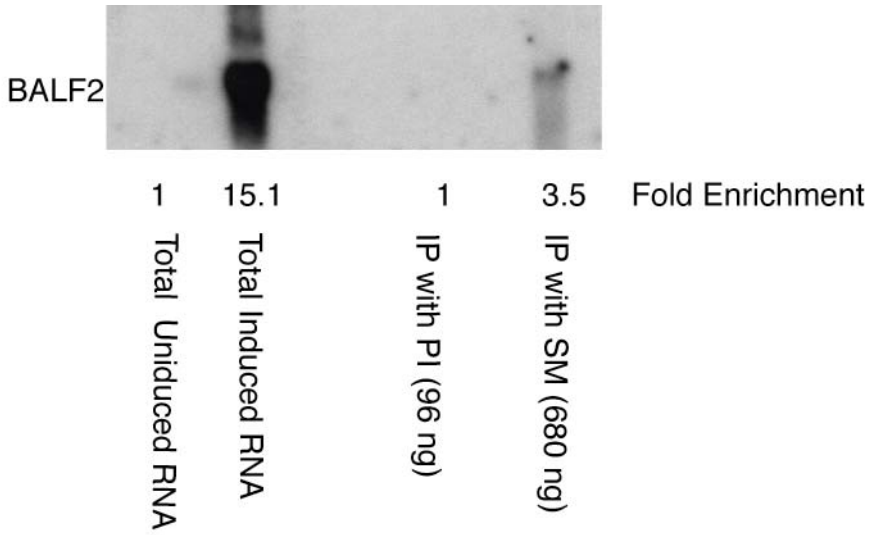

Figure 3-5. Enrichment of BALF2 by SM in B958-ZHT cells. Total RNA from induced and uninduced B958-ZHT cells were prepared. Preimmune serum (control) or SM co-immunoprecipitated RNAs were isolated. Enrichment of BALF2 was measured by Northern blotting. Fold induction and SM-enrichment of BALF2 was determined by phosphorimaging detection.

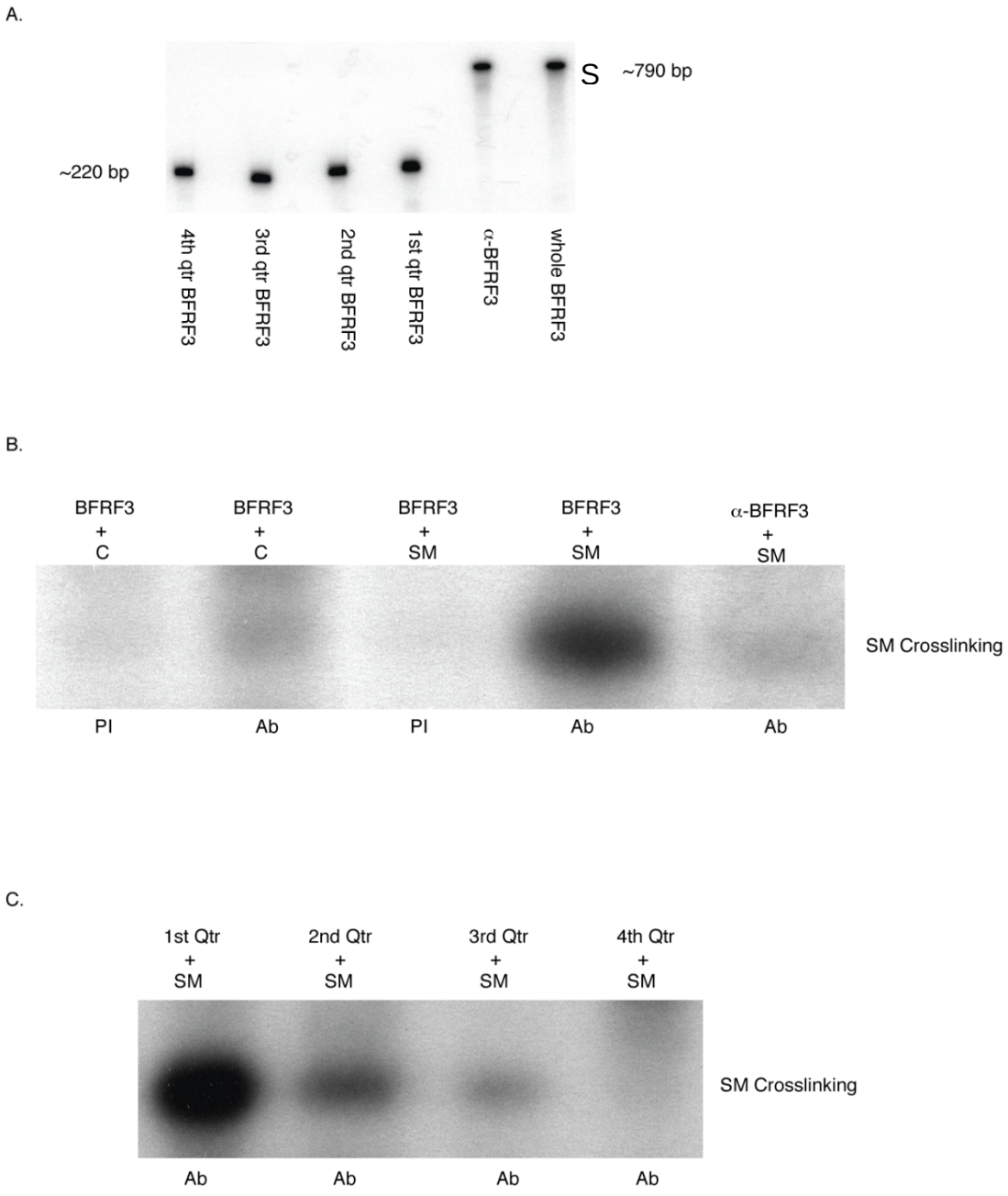

Figure 3-6. Crosslinking of SM protein to BFRF3 RNA. A) RNA transcripts of BFRF3, α-BFRF3 and regions representing one-fourth portions of BFRF3 were generated using ^{32}P-UTP. B) Lysates from cells transfected with pcDNA3 (control, C) or SM were incubated with the entire BFRF3 transcript or α-BFRF3 and UV crosslinked. Samples were treated with RNAse to hydrolyze unbound RNA and SM was immunoprecipitated with SM-specific antibody (Ab) or preimmune serum (PI). Samples were visualized by SDS-PAGE and autoradiography. SM was visualized only when covalent linkage of RNA was detected. C) Crosslinking of SM to various regions of BFRF3. Subclones of BFRF3 were constructed. Each subclone represents one-fourth of the entire BFRF3.

CHAPTER 4
DISCUSSION AND FUTURE AIMS

Discussion

The EBV SM protein is a post-transcriptional regulator of gene expression. Although a number of studies have demonstrated that SM directly binds RNA, the basis for SM/RNA interactions have not been completely characterized, and the nature of SM's specificity for binding RNA, such as any RNA sequence or structural motifs and binding factors have yet to be identified. Several studies have shown that SM increases expression of certain reporter and EBV lytic genes and exhibits gene specificity (12, 20, 22). For example, SM enhances the expression of CAT while little to no effects on firefly luciferase and human growth hormone were reported (20, 22). Possible explanations of specificity in gene expression by SM could be the existence of high affinity RNA sequence motifs or secondary structures that act as SM binding sites. Since SM does demonstrate preference for activation of intronless mRNAs *in vitro*, other determinants in gene-specific activation by SM could include the stability of the transcript and other cellular export factors which may compete with SM for binding of target RNAs. In this study, we have demonstrated that SM associates with many EBV RNAs. This correlates with previous data which showed that expression of approximately 50% of the EBV ORFs was dependent on SM (12). However, in addition to the overall enrichment of RNA in SM IPs, which suggests a nonspecific RNA binding capability, preferential association with several RNAs was demonstrated. This hierarchy of affinities for SM suggests the presence of specific SM binding sites.

Interestingly, among the bound RNAs, we found several transcripts in both types of EBV infected cell lines examined, which were highly associated with SM. Regardless of their abundance in the EBV infected cells, SM exhibited a higher preference for these RNAs

suggesting that these transcripts may contain specific SM-binding elements. During our analysis we noticed that many of the highly enriched RNAs were expressed from late genes. For example BFRF3 codes for a component of the minor capsid antigen, BBRF3 codes for the envelope glycoprotein gM, BTRF1 codes for the tegument protein and BDLF3 codes for the glycoprotein gp150. These RNAs were highly enriched in both EBV infected cell lines, while the RNA expressed by BALF2, an early gene which was less highly enriched in both cell lines, codes for the single-stranded DNA binding protein. This preference of SM to bind late gene RNAs may be an indication that a specific time in RNA processing could also be an important factor for binding specificity of SM. For example, SM may bind nascent RNAs shortly after transcription. In order to address this issue, repeat IP/RT-qPCR experiments would need to be employed at different time points after induction of lytic replication.

Although several studies have shown binding of SM to RNA *in vitro* and *in vivo*, relatively little is known about the specificity of its targets. No specific RNA sequence motif or structural element has been identified to be required for SM binding. In our study, we showed that SM specifically binds the BFRF3 RNA sequence containing the 5' and 3' UTR but not its anti-sense sequence, although the anti-sense contains the same GC content. Therefore, it is likely that a specific sequence or structural element exists in RNAs to which SM binds. To further delineate the SM binding site of BFRF3, we constructed four different RNA transcripts representing one-fourth regions of BFRF3 and performed an *in vitro* cross-linking/IP assay. We found that SM specifically bound the first 189 nucleotides of BFRF3 (the 5'UTR) most highly and some binding were also detected in the second 186 nucleotides. Little binding was seen in the third quarter region of BFRF3 and no binding was found in the last quarter region of BFRF3. This

data suggests that there may be a specific sequence or structural motif in the first quarter of the BFRF3 sequence that is required for SM to bind.

Future Aims

Although we have identified a 189 nt region in which SM specifically binds *in vitro*, further verification is needed. The BFRF3 subclones used in the *in vitro* cross-linking assay may not represent the actual secondary structures of those BFRF3 regions. In order to test whether the first 189 nts of BFRF3 is required for SM binding, it should be possible to construct a plasmid containing the SM-binding portion of BFRF3 fused to a non SM-binding sequence, such as the anti-sense BFRF3 sequence, to determine if in fact the incorporation of our SM-binding region will allow SM to bind to the anti-sense BFRF3. If in fact the 5' UTR region of BFRF3 is required for SM binding, one could continue to further map down the region by repeating the crosslinking/IP assay using RNA targets representing various smaller regions of the 5'UTR. Further research such as repeating the *in vitro* crosslinking/IP assay with other highly enriched sequences such as BDLF3 and BBRF3 and mapping down their binding sites would also be helpful in identifying specific sequence binding motifs. Finally, an *in vivo* study looking for SM binding sites on RNAs should yield further insight into SM binding specificity and correlate the data presented here with *in vivo* SM response elements.

LIST OF REFERENCES

1. **Adams, A.** 1987. Replication of latent Epstein-Barr virus genomes in Raji cells. J. Virol. **61:**1743-1746.

2. **Baer R., Bankier A.T., and Biggin M.D.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature **310:**207-211.

3. **Bello, L. J., A. J. Davison, M. A. Glenn, A. Whitehouse, N. Rethmeier, T. F. Schulz, and J. Barklie Clements.** 1999. The human herpesvirus-8 ORF 57 gene and its properties. J. Gen. Virol. **80 (Pt 12):**3207-3215.

4. **Boyle, S. M., V. Ruvolo, A. K. Gupta, and S. Swaminathan.** 1999. Association with the cellular export receptor CRM 1 mediates function and intracellular localization of Epstein-Barr virus SM protein, a regulator of gene expression. J. Virol. **73:**6872-6881.

5. **Buisson, M., F. Hans, I. Kusters, N. Duran, and A. Sergeant.** 1998. The C-terminal region bu not the Arg-X-Pro repeat of Epstein-Barr virus protein EB2 is required for its effect on RNA splicing and transport. J. Virol. **73:**4090-4100.

6. **Chen, L., G. Liao, M. Fujimuro, O. J. Semmes, and S. D. Hayward.** 2001. Properties of two EBV Mta nuclear export signal sequences. Virology **288:**119-128.

7. **Chevallier-Greco, A., E. Manet, P. Chavrier, C. Mosnier, J. Daillie, and A. Sergeant.** 1986. Both Epstein-Barr virus (EBV)-encoded trans-acting factors, EB1 and EB2, are required to activate transcription from an EBV early promoter. EMBO J. **5:**3243-3249.

8. **Defechereux, P., L. Melen, L. Baudoux, M. P. Merville-Louis, B. Rentier, and J. Piette.** 1993. Characterization of the regulatory functions of varicella-zoster virus open reading frame 4 gene product. J. Virol. **67:**4379-4385.

9. **Fields, B. N., D. M. Knipe, and P. M. Howley.** 2007. Fields' virology. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia.

10. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23:**175-182.

11. **Gupta, A. K., V. Ruvolo, C. Patterson, and S. Swaminathan.** 2000. The human herpesvirus 8 homolog of Epstein-Barr virus SM protein (KS-SM) is a posttranscriptional activator of gene expression. J. Virol. **74:**1038-1044.

12. **Han, Z., E. Marendy, Y. D. Wang, J. Yuan, J. T. Sample, and S. Swaminathan.** 2007. Multiple roles of Epstein-Barr virus SM protein in lytic replication. J. Virol. **81:**4058-4069.

13. **Hiriart, E., L. Bardouillet, E. Manet, H. Gruffat, F. Penin, R. Montserret, G. Farjot, and A. Sergeant.** 2003. A region of the Epstein-Barr virus (EBV) mRNA export factor EB2 containing an arginine-rich motif mediates direct binding to RNA. J. Biol. Chem. **278:**37790-37798.

14. **Johannsen, E., M. Luftig, M. R. Chase, S. Weicksel, E. Cahir-McFarland, D. Illanes, D. Sarracino, and E. Kieff.** 2004. Proteins of purified Epstein-Barr virus. Proc. Natl. Acad. Sci. U. S. A. **101:**16286-16291.

15. **Marshall, M., U. Leser, R. Seibl, and H. Wolf.** 1989. Identification of proteins encoded by Epstein-Barr virus trans-activator gene. J. Virol. **63:**938-942.

16. **Nemerow, G. R. and N. R. Cooper.** 1984. Early events in the infection of human B lymphocytes by Epstein-Barr virus: the internalization process. Virology **132:**186-198.

17. **Papin, J., Vahrson, W., Hines-Boykin, R., Dittmer, D.P.** 2004. Real-Time Quantitative PCR Analysis of Viral Transcription. Methods in Molecular Biology **292**:449-480.

18. **Ragoczy, T., L. Heston, and G. Miller.** 1998. The Epstein-Barr virus Rta protein activates lytic cycle genes and can disrupt latency in B lymphocytes. J. Virol. **72:**7978-7984.

19. **Rooney, C. M., D. T. Rowe, T. Ragot, and P. J. Farrell.** 1989. The spliced BZLF1 gene of Epstein-Barr virus (EBV) transactivates an early EBV promoter and induces the virus productive cycle. J. Virol. **63:**3109-3116.

20. **Ruvolo, V., A. K. Gupta, and S. Swaminathan.** 2001. Epstein-Barr virus SM protein interacts with mRNA *in vivo* and mediates a gene-specific increase in cytoplasmic mRNA. J. Virol. **75:**6033-6041.

21. **Ruvolo, V., L. Sun, K. Howard, S. Sung, H. J. Delecluse, W. Hammerschmidt, and S. Swaminathan.** 2004. Functional analysis of Epstein-Barr virus SM protein: identification of amino acids essential for structure, transactivation, splicing inhibition, and virion production. J. Virol. **78:**340-352.

22. **Ruvolo, V., E. Wang, S. Boyle, and S. Swaminathan.** 1998. The Epstein-Barr virus nuclear protein SM is both a post-transcriptional inhibitor and activator of gene expression. Proc. Natl. Acad. Sci. U. S. A. **95:**8852-8857.

23. **Sandri-Goldin, R. M.** 2008. The many roles of the regulatory protein ICP27 during herpes simplex virus infection. Front. Biosci. **13:**5241-5256.

24. **Sergeant, A., H. Gruffat, and E. Manet.** 2008. The Epstein-Barr virus (EBV) protein EB is an mRNA export factor essential for virus production. Front. Biosci. **13:**3798-3813.

25. **Swaminathan, S.** 2005. Post-transcriptional gene regulation by gamma herpesviruses. J. Cell. Biochem. **95:**698-711.

26. **Takada, K. and Y. Ono.** 1989. Synchronous and sequential activation of latently infected Epstein-Barr virus genomes. J. Virol. **63:**445-449.

27. **Tanner J, Weis J, Fearon D, Whang Y, Kieff E.** 1987. Epstein-Barr virus gp350/220 binding to the B lymphocyte C3d receptor mediates adsorption, capping, and endocytosis. Cell **50:**203-213.

28. **Tsurumi, T.** 1997. Molecular mechanism of lytic phase of Epstein-Barr virus DNA replication. Nippon Rinsho **55:**321-327.

29. **Tsurumi, T.** 1991. Primer terminus recognition and highly processive replication by Epstein-Barr virus DNA polymerase. Biochem. J. **280 (Pt 3):**703-708.

30. **Winkler, M., S. A. Rice, and T. Stamminger.** 1994. UL69 of human cytomegalovirus, an open reading frame with homology to ICP27 of herpes simplex virus, encodes a transactivator of gene expression. J. Virol. **68:**3943-3954.

31. **Zalani, S., E. Holley-Guthrie, and S. Kenney.** 1996. Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. Proc. Natl. Acad. Sci. U. S. A. **93:**9194-9199.

BIOGRAPHICAL SKETCH

Zhao Han was born in Dalian, China and is the only child of Liping Zhang and Yichun Han. Zhao attended the University of Florida from 1999 to 2004 where she pursued a Bachelor of Science degree in microbiology and cell science with minors in chemistry and piano performance. During that time, Zhao was part of the University Scholars Program where she performed research measuring the ratios of phospholamban and serca levels in hypertensive rats under the guidance of Dr. Harm Knot. In August 2007, Zhao began her Master of Science degree in medical sciences under the supervision of Dr. Sankar Swaminathan. She completed her master's work and thesis in December 2008 and plans to continue her education in medical school.