CHARACTERIZING THE CRD-BP-RNA INTERACTION *IN-VITRO* AND IN CELLS

by

Mark Barnes

B.Sc. (Biochemistry and Molecular Biology), University of Northern BC, 2011

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
MATHEMATICAL, COMPUTER AND PHYSICAL SCIENCES
(CHEMISTRY)

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

April 2013

Your file Votre référence
ISBN: 978-0-494-94138-6

Our file Notre référence
ISBN: 978-0-494-94138-6

Abstract

The highly conserved family of RNA-binding proteins known as the VICKZ RNA-binding proteins play an integral role in the formation of cytoplasmic RNPs which leads to the stabilization, localization and translational control of many mRNA transcripts in the cell. The key investigation of this thesis was to analyse the binding ability of the VICKZ protein family member, the coding region determinant-binding protein (CRD-BP), both *in-vitro* and in cells. CRD-BP has four K-homology (KH) domains and two RNA-recognition motif (RRM) domains. Deletion studies in CRD-BP orthologs have shown that the KH domains, and not the RRM domains, are predominantly responsible for binding to RNA substrates. However, it is still unclear to what extent each of the KH domains play in their physical interaction with RNA molecules, nor is it known if each of the KH domains an play equal role in interacting with different RNA substrates. In an effort to address the above questions, we used site-directed mutagenesis to mutate the first glycine of the G-X-X-G motif in each KH domain separately, and in combinations. We mutated the glycine to an aspartate to introduce both physical and electrostatic hindrance for binding at the G-X-X-G motif. The goal was to determine if such a mutation can disrupt CRD-BP's ability to bind its RNA substrates both *in-vitro* and in cells. Our results showed that KH single mutants KH2, KH3 and KH4 did not disrupt the CRD-BP-c-myc CRD RNA interaction *in-vitro*. CRD-BP KH1 single mutant exhibited a modest reduction in binding to the c-myc CRD RNA substrate *in-vitro*. However, double KH domain mutations (KH1-2, KH1-3, and KH2-4) resulted in a complete abrogation of CRD-BP's ability to bind the c-myc CRD RNA substrate, suggesting these KH domains work in tandem to bind to the c-myc CRD RNA substrate *in-vitro*. Interestingly, the CRD-BP KH domain double mutant, KH3-4, showed only a modest reduction in the c-myc CRD RNA substrate binding, suggesting that the

first glycine in the G-X-X-G motif of KH3 and KH4 does not play a significant role in binding the c-myc CRD RNA substrate *in-vitro.*

The mRNA-binding ability of the CRD-BP KH single mutants, KH1, KH2, KH3 and KH4, were then analyzed in HeLa cells. Our results showed that KH1, KH2 and KH4 single mutants did not affect CRD-BP binding to β-actin mRNA in cells, but the KH3 single mutant showed a significant decrease in the ability to bind β-actin mRNA in cells. CRD-BP KH1, KH2, and KH3 single mutants showed a significant decrease in the ability to bind c-myc mRNA in cells, but the KH4 single mutant did not affect the CRD-BP-c-myc mRNA interaction in cells. Finally, all four CRD-BP KH single mutants showed a significant decrease in the ability to bind CD44 mRNA in cells. These results demonstrated, for the first time, that the KH domains of CRD-BP KH play unequal roles in binding different mRNAs in the cells. Such findings have important implications for future studies aimed at developing potential anti-cancer compounds that act by breaking CRD-BP-RNA interactions.

A secondary goal of this thesis was to test the ability of specific oligonucleotides and antibiotics to break the CRD-BP interaction in cells. Previous research showed that DD7 oligonucleotide and neomycin small molecule antibiotic can disrupt the CRD-BP-CD44 RNA interaction *in-vitro.* Results from this thesis showed that neither DD7 oligonucleotide nor neomycin had ability to disrupt CRD-BP-CD44 mRNA interaction in HeLa cells. Such findings have important implications for future studies aimed at developing potential anti-cancer compounds that act by breaking CRD-BP-RNA interactions.

Table of Contents

Abstract.....i

Table of Contents.....iii

List of Tables.....vi

List of Figures.....vii

Acknowledgements.....viii

Academic Discussions.....ix

References.....x

Chapter 1- Introduction

1.1 mRNA degradation in the control of gene expression.....1
1.2 *cis*-elements.....3
1.3 Ribonucleases.....4
1.4 microRNA.. 4
1.5 RNA-binding protein.....5
1.6 Coding region determinant-binding protein (CRD-BP) orthologs.....8
1.7 Structure and Function of CRD-BP and its orthologs.....9
1.8 CRD-BP and its orthologs control mRNA degradation.....13
1.9 CRD-BP in cancer.....14
1.10 Using oligonucleotides and antibiotics to disrupt protein-RNA interaction.....16
1.11 Aims of this research.....18

Chapter 2- Development of fluorescent based method to study CRD-BP-RNA interaction

2.1 Methodology.....21
2.1.1 Site-directed mutagenesis to generate CRD-BP mutant variants.....21
2.1.2 Transformation of competent *E.coli* BL21 (DE3) cells.....23
2.1.3 Cell Growth and Induction.....23
2.1.4 CRD-BP Purification under denaturing conditions.....24
2.1.5 Protein Dialysis.....24
2.2 Results.....25
2.3 Discussion.....28

Chapter 3- Development of a fluorescent based method to study the CRD-BP-RNA interaction *in-vitro*

3.1 Methodology.....30
3.1.1 PCR Generation of linear templates for in-vitro transcription.....30
3.1.2 Generation of unlabeled RNA acceptor molecule via *in-vitro* transcription.....31
3.1.3 Ligation of fluorescently labeled oligonucleotide to IVT acceptor.....32
3.1.4 Generation of 3'-Fluorescein-5-Thiosemicarbazide labeled RNA oligonucleotides.....32
3.1.5 Internally labeled RNA oligonucleotide using Fluorescein-UTP.....33
3.1.6 Fluorescent electrophoretic mobility shift assays.....33
3.2 Results.....34
3.3 Discussion.....40

Chapter 4- Assessing the ability of CRD-BP variants to bind c-myc CRD RNA *in-vitro*

4.1 Methodology.....44
4.1.1 Generation of internally ^{32}P-labeled c-myc RNA substrate.....44
4.1.2 Radioactive electrophoretic mobility shift assays.....45
4.1.3 Calculation of CRD-BP dissociation constant using [^{32}P] c-myc CRD 1705-1886 RNA substrate.....46
4.2 Results.....47
4.3 Discussion.....52

Chapter 5- Characterization of the CRD-BP-RNA Interaction in Cells

5.1 Methodology.....55
5.1.1 Immuno-precipitation of FLAG-CRD-BP-RNA complex.....55
5.1.2 cDNA synthesis and RT-qPCR.....57
5.1.3 Total RNA extraction from antibiotic-treated HeLa cells.....58
5.1.4 Oligonucleotides and antibiotics.....59
5.2 Results.....60
5.2.1 Optimizing the experimental conditions for immuno-precipitation coupled quantitative real-time PCR.....60
5.2.2 Assessing the ability of wild-type CRD-BP and its mutants to physically associate with mRNAs in HeLa cells.....63
5.2.3 Assessing the ability of oligonucleotides and small molecule antibiotics to disrupt CRD-BP-RNA interaction in cells.....66
5.3 Discussion.....68
5.3.1 Optimization of experimental conditions.....69
5.3.2 Effect of point mutation at KH domains G-X-X-G motif on the ability of CRD-BP to interact with mRNAs in cells.....71

5.3.3 Effect of specific oligonucleotides and small molecule antibiotics on CRD-BP-RNA interaction in cells.....73

Chapter 6-General Discussion

6.1 General Overview.....77
6.2 Generation and Purification of the Coding Region Determinant-Binding Proteins.....78
6.3 Development of a fluorescent based method to study the CRD-BP-RNA interaction.....79
6.4 Assessing the ability of CRD-BP mutant variants to bind c-myc CRD RNA *in-vitro*.....80
6.5 Characterization of CRD-BP-RNA Interaction in Cells.....84
6.6 Explaining the discrepancy between the CRD-BP-RNA interaction *in-vitro* and in cells.....88
6.7 Assessing the ability of oligonucleotides and small molecule antibiotics to disrupt CRD-BP-RNA interaction in cells.....90
6.8 Future Studies.....91
6.9 Concluding Remarks.....92

List of Tables

Table 1. Primers used in the site-directed mutagenesis of CRD-BP.....22
Table2. Primers used in the sequencing of the mutated *Mus musculus* CRD-BP DNA.....23
Table 3. Site-directed mutagenesis sequence information.....25
Table 4. Reagents used in the generation of linearized DNA template for IVT.....31
Table 5. Reagents used in the generation of unlabeled RNA substrates.....31
Table 6. RNA fragments and reagents used in the RNA ligation reactions.....32
Table 7. Reagents used for radioactive EMSA experiments 34
Table 8. RNA Labeling efficiency data. 37
Table 9. Reagents used for radioactive EMSA experiments.....46
Table 10. CRD-BP variant dissociation constants (K_d) for c-myc CRD RNA.....53
Table 11. 2'-O-methyl oligonucleotides used in cell-based experiments.....59
Table 12. Optimal antibiotics and oligonucleotide treatment concentrations.....60
Table 13. Quantification and quality of immuno-precipitated RNA.....61
Table 14. Primer sets used in real-time RT-PCR experiments.....61

List of Figures

Figure 1. Eukaryotic mRNA degradation pathways.....2
Figure 2. A variety of sequence motifs within mRNAs recognised by RNA-binding proteins.....6
Figure 3. Schematic diagram of CRD-BP structural domains.....9
Figure 4. Schematic of Type I KH domain structure.....11
Figure 5. Conserved structural features within Type I KH domains responsible for nucleic acid interactions.....12
Figure 6. Locations of point mutations within CRD-BP.....19
Figure7. SDS-PAGE analysis of recombinant 6xHis-tagged CRD-BP purified from *E.coli.* BL21 cells.....26
Figure 8. Post-dialysis SDS-PAGE analysis of purified 6xHis-tagged CRD-BP elution fractions chosen for binding experiments.....28
Figure 9. Standard curve for BSA protein assay.....28
Figure 10. Agarose gel analysis of linearized c-myc 167 nt DNA template to be used in *in-vitro* transcription.....34
Figure 11. Agarose gel analysis of *in-vitro* transcribed 167nt RNA fragment.....35
Figure 12. RNA ligation reaction analysis on 12% denaturing poly-acrylamide gel.....35
Figure 13. Analysis of 3'-fluorescein-5-thiosemicarbazide end-labeling of *in-vitro* transcribed 182nt c-myc and 39nt CD44 RNA substrates.....36
Figure 14. Analysis of internally labeled RNA substrates using Fl-UTP.....37
Figure 15. Electrophoretic Mobility Shift Assay of WT-CRD-BP and c-myc CRD (1705-1886) RNA substrate.....38
Figure 16. Electrophoretic mobility shift assay of WT-CRD-BP and c-myc CRD (1705-1886) RNA substrate.....39
Figure 17. Electrophoretic Mobility Shift Assay of WT-CRD-BP and CD44 3'UTR RNA substrate.....40
Figure 18. Schematic diagram of the RNA ligation reaction to generate the fluorescently-labeled RNA substrate.....41
Figure 19. Audioradiograph to calculate the dissociation constant (K_d).....46
Figure 20. The generation of ^{32}P internally labeled c-myc CRD (nts 1705-1886) RNA.....47
Figure 21. Electrophoretic mobility shift assay of wild-type and mutant CRD-BP.....48
Figure 22. CRD-BP variants binding curves as generated by the Hill Equation.....50
Figure 23. Comparison of CRD-BP variants dissociation constants (K_d).....51
Figure 24. HeLa cells visualised under a light microscope.....60
Figure 25. Melt-curve analysis of RT-qPCR primers.....62
Figure 26. Standard curves generated to assess PCR reactions efficiency.....63
Figure 27. Expression of exogenous FLAG-WT-CRD-BP in HeLa cells.....63
Figure 28. Analysis of specific mRNAs physically associated with CRD-BP and its mutants...65
Figure 29. Effect of antibiotic treatment on mRNA and protein expression in WT-CRD-BP transfected HeLa cells.....67
Figure 30. Effect of oligonucleotide treatment on protein and gene expression in WT-CRD-BP transfected HeLa cells.....68
Figure 31. Structures of antibiotics used in protein binding experiments.....74
Figure 32. IMP1 KH34 pseudodimer configuration of reveals β-actin RNA-binding surfaces..81

Acknowledgements

I would like to thank the following individuals who have supported, encouraged and assisted me throughout my M.Sc. journey: Dr. Chow Lee, Dr. Maggie Li and Gerrit van Rensburg. I would also like to thank all current and past Lee lab members who have helped me. A special thanks goes out to my M.Sc. supervisory committee members Dr. Stephen Rader and Dr. Rob Olson; your time was generously donated at my request, which I am truly grateful for. Also, thank you to the members of the Gorrell, Rader and Gray labs that have helped me. Most importantly I would like to thank my wife, Christina Barnes, for her constant encouragement and financial support throughout my time in Graduate school. I would also like to thank Bobby LeBruno and Apollo for all of their ongoing support throughout my research experience.

Academic Discussions

Barnes M, Van Rensburg G and Lee CH. (2013) Determining the Functional Domains of the RNA-Binding Protein CRD-BP. 8th Annual UNBC Graduate Conference, UNBC, Prince George.

Barnes M, Van Rensburg G and Lee CH. (2012) Analyzing the RNA-binding Function of an Oncogenic Protein. Realities of Northern Oncology Conference 2012.

Mehmood K, **Barnes M** and Lee CH. (2012) Development of fluorescent-based method to study the CRD-BP-RNA interaction. 8th Annual Annual Western Canadian RNA conference (RiboWest), University of Lethbridge, Lethbridge.

Barnes M, Lee CH. (2011) Breaking the CRD-BP-RNA Interaction. 7th Annual Annual Western Canadian RNA conference (RiboWest), Prince George.

Chapter 1- Introduction

Cancer genetics is extremely complex; many cellular events simultaneously take place that give rise to observed cancer phenotypes. There are six distinct, but complementary, hallmarks that have been identified to be necessary for tumor progression; sustaining proliferative signaling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis and resisting cell death (Hanahan and Weinberg., 2011). The underlying mechanisms that enable each of these hallmarks to materialize can be explained by The Central Dogma of Molecular Biology. The Central Dogma of Molecular Biology states that genetic material flows from DNA to RNA to protein. There are numerous steps to go from the DNA molecule in the nucleus, to a functional protein within the cell. Each step in the generation of a functional protein is a possible site for the regulation of a gene; gene expression is regulated at the transcriptional, post-transcriptional or translational levels. The abnormal regulation of a gene at any of these levels can give cells the capabilities to alter their gene expression that ultimately leads to neoplasms (Diaz-Cano, 2012). To understand the complexity of tumor heterogeneity at the translational level, one must first understand the factors that are involved in the normal expression of a gene.

1.1 mRNA degradation in the control of gene expression

The various gene regulatory mechanisms within the cell are quite complex (Farina et al., 2008). For instance, mRNA can be degraded by many mRNA decay pathways, which is dependent on its *cis*-element sequences and availability of various trans-acting factors (Farina et al., 2008). The addition of the 5' methyl-guanasine cap and the 3' poly(A) tail consisting of multiple adenosine monophosphates are post-transcriptional modifications that are done to protect the mRNAs from nucleolytic attack from degradative enzymes at the 5' and 3' ends of

mRNA, respectively. There are three mRNA decay pathways; the deadenylation-dependent pathway, deadenylation-independent 5'decapping pathway and endonucleolytic pathway. The current dogma favours the 5' decapping and the 3' deadenylation-dependent pathways as the main mode to degrade mRNA molecules. The cell employs various mechanisms to metabolize both normal and abnormal mRNA transcripts. Exosomes, which are complexes of ribonucleases, are involved in the degradation of aberrant mRNA, 3'-5' and 5'-3' exonucleolytic decay pathways, and ARE-mediated decay (indicated by the Ex in Figure 1) (Li et al., 2010).

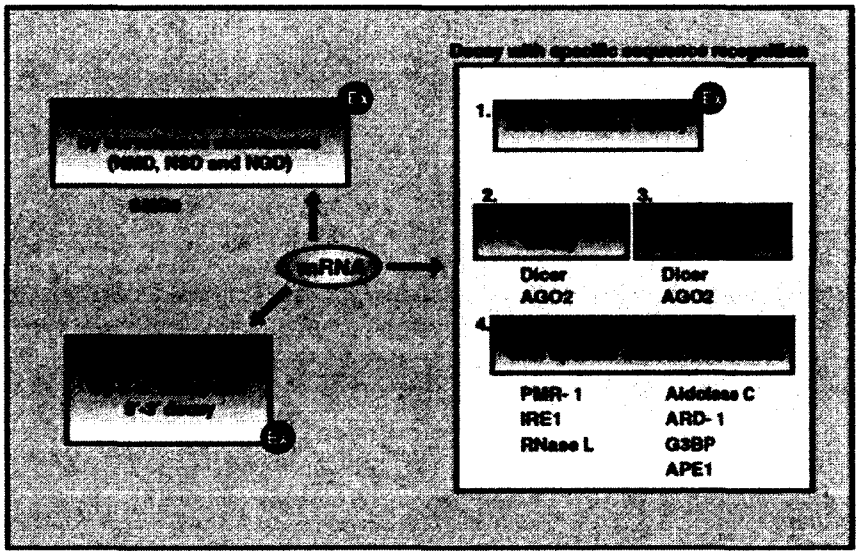

Figure 1. Eukaryotic mRNA degradation pathways (Li et al., 2010).

The endonucleolytic pathway is a much less studied, and perhaps less understood mRNA degradation pathway, but it does seem to have a place in the control of gene expression (Barnes et al., 2009). Recent research has identified that the role of endoribonucleases in eukaryotic mRNA metabolism is more significant than previously believed (Li et al., 2010). For instance, Rrp44 (Dis3), a component of the eukaryotic exosome complex, contains multiple structural

motifs that are responsible for both its RNA-binding activity and endoribonucleolytic activity, more specifically, the PIN domain contains a conserved amino acid sequence that is responsible for its endoribonucleolytic activity (Li et al., 2010). Rrp44 has been shown to play a role in each of the mRNA decay pathways described above. Furthermore, it has been shown that the surveillance mechanisms within mammalian cells that are responsible for nonsense-mediated mRNA decay involves the cleavage of the mRNA transcript by SMG6, which contains a PIN domain that has shown to have endonucleolytic activity *in-vitro* (Li et al., 2010). One commonality amongst each of these pathways is that RNA-binding proteins play a vital role in each of the mRNA decay pathways by binding to the transcripts and mediating these steps; thus, leading to the regulation of the genes involved.

1.2 *cis*-elements

cis-regulatory elements are RNA sequences that have the ability to regulate the expression of the transcript on which the sequence is found (Alberts et al., 2008). Theses sequences are often binding sites for RNA-binding proteins that either protect mRNA transcripts, or promote the degradation of the transcript. Iron response elements (IREs) are an example of a *cis*-regulatory element. IREs are sequences found in the 5' or 3' ends of mRNAs. When iron regulatory proteins are bound to these sequences, they alter the expression of the target gene (Nelson and Cox., 2008). Other important regulatory sequences within mRNA transcripts are the AU-rich elements (ARE). These *cis*-regulatory elements are located within in the 3' UTR of mRNA transcripts and they have been found to both stabilize and target mRNAs for degradation (Li et al., 2010). The *cis*-regulatory element, the coding region determinant (CRD), found within the c-*myc* mRNA transcript, has been shown to be the binding site for CRD-BP, and when bound to c-

myc mRNA, increases the mRNA half-life by stabilizing the transcript and protecting it from endonucleolytic attack by nucleases (Barnes et al., 2009).

1.3 Ribonucleases

Ribonucleases are enzymes that have the ability to catalyze the degradation of RNA molecules both in an exoribonucleolytic and endoribonucleolytic fashion. There are numerous ribonucleases and they are classified as according to the mechanism used to degrade the RNA molecules. Exoribonucleases are enzymes that degrade the RNA molecules externally from either the 5' or 3' end (Farina et al., 2008). Conversely, endoribonucleases are enzymes that catalyze the degradation of RNAs by attacking the molecule internally. RNaseA is a well-studied endoribonuclease that is commonly used in research laboratories. RNaseA prefers to cleave single stranded RNA molecules predominately at unpaired C and U nucleotides on the 3' end of mRNAs (Richards, 1972). RNaseH is another ribonuclease that cleaves RNA in an RNA/DNA duplex, resulting in a single stranded RNA molecule (Schmitt 2009). RNA-binding proteins have the ability to protect mRNA transcripts from ribonucleases by binding to mRNA targets and shielding them from ribonucleolytic attack.

1.4 microRNA

It has been shown that there are many factors that have the ability to regulate genes within the cell. Although coding regions in both DNA and RNA play essential roles in the expression of genes, there are noncoding RNA transcripts that regulate many animal genes (Alberts et al., 2008). Short noncoding RNA called microRNA (miRNA) appear to regulate the expression of as many as one third of the protein coding genes in humans (Alberts et al., 2008). These single stranded RNA transcripts locate and bind to complementary mRNA transcripts

(Alberts et al., 2008). Once synthesized, the pre-miRNAs are processed in the nucleus and then the mature miRNA interact with the *RNA-induced silencing complex* (RISC) which degrades mRNA transcripts (Alberts et al., 2008).

Many miRNAs are known to base pair with mRNAs, usually in the 3'-UTR of the mRNAs, which targets the mRNAs for degradation as mention above (Elcheva et al., 2008). The ubiquitin ligase β*TrCP1* has an unstable mRNA transcript that is rapidly degraded in the cell due to its *cis*-acting destabilizing element located within the β*TrCP1* mRNA coding region (Elcheva et al., 2009). miR-183 has been shown to bind to the coding region of the β*TrCP1* mRNA, thus targeting the transcript to miRNA-dependent degradation (Elcheva et al., 2009). The RNA-binding protein CRD-BP can bind to the β*TrCP1* coding region, protecting it from the miRNA dependent mRNA decay (Elcheva et al., 2009).

1.5 RNA-Binding Proteins

The post-transcriptional modifications of mRNA transcripts at the 5' and 3' ends stabilize and protect the mRNAs from nucleolytic attack from degradative enzymes. Another key protector of mRNAs, as indicated above, is RNA-binding proteins. RNA-binding proteins bind to and protect mRNAs from nucleases (Mullen and Marzluff., 2008). It is well documented that RNA-binding proteins play a role in the translational regulation of genes; however, little is known about the specific interactions that allow for this regulation (Hogan et al., 2008). It is for these reasons that further study of RNA-binding proteins is intriguing to researchers. Recent studies have shown that there is extensive RNA-binding protein involvement in the post-transcriptional processing of mRNAs. These proteins recognize specific features in the mRNAs and they bind nascent transcripts at specific times during the processing of the mRNA (Hogan et

al., 2008). Hogan *et al* (2008) showed that there is a variety of motifs that the RNA-binding proteins are able to recognize and subsequently bind to (Figure 2). The consensus among sequences recognized by many different RNA-binding proteins may assist the identification of new RNA motifs for other RNA-binding proteins (Hogan et al., 2008).

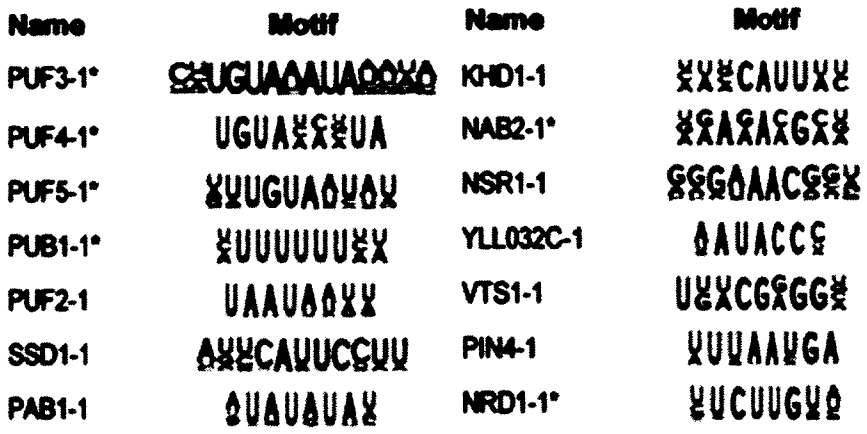

Figure 2. A variety of sequence motifs within mRNAs recognised by RNA-binding proteins (Hogan et al., 2008).

Due to the gene regulatory nature of RNA-binding proteins, their overexpression has been identified in many diseases; such as diabetic retinopathy, coronary artery disease and tumor angiogenesis in cancers (Levy et al., 1998). It is well known that tumor angiogenesis is key to tumor growth and metastasis of cancers. Vascular endothelial growth factor (VEGF) is an angiogenic factor that, when expressed, signals angiogenesis (Levy et al., 1998). The ubiquitous RNA-binding protein Hu-antigen (HuR) is comprised of three RNA recognition motifs (RRM) which have been shown to bind AREs in the 3' UTR of VEGF mRNA and stabilize the mRNA transcript (Benoit, R.M., et al, 2010; Levy, N., et al., 1998). Of the many ARE-binding proteins, HuR is the only one known to use its first two tandem RRMs to bind AREs and silence genes via

an ARE-mediated decay pathway (Benoit et al., 2010). The crystal structure of the first RRM domain of HuR has been resolved and it revealed that the two α-helices and a four-stranded β-sheet structure contained hydrophobic residues that resembled RRM domains of other RNA-binding proteins. These common structural features between HuR and other RNA-binding proteins indicate their functional significance and may reveal potential targets for the synthesis of anticancer therapeutics (Benoit et al., 2010).

The neuro-oncologic ventral antigen (Nova) RNA-binding protein has been shown to be present in the autoimmune neurological disorder paraneoplastic opsoclonus-myoclonus ataxia and highly expressed in tumors that grow outside the central nervous system (Lewis et al, 1999). Nova contains three highly conserved KH domains connected by a flexible G-X-X-G linker region, similar to that of hnRNPs (Lewis et al, 1999). The third KH domain contains an RNA binding scaffold that recognizes sequences rich in UCAU tetrads which is responsible for Nova RNA binding (Lewis et al, 1999). Mutagenesis studies of proteins that share KH domain structural features with Nova revealed common amino acids that are involved in RNA-binding. For example, the GLD-1 single KH domain from *C. elegans* shares structural features with Nova and has been studied to determine the specifics of the binding surfaces. Loss-of-function mutations (Gly227-Asp and Gly227-Ser) were incorporated in to the G-X-X-G loop of GLD-1 in the attempt to disrupt RNA binding (Lewis et al, 1999). It was shown that the amino acids in the G-X-X-G loop of GLD-1 important for RNA-binding were also important for RNA-binding in Nova; this suggest similar binding characteristics between conserved protein domains (Lewis et al, 1999). Due to the structural similarities between KH domains of different proteins, therapeutic approaches that target one RNA-binding protein may be ideal for other RNA-binding proteins.

1.6 Coding region determinant binding protein (CRD-BP) orthologs

The mouse coding region determinant binding protein (CRD-BP) is a member of the VICKZ RNA-binding proteins (Vg1RBP/Vera, IMP-1,2,3, CRD-BP, KOC, ZBP-1) (Yisraeli, 2005). Multiple studies have been performed to determine the mRNAs that VICKZ proteins are known to bind. The *Xenopus* Vg1 RNA-binding protein (Vg1RBP, also known as Vera) localizes Vg1 mRNA within the cell during *Xenopus* oogenesis (Git and Standart, 2002). Vg1RBP recognizes the specific 3'UTR vegetal localization element (VLE) and localizes the mRNA to the vegetal cortex; therefore, *Xenopus* Vg1RBP is key in cellular developmental and differentiation (Git and Standart, 2002). The human form of CRD-BP, IMP-1, has been shown to bind to at least 5 RNA targets, many of which are implicated in cancer (Liao et al., 2005). IMP-1 binds to the 5'-UTR of IGF-II mRNA, inhibiting the translation of the transcript (Liao et al., 2005). Transgenic mice that overexpress IMP-1 have shown increases in IGF-II protein levels, which was accompanied by the induction of mammary tumors (Liao et al., 2005). IMP-1 plays a role in the localization of H19 RNA by binding to the 3'-UTR of H19 mRNA; it also binds to the coding region determinant and stabilizes c-myc mRNA, promoting translation; Tau mRNA, which is mainly localized within axons, is also bound by IMP-1 in tau's 3'-UTR (Liao et al., 2005). These studies have shown that IMP-1 is implicated in the regulation of many oncogenic genes. The expression pattern of the KOC RNA-binding protein (also known as IMP-3) in benign urothelium and urothelial tumors is absent in low-grade urothelial tumors; however, KOC overexpression was observed in high-grade papillary urothelial carcinoma (Li et al., 2008). It shows that the VICKZ protein family play extensive roles in the regulation of many genes, many that lead to the development of various cancers (Doyle et al., 1998).

1.7 Structure and function of CRD-BP and its orthologs

The mouse coding region determinant binding protein (CRD-BP) is a member of the highly conserved VICKZ RNA-binding protein family (Yisraeli, 2005). CRD-BP is a 577 amino acid protein which contains two RNA recognition motifs (RRMs) and four K-homology (KH) domains (Figure 3) (Chao, J., et al., 2010).

Figure 3. Schematic diagram of CRD-BP structural domains (Chao, J.A., et aL., 2010).

Many RNA targets of CRD-BP have been identified, and it has been shown that the KH domains, not the RRMs, are responsible CRD-BP binding to its RNAs (Git, A., and N. Standart, 2002; Oberman, F., et al., 2002; Farina, K.L., et al, 2003; Chao, J., et al., 2010). Truncation variants of Vg1RBP, the ortholog of CRD-BP in frogs, were generated to determine the important RNA-binding domains (Git and Standart, 2002). It was shown that variants spanning both KH didomains, R-K12-K34 and K12-K34, showed the greatest affinity for its vegetal localization element (VLE) with a K_d of 15 nM. Variants containing single KH didomains (R-K12, K12, -K34, K34) showed a reduction in the affinity for VLE with a K_d of 300 nM; and the variant that only contained the RRM didomain (R-) showed the lowest affinity for VLE with a K_d of 750 nM (Git and Standart, 2002). Therefore, it was concluded that the RRM domains of Vg1RBP do not play a significant role in binding VLE (Git and Standart, 2002). Furthermore, Oberman *et al* (2002) showed that when Vg1RBP was truncated with the missing KH4 domain,

it was no longer able to bind the VLE RNA probe; indicating that the KH4 domain plays an important role in binding VLE.

To determine the importance of structural domains of ZBP1, the ortholog of CRD-BP in chickens, in binding β-actin mRNA, Farina *et al* (2003) performed GST pull-down assays using truncated versions of ZBP1. It was shown that full length ZBP1 was able to bind to the zipcode region of β-actin mRNA with a K_d of 1-10 nM. The truncated versions containing only the RRM domains or only the KH1-KH2 didomain, were unable to bind the zipcode region of β-actin mRNA (Farina et al., 2003). This showed that KH3 and KH4 domains, not the RRMs nor KH1 and KH2 domains, are important in binding the β-actin mRNA zipcode region. A recent structural study showed that ZBP1 induces the RNA looping of β-actin mRNA upon KH3 and KH4 domain recognition, suggesting the importance of the KH3 and KH4 domains in binding the β-actin mRNA (Chao, J., et al., 2010). In the same study, KH34 didomain crystal structure of IMP1, the human ortholog of CRD-BP, was determined to a resolution of 2.75 Å using molecular replacement (Chao, J., et al., 2010). It was revealed that both the KH3 and KH4 domains adopted the $\beta_1\alpha_1\alpha_2\beta_2\beta_2\alpha_3$ orientation with the highly conserved flexible G-X-X-G linker connecting α_1 and α_2 helices; it was concluded that KH3 and KH4 played a significant role in the binding of β-actin mRNA zipcode region (Chao, J., et al., 2010). It should be noted that the G-X-X-G linker is present in all KH domains of the VICKZ RNA-binding proteins.

This is not the first time a G-X-X-G motif of a KH domain has been shown to play an important role in binding RNA substrates. The hnRNP K homology (KH) domain was first identified in the human heterogeneous nuclear ribonucleoprotein K (hnRNP K) and it was shown that the KH domains were directly associated with binding RNA and ssDNA (Valverde et al., 2008). As with CRD-BP, KH domains are typically found in multiple copies that are separated

by a flexible Gly-rich linker (G-X-X-G motif) (Valverde et al., 2008). The eukaryotic type I fold KH domains contain a β-sheet which is composed of three antiparallel β-strands abutted by three α-helices ($\alpha_1\alpha_2\alpha'$) (Figure 4); this secondary structural arrangement seems to be vital for physical interactions of KH domains with its RNAs (Valverde et al., 2008).

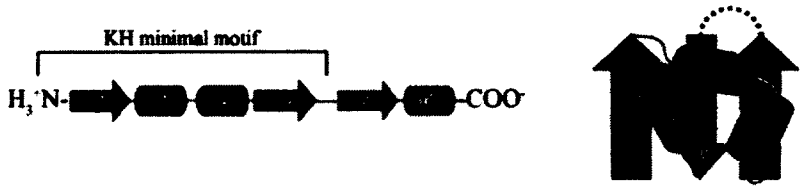

Figure 4. Schematic of Type I KH domain structure (Valverde et al., 2008).

KH domains between proteins show a high level of similarity. For instance, KH1 domain in CRD-BP is more similar to KH1 in another protein, than it is to KH2 of CRD-BP (Valverde et al., 2008). This is indicative of the highly conserved functionality of each the KH domains (i.e. the KH domains play similar roles between different orthologs). Common secondary structural features are found in all KH domains that assist in the recognition and interaction of four nucleic acids. The α_1-helix, G-X-X-G linker, α_2-helix, β_2-strand and the variable looped regions are all present in KH domains and play a vital role in binding (Figure 5) (Valverde et al., 2008). Poly(C)-binding proteins (PCBPs), is a nucleic acid-binding protein involved in multiple cellular processes, including mRNA stabilization and translational control (Valverde et al., 2008). PCBP contains three type I KH domains that has been shown to bind the poly(C)-rich nucleic acid sequence 5'-AACCCU/TAACCCU/T-3' and through crystallization, it was identified that the PCBPs recognize the tetranucleotide motif consisting of CCCU/T nucleotides (Valverde et al., 2008).

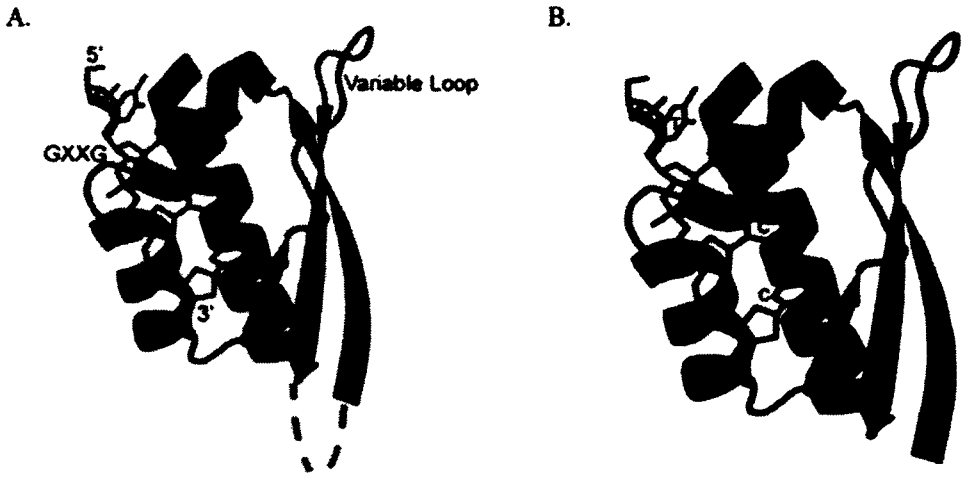

Figure 5. Conserved structural features within Type I KH domains responsible for nucleic acid interactions. The secondary structural configuration (A) promote the energetically favorable interactions with nucleic acid motifs (B) (Valverde et al., 2008).

Furthermore, tetranucleotides have been shown to interact with the KH4 domain of KSRP and the KH3 domain of Nova-2 in a similar fashion, both of which interact with the hydrophobic binding platform contacting the G-X-X-G motifs of the KH domains (Valverde et al., 2008). It has also been shown that the KH domains bind RNAs in a tandem, but independent fashion. For instance, the KH3 and KH4 domains of the FUSE-binding protein are connected by a Gly-rich linker and have been shown to regulate c-myc expression (Valverde et al., 2008). Although the KH domains of FUSE-binding protein are closely linked together, they behave independently of each other. The ssDNA binds in the groove between the α_1-helix and α_2-helix adjacent to the G-X-X-G loop of one KH domain; the charged hydrophilic edges of the 10Å binding site favor the binding of pyrimidines over purines (Valverde et al., 2008). These binding features found in the FUSE-binding protein have been found to be common to KH domain binding proteins, including CRD-BP (Valverde et al., 2008).

1.8 CRD-BP and its orthologs control mRNA degradation

The role of RNA-binding proteins in mRNA decay pathways has been well documented. CRD-BP is an RNA-binding protein that has been shown to be involved in the control the many mRNAs expression within the cell. For instance, CRD-BP has been shown to bind to the coding region of βTrCP1mRNA, protecting it from degradation (Elcheva et al., 2009). Using immuno-precipitation and northern analysis, it was shown that upon overexpression of CRD-BP, βTrCP1 mRNA and protein levels increased significantly, suggesting that CRD-BP plays a direct role in the expression of βTrCP1 at the translational level (Noubissi et al., 2006). It was shown that βTrCP1expression was β-catenin/Tcf-dependent and the CRD-BP promoter region directly interacts with β-catenin, which leads to the up-regulation of βTrCP1 (Noubissi et al., 2006). IMP1 was originally discovered due to its ability to associate with and stabilize insulin-like growth factor II (IGF-II) mRNA in mouse embryonic development (Zhou et al., 2009). Using mobility shift analysis and cellular localization experiments, it was determined that IGF-II leader 3 mRNA contains six IMP1 binding sites and once IMP1is bound, it stabilizes and localizes IGF-II mRNA (Nielsen et al., 1999). IMP1 has been shown to mediate CD44 mRNA degradation. Using immuno-precipitation and qRT-PCR analysis, FLAG-IMP1 was shown to associate with CD44 mRNAs in cells (Vikesaa et al., 2006). CD44 has been shown to be important in the formation of invadopodia and in IMP-depleted cells there was a significant reduction the formation of invadopodia; suggesting a direct physiological role of the IMP-CD44 interaction in cells (Vikesaa et al., 2006). Furthermore, IMP1/CRD-BP has been associated with H19 RNA, Tau mRNA, c-myc mRNA, MDR1 mRNA and PABP mRNA and it has been suggested that IMP potentially associates with 2-3% of the cellular transcriptome, which was determined by a

comprehensive microarray analysis of IMP1-containing RNP (ribonucleoprotein) granules (Zhou, Y., et al, 2009; Jonson, L., et al, 2007; Sparanese, D and C, Lee. , 2007)

1.9 CRD-BP in Cancer

CRD-BP has been shown to play roles in human cancers by stabilizing oncogenic mRNAs; which include c-myc and CD44 mRNAs. c-myc, a helix-loop-helix/leucine zipper type transcription factor, induces cell proliferation, differentiation and neoplastic transformation when abundantly expressed (Doyle et al., 1998). CD44 protein belongs to a family of cell surface adhesion molecules that are involved in both cell-cell and cell-matrix communications (Goodison et al., 1999). The CD44 gene, which has only been shown to be expressed in higher organisms, codes for a transmembrane protein that contains an extracellular N-terminal and cytoplasmic tail (Goodison et al., 1999). CRD-BP binds CD44 mRNA at multiple sites at its 3' UTR leading to the stabilization of CD44 mRNA. This in turn increases the translation of the CD44 proteins necessary for cell adhesion, cytoplasmic spreading and invadopodia formation (Vikesaa et al., 2006). CRD-BP regulates the expression of CD44, therefore, it orchestrates cellular invasion of surrounding tissues.

As mentioned above, the degradation of βTrCP1 is regulated by CRD-BP. βTrCP1is an F-box protein that is involved in cell-cycle regulation and it has been shown to target β-catenin and IκB for ubiquitylation and degradation; βTrCP1 has also been shown to control the degradation of oncogenic RNAs upon ubiquitylation (Nakayama and Nakayama, 2006). Consequences of the overexpression of βTrCP1 include the increased ubiquitylation of both β-catenin and IκB which leads to cellular proliferation due to the Wnt pathway inactivation; genetic studies show alterations in βTrCP genes that result in cancer phenotypes (Nakayama and Nakayama, 2006). CRD-BP has been shown to bind and stabilize the zinc finger transcription

factor, GLI1, which has been implicated in the Wnt and Hedgehog (Hh) pathways that are critical to embryonic development and the progression of tumors (Noubissi et al., 2009). Upon CRD-BP knockdown, β-catenin/Tcf–dependent GLI1 up-regulation was prevented in 293T cells, implying that CRD-BP stabilizes GLI1 which then acts on β-catenin and Tcf (Noubissi et al., 2009). Increased levels of *K-Ras* mRNA have been observed in colon cancer cells and this overexpression is facilitated by the overexpression of IMP1 (Mongroo et al., 2011). IMP1 binds to two regions located on *K-Ras* mRNA; both the coding region and the 3' UTR of *K-Ras* mRNA are binding sites for IMP1 (Mongroo et al., 2011). It is hypothesized that IMP1 regulates the expression of K-Ras by directly interacting with *K-Ras* mRNA which results in cellular proliferation due to K-Ras involvement in the Wnt/Hh pathway (Mongroo et al., 2011). Microphthalmia-associated transcription factor (MITF) is a helix-loop-helix leucine zipper transcription factor that belongs to the MYC superfamily of proteins and it is involved in melanocyte development and melanogenesis (Goswami et al., 2010). Knockdown of CRD-BP in Normal Human Melanocytes (NHM) resulted it a reduction of MITF-dependent luciferase activity; however, the overexpression of CRD-BP in NHMs resulted in a 3-fold increase in the MITF mRNA levels (Goswami et al., 2010). MITF mRNA evades the miRNA-regulated degradation upon CRD-BP binding to MITF mRNA and counteracting the miR-340-mediated degradation (Goswami et al., 2010). The Ras-ERK/MAP (Mitogen-Activated Protein) kinase signaling pathway plays an extensive role in the growth and proliferation of cells and it has been shown that this pathway is tightly regulated by KH domain RNA-binding proteins (Whelan et al., 2012). Therefore, it is well documented that CRD-BP orthologs play an extensive role in the regulation of genes within the cell, many of those who are involved in aberrant cell growth.

CRD-BP has also been implicated in the development of mammary tumors in adult transgenic mice; transgenic mice expressing WAP-CRD-BP showed a 95% correlation with WAP-CRD-BP expression and the formation of mammary tumors (Tessier et al., 2004). Furthermore, the known CRD-BP mRNA target IGF-II showed a 100-fold increase in mammary glands expressing WAP-CRD-BP; indicating the role in CRD-BP tumor progression (Tessier et al., 2004).

As mentioned above, CRD-BP binds and protects mRNAs from nucleolytic attack of nucleases. CRD-BP expression levels are up-regulated in many human cancers including breast, colon, skin, ovary and brain, among others (Doyle, G., et al., 1998; Ioannidis, P., et al., 2005; Dimitriadis, E., et al., 2007; Kobel, M., et al., 2007; Doyle, G., 2000, Ioannidis, P., 2003; Ross, J., 2001; Elcheva 2008). CRD-BP over-expression in human breast cancers, and the rates that it is detected in other cancers, shows a strong correlation between CRD-BP expression and cancer (Doyle, G., 2000; Ioannidis, P., et al., 2005). The above discussion provided extensive evidence for the role of CRD-BP implication in cancers. CRD-BP expression is extremely low in normal adult tissues; it is however, abundantly expressed in many types of human cancers. Hence, CRD-BP is an ideal target for cancer therapy.

1.10 Using oligonucleotides and antibiotics to disrupt protein-RNA interaction

The expression of genes often depends on the closely regulated interactions between proteins and RNAs which form complexes to function. As a result, these complexes can be targeted to prevent expression of genes. There are few possible approaches to targeting protein-RNA interactions in cells. Whether it is targeting the protein or mRNAs, inhibiting proteins from interacting with certain RNAs can be a useful way of inhibiting gene expression.

RNA interference (RNAi) is a ubiquitous mechanism that is used in cells to reduce the expression of genes. The RNAi mechanism relies on double-stranded RNAs being processed into in single-stranded RNAs which then bind to complementary sequences on mRNA transcripts to trigger degradation by RISC (Dykxhoorn et al., 2006). This is an extremely useful mechanism when it comes to regulating gene expression in cells. Researchers use the RNAi mechanisms to their advantage by designing oligonucleotides to bind to certain mRNAs resulting in reduced protein expression. Furthermore, oligonucleotides have been used in clinical settings to yield the same result. Zellweger *et al* (2002) report using antisence oligonucleotides (AON) against clusterin mRNA, in conjunction with radiation therapy, to improve pateint prognosis. Clusterin expression plays a protective role against apoptic cell death and it is associated with androgen-independent recurrence of prostate cancer; the addition of the AON to the treatment regimen resulted in increased cell death (Zellweger, T., et al, 2002).

The use of molecules to break protein-RNA interaction as a therapeutic approach has been largely un-explored. There are, however, some examples of this approach. The aminoglycoside antibiotic neomycin B, has been reported to break protein-RNA interactions (Zapp et al., 1993). During early infection of HIV, the small accessory protein Rev binds to the Rev Response Element (RRE) of mRNAs, both protecting the mRNA and localizing the transcript into the cytoplasm where it is translated (Van Ryk and Venkatesan, 1999). Neomycin B was used to inhibit the binding of Rev to the RRE, thus preventing the localization of mRNAs to the cytoplasm (Van Ryk and Venkatesan, 1999). Furthermore, using a high-throughput fluorescence polarization method, low-molecular-weight inhibitors have been identified for the RNA-binding protein HuR (Meisner et al., 2007). HuR homodimerization formation has been observed when HuR binds to its ARE-controlled mRNAs and low-molecular-weight inhibitors

have been used to disrupt this dimerization and prevent HuR from binding its mRNA (Meisner et al., 2007). Using fluorescence polarization, it was shown that the low-molecular-weight molecules Dehydromutactin and MS-444 were able to disrupt the HuR-ARE-RNA interactions at concentrations of 1-20 uM of molecule (Meisner et al., 2007). These inhibitors were also shown to inhibit HuR trafficking in cells and two HuR-mediated phenotype, cytokine expression and T-cell activation (Meisner et al., 2007).

Rapid microplate-based fluorescence anisotropy methods have also been developed to study other protein-RNA interactions. Vigilin (HDL-BP) is a ubiquitous protein contains 14 KH domains and it has been shown to bind the 3' UTR of vitellogenin mRNA, thus, stabilizing the mRNA (Dodson and Shapiro, 1997). Mao *et. al.* (2006) used fluorescence anisotropy to analyze the binding characteristics of both vigilin and CRD-BP. It was shown that vigilin bound to c-myc CRD with a 2-3 times lower affinity than it did to the 3' UTR of vitellogenin mRNA (Mao et al., 2006). Furthermore, it was reported that neomycin and H33342 were able to disrupt the vigilin-vitellogenin-RNA interaction (Mao et al., 2006). However, it is still unknown whether neomycin and H33342 are effective in inhibiting vigilin-mediated phenotype in cells.y

1.11 Aims of this research

The overall goal of this thesis is to develop a better understanding of the molecular interaction between CRD-BP and its target RNA, specifically on c-myc and CD44 mRNA. It is hoped that such knowledge could help in future design of new classes of anti-cancer compounds that act by breaking the protein-RNA interaction. As discussed in sections 1.6-1.9, the structure of CRD-BP orthologs has been extensively studied. Using deletion/truncation studies, it has been shown that the KH domains, and not the RRM domains, are involved in the physical interaction between CRD-BP and its target RNAs. However, it is still unclear what significant role each of

the KH domains plays in interacting with target RNA molecules. It is also unknown whether each of the KH domains plays similar role in binding to different RNA molecules. A recent study has shown that KH3 and KH4 domains of IMP1 possess a flexible Gly-rich linker which is important in binding nucleic acids (Chao, J., et al., 2010). In this study, we mutated the first Gly residue within the G-X-X-G motif in each CRD-BP KH domain as an approach to address the above questions. We used site-directed mutagenesis to generate point mutations within the G-X-X-G motifs of the KH domains (Figure 6). We mutated the first glycine of each G-X-X-G motif in the four CRD-BP KH domains to an aspartate (Figure 6). We hypothesized that introduction of an aspartate to any known binding region will create both steric hindrance and unfavorable electrostatic interactions; thus, disrupting the CRD-BP-RNA interaction. Four negative control variants were also generated by introducing mutations within regions of CRD-BP that are not expected to play roles in binding RNAs (Y5A, Q85A, E446D and D527E, Figure 6).

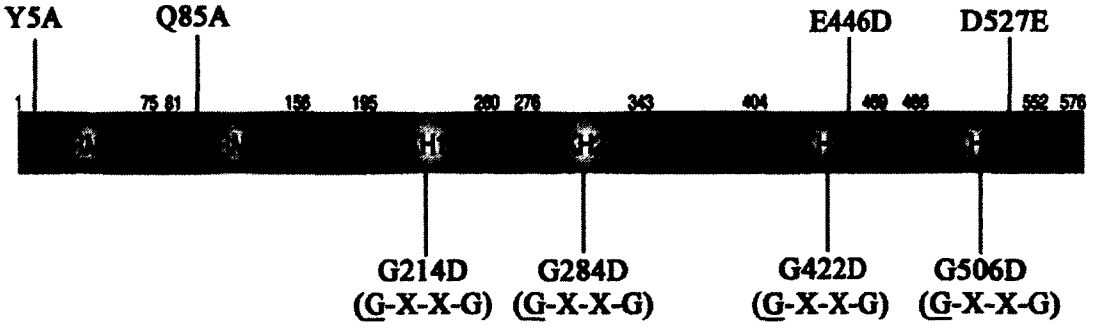

Figure 6. Locations of point mutations within CRD-BP (Christiansen et al., 2009).

The specific aims of this thesis are four-fold: (i) to develop a fluorescence-based method to study the CRD-BP-RNA interaction using electrophoretic mobility shift assays; (ii) to analyze the CRD-BP-RNA interaction *in-vitro* using radioactive electrophoretic mobility shift assays;

(iii) to analyze the CRD-BP-RNA interaction in cells; (iv) to test if specific antisense oligonucleotides and antibiotics can disrupt the CRD-BP-RNA interaction in cells.

Chapter 2

Generation and Purification of the Coding Region Determinant-Binding Proteins

The following chapter describes the methodology used to generate and purify recombinant WT-CRD-BP and the mutant-CRD-BP variants. The purified recombinant CRD-BP proteins will then be used to assess CRD-BP protein variants ability to bind target RNAs *in-vitro.* The goal is to determine the specific role of amino acid residues within known CRD-BP binding regions that are responsible for its physical interaction with target RNA substrates. This chapter will also discuss the results of the purification of CRD-BP variants, followed by a brief discussion.

2.1 Methodology

2.1.1 Site-directed mutagenesis to generate CRD-BP mutant variants

In order to generate mutant variants of CRD-BP, PCR based site-directed mutagenesis was used to incorporate desired mutations in the pet28b-plasmids to be transformed into *E.coli* BL21 cells to subsequently generate recombinant CRD-BP. It should be noted that at the same time site-directed mutagenesis was used to incorporate the same mutations into pcDNA-plasmids to be transfected into mammalian cells. Mammalian cell experiments will be discussed in Chapters 5 and 6. Typical site-directed mutagenesis reactions contained approximately 100 ng of plasmid DNA, 5xPhusion DNA polymerase buffer, 0.28 mM dNTPs, 1.8 uM primers (Table 1) and 1 uL Phusion DNA polymerase. The reactions were then placed in a thermocycler using the following program - 95°C-45 sec, 55°C-45 sec, 68°C-10 min for 18 cycles followed by incubation at 68°C for 20 min. Following the incubation steps, the reactions were transferred to a 1.5 mL eppendorf tube and the tube was incubated for 16 hours at 37°C with the restriction enzyme Dpnl to digest the methylated DNA. The following day, the DNA was precipitated

using standard ethanol precipitation [2.5 x volumes absolute ethanol, 5 uL 5 M NaCl, 1 uL glycogen (20 mg/mL)]. DNA products size were analysed on an agarose gel to ensure the presence of the full length plasmid. Mini-prep (QIAgen) was performed on precipitated DNA and then the purified DNA was then sent for DNA sequencing for verification (Macrogen, South Korea). A total of three sequencing reactions were performed for a complete evaluation of the CRD-BP sequence within the plasmid (Table 2). Once sequences were confirmed, purified DNA was transformed in to *E.coli* cells to generate the recombinant proteins.

Table 1. Primers used in the site-directed mutagenesis of CRD-BP.

Mutation	**CRD-BP structural domain**	**Primer Sequences**[a]
F42A	RRM1	Forward, 5'-TCCGGCTACGCCGCAGTGGATTGCCCC-3' Reverse, 5'-GGGGCAATCCACTGCGGCGTAGCCGGA-3'
Q85A	RRM2	Forward, 5'-AGTCGGAAAATAGCAATCCGCATTATT-3' Reverse, 5'-AATATTGCGGATTGCTATTTTCCGACT-3'
D527E	KH4 Variable Loop	Forward, 5'-GTAGTGCCAAGAGAACAGACCCCGGAT-3' Reverse, 5'-ATCCGGGGTCTGTTCTCTTGGCACTAC-3'
G284D	G-X-X-G of KH2	Forward, 5'-GGGCGACTCATTGACAGGGAAGGGCGG-3' Reverse, 5'-CCGCCCTTCCTTGTCAATGAGTCGCCC-3'
G506D	G-X-X-G of KH4	Forward, 5'-CCGCGTCATCGACAAAGGCGGCAAAAC-3' Reverse, 5'-GTTTTGCCGCCTTTGTCGATGACGCGG-3'

[a] Eight primer sets were used to generate both the single and double-base mutations in the RRMs, variable looped regions and the KH domains.

Table2. Primers used in the sequencing of the mutated *Mus musculus* CRD-BP DNA.

Primer[a]	Sequence
CRD-BP Custom Primer	GGTCAACGTCACCTACTCT

[a] The three primers used covered the 5'-region (T7 Promoter), the 3'-region (T7 Terminator) and the region between the two (CRD-BP Custom Primer).

T7 Promoter and T7 Terminator primers were provided by Macrogen. CRD-BP CR primers were custom synthesized by IDT Inc.

2.1.2 Transformation of competent *E.coli* BL21 (DE3) cells

Approximately 100 ng of purified pet28B-WT-CRD-BP plasmid was added to 100 uL of thawed BL21 cells and then placed on ice for 25 min. The transformation mixture was heat shocked at 42°C for 90 seconds, before being placed on ice for 1 min. LB-broth (~300uL) was added to the mixture and the mixture was incubated at 37°C for 30 min. Following incubation, the cells were plated onto LB-Kanamycin plates and colonies were allowed to grow overnight (~16 hours) at 37°C.

2.1.3 Cell Growth and Induction

Following the overnight incubation at 37°C, 20 bacterial colonies were picked from the plate and placed in 100 mL of LB-Kanamycin broth in a 250 mL flask and were placed on a shaker to grow (200 rpm, 37°C) for ~3 hours. The 100 mL of bacteria were placed in 900 mL of LB-Kanamycin broth in a 3 L flask. The flask was placed back in the shaker until an OD_{600}= 0.5 was reached (~2-3 hours). Once the cells reached OD_{600}= 0.5, 1 mL 1 M IPTG was added to induce recombinant protein production. Following a 6-hour growth period in the shaker, the bacterial cultures were spun down (3000 x g, 4°C, 15 min) and the bacterial pellet was frozen overnight at -80°C.

2.1.4 CRD-BP Purification under denaturing conditions

Protein purification began with the thawing of the frozen bacterial pellet on ice for 15 min. Once thawed, the pellet was resuspended in 12 mL of Buffer B (100 mM NaH_2PO_4, 10 mM Tris-Cl, 8 M Urea, adjusted to pH8) and placed in a 50mL falcon tube. The solution was then placed gently agitated at 4°C until the solution was semi-translucent (~60 min). While waiting, the remaining buffers were pH adjusted using the stock Buffer B (C-pH 6.5, D-pH 6.3, E-pH 5.9). The bacterial solution was then spun down (13,200 rpm, 4°C, 30 min) and the supernatant was retained. The supernatant was then loaded on to a Nickel-NTA gravity drip column (column was equilibrated with Buffer B prior to loading supernatant). Each buffer was passed over the supernatant loaded on the column (Buffers B-E, with the desired protein eluting in Buffer E). Fractions from each buffer condition were collected to check protein purity. Samples of the fractions (~16 uL) were mixed with 4 uL 5 x Sample buffer and 1 uL of β-mercaptoethanol before being boiled for 10 min. Samples were then loaded on to a 12% SDS-PAGE gel; 19:1 acrylamide:N,M-methylenebisacrylamide and resolved at 120 V for ~35 min. The protein fractions used for subsequent experiments were from elution fraction E.

2.1.5 Protein Dialysis

All purified proteins were dialysed using a two-step gradient to remove the urea from the Buffer E elution buffer and to promote renaturation of the protein. Approximately 100 uL of the purified protein was placed into a Slide-A-Lyzer mini dialysis unit (Pierce, Rockland IL, USA) and the unit was placed in a 400 mL beaker containing 50mL of dialysis buffer A (1 M Tris-Cl, 0.2 M glutathione-reduced, 0.1 M glutathione-oxidized, 1.33 M glycerol, 2 M urea, 0.0001% Triton-X, adjusted to pH 7) and buffer exchanged was allowed for 24 hours. After 24 hours, the dialysis units were transferred to 1 L beaker containing 800mL of dialysis buffer B (1 M Tris-Cl,

1.33 M glycerol, 0.0001% Triton-X, adjusted to pH 7). Following protein dialysis, the concentration of the protein was determined using Bradford protein quantification assay (BioRad). The dialyzed proteins were now ready to be used in the protein binding experiments.

2.2 Results

Following successful site-directed mutagenesis (Table 3) and transformation of DNA, the kanamycin selective plates produced viable *E.coli* BL21 colonies. The single bacterial colonies were transferred from the plate to 3 L flask for the large scale production of CRD-BP variants.

Table 3. Site-directed mutagenesis sequence information.

Mutation	Original sequence	Mutated sequence
F42A	5'-GGCTACGCC**TTC**GTGGATTGC-3'	5'-GGCTACGCC**TTC**GTGGATTGC-3'
Q85A	5'-CGGAAAATA**CAG**ATCCGCAAT-3'	5'-CGGAAAATA**GCA**ATCCGCAAT-3'
D527E	5'-GTGCCAAGA**GAC**CAGACCCCG-3'	5'-GTGCCAAGA**GAA**CAGACCCCG-3'
KH2	5'-CGACTCATT**GGC**AGGGAAGGG-3'	5'-CGACTCATT**GAC**AGGGAAGGG-3'
KH4	5'-CGTCATC**GGC**AAAGGCGGCAA-3'	5'-CGTCATC**GAC**AAAGGCGGCAA-3'

Following PCR-based site-directed mutagenesis, desired mutated sequences were confirmed (bold and underlined). Generation of mutations in RRM1 (F42A) and RRM2 (N124A) were unsuccessful and therefore were not used in either *in-vitro* or *in-vivo* experiments. KH double mutants (KH1-2, KH1-3, KH2-4, KH3-4) were generated by a two-step site-directed mutagenesis of the single mutants. Single-base mutations were first generated in the desired KH domains. Once those sequences were confirmed, site-directed mutagenesis was again used to incorporate the second base mutation within the mutated construct. All double mutant sequences were verified by sequencing (data not shown, Gerrit van Rensburg was responsible for generating the KH1, KH2, KH3, KH4, KH1-2,KH1-3, KH2-4 and KH3-4 mutants).

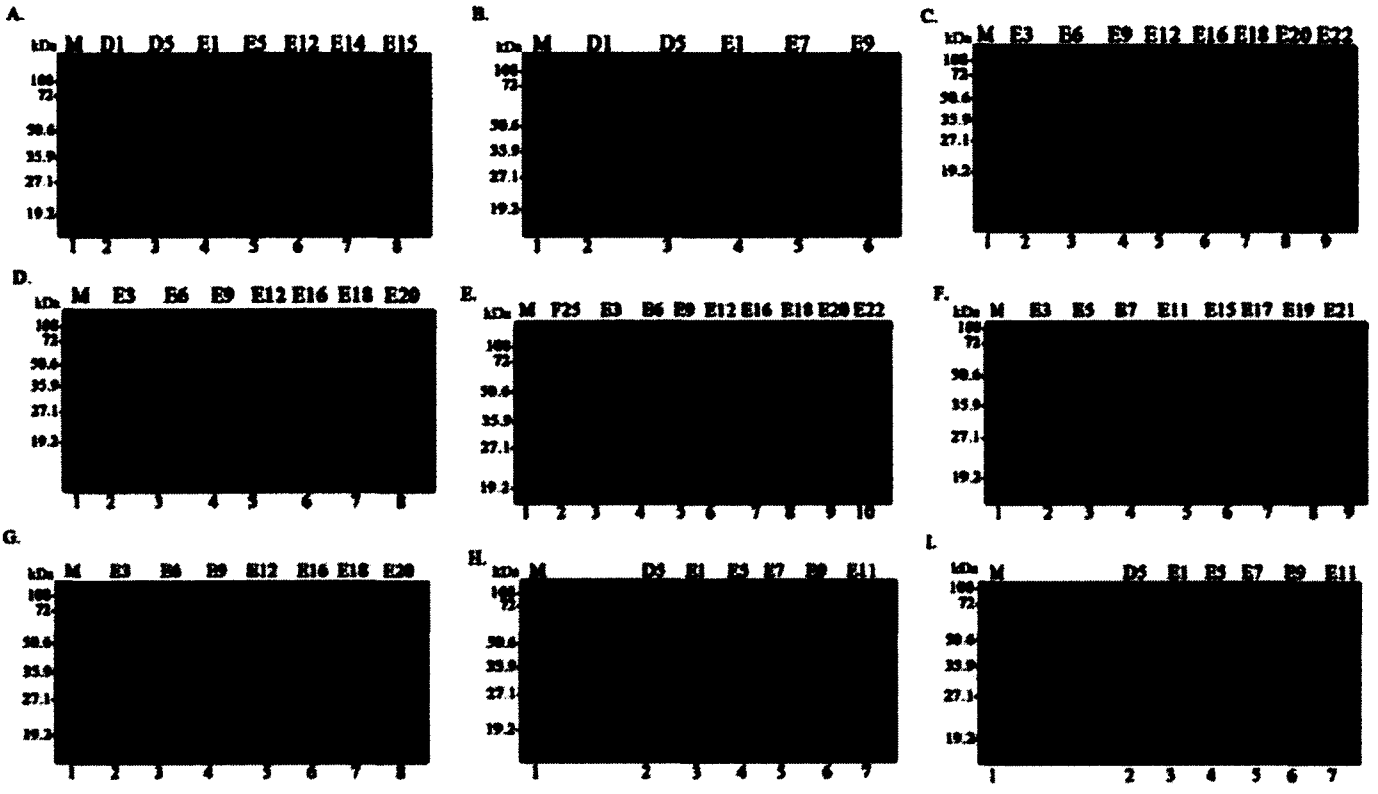

Figure 7. SDS-PAGE analysis of recombinant 6xHis-tagged CRD-BP purified from *E.coli.* BL21 cells. To ensure CRD-BP generation and purification was successful, various D-elution and E-elution fractions (16 uL of each fraction) were resolved on 12% SDS-PAGE gels to assess fraction purity. Gels were visualized using Coomassie blue stain. (**A**) CRD-BP-Wild-type, (**B**) CRD-BP-E446D, (**C**) CRD-BP-Y5A, (**D**) CRD-BP-Q85A (**E**) CRD-BP-D527E (**F**) CRD-BP-KH1-2 (**G**) CRD-BP-KH1-3 (**H**) CRD-BP-KH2-4 (**I**) CRD-BP-KH3-4.

The large scale purification of the CRD-BP variants using Ni-NTA purification columns is illustrated in Figure7, A-I. The induction of CRD-BP via IPTG was successful based on the presence of the band at ~68 kDa (Figure 7, A-I), which agrees with the literature molecular weight of CRD-BP (Doyle et al., 1998). Contaminating proteins within each of the CRD-BP variant cell lysates were washed off the column in the A-D wash fractions. However, due to

experimental conditions, each pure CRD-BP variant eluted off the Ni-NTA column at different E-elution fractions. WT-CRD-BP began to elute of the column at approximately elution fraction E5, with the pure product eluting at approximately elution fraction E12 (Figure 7, A). CRD-BP-E446D mutant began eluting at elution fraction E1, with the pure product eluting at approximately elution fraction E7 (Figure 7, B). CRD-BP-Y5A mutant began to elute of the column at approximately elution fraction E3, with the pure product present in the elution fractions around E9 (Figure 7, C). CRD-BP-Q85A mutant protein began to elute at E3 elution fractions, however the pure protein was not eluted until approximately elution fraction E16 (Figure 7, D). CRD-BP-D527E mutant protein showed heavy contaminants in elution fractions around E3 and the pure protein eluted at approximately elution fraction E18 (Figure 7, E). CRD-BP-KH1-2 double mutant began to elute off the column in elution fractions around E3 with the pure product eluting in elution fraction E11 (Figure 7, F). CRD-BP-KH1-3 double mutant began eluting off the column at elution fraction E3 with the pure product eluting at elution fraction E16 (Figure 7, G). CRD-BP-KH2-4 double mutant began to elute off the column at elution fraction E1 with the pure product eluting at elution fraction E5 (Figure 7, H). Finally, CRD-BP-KH3-4 double mutant began to elute off the column at elution fraction E1 with the pure product appearing at elution fraction E5.

Dialysis fraction candidates were chosen based on their purity and correct band size. Typically, purified proteins in elution fractions E5-E16 were used for dialysis due to their high purity. Once the chosen fractions were dialyzed, they were analysed again on a denaturing polyacrylamide gel to ensure all fractions to be used in subsequent binding reactions were of relative equal purity (Figure 8).

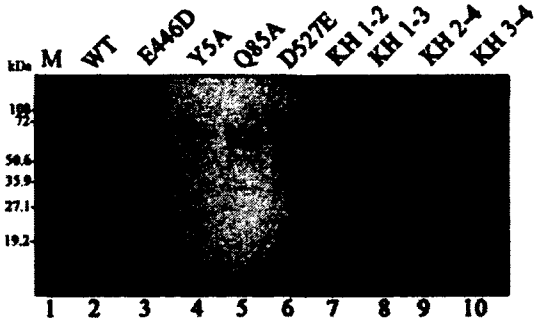

Figure 8. Post-dialysis SDS-PAGE analysis of purified 6xHis-tagged CRD-BP elution fractions chosen for binding experiments. Based on purity from Figure above, E-elution fractions (~2 ug) were analyzed to ensure equal quantities of protein were being used. Gel was stained using Coomassie blue.

The protein fractions were then quantified using BSA protein assay (Figure 9). Once dialyzed and quantified, the proteins were ready for use in the protein binding experiments (Chapter 3 & 4).

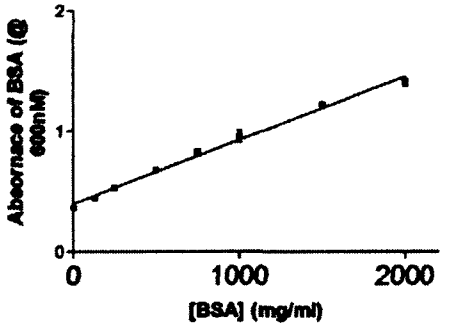

Figure 9. Standard curve for BSA protein assay. Post dialysis quantification of CRD-BP variants was done using the standard curve method and linear regression ($Y=0.00049x + 0.3690$, $R^2= 0.984$). BSA protein was dissolved in dialysis Buffer E.

2.3 Discussion

Previous studies have shown that truncated versions of CRD-BP are able to bind its RNA transcripts (Farina et al., 2003; Chao et al., 2010). The current study investigated the effects of

generating specific mutated versions of CRD-BP using site-directed mutagenesis. Specific amino acid residues within known binding regions were successfully mutated, as confirmed by sequencing. The purpose of generating the mutations in the known binding regions was to determine the role played by specific amino acid residues in CRD-BP in binding to its target RNAs. Furthermore, mutations (Y5A, Q85A, E446D and D527E) were incorporated in to regions of CRD-BP that were not expected to affect binding. Such mutants are meant to serve as negative controls. The mutational effects on CRD-BP-RNA binding will be discussed in subsequent chapters.

Due to the fact that 6xHis-tagged CRD-BP was purified under denaturing conditions in 8 M urea, a dialysis method was chosen to dialyze out the urea and refold the protein to its functional native form to minimize protein aggregation during the renaturation process. Purification under denaturing conditions can cause a higher denatured:native polar amino acid residues exposed to solvent causing both decreased solubility and increased intermolecular interactions between proteins, both which increase protein aggregation (Wetlaufer and Xie, 1995). The purified proteins were subsequently going to be used in protein binding experiments to assess CRD-BP's binding ability and because protein aggregation can be problematic for protein binding experiments, ensuring minimal aggregation would ensure optimal binding experiment results. The two-step dialysis method chosen did promote slow renaturation of the CRD-BP proteins. The slow removal of the denaturing urea from the purification buffer provided minimal aggregation of mis-folded proteins. This was confirmed by the analysis of CRD-BP using native polyacrylamide gels.

Chapter 3

Development of a fluorescent based method to study the CRD-BP-RNA interaction *in-vitro*

The following chapter describes the experimental techniques used in the successful generation of fluorescently labeled CD44 and c-myc RNA substrates to be used in protein-RNA binding experiments. The reason for the development of a fluorescent-based method to study the CRD-BP-RNA interaction was two-fold: (i) the use of fluorescently labeled substrates is safer and more economical than radioisotope [^{32}P] labeled RNA substrates, and (ii) developing the fluorescent-based method to study the interaction is a precursor experiment to the future goal of using fluorescent polarization method to study the protein-RNA interaction. Moreover, this chapter will cover data analysis of the fluorescent based protein binding experiments. A brief discussion of the results will conclude this chapter.

3.1 Methodology

3.1.1 PCR Generation of linear templates for in-vitro transcription

The generation of linearized 167 nt c-myc DNA template to be used for *in-vitro* transcription (IVT) was done using PCR. Reaction tubes were set up based on Table 4. PCR reactions were performed in a thermocycler using the following program – 94°C- 30 sec, 50°C-30 sec,72°C - 45 sec for 30 cycles. PCR products were subsequently checked on a 1% agarose gel. The desired PCR product was then purified from the agarose gel (Qiagen QIAEX®II gel extraction kit). Following gel purification, the PCR reactions were subjected to standard ethanol precipitation (2.5 x volume ethanol, 1/10 volume NaAc (3 M, pH 5.2), 10x glycogen). DNA concentrations were determined using NanoDrop.

Table 4. Reagents used in the generation of linearized DNA template for IVT

Reagent	Amount
10 x PCR buffer	3.5 uL
Reverse primer (T7myc 1871) 100ng/uL	1 uL
Taq polymerase (NEB)	0.5 uL
Total	35 uL

3.1.2 Generation of unlabeled RNA acceptor molecule via *in-vitro* transcription

Unlabeled RNA was generated using T7 polymerase-directed IVT from DNA templates (above). Each 100 uL reaction was set up based on Table 5. The IVT reactions were incubated at 37°C overnight. Following incubation at 37°C, reactions were treated with DNase to ensure a complete degradation of the DNA template. Reactions were subsequently phenol:chloroform (1:1 volumes) extracted to remove residue reaction proteins and further purify the RNA substrate. The RNA was then passed over G50 columns to remove residual free nucleotides. Again, standard ethanol precipitation was performed to concentrate the RNA substrate before it was resuspended in DEPC H_2O and quantified using NanoDrop. PCR products were analysed using a 2.5% agarose gel to ensure the correct size of product was produced and to confirm effective DNase treatment.

Table 5. Reagents used in the generation of unlabeled RNA substrates

Reagent	Amount
10 x T7 buffer (w/ Triton X)	10 uL
100 mM CTP	5 uL
100 mM UTP	5 uL
T7 polymerase	5 uL
Total	100 uL

3.1.3 Ligation of fluorescently labeled oligonucleotide to IVT acceptor

To generate a fluorescently labeled RNA substrate for subsequent binding experiments, 2-part ligation reactions were used. In the first part of the ligation reactions, the *in*-vitro transcribed 167 nt acceptor RNA oligonucleotide was aligned with the 15nt IDT fluorescein-labeled donor molecule using a splint molecule (Table 6). Annealing reactions were carried out in the thermocycler using the following program -65°C- 3 min, 25°C- 5 min. Following incubation in the thermocycler, one unit of T4 RNA ligase (20 U/uL) was added to the reactions and they were incubated at 37°C until the reaction was complete. Various ratios of Acceptor:Splint:Donor molecules were used to optimize ligation reactions. Products were check on 12% denaturing polyacrylamide gels with and without staining with SYBR® Safe DNA gel stain. The experiment described above was also done using a both a DNA and RNA splint to optimize ligations.

Table 6. RNA fragments and reagents used in the RNA ligation reactions

Nucleotide Fragment	Volume (uL)
IDT 15nt c-myc (nts 1872-1886-Fam) [6.6 uM]	6
IDT DNA splint 5'-CTTGTCGATGCCTTGAGA-3' [6.6 uM]	8

3.1.4 Generation of 3'-Fluorescein-5-Thiosemicarbazide labeled RNA oligonucleotides

To improve labeling efficiency of the c-myc, 3'-end labeling using Fluorescein-5-Thiosemicarbazide was done. Reactions were carried out in a two-step process. First, a 181 nt c-myc RNA fragment was synthesized using IVT as described above. The vicinal hydroxyls present at the 3'-end of the RNA oligonucleotide were oxidized to aldehydes using the following reaction solution -H_2O, 0.5 M NaOAc [pH 5.1], 10 uM RNA oligo, 500 uM $NaIO_4$-incubate at

room temperature for 90 min. Following the incubation, a standard ethanol precipitation was performed. Secondly, the now oxidized RNA oligonucleotide was incubated with fluorescein-5-thiosemicarbazide at 4°C overnight (0.5 M NaOAc [pH 5.1], 200 mM fluorescein-5-thiosemicarbazide, H_2O). Following the overnight incubation, again ethanol precipitation was done. The re-suspended RNA oligonucleotide was column purified using G-25 column. Concentration and labeling efficiency was determined using NanoDrop.

3.1.5 Internally labeled RNA oligonucleotide using Fluorescein-UTP

Again to improve labeling of the RNA oligonucleotide substrate, fluorescein-labeled UTP was used to internally label the RNA substrate via MAXIscript® kit (Invitrogen) (Nuclease free H_2O, 10 x buffer, 10 mM NTPs, Fl-UTP, 1 U T7 polymerase-37°C for 60 min). Following the incubation, the RNA was ethanol precipitated and column purified as described above. Standard phenol:chloroform extraction was also performed to remove any residual proteins from the IVT reaction.

3.1.6 Fluorescent electrophoretic mobility shift assays

To assess the ability of WT-CRD-BP to bind the fluorescently labeled substrate, electrophoretic mobility shift assays (EMSA) were used. EMSA binding buffer (Table 7) was made fresh before each experiment and was placed on ice until it was needed. Fluorescently labeled RNA substrates were subjected to a denaturation and renaturation 2-step process that involved heating the RNA to 55°C for 5 min and then cooling at room temperature for 7 min. EMSA reaction tubes (binding buffer, RNA, protein) were then incubated at 37°C for 10 min, before being placed on ice for 5 min. This heating cooling cycle to facilitate binding was repeated, for a total of two times. EMSA loading dye (3 uL) was then added to the reaction and

17 uL of the reaction was loaded on to an 8% native PAGE gel. Complexes were resolved at 20 mA for 90 min. The gels were visualized using a Kodak Imager.

Table 7. Reagents used for radioactive EMSA experiments

Reagent Identity	Reagent Composition
EMSA loading dye	250 mM Tris-Cl (pH 7.4), 0.2% bromophenol blue, 0.2% xylene cyanol, 40% sucrose

3.2 Results

Prior to generating the *in-vitro* transcribed RNA acceptor molecule for ligation reactions, linearized DNA template was generated and gel purified (Figure 10).

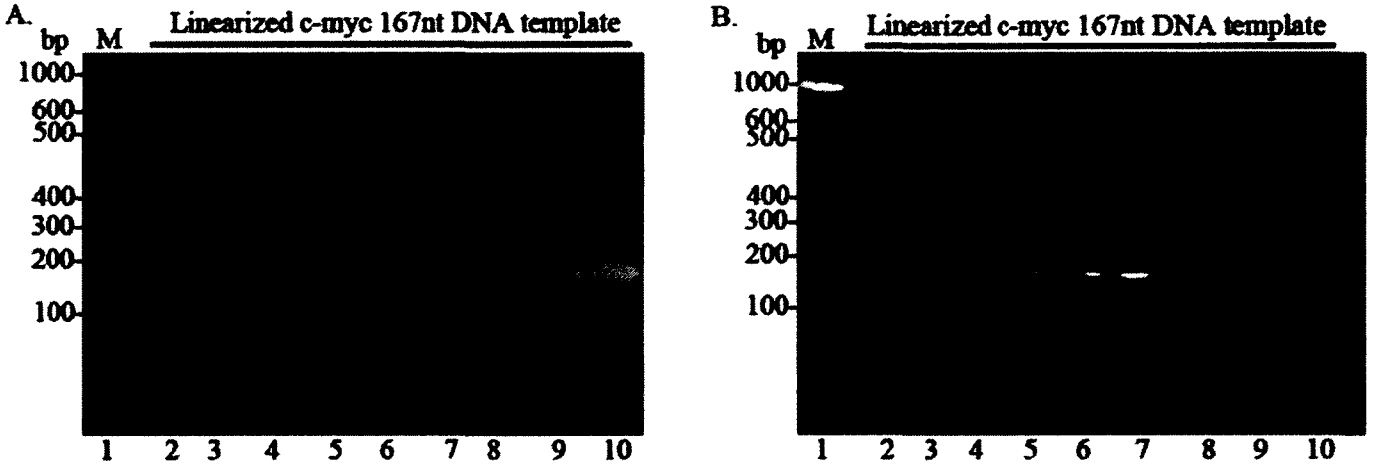

Figure 10. Agarose gel analysis of linearized c-myc 167 nt DNA template to be used in *in-vitro* transcription. PCR reactions were loaded on to a 2% agarose gel. Successful generation of 167 nt DNA template evident by the intense band at approximately 200 bp (A.). Following visualization, appropriate bands were excised from gel (B.) and the RNA substrate was gel purified.

Following gel purification of the DNA template, *in-vitro* transcription of the 167nt RNA acceptor molecule was generated and DNase treated to remove residual DNA template (Figure 11).

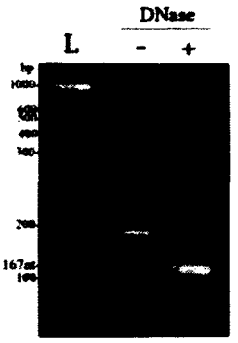

Figure 11. Agarose gel analysis of *in-vitro* transcribed 167nt RNA fragment. Samples (2 uL of IVT reaction, 6 uL H_2O, 2 uL gel loading dye) from pre- and post-DNase treatments were loaded on a 2.5% agarose gel. The dissapearance of the band around 200 bp in the (+) DNase lane shows the treatment was successful and the DNA template was degraded.

The *in-vitro* transcribed 167nt c-myc RNA substrate was incubated with various ratios of Acceptor:Splint:Donor (1:0.5:1, 1:1:1, 1.5:1:1, 1:1:1.5, 2.5:1:1.6, 1:1.5:2) molecules to ensure optimal ligation reaction yield. However, ligation reactions yielded very low full length labeled product which were mainly only visible following staining with SYBR® Safe DNA gel stain (Figure 12) suggesting a low efficiency RNA ligation reaction was occurring and no useable substrate was generated. I therefore abandoned this method for labeling larger RNA fragment.

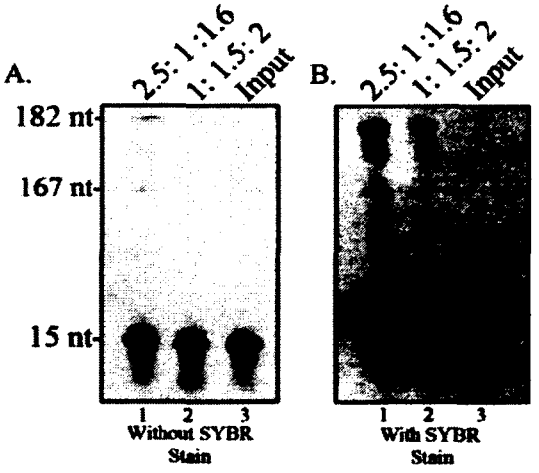

Figure 12. RNA ligation reaction analysis on 12% denaturing poly-acrylamide gel. RNA ligation reactions with various ratios of Acceptor:Splint:Donor, as shown, were loaded on the gel. The Input lane (3) contained only the 15nt donor molecule. They were first visualized at λex 280 nm, λem 530 nm in the absence of SYBR Safe stain (A) and then again visualized at λex 280 nm, λem 530 nm following a 30 min incubation in the SYBR Safe stain (B).

After multiple failed attempts to improve the efficiency and yield of the RNA ligation reactions, I then proceeded to 3'-end labeling of RNA method using fluorescein-5-thiosemicarbazide. A 3' fluorescent end-labeling of RNA oligonucleotides protocol from the Ryder Lab at the University of Massachusetts Medical School was adjusted to fit the needs of the current study. Using fluorescein-5-thiosemicarbazide and *in-vitro* transcribed RNA substrates (181 nts c-myc and 39 nts CD44 RNAs), labeling reactions were carried out and visualized on denaturing PAGE gels. The efficiency of the labeling reactions was low based on band intensity (Figure 13) and calculation of % efficiency (Table 8). Again, the desired 181 nt c-myc products only appeared following incubation in SYBR® Safe DNA gel stain, indicating poor signal from the fluorescein labeling. CD44 39 nt RNA labeling was more efficient.

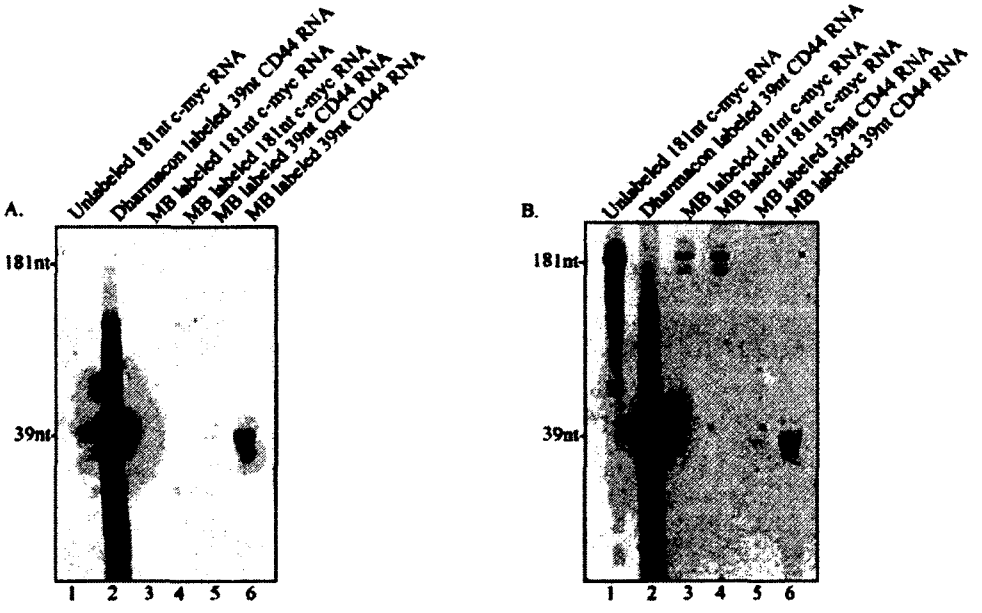

Figure 13. Analysis of 3'-fluorescein-5-thiosemicarbazide end-labeling of *in-vitro* transcribed 182nt c-myc and 39nt CD44 RNA substrates. Reaction samples were loaded on a 12% polyacrylamide gel. Reactions were first visualized at λex 280 nm, λem 530 nm in the absence of SYBR Safe stain (A) and then again visualized at λex 280 nm, λem 530 nm following a 30min incubation in the SYBR Safe stain (B). Unlabeled 181nt c-myc RNA (10.5 uM) was loaded to act as a positive control marker for the c-myc labeling reactions (Lane 1).

Dharmacon labeled 39nt CD44 RNA (48uM) was loaded to act as a positive control marker for the CD44 RNA labeling reactions. Various concentrations of fluorescein-5-thiosemicarbazide were used in the reactions and subsequent samples of the reactions were loaded on the gel (200 uM dye used in reactions in Lanes 3&5, 200 mM dye used in reactions in lanes 4&6).

Table 8. RNA labeling efficiency data.

RNA Substrate	A_{260}	A_{490}	RNA/ F-5-T (%)
182nt c-myc (200 mM)	0.280	0.012	5%
39nt CD44 (200 mM)	0.082	0.020	25%

Maximal absorbencies of RNA (260nM) and fluorescein (490nm) were determined and using the Beer-Lambert law (A=ɛcl) with an extinction coefficient of 70,000 $M^{-1}cm^{-1}$ the ratio of RNA to fluorescein-5-thiosemicarbazide was determined. This ratio, expressed as a percentage, gives the efficiency of the labeling reactions (efficiencies of 70%-95% are considered to be sufficient).

Due to multiple failed attempts at increasing 3'-fluorescein-5-thiosemicarbazide labeling efficiency, I then proceeded to using internal fluorescent labelling of RNA. Using MAXIscript® kit (Invitrogen) with Fluorescein-UTP (Roche). Both the 181 nt c-myc RNA and 39 nt CD44 RNA substrates were internally labeled with high yield and efficiency (Figure 14). The IVT products were intact with little or no degradation.

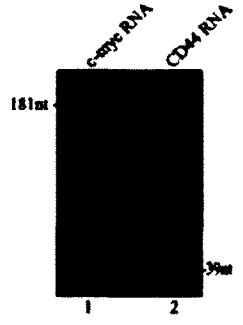

Figure 14. Analysis of internally labeled RNA substrates using Fl-UTP. *In-vitro* transcription using Fluorescein-labeled-UTP was done to generate both c-myc and CD44 RNA substrates. Samples (100nM of RNA) of the IVT were loaded on to a 12% denaturing polyacrylamide gel and visualized at λex 280 nm, λem 530 nm. The products were not visualized with SYBR safe gel stain due to the efficiency of the internal labeling.

Following successful internal labeling of the RNA substrates, protein binding experiments was then performed. Fluorescent electrophoretic mobility shift assays were conducted in order to develop a fluorescent based method to analyze the CRD-BP-RNA interaction *in-vitro*. It was shown that purified WT-CRD-BP (Chapter 2) has the ability to bind the 181 nt c-myc RNA substrate indicated by the bound complex formation (Figure 15).

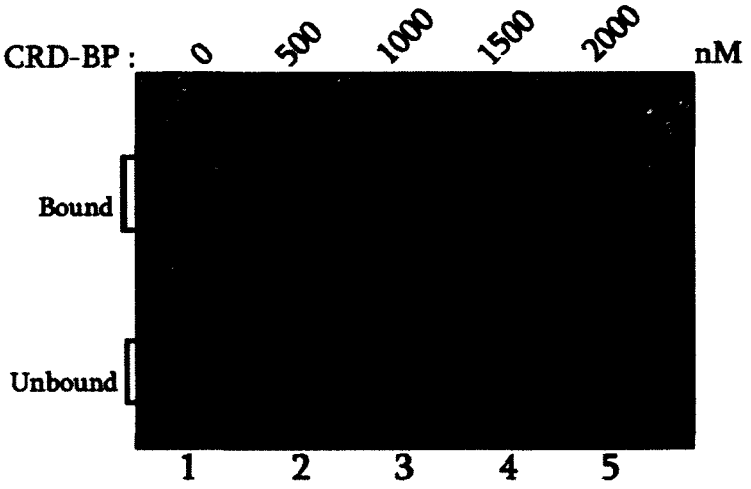

Figure 15. Electrophoretic Mobility Shift Assay of WT-CRD-BP and c-myc CRD (1705-1886) RNA substrate. Various concentrations of WT-CRD-BP (as indicated) was incubated with 100 nM of fluorescein internally labeled 181 nt c-myc RNA substrate. Binding reactions (17 uL) were loaded on the 8% native polyacrylamide gel and they were resolved at 20 mA for 90mins. The ratio of Fl-UTP:UTP was ~1:1 in the labeling reactions. The RNA substrate was body labeled using fluorescein-UTP.

It was noted that some of the complex remained stuck in the well; therefore the signal from the complex within the gel was weak. This fact greatly affects the analysis of CRD-BP's ability to bind its RNAs. Due to the weaker than expected CRD-BP-RNA complex signal, the ratios of Fl-UTP: UTP in the labeling reactions were adjusted to compensate for complex retardation in the wells of the gel. By reducing the number of fluorescein-labeled UTP nucleotides within the RNA substrate, and reducing the molecular weight of the complex, more of the complex was thought to enter the gel. The binding profile CRD-BP with the adjusted Fl-

UTP:UTP ratio within the c-myc RNA substrate showed a slight improvement of the signal from the CRD-BP-RNA complex within the gel (Figure 16). However, the complex within the well was still present, especially at the higher concentrations of CRD-BP within the binding reactions, indicating possible protein aggregation issues.

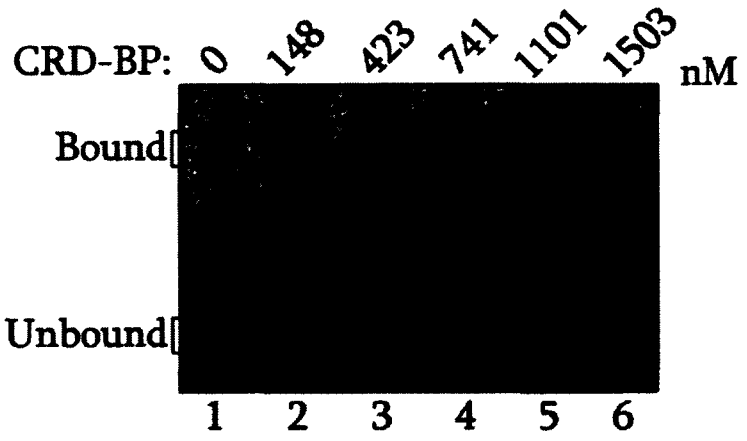

Figure 16. Electrophoretic mobility shift assay of WT-CRD-BP and c-myc CRD (1705-1886) RNA substrate. Various concentrations of WT-CRD-BP (nM indicated) were incubated with 72 nM fluorescein internally labeled 182nt c-myc RNA substrate. The ratio of Fl-UTP:UTP was ~1:9 in the labeling reactions. The reduced RNA concentration used in the binding reaction and the reduced number of labeled UTP nucleotides in the *in-vitro* transcription of the RNA substrate were done to promote CRD-BP-RNA complex entrance in to the 8% native polyacrylamide gel. Complexes were resolved at 20 mA for 90 min. The RNA substrate was body labeled using fluorescein-UTP.

In the fluorescent EMSA reactions to analyze the CRD-BP-RNA interaction, 39nt CD44 Fl-UTP internally labeled substrate was also used in the binding experiments. Although the 39 nt CD44 RNA substrate was 143 nt smaller than the 181 nt c-myc RNA substrate, WT-CRD-BP still showed the ability to bind the substrate at a concentration range of 20-530 nM. The binding profile of CRD-BP with the 39 nt CD44 substrate showed a reduced signal from the protein-RNA complex within the gel at the higher protein concentrations (740-950 nM) (Figure 17), indicating a common problem experienced in both binding reactions (i.e. protein aggregation problem).

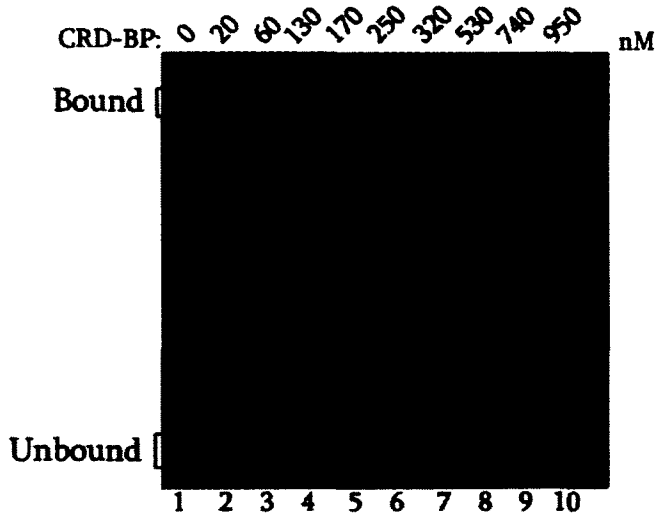

Figure 17. Electrophoretic Mobility Shift Assay of WT-CRD-BP and CD44 3'UTR RNA substrate. Various concentrations of WT-CRD-BP (nM indicated) were incubated with 50 nM fluorescein internally labeled 39nt CD44RNA substrate. CRD-BP-RNA complexes were resolved at 20 mA for 90 min on an 8% native polyacrylamide gel. The RNA substrate was 3'-fluorescein-5-thiosemicarbazide end-labeled.

3.3 Discussion

In the development of a fluorescent-based method to study the CRD-BP-RNA interaction there were many experimental problems that arose both with the generation of a fluorescently-labeled RNA substrate and also with the resolution of CRD-BP-RNA complexes in the electrophoretic mobility shift assays. Firstly, generating the fluorescein-labeled 181 nt c-myc and 39 nt CD44 RNA substrates with high yield and efficiency proved to be challenging. RNA ligation reactions were the first method used. This involves ligating an unlabeled 167nt *in*-vitro transcribed c-myc acceptor RNA oligonucleotide with a fluorescein-labeled 15nt c-myc donor RNA oligonucleotide. Splints were used to hold the acceptor and donor molecules in a position that favors the RNA ligase catalyzed ligation of the fragments (Figure 18).

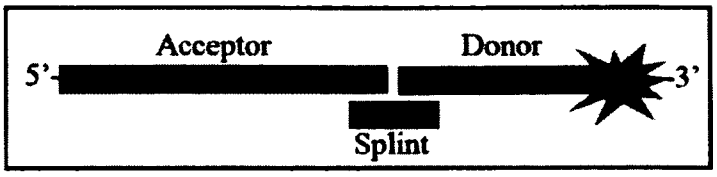

Figure 18. Schematic diagram of the RNA ligation reaction to generate the fluorescently-labeled RNA substrate. DNA and RNA splints were used to orient the acceptor and donor molecules for RNA ligase to ligate the fragments together.

To optimize the yield and efficiency of the RNA ligation reactions, one can adjust the ratios of acceptor:splint:donor molecules (Stark et al., 2006). Although these ratios were adjusted in the attempt to improve the yield of the labeling reactions, optimal yields of full length fluorescently-labeled substrates was not achieved. The size of the acceptor molecule may have been the culprit for the reduced yield. Optimal oligonucleotides are 30-50 nucleotides in length (Stark et al., 2006). The current study used a 67 nt acceptor molecule. Do to the low yield of full length labeled ligation product, another labeling technique was then used.

Exploiting the 3'-hydroxyl group present within an RNA molecule for research purposes is not a new method used to generate labeled RNAs (Reines and Cantor, 1974). Addition of fluorescent molecules to the 3' end of RNA molecules allows for the development of a non-radioactive probe that can be used in quantitative analyses of protein-RNA interactions (Pagano et al., 2011). This two-step method first involves the oxidation of the vicinal 3'hydroxyl group on the RNA molecule using an oxidizing agent such as sodium periodate ($NaIO_4$). The second step involves the formation of a covalent linkage between the fluorescein-5-thiosemicarbazide (FTSC) and the RNA (Pagano et al., 2011). This method was employed in the current study to generate the fluorescently labeled substrate. However, yields that resulted were deemed too low to further proceed to protein binding experiments. The reason for the low yield may be due to the incomplete oxidation of the vicinal hydroxyl groups within the RNAs or an unfavorable FTSC:RNA oligonucleotide ratio (Pagano et al., 2011). $NaIO_4$ incubation times were adjusted to

ensure complete oxidation of the vicinal hydroxyl groups, but to no avail. Both molar and end equivalents of RNA molecules to FTSC molecules were adjusted to improve labeling. But unfortunately, the desired yield was not achieved. Therefore, a third labeling method to generate the labeled RNA substrate was attempted.

There are multiple labeling methods that generate a suitable substrate, including both end-labeling and body labeling (Pagano et al., 2011). Due to the fact that RNA ligation and end-labeling methods were shown to result in poor yields of labeled RNA substrate, body labeling using *in-vitro* transcription with fluorescein-labeled Uridine triphosphate (UTP) was used. There are reasons this labeling method was not the initial method selected for the generation of the fluorescein-labeled RNA substrate. Fluorescein has a molecular weight of 332.31 g mol^{-1} which is very close to the average molecular weight of a nucleotide base 330.0 g mol^{-1} (Sun et al., 1997). Also, the 181 nt contains an 18.1% UTP composition, which means that a possible 33 fluorescein-UTP nucleotides could have been incorporated in to the substrate. The concern was that this possible 11 kDa increase in the molecular mass may not only interfere with the physical interaction with CRD-BP, but also affect the CRD-BP-RNA complex from entering the native polyacrylamide gel in the binding experiments. This was not a concern in the previous two methods because the RNA substrates were being end-labeled with a single fluorescein molecule, as opposed to body labeling with multiple fluorescein-UTP nucleotides. For this reason, ratios of labeled:unlabeled UTP were adjusted until there was a sufficient signal from the RNA substrate and the CRD-BP-RNA complex was not hindered in the electrophoresis. Optimization of this method resulted in a successful generation of fluorescein labeled 181 nt c-myc RNA substrate which was used in the fluorescent EMSA. However, although the RNA substrate was generated,

there were still challenges faced once the RNA-binding experiments were performed (i.e. protein aggregation).

The second major issue in the development of the fluorescent-based method to analyze the CRD-BP-RNA interaction was the EMSA. Low signals were obtained and immobilization of the CRD-BP-RNA complex within the wells of the native polyacrylamide gels was observed. After the addition of the multiple fluorescein-UTP nucleotides was ruled out as the culprit (based on free RNA entering the gel with no issue), it was determined that the problem was with the protein. Although the dialysis method described in Chapter 2 states that a functional folded CRD-BP would result, for the fluorescent EMSA experiments above, a multiple pH step gradient dialysis method was used to refold the protein following purification. This resulted in protein aggregation within the sample used in the EMSA; therefore, protein-RNA complexes were partially hindered. Prior to the conclusion that the dialysis method was not optimal and was the cause of protein aggregation, the switch to radioactive EMSA using [^{32}P]-labeled RNA substrates was made. Further explanation of the events that lead to the conclusion of the dialysis method being the issue with the EMSAs will be discussed in Chapter 4.

Chapter 4

Assessing the ability of CRD-BP mutant variants to bind c-myc CRD RNA *in-vitro*

Previous studies have used deletion studies to determine the importance of RNA-binding domains of CRD-BP orthologs. It has been shown that the KH domains, not the RRM domains, are responsible for binding RNA substrates (Chao, J.A., et aL., 2010; Farina, K.L, et al., 2003; Git, A., and N. Standart, 2002; Oberman, F., et al., 2002). To date, there has been no study using site-directed mutagenesis to determine important amino acid residues within these binding domains. Specifically, the relative significance of the G-X-X-G motif in each of the four KH domains of CRD-BP has not been investigated.

The following chapter describes the experimental techniques used to analyze CRD-BP's ability to bind c-myc CRD RNA (nts 1705-1886) *in-vitro*. Furthermore, this chapter will describe the experimental approach used to determine the specific role of the first glycine within the G-X-X-G motif located in the four K-homology (KH) domains of CRD-BP. The generation and purification of both wild-type and mutant CRD-BP proteins were described in Chapter 2. The generation and purification of [^{32}P]-UTP c-myc CRD RNA substrate is described below. A presentation of the results and observations will preclude a discussion and interpretation of the results.

4.1 Methodology

4.1.1 Generation of internally ^{32}P-labeled c-myc RNA substrate

Radio-isotope [^{32}P]-labeled c-myc CRD RNA substrate was generated via *in-vitro* transcription. Linearized DNA template was prepared as above. Approximately 100 ng of linearized DNA template (c-myc CRD RNA nts 1705-1886) was added to the IVT reaction and using the MAXIscript® kit (Invitrogen) (Nuclease free H_2O, 10 x buffer, 10 mM NTPs, [^{32}P]-

UTP, 1 U T7 polymerase) the reactions were incubate at 37°C for 60 min. Following the incubation, 1uL Turbo DNase (5 U/uL) was added and the reaction was incubated at 37°C for 10 min. The reaction mixture was loaded onto an 8% denaturing PAGE gel and gel extraction of the desired product was performed. Gel homogenate was loaded on a DTR column to remove gel fragments and free nucleotides. Standard phenol:chloroform extraction was then used to remove any possible residual proteins. Ethanol precipitation was then carried out to concentrate the RNA substrate. The radioactivity of the substrate was checked using a scintillation counter. The RNA substrates were then diluted and subsequent binding reactions used ~20,000 - 50,000 cpm (counts per min) per reaction.

4.1.2 Radioactive electrophoretic mobility shift assays

To assess CRD-BP variant proteins binding ability to the 181 nt [^{32}P]-labeled c-myc RNA substrate, radioactive electrophoretic mobility shift assays (EMSA) were used. This technique was chosen due to the increased sensitivity over fluorescent EMSA. EMSA binding buffer (Table 9) was made fresh before each experiment and was placed on ice until it was needed. [^{32}P]-labeled c-myc CRD (nts 1705-1886) RNA substrate was subjected to a denaturation and renaturation 2-step process that involved heating the RNA to 55°C for 5 min and then cooling at room temperature for 7 min. Various amounts of the CRD-BP proteins, binding buffers and RNA substrate were added to the EMSA reaction tubes. EMSA reactions were then incubated at 37°C for 10 min, before being placed on ice for 5 min. This heating-cooling cycle to facilitate binding was then repeated, for a total of two times. EMSA loading dye (3 uL) was then added to the reactions and 17 uL of the reaction was loaded on to a 4% native PAGE gel. Complexes were resolved at 20 mA for 50 min. Audioradiographs were generated using a Cyclone Storage Phosphor Screen System and Optiquant Software.

Table 9. Reagents used for radioactive EMSA experiments

Reagent Identity	Reagent Composition
EMSA loading dye	250 mM Tris-Cl (pH 7.4), 0.2% bromophenol blue, 0.2% xylene cyanol, 40% sucrose

4.1.3 Calculation of CRD-BP dissociation constant using [^{32}P] c-myc CRD 1705-1886 RNA substrate

EMSA audioradiographs that were generated were used to determine the dissociation constant of CRD-BP using the [^{32}P] c-myc CRD 1705-1886 RNA substrate. Optiquant Software was used to quantify the bound vs unbound values (blue boxes, Figure 19). The software will calculate the total signal in each lane. Once the bound/total values were generated using the software, the values were plugged in to the Hill equation to calculate the K_d (Figure 19). Figure 19 illustrates the quantification values from the free RNA lane and the WT-CRD-BP concentration of 147 nM (Lane 1 &2, respectively). The software calculated that 8.5 % of the signal was bound. This value is the bound value in the Hill equation and the total would be the free RNA signal (100%).

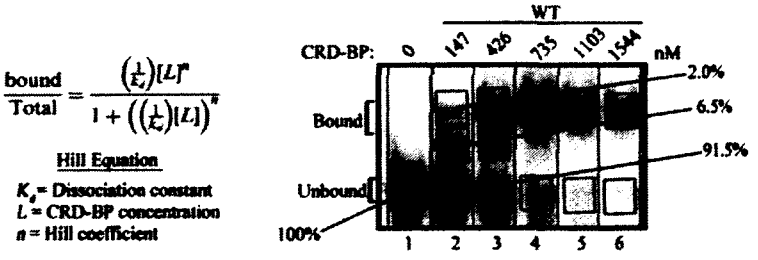

Figure 19. Audioradiograph to calculate the dissociation constant (K_d). Each audioradiograph that was generated was then used to calculate the K_d for CRD-BP. Total signal from each lane (red lines) was calculated by drawing quantification boxes (blue boxes) around the signal generated from the [^{32}P] c-myc CRD 1705-1886 RNA substrate. K_d values were then calculated using the Hill equation.

4.2 Results

Generating the [^{32}P]-labeled c-myc CRD RNA (nts 1705-1886) was done via *in-vitro* transcription using [^{32}P]-UTP as described above. The RNA substrate was resolved on a denaturing-PAGE gel and the appropriate 181 nt size was evident by its relative position to the internal marker, xylene cyanol (Figure 20). The band was then excised and the RNA substrate was gel purified. RNA substrates were successfully labeled with an average radioactivity of ~300,000 cpm.

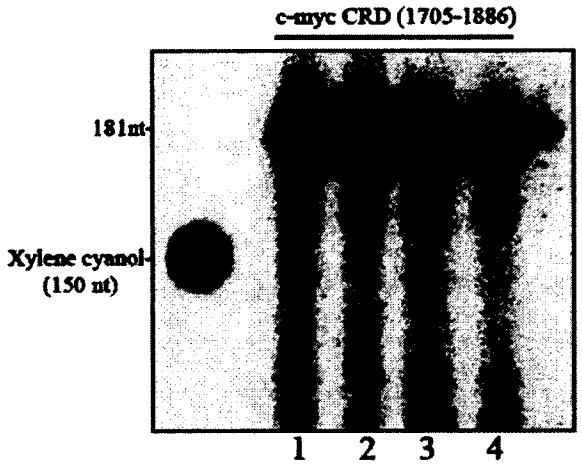

Figure 20. The generation of ^{32}P internally labeled c-myc CRD (nts 1705-1886) RNA. Audioradiograph of an 8% denaturing-PAGE gel illustrating the [^{32}P]c-myc CRD RNA substrate (lanes 1-4).

Following gel purification of the RNA substrate and the dialysis of the CRD-BP variants, radioactive EMSAs were performed to assess the binding affinity of the proteins to c-myc CRD RNA. The CRD-BP mutants were compared to the WT-CRD-BP. The binding profiles of each of the CRD-BP variants are shown in Figure 21.

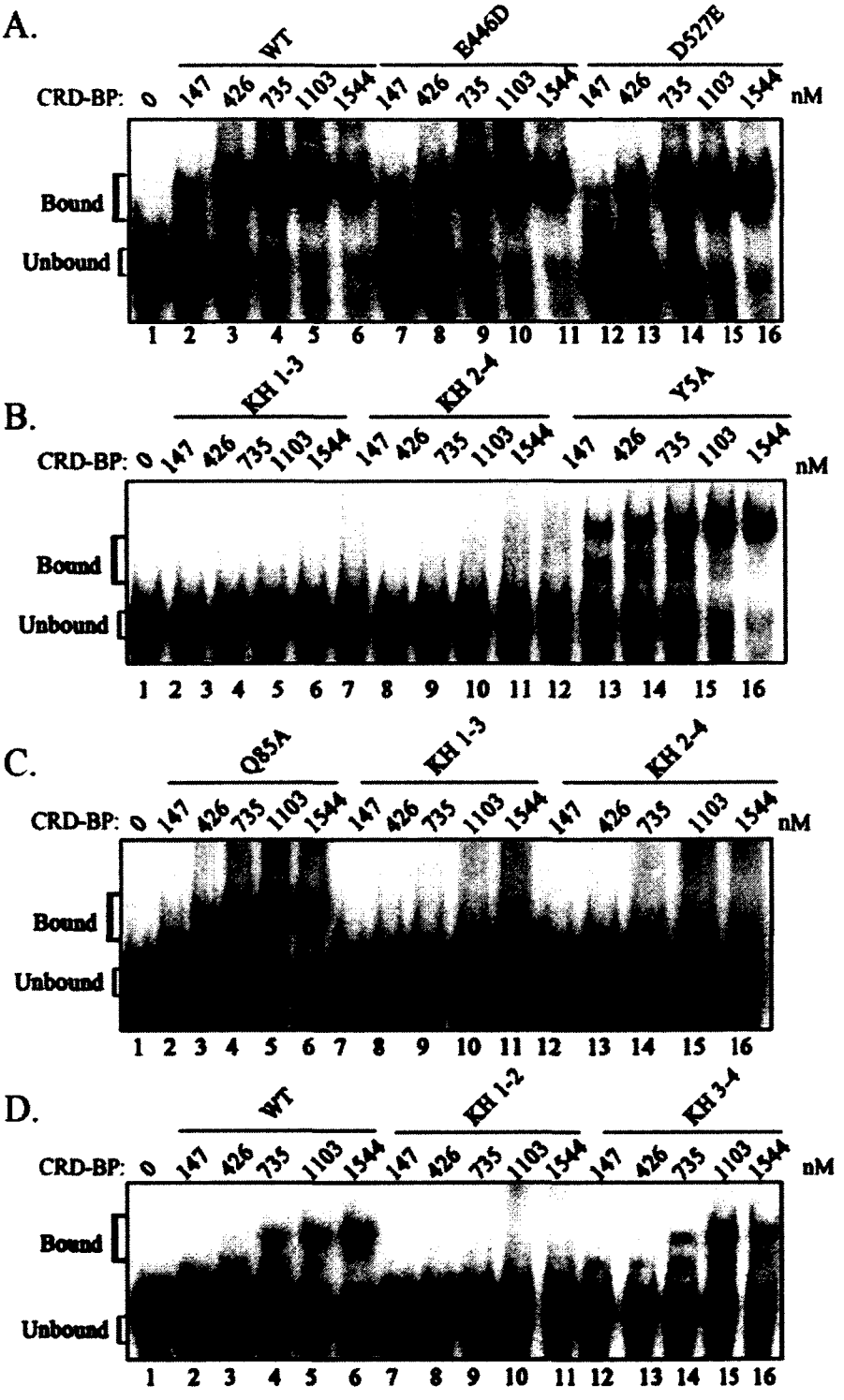

Figure 21. Electrophoretic mobility shift assay of wild-type and mutant CRD-BP. EMSA reactions (17 uL) were loaded on a 4% native PAGE gel and complexes were resolved at 20 mA for 60 min. [^{32}P] c-myc RNA substrate (20,000 - 50,000 cpm/reaction) was used in the binding reactions. The above images are representative of at least two biological replicates.

WT-CRD-BP, CRD-BP-E446D and CRD-BP-D527E all showed comparable binding profiles (Figure 21, A). CRD-BP-Y5A and CRD-BP-Q84A also showed comparable binding profiles to that of the WT-CRD-BP (Figure 21, B & C, respectively). These CRD-BP variants were expected to have no effect and meant to serve as negative controls; therefore they were expected to have a binding profile that was comparable to WT-CRD-BP. Interestingly, CRD-BP-KH2-4, CRD-BP-KH1-3 and CRD-BP-KH1-2, showed a complete loss in their ability to bind the c-myc CRD RNA substrate (Figure 21, B,C&D, respectively). Perhaps the most intriguing binding profile was that of the CRD-BP-KH3-4 variant. Although it contains the two mutations located within the G-X-X-G motif in both KH3 and KH4, the CRD-BP-KH3-4 variant was still able to bind the c-myc CRD RNA substrate (Figure 21, D). Following the qualitative analysis of the EMSA gels, a more thorough quantitative analysis was done. Determining the dissociation constant (K_d) for a ligand binding to a macromolecule (i.e. RNA substrate binding to a protein) gives a quantitative analysis of the affinity the ligand has for the macromolecule. To determine the affinity that CRD-BP has for the 181nt c-myc CRD RNA substrate, the K_d was determined using the Hill equation. The binding data for each of the CRD-BP variants (Figure 21, A-D) was fit to the Hill equation and showed a sigmoidal curve when graphed (Figure 22).

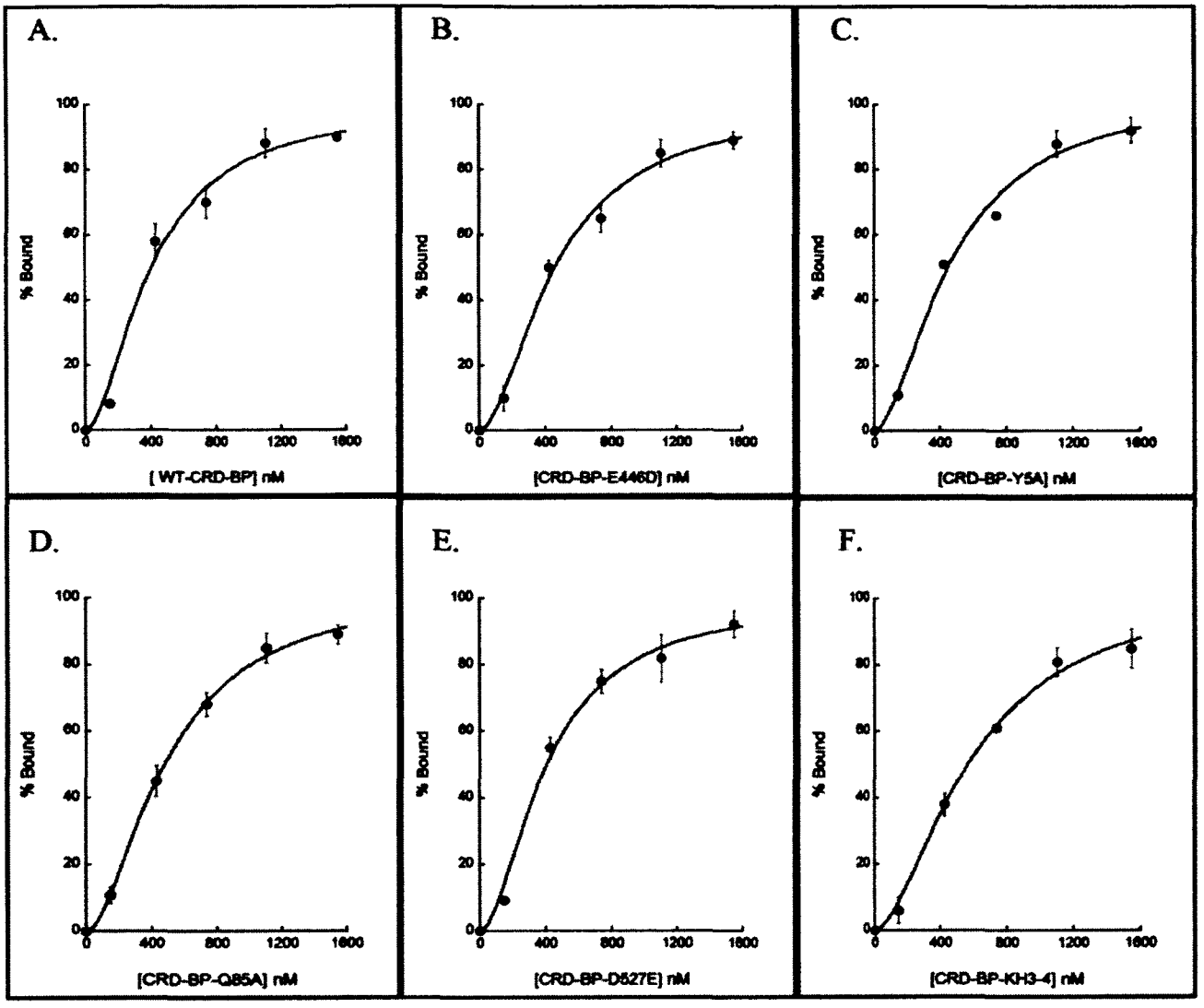

Figure 22. CRD-BP variants binding curves as generated by the Hill Equation. Binding curves were generated using densitometry and the data from Figure ?A-D. CRD-BP variants were incubated with [^{32}P] c-myc CRD 1705-1886 and the binding profiles were analysed using densitometry. % bound vs % unbound values were used to generate the binding curves and calculate the dissociation constant.

Each of the following dissociation constants was determined using the [^{32}P] c-myc CRD 1705-1886 RNA substrate. Binding profiles were analysed using densitometry and binding values were fit to the Hill equation. The dissociation constant for WT-CRD-BP was determined to be 398.03 ± 52.79 nM (Figure 22, A). The dissociation constant for CRD-BP-E446D was determined to be 459.54 ± 42.18 nM (Figure 22, B). The dissociation constant for CRD-BP-Y5A

was determined to be 467.57 ± 45.68 nM (Figure 22, C). The dissociation constant for CRD-BP-Q85A was determined to be 490.56 ± 20.01 nM (Figure 22, D). The dissociation constant for CRD-BP-D527E was determined to be 402.43 ± 37.25 nM (Figure 22, E). Finally, the dissociation constant for CRD-BP-KH3-4 was determined to be 586.89 ± 42.89 nM (Figure 22, E). Binding curves were not generated and dissociation constants were not calculated for the KH double mutants KH1-2, KH1-3 and KH2-4 because there was no binding exhibited. Further analysis was done to determine if there were any significant differences in the CRD-BP variants K_d, as compared to that of WT-CRD-BP (Figure 23). Using a two-tailed Student's t-test ($P < 0.05$), only the K_d of the CRD-BP-KH3-4 double mutant showed a significant difference when compared to the K_d of WT-CRD-BP (Figure 23).

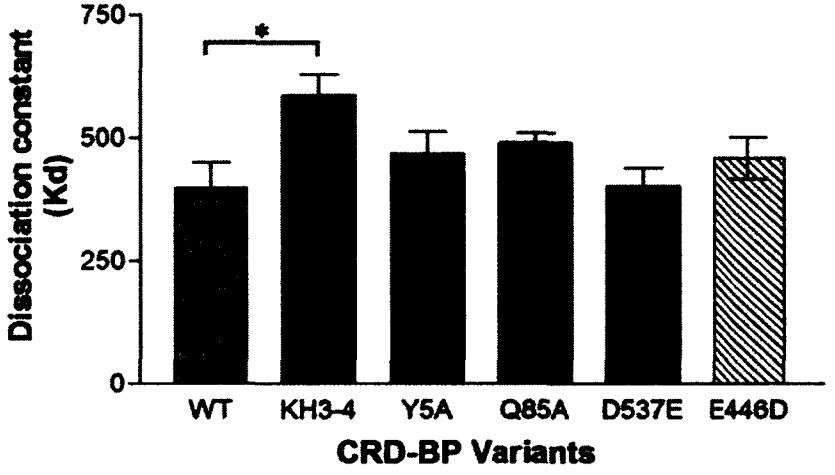

Figure 23. Comparison of CRD-BP variants dissociation constants (K_d). Statistical significance between the K_d values of the CRD-BP mutant variants and the K_d of WT-CRD-BP was assessed by the unpaired Student's t-test (n=4). Only the CRD-BP-KH3-4 double mutants K_d (P= 0.0321) showed a significant difference when compared to WT-CRD-BP. The asterisk indicates that the p-value is less than 0.05.

4.3 Discussion

To determine the role of the G-X-X-G motif located within the four KH domains of CRD-BP in binding RNA substrates, site-directed mutagenesis was used to incorporate a missense mutation (G>D) at the first glycine in each of the four G-X-X-G motif within CRD-BP. Once the mutations were confirmed using sequencing, electrophoretic mobility shift assays (EMSA) were used to analyze the protein-RNA interactions. Gel shift assays are a simple way to determine both qualitative and quantitative relationships between proteins and nucleic acids (Ryder, et al. , 2008). This powerful technique allows for the *in-vitro* analysis of how proteins and their RNA substrates interact with each other. Here, EMSA were used to analyze the CRD-BP-RNA interaction. [^{32}P]-c-myc CRD 1705-1886 RNA substrate was incubated with the various CRD-BP variants and the complexes were resolved on a PAGE gel. The gels were then analyzed using densitometry. The shift assays were then used to determine the dissociation constant K_d by fits to the Hill equation. The negative control mutants (E446D, Y5A, Q85A, and D527E) exhibited similar binding affinity to that of WT-CRD-BP; showing no statistical significant difference in the K_d. No effect was exhibited following the incorporation of the missense mutations in the RRM domains and in the variable looped regions of the KH domains, indicating that the amino acids involved do not play any role in the binding of CRD-BP to the c-myc CRD RNA substrate. Mutations within the G-X-X-G motifs of the CRD-BP KH domains did, however, affect binding.

Previous data (van Rensburg, unpublished data) showed that only single point mutations (G>D) at the first glycine on the G-X-X-G motif within CRD-BP KH1 domains has a modest effect (about 2-fold decrease) on the binding ability of the CRD-BP mutants (Table 10). Similar

point mutations in the G-X-X-G motif in each of the KH2, KH3 and KH4 domain had no effect (Table 10).

Table 10. CRD-BP variant dissociation constants (K_d) for c-myc CRD RNA.

CRD-BP Variant	K_d
G214D (KH1)	723.07 ± 75.13*
G422D (KH3)	320.81 ± 42.75
G214D-G284D (KH1-2)	No Binding
G284D-G506D (KH2-4)	No Binding

Student's t-test (n=4, $P^* < 0.05$) was used to determine whether there was a statistically significant difference between the K_d values of the KH single mutants compared to the WT-CRD-BP K_d. The asterisk indicates that the p-value is less than 0.05.

To follow up on the previous data generated in Dr. Lee's Lab (van Rensburg, unpublished data), a further investigation of the role of the first glycine of the KH domains was done. Different combinations of missense mutations within multiple KH domains were generated. Surprisingly, KH1-2, KH 1-3 and KH2-4 double mutants showed a complete loss in the ability to bind the c-myc RNA substrate (Figure 21). These results indicate that when the first glycine of KH1 & 2, KH1 & 3 and KH2 & 4 are mutated together, they have a detrimental effect on CRD-BP's ability to bind the c-myc CRD RNA substrate. Therefore, it shows that when the first glycine of the G-X-X-G motif from multiple KH domains is mutated together, binding can be disrupted. The G-X-X-G motifs of the different KH domains show cooperative binding. These results suggest that the KH domains in CRD-BP bind in tandem to c-myc CRD RNA.

Perhaps the most interesting observation made was with the KH3-4 double mutant. This mutant was still able to bind the c-myc RNA substrate. The binding affinity to the c-myc RNA

substrate was significantly lower than that of the WT-CRD-BP, however binding was not completely abolished as with the other double mutants. This indicates that the combination of mutations within the KH3 and KH4 G-X-X-G motif has less of an effect on the binding ability of CRD-BP. This suggests that regions other than the G-X-X-G motif within KH3 and KH4 are important for binding to c-myc CRD RNA. A comprehensive discussion of the biological significance that the missense mutations within the G-X-X-G motif of CRD-BP has on binding, and the role that the VICKZ protein KH domains play with RNA binding, will be discussed in Chapter 6.

Chapter 5

Characterization of the CRD-BP-RNA Interaction in Cells

This chapter describes the studies performed to analyze CRD-BP-RNA interactions in HeLa human cervical cancer cells. In Chapter 4, various CRD-BP mutants including mutants with point mutation at the KH domains were directly assessed for ability to bind c-myc CRD RNA *in-vitro*. In the first part of this Chapter, immuno-precipitation method coupled with qPCR was used to assess the ability of WT-CRD-BP and its various mutants to physically associate with c-myc, CD44 and β–actin mRNAs in HeLa cells.

Previous work done in Dr. Lee's lab has shown that oligonucleotides designed to target the 3'-UTR of hCD44 mRNA have the ability to break the WT-CRD-BP-RNA interaction *in-vitro*. Furthermore, previous data has shown that the antibiotic neomycin has the ability to break the CRD-BP-RNA interaction *in-vitro*. The second part of this chapter describes the experimental techniques used to test whether specific oligonucleotides and antibiotics have the ability to break the CRD-BP-RNA interactions in cells. The experimental approach used was similar to that described in Chapter 4, however, there were some differences that will be described below.

5.1 Methodology

5.1.1 Immuno-precipitation of FLAG-CRD-BP-RNA complex

To analyse the ability of both wild-type and mutant CRD-BP variants to bind mRNA in cells, immuno-precipitation method using anti-FLAG antibody was done. HeLa human cervical cancer cells, purchased from American Tissue Culture Collection (ATCC), were plated on 100 mm dishes at 10 x 10^4 cells/ ml. The cells were cultured in Minimum Essential Media (Gibco) containing 10% Fetal Bovine Serum. On the second day of growth, the cells were transfected

with FLAG-CRD-BP plasmids (Wild Type and mutants) using Lipofectamine 2000 reagent (Invitrogen). On the fourth day of growth, the cells were lysed using TCLB-L buffer [TCLB (50 mM Tris, 130 mM NaCl, 625 uM EDTA, 2.5% Triton x-100), 7 x TCLB-C (TCLB, complete tablet), VRC (800 uM), DTT (400 uM), RNasin (40 U/ uL)]. The TCLB-L was added to the plate and the cells were incubated on ice for 5 min. The cells were then removed from the plates, using cell scrapers, and transferred to an eppendorf. The lysed cells were then aspirated using a 26 gauge needle 5 times to break up the nuclei before incubation on ice for an additional 30 min.

While the cells were on ice, the Protein-G agarose beads for the immuno-precipitaion were equilibrated. The beads were first re-suspended until homogenous, then for each plate of cells, 150 uL of homogenized beads were added to an eppendorf. The beads were spun at 3000 x g for 1 min and the supernatant was removed. The beads were re-suspended in 500 uL TCLB-L, spun at 3000 x g for 1 min and the supernatant was removed. This TCLB-L wash step was repeated once more before the beads were re-suspended in 75 uL of TCLB-L.

After the 30 min incubation on ice, the cells were spun at 16,100 x g for 10 min at 4°C. The supernatant was collected in a new eppendorf and the remaining cell lysate was pre-cleared. To the cell lysate, 50 uL of the 50% Protein-G mixture was added. The cell lysate was then rotated at 4°C for at least an hour before being spun at 3000 x g for 1 min. The supernatant was discarded and the pre-clearing was repeated for a total of at least 3 hours. After pre-clearing 2 times, the pre-cleared lysate was collected and 5 uL of Anti-FLAG antibody (Sigma) was added. The lysate containing the antibody was rotated at 4°C overnight.

The following morning, the remaining resin was added to the lysate and the lysate was rotated for 4 hours at 4°C. The lysate/resin was spun down at 3000 x g for 1 min. The resin was

washed 4 times with TCLB-W buffer [TCLB, DTT (400 uM), VRC (800 uM)] for 5 min at 4°C. After the TCLB-W wash steps, the resin was washed 5 times with TCLB-U [TCLB, DTT (800 uM), VRC (1.60 mM), 0.46 M Urea], each for 15 min at 4°C. After re-suspending the resin following the 4th TCLB-U wash step, 50% of the resin was removed, spun down at 3000 x g for 1 min, and re-suspended in 16 uL of H_2O to be used for Western analysis. The remaining resin was washed, and then the supernatant was removed. The resin was then incubated in 100 uL of TCLB-PK [500 uL TCLB-W, Proteinase K (150 ug), 0.1 % SDS] for 30 min at 50°C. Following the incubation step, the RNA was phenol/chloroform extracted twice. To the extracted RNA, 4 uL yeast tRNA (5 mg/ml), 24 uL 3 M sodium acetate and 500 uL 100% ethanol was added and the eppendorf was placed at -80°C to precipitate the RNA.

The precipitated RNA was then re-suspended in 8 uL of nuclease-free H_2O. The RNA was DNase treated using a DNA-free kit (Ambion). The 10 x buffer (1 uL) and 1 uL of DNaseI was added to the RNA and then the mixture was incubated at 37°C for 30min. Following the incubation, 2 uL of the inactivating reagent was added and the mixture was incubated at RT for 2 min and then spun at 10,000 x g for 1.5 min. The purified RNA was then used to make cDNA for real-time quantitative polymerase chain reaction (RT-qPCR) analysis.

5.1.2 cDNA synthesis and RT-qPCR

The BioRad iScript kit was used for cDNA synthesis. Equivalent amounts of RNA were used for each cDNA synthesis reactions. RNA (~1 ug), 4 uL of 5 x iScript mix, 1 uL of reverse transcriptase and H_2O (up to 20 uL) was placed in the thermocyler and the first strand synthesis of cDNA was done using the following program - 25°C 5 min, 42°C 30 min, 85°C 5 min. Following cDNA synthesis, qPCR reactions were mixed by adding 2 uL cDNA, 12.5 uL iQ™

SYBR® Green Supermix (Bio-Rad), 1.0 uL forward primer (10 uM), 1.0 uL reverse primer (10 uM) and 9.5 uL nuclease-free water. The reaction mixtures (25 uL) were added to 96-well PCR plates and placed in a BioRad iQ5 cycler where the cDNA was amplified using a two-step PCR reaction. Prior to all qPCR experiments, optimal temperature for each primer set used (β-actin, c-myc and CD44) were determined using melt-curve analysis. The efficiency of the qPCR reactions was also established using the standard curve method.

5.1.3 Total RNA extraction from antibiotic-treated HeLa cells

To assess the effect of antibiotics on target mRNA levels, HeLa cells (ATCC) were treated with neomycin and lincomycin. Cells were plated on 100 mm dishes at 10 x 10^4 cells/ml. The cells were cultured in Minimum Essential Media (Gibco) containing 10% Fetal Bovine Serum. On the second day of growth, the cells were transfected with pcDNA-FLAG-WT-CRD-BP plasmid using Lipofectamine 2000 reagent (Invitrogen). On the third day, cells were treated with neomycin and lincomycin. Neomycin was used because it has previously shown to have the ability to break the WT-CRD-BP interaction *in-vitro*. Lincomycin was unable to break the interaction *in-vitro*; therefore, it was meant to serve as a negative control. Initially, a concentration dependent reaction was done to determine the optimal antibiotic treatment amount. The optimal concentration was determined to be 750 uM based on the lack of cellular death following treatment. This was the concentration used for all subsequent antibiotic experiments.

On the fourth day, cells were harvested and the RNA was extracted from cells using the *mir*Vana™ miRNA Isolation Kit (Life Technologies). The cells were first washed with 3 mL of PBS and 200 uL of lysis buffer was added and pipetted up and down to mix lysate. The lysate was further mixed using SP Vortex. Then 60 uL of homogenate additive was added to the lysate it was incubated on ice for 10 min. Acid-phenol:chloroform (600 uL) was added to the lysate,

vortexed, and spun at 10,000 x g for 5 min. The aqueous layer was collected and pure ethanol was added to it. The solution was then loaded on to the *mir*Vana Filter Cartridge, which immobilizes the RNA. Heated (95°C) elution buffer was added to the center of the column and the RNA was eluted with a 10,000 x g spin for 1.5 min. RNA integrity was checked on 1% agarose gels. RNA was quantified using NanoDrop.

5.1.4 Oligonucleotides and antibiotics

Previous work done in Dr' Lee's lab has shown that specific oligonucleotides and antibiotics have the ability to break the WT-CRD-BP-CD44 RNA interaction *in-vitro*. To assess the ability of oligonucleotides and antibiotics to break the CRD-BP interaction in cells, immuno-precipitation experiments were performed. Oligonucleotides used were designed to target the 3'-UTR of hCD44 mRNA (Table 11). The oligonucleotides with a modified 2'-O-methyl were commercially synthesized (Integrated DNA Technologies, Inc). Scrambled-negative (SN) oligonucleotide (Integrated DNA Technologies, Inc) was used as a negative control. Immuno-precipitation experiments were performed as described above; with the additional treatment of oligonucleotides or antibiotics 24 hours post transfection with pcDNA-FLAG-WT-CRD-BP. Concentration-dependent experiments of each oligonucleotide and antibiotic used were performed to determine the optimal treatment concentrations (Table 12). The extraction of the FLAG-CRD-BP-RNA complex was described above.

Table 11. 2'-O-methyl oligonucleotides used in cell-based experiments.

Oligonucleotide Name	Sequence Identity	Nucleotide location at 3' UTR on hCD44 mRNA
DD7	5'-CCTTTCCAACCGCTAGTCCTTA-3'	2976-2997

Table 12. Optimal antibiotics and oligonucleotide treatment concentrations.

Molecule	Optimal treatment Concentrations
Lincomycin	750 uM
DD7-oligonucleotde	500 nM

5.2 Results

5.2.1 Optimizing the experimental conditions for immuno-precipitation coupled quantitative real-time PCR

Before proceeding to the immuno-precipitation (IP) of the CRD-BP-RNA complex, HeLa cell growth was optimized to ensure consistent cellular growth among each treatment group. Cells were visually inspected prior to harvesting to ensure optimal cell growth was achieved (Figure 24).

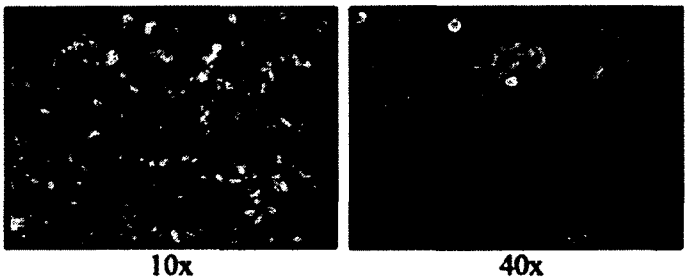

Figure 24. HeLa cells visualised under a light microscope. Cell density was determined by visually inspecting the cells prior to harvesting the cells. An initial magnification (10x) was used to inspect the entire 100 mm dish to ensure a consistent cellular growth throughout. Further magnification was done (40x) to inspect the media for bacterial contamination.

Once optimal cell growth was achieved, the cells were transfected with the appropriate FLAG-pcDNA-CRD-BP plasmids. The FLAG-CRD-BP-RNA complexes were immuno-precipitated and the RNA was isolated from the complex. RNA integrity and purity was determined using the 260/280 and 260/230 ratios. The integrity/purity ratios from each treatment groups were in the acceptable range (~ 2.0) (Table 13).

Table 13. Quantification and quality of immuno-precipitated RNA.

Plasmid Transfected	[ng/uL]	260/280	260/230
pcDNA-WT-CRD-BP	1463	1.94	2.06
pcDNA-CRD-BP-Q85A	1419	1.90	2.00
pcDNA-CRD-BP-D527E	1683	2.01	2.04
pcDNA-CRD-BP-KH2	1639	1.95	2.03
pcDNA-CRD-BP-KH4	1525	1.88	2.20

Following purification of the RNA, synthesis and amplification of cDNA, melt-curve analysis was performed to analyze the specificity of the primer sets used (Table 14). c-myc primers produced a single band at ~88°C, CD44 primers produced a single band at ~84°C and β-actin primers produced a single band at ~86°C (Figure 25, A,B,C, respectively). These results indicated a single PCR product for each of the primer sets (Table 14).

Table 14. Primer sets used in real-time RT-PCR experiments

Target Gene	Primer Sequences
c-myc	Forward, 5'-ACGAAACTTTGCCCATAGCA-3' Reverse, 5'-GCAAGGAGAGCCTTTCAGAG-3'

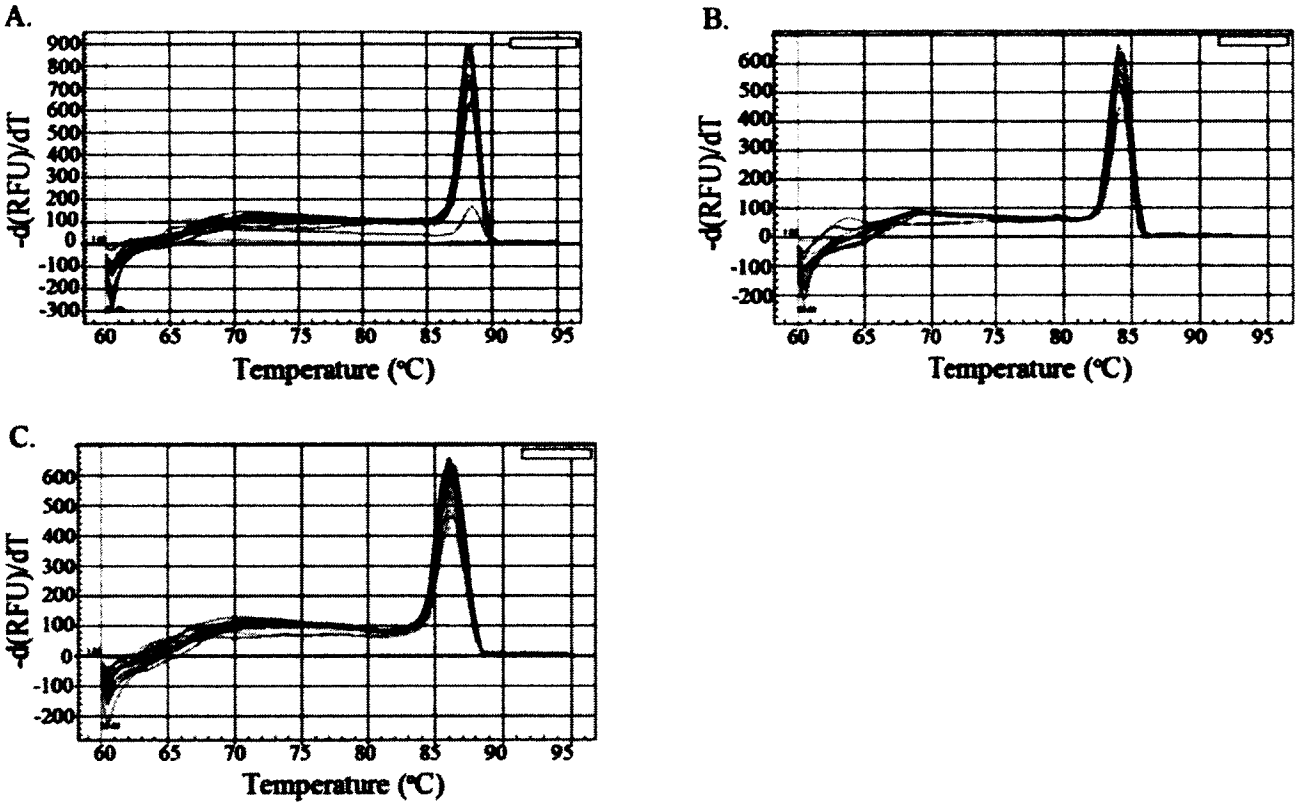

Figure 25. Melt-curve analysis of RT-qPCR primers. c-myc (A), CD44 (B) and β-actin (C) primer sets were analysed for primer specificity which is indicated by the single peaks above. The optimal annealing temperature for all three primer sets was 60°C, which was subsequently used with all RT-qPCR experiments.

The efficiency of the qPCR reactions were evaluated using standards curves generated by performing a dilution series experiment for each gene used. The c-myc standard curve exhibited an efficiency of 100.2% (Figure 26, A), the CD44 standard curve exhibited an efficiency of 101.1% (Figure 26, B), and the β-actin standard curve exhibited an efficiency of 95.0% (Figure 26, C).

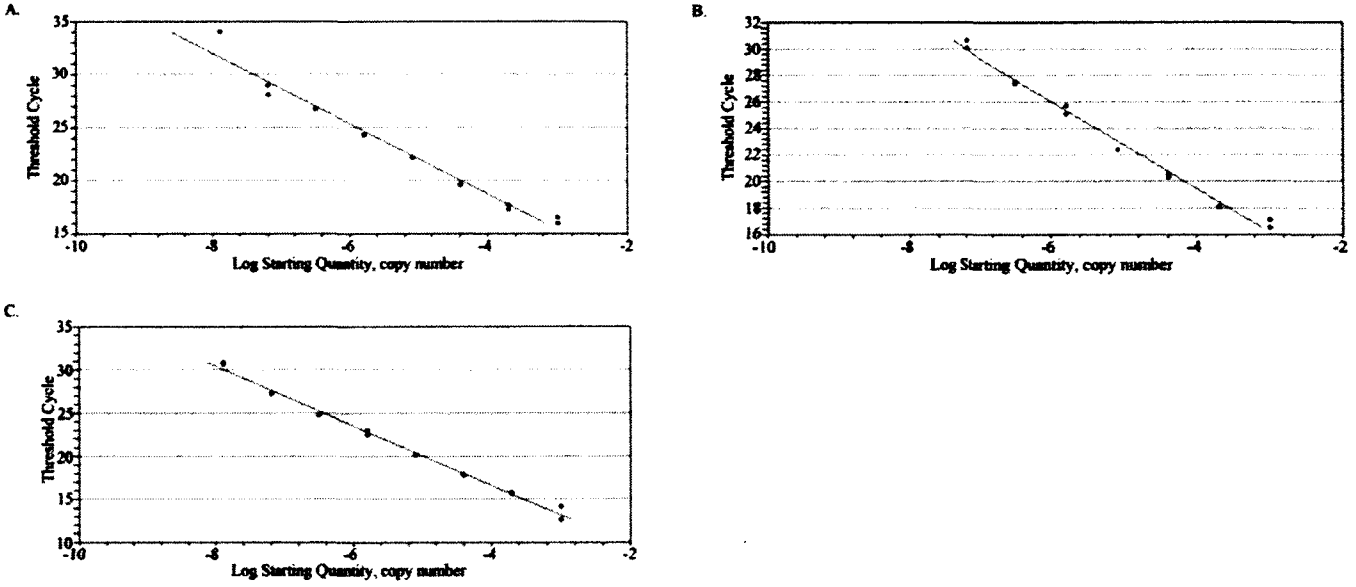

Figure 26. Standard curves generated to assess PCR reactions efficiency. A serial dilution series was used to generate standard curves for c-myc (A, R^2= 0.975, slope= -3.317), CD44 (B, (R^2= 0.991, slope= -3.296), and β-actin (C, R^2= 0.993, slope= - 3.456)

5.2.2 Assessing the ability of wild-type CRD-BP and its mutants to physically associate with mRNAs in HeLa cells.

Immuno-precipitation of the Flag-CRD-BP-RNA complexes was then performed following the optimization of HeLa cell growth and RT-qPCR conditions. Based on Western blot analysis, it was determined that the optimal amount of pcDNA-FLAG-CRD-BP plasmid to transfect into HeLa cells was 10 ug (Figure 27).

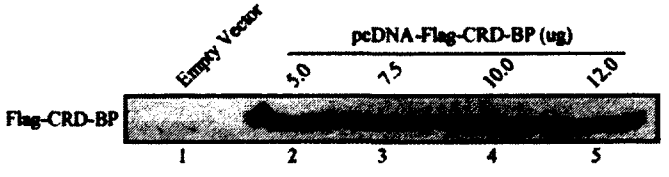

Figure 27. Expression of exogenous FLAG-WT-CRD-BP in HeLa cells. Western blot analysis was used to detect the exogenous FLAG-WT-CRD-BP protein (60 ug) levels at different amount of pcDNA-FLAG-WT-CRD-BP plasmid transfected (Lanes 2-4). The exogenous FLAG-CRD-BP was probed with anti-flag antibody (Promega).

Western Blot analysis was then performed on the immuno-precipitated FLAG-CRD-BP variants to ensure equivalent protein expression; this confirms that potential differences in target gene expression are due to differences in CRD-BP binding to its RNAs and not due to differences in protein expression levels. The negative controls Y5A and D527E showed comparable protein expression (Figure 28, A) and target mRNA levels (Figure 28, C,E,G) to that of WT-CRD-BP, while no protein expression was detected with the negative controls Q85A and E446D (Figure 28, A). The cause for the lack of protein expression in Q85A and E446D mutants is currently unknown. It is possible that the point mutation resulted in disruption of protein structure which prevented the FLAG-tag to be exposed. Given the lack of expression of FLAG-CRD-BP from cells transfected with plasmids carrying these mutations, the mRNA associated with these FLAG-CRD-BP mutants were not analysed. The KH single mutants KH1, KH2, KH3 and KH4 all showed relatively comparable protein expression to that of WT-CRD-BP (Figure 28, B). Following the confirmation of equal FLAG-CRD-BP protein expression, mRNA levels of β-actin, c-myc and CD44 which physically associated with FLAG-CRD-BP were then measured using real-time qPCR. Cells transfected with CRD-BP-KH3 or CRD-BPKH4 showed significant decrease in β-actin mRNA; while the CRD-BP-KH1 and CRD-BP-KH2 KH2 treatment groups had β-actin mRNA levels comparable to that of WT-CRD-BP (Figure 28, D). There was a significant decrease in c-myc mRNA levels in cells transfected with CRD-BP-KH1, CRD-BP-KH2 and CRD-BP-KH3; while the levels of c-myc mRNA associated with CRD-BP-KH4 and WT-CRD-BP were comparable (Figure 28, F). Each of the KH single mutants showed a significant decrease in the level of CD44 mRNA associated with it (Figure 28, H).

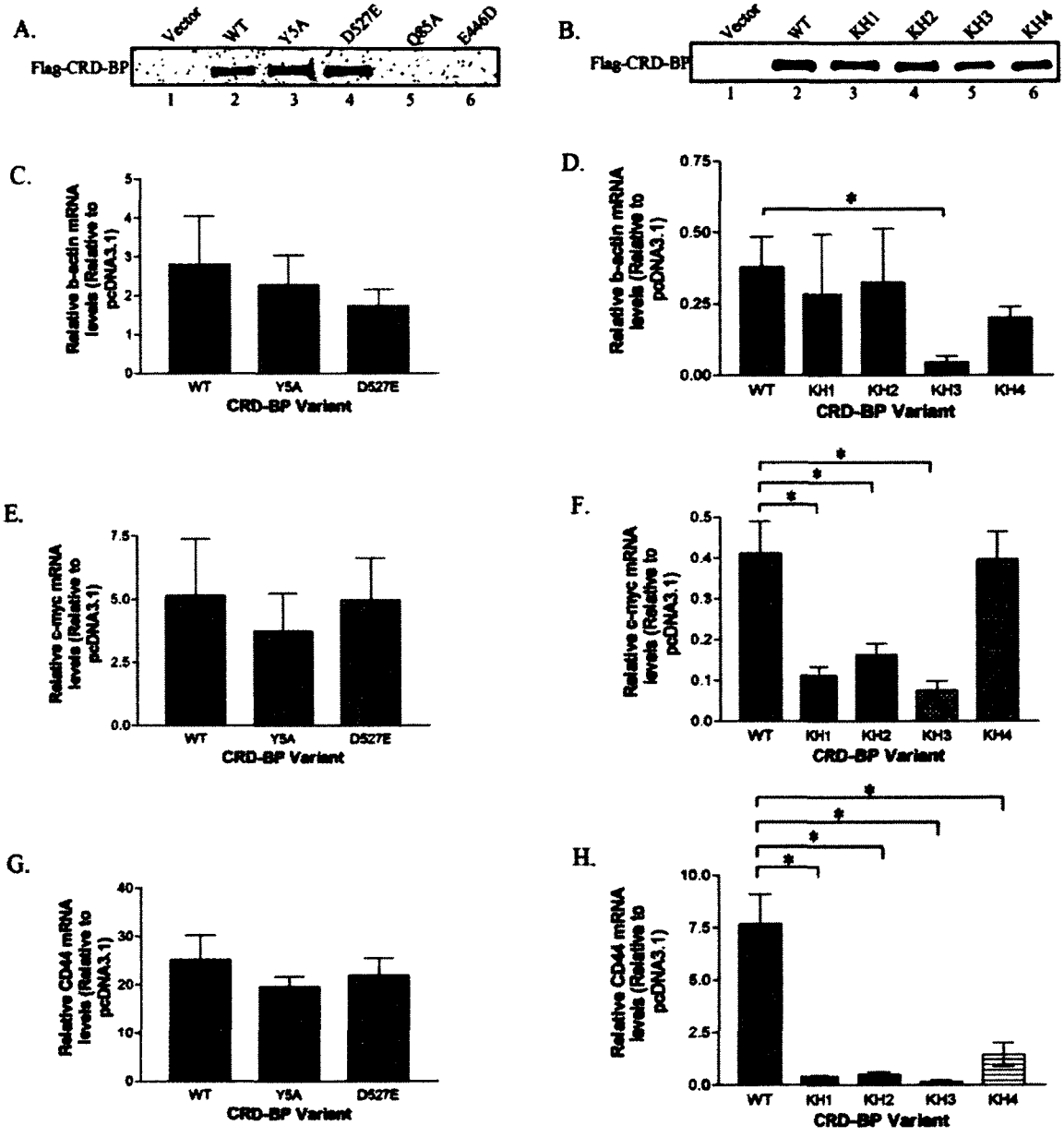

Figure 28. Analysis of specific mRNAs physically associated with CRD-BP and its mutants. FLAG-CRD-BP-RNA complexes isolated from HeLa cells were pulled down with anti-FLAG antibody. 16 uL of immuno-precipitated lysate was used in Western blot analysis to detect and confirm comparable exogenous FLAG-CRD-BP protein expression (A,B). The immuno-precipitated FLAG-CRD-BP was detected using anti-FLAG antibody. β-actin, CD44 and c-myc mRNAs isolated from the complexes were measured using qRT-PCR. mRNA levels were expressed over the pcDNA-FLAG control and compared to WT-CRD-BP (C-H). Statistical significance between the expression of target genes in WT-CRD-BP treated cells and CRD-BP mutants was assessed by the unpaired Student's t-test (n=4). The asterisk indicates that the p-value is less than 0.05.

5.2.3 Assessing the ability of oligonucleotides and small molecule antibiotics to disrupt CRD-BP-RNA interaction in cells.

To determine the effect of neomycin and lincomycin on β-actin, c-myc and CD44 mRNA levels in pcDNA-FLAG-WT-CRD-BP transfected cells, total cellular protein and total RNA were extracted from the HeLa cells. As shown in Figure 29B, the exogenous protein levels were reduced in the antibiotic-treated cells as compared to the control (non-treated). There was no significant change in the mRNA levels of the β-actin, CD44 or c-myc transcripts in anti-biotic-treated cells compared to non-treated cells (Figure 29, A).

An immuno-precipitation experiment to assess the ability of the antibiotics to break the WT-CRD-BP-RNA interaction in cells was then performed. The immuno-precipitated levels of WT-CRD-BP protein from antibiotic-treated cells were found to be significantly reduced (Figure 29, D). The assessment of the immuno-precipitated target mRNA levels showed that there was no significant effect on the mRNA levels of the target genes (Figure 29, C).

To determine if the oligonucleotides possess the ability to break the WT-CRD-BP-RNA interaction *in-vivo*, immuno-precipitation experiments on oligonucleotide-treated cells were performed. Protein expression levels were found to be unaffected by the treatment of the oligonucleotides. The scrambled-negative, DD1 and DD7 oligonucleotide treated cells all showed equivalent FLAG-CRD-BP expression levels (Figure 30, A). Furthermore, there was no significant effect on the target RNA transcript levels when cells were treated with the oligonucleotides (Figure 30, B). These results suggest that DD1 and DD7 oligonucleotides do not have the ability to break the WT-CRD-BP-CD44 and c-myc RNA interactions in cells.

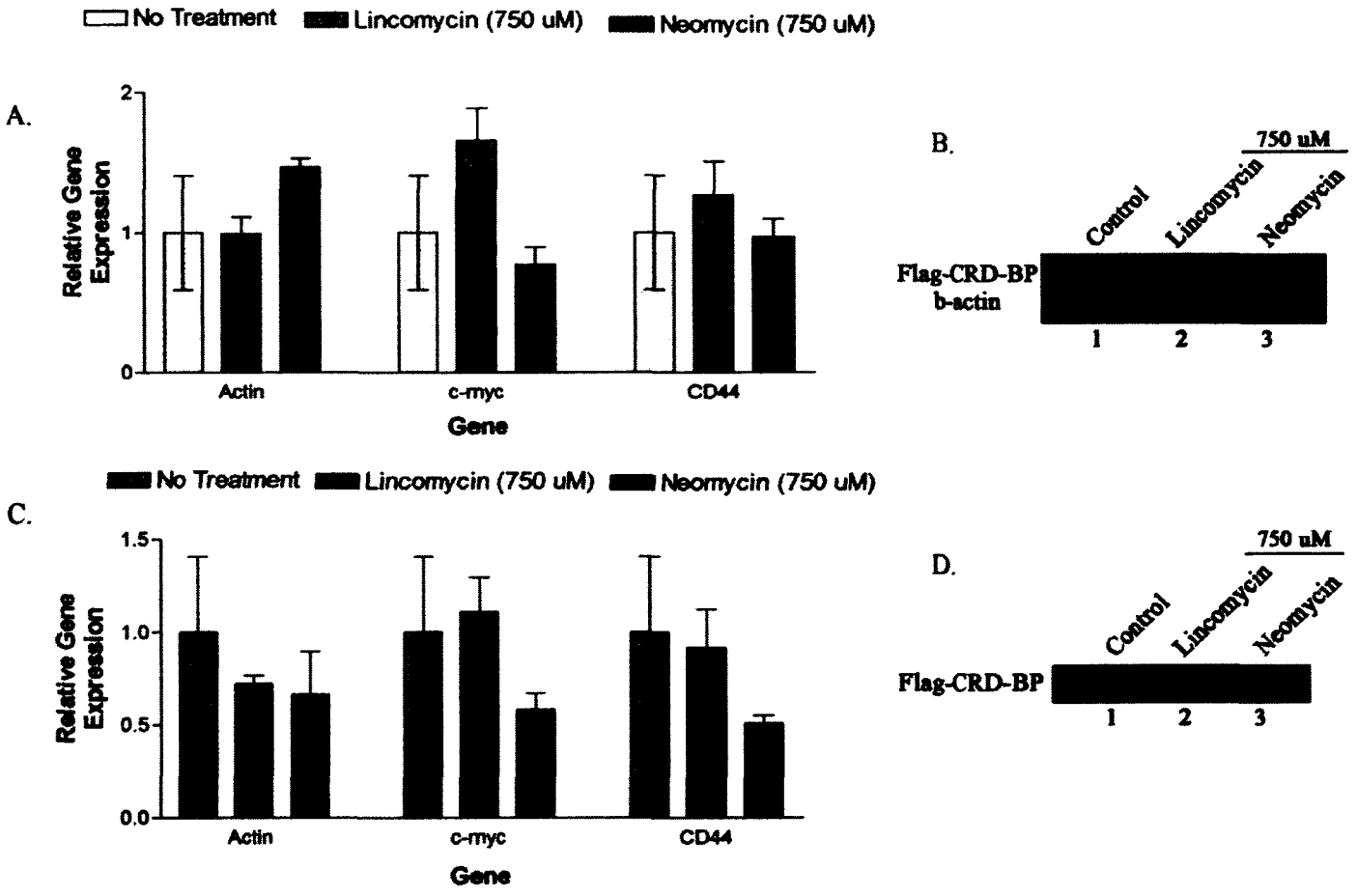

Figure 29. Effect of antibiotic treatment on mRNA and protein expression in WT-CRD-BP transfected HeLa cells. Total cellular RNA (A) and total cellular proteins (60 ug) (B) were extracted from HeLa cells. Immuno-precipitated RNA (C) and 16 uL of immuno-precipitated FLAG-CRD-BP lysate (D) from WT-CRD-BP transfected HeLa cells. RT-PCR and Western analysis was used to assess the effect of antibiotic treatment (750 uM) on protein and relative mRNA levels. Statistical significance between gene expression from no treatment cells versus antibiotic treated cells was assessed by the unpaired Student's t-test (n= 6, p= 0.05).

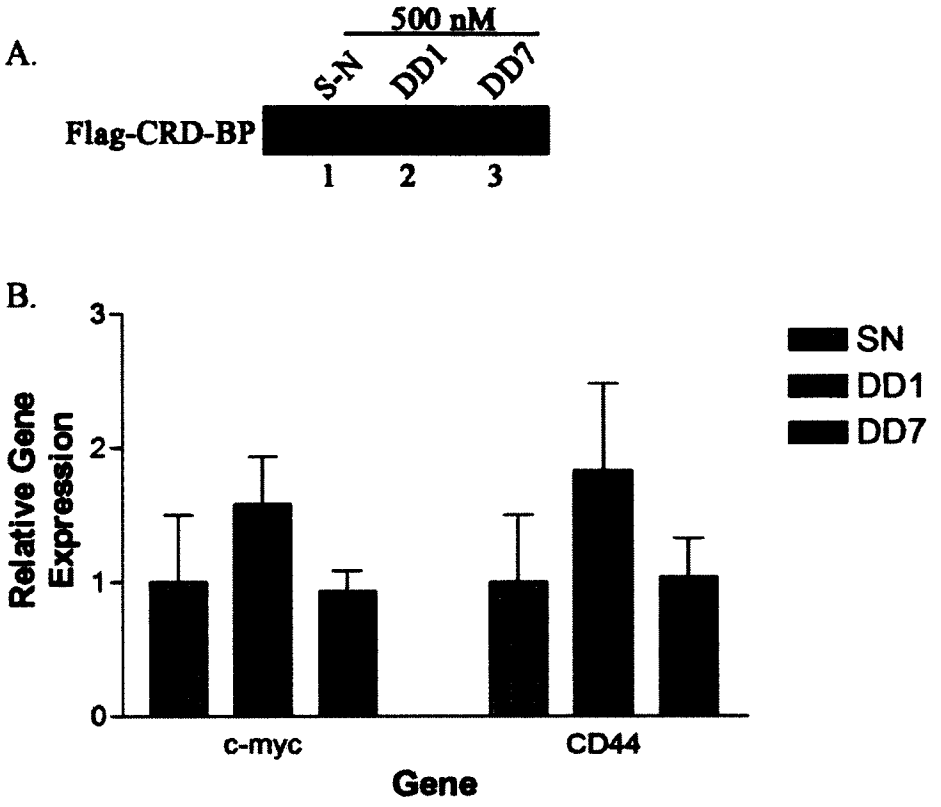

Figure 30. Effect of oligonucleotide treatment on protein and gene expression in WT-CRD-BP transfected HeLa cells. 16 uL of Immuno-precipitated FLAG-CRD-BP lysate (A) and immuno-precipitated RNA (B) from WT-CRD-BP transfected HeLa cells was analysed using western blot analysis and RT-PCR, respectively, to assess the effect of oligonucleotide treatment (500 nM) on relative gene expression. Two-way ANOVA statistical analysis (n=4, p=0.05) was used to assess the statistical differences between control and test groups.

5.3 Discussion

A further investigation of the role of the G-X-X-G motif located within the four KH domains of CRD-BP in binding target RNAs in cells was done. pcDNA-FLAG-CRD-BP plasmids with the first glycine of the G-X-X-G motif mutated to an aspartate were transfected into HeLa cells to assess the binding ability of the CRD-BP variants. Immuno-precipitation (IP) method to analyse protein-RNA interactions in cells has been previously used to identify mRNA

targets for the RNA-binding protein SF2/ASF (Sanford et al., 2008). Such powerful technique allows for the quantitative analysis of protein-RNA interactions within cells. Here, IP was used to compare the binding ability of CRD-BP variants.

5.3.1 Optimization of experimental conditions

CRD-BP-RNA complexes were precipitated using anti-FLAG antibodies and total RNA were isolated and purified prior to qPCR. To measure specific mRNAs that physically associate with FLAG-CRD-BP using RT-PCR, the first step was to synthesize cDNA using PCR. When analysing mRNA levels from treated cells, the mRNA is typically normalized to ensure equal amounts of template for each cDNA synthesis reaction. However, with immuno-precipitation, the normalization step is achieved by analysing the precipitated protein expression levels via Western analysis.

A necessary step with cDNA synthesis is to include a no-template-control (NTC) in the cDNA synthesis reactions (Bustin and Nolan, 2004). An NTC contains no RNA template in the reverse-transcription reaction. This step was to confirm that the DNase treatment following the isolation of RNA transcripts from the CRD-BP complexes was successful. Each of the cDNA synthesis reactions in the current study included a NTC group to confirm that the samples were free of genomic DNA. The NTC samples were always included in the RT-PCR reactions to confirm that the cDNA template did not contain any genomic DNA. The C_t values of the NTC samples were > 5 C_t values from the test samples analysed which indicated that there was no DNA contaminants present (Bustin and Nolan, 2004). This is an essential step when performing RT-PCR because any genomic DNA contamination within the cDNA template samples would result in an inaccurate interpretation of the gene expression levels analysed.

Exploratory screening of various annealing temperatures for each primer set used is a normal optimization step with RT-PCR reactions (Bustin and Nolan, 2004). This step was used to ensure there are no reaction artifacts and no non-specific signal amplification which could interfere with the signals produced.

The maximal PCR reaction efficiency (100%) is where every PCR product is replicated each cycle (Pfaffl et al., 2002). For each primer set specific to each gene being analyzed in a RT-PCR reaction the reaction efficiency must be validated in a separate experiment prior to the actual analysis of gene expression (Pfaffl et al., 2002). This step is to ensure that the amount of cDNA used in each reaction, and the primer concentrations used, will yield a PCR reaction efficiency of ~100% (ideally between 90-110%) (Life Technologies Corporation, 2011). Furthermore, the slope of the linear regression line generated must be between -3.60 and -3.10 (Life Technologies Corporation, 2011). Efficiency curves are generated using a serial dilution of the cDNA template and they are used to determine if the starting material being used in the RT-PCR reaction falls within the linear dynamic range. The linear dynamic range is the range of starting template amount that will allow for the detection of a range of transcripts within the sample. After generating a standard curve with the appropriate efficiency and slope, the cDNA starting material used in all subsequent reactions must fall within the linear dynamic range (Quellhorst and Rulli, 2008). Failure to complete this preliminary experiment can lead to RT-PCR quantitative results that may be interpreted incorrectly. For all the RT-PCR reactions used in the current study, efficiencies were calculated with the correct slopes and the RT-PCR starting material fell within the linear dynamic range. Therefore, the RT-PCR results can be considered reliable and accurate.

5.3.2 Effect of point mutation at KH domains G-X-X-G motif on the ability of CRD-BP to interact with mRNAs in cells

To determine the role of the G-X-X-G motif at each of the KH domains of CRD-BP, immuno-precipitated RNA was analysed using RT-PCR. The results show that mutating the first glycine of the G-X-X-G motif to an aspartate in the different KH domains has an effect on CRD-BP's ability to bind mRNA in cells. The previous *in-vitro* work done in Dr. Lee's lab showed that all the single KH mutants (KH2, KH3, and KH4) with the exception of KH1, had comparable affinities for c-myc CRD RNA substrate as did WT-CRD-BP (Table 10). KH1 mutant was found to have modest decrease in affinity for c-myc CRD RNA as compared to the WT CRD-BP (Table 10). This in contrast to the effect seen in cells; KH1, KH2, KH3 mutants had reduced physical association while KH4 had comparable association with c-myc mRNA (Figure 28, F). Such discrepancy can be explained by discussing the differences between protein-RNA interactions *in-vitro* versus in cells. When performing a protein binding experiment *in-vitro* there are limited factors that are involved in the direct orchestration of the protein interacting with the RNA substrate. The *in-vitro* EMSA binding reaction contains essentially two molecules involved in the binding reaction; the protein and the RNA substrate. The reactions are carried to ensure protein-RNA interactions will occur. However, within the cellular environment, the protein-RNA interactions are highly complex; therefore, even a small change (i.e. single mutation in the RNA-binding protein) may yield dramatic results.

The CRD-BP ortholog, IGF2BP1 (IMP1) has been shown to cooperate with other proteins within the cell to stabilize c-myc mRNA in a CRD-dependent fashion (Weidensdorfer et al., 2009). HNRNPU, SYNCRIP, YBX1, and DHX9 proteins, along with IMP1, were shown to form a RNP complex within the cell which triggered the RNA-dependent co-localization of the c-myc RNA/IMP1 complex (Weidensdorfer et al., 2009). This six protein complex binds the c-

myc CRD and protects c-myc RNA from degradation (Weidensdorfer et al., 2009). More specifically, as a component of the complex, IMP1 is considered to be responsible for directly interacting with c-myc CRD and protecting it from endonucleolytic cleavage (Weidensdorfer et al., 2009). Previous proteomic techniques have identified numerous proteins that assemble with IMP1 to form granule-like complexes which function to protect c-myc mRNA (Weidensdorfer et al., 2009). Knockdown analysis was used to confirm that the IMP1-RNP, assisted by co-factors (HNRNPU, SYNCRIP, YBX1, and DHX9), forms a RNP complex which stabilizes c-myc mRNA (Weidensdorfer et al., 2009). Therefore, it has been shown that within cells, RNA-binding proteins function with the assistance of co-factors. This may explain why a single mutation within CRD-BP does not play a large role in binding *in-vitro*, but in cells the effect of the mutation is amplified due to the various protein-protein and protein-RNA interactions.

It is possible that CRD-BP interacts with its RNA substrates in cells in a complex arrangement so that only RNA binding surfaces exposed to the RNAs is the G-X-X-G motifs; therefore, any mutations would cause a change in the affinity CRD-BP has for its RNAs. This would not be a problem *in-vitro* because the only protein that is present in the EMSA reactions is CRD-BP, therefore, multiple RNA binding surfaces are available for the RNA to interact with, and so a single mutation within one KH domain may not have an effect on binding. It should be noted that RNAsin is present in the reaction mixture, but it is assumed not to interact with CRD-BP in any way that may affect that way it binds its RNAs.

The RNA substrate used in the *in-vitro* experiments is a highly truncated version of the c-myc transcript that exists within cells. This could also explain the difference between the CRD-BP binding *in-vitro* versus *in-vivo*. The EMSA reaction contains a truncated version of the c-myc RNA substrate; therefore it is a smaller molecule that may have the ability to exploit other

binding clefts that are too small for the full length transcript to take advantage of. KH domains have a known binding cleft that is 10Å and therefore it favors the binding of pyrimidines (Valverde et al., 2008). The 181nt c-myc CRD RNA substrate used in the EMSA reactions is nearly 50% pyrimidines. It is possible that the substrate is exploiting regions of the proteins that are not affected by the point mutation introduced. This mechanism would not be possible *in-vivo* due to the full length RNA transcript and the possible other proteins involved in binding. A further discussion on the effect of the mutated KH domains will be addressed in the general discussion.

5.3.3 Effect of specific oligonucleotides and small molecule antibiotics on CRD-BP-RNA interaction in cells

Previous data generated in Dr. Lee's lab showed that some antibiotics were able to break the CRD-BP-RNA interaction *in-vitro*. Neomycin was able to break the CRD-BP-RNA interaction with a Ki of 3.44 uM; whereas, lincomycin had no effect on the CRD-BP-RNA interaction based on the Ki of >1000 uM. These results were generated using an EMSA that had the antibiotics added to the reaction mixture prior to loading the reactions on a native PAGE gel. It comes as no surprise that neomycin would inhibit CRD-BP from binding its RNA substrate *in-vitro*. Aminoglycosides have been used to break protein-RNA interactions because of the positive charge they carry at physiological pH (Zapp et al., 1993).

Neomycin is an aminoglycoside with a molecular weight of 614.644 g mol^{-1} (Barbieri et al., 2007). It contains six amino groups (Figure 31) with pK_a values that range from 5.7 to 8.8. At physiological pH, neomycin carries a predominantly positive charge which interacts favorably with the negative charge of RNA. Therefore, when added to an EMSA reaction (which is carried out at physiological pH), the basic chemical charges possessed by both the RNA and neomycin

promote the physical interaction. This in turn would prevent CRD-BP from binding its c-myc CRD RNA solely based on the fact that the neomycin prevents the physical interaction from occurring. Lincomycin contains a single amine group (Figure 31) with a pK_a of 7.5 (Herr and Slomp., 1967). At physiological pH, lincomycin possesses a small positive charge which leads to weak interactions with nucleic acids; therefore, it is an ideal negative control. This is the rational that lead to using antibiotics as protein-RNA interaction inhibitors and can explain why it was able to break the CRD-BP-RNA interaction *in-vitro*.

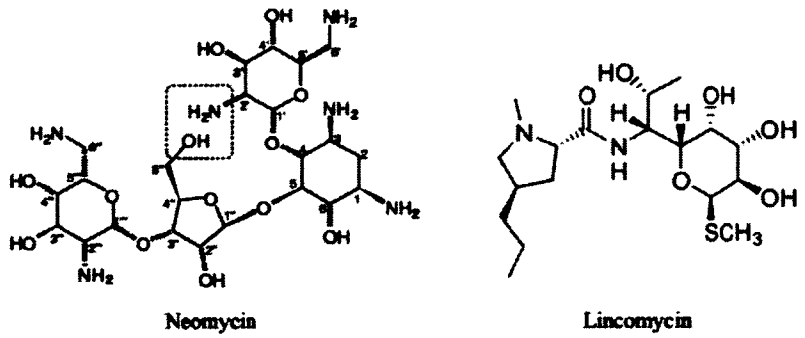

Figure 31. Structures of antibiotics used in protein binding experiments (Barbieri et al., 2007).

To assess if antibiotics have an effect of CRD-BP's ability to bind mRNA *in-vivo*, HeLa cell experiments were performed. Neomycin was used based on its ability to break the CRD-BP-RNA interaction *in-vitro*; lincomycin was used to serve as a negative control based on its lack of ability to break the CRD-BP-RNA interaction *in-vitro*. The results from the cell experiments showed that both antibiotics did not interfere with the CRD-BP-RNA interaction. The fact that the antibiotics had no effect on the CRD-BP-RNA interaction is likely due to the non-specific interactions that out-weighed any possible specific interactions. Once in the cells, the antibiotics

could have interacted with any molecule with a negative charge. The specific CRD-BP and RNA (actin, c-myc and CD44) interactions are simply too specific for the antibiotics to affect.

The addition of the antibiotics, however, did affect the CRD-BP protein expression. Both neomycin and lincomycin are clinically used to inhibit the growth of bacteria. The mechanism of action of both antibiotics is to inhibit protein translation. Neomycin irreversibly binds to 30S ribosome and inhibiting initiation, slowing translation (Mayer, 2013). Lincomycin binds to the 50s ribosome inhibiting the activity peptidyl tansferase which slows translation (Mayer, 2013). Therefore, it is not surprising to see the reduced protein expression when cells were treated with these antibiotics.

Previous data in Dr. Lee's lab have shown that specific 2'-OCH_3 oligonucleotides designed to target the 3'UTR of hCD44 mRNA were able to break the CRD-BP interaction *in-vitro*. DD1 oligonucleotide was unable to prevent the CRD-BP-RNA complex from forming *in-vitro*. DD7 however, was able to disrupt the formation of the CRD-BP-RNA interaction *in-vitro*.

To further investigate if specific oligonucleotides could break the CRD-BP-RNA interaction in cells, immuno-precipitation experiments were performed. The results showed that the DD1 and DD7 oligonucleotides did not exhibit the ability to disrupt the CRD-BP-CD44 and c-myc RNA interactions in cells. The ability of the DD oligonucleotides to break the CRD-BP-RNA interaction *in-vitro* is greatly increased as compared to that in cells. It is possible that *in-vitro* there is limited restraints that prevent the oligonucleotide from interacting with its specific sequence along the mRNA strand. In the EMSA binding reaction mixture, the physical and chemical barriers that exist within the cell are completely absent. This enables the

oligonucleotide to interact with its specific RNA and prevent RNA-binding proteins from binding.

RNA is known to have extensive secondary structures based on their primary sequence and the various interactions it takes part within the cell. The lack of an effect exhibited by the DD oligonucleotides within the cell may have been due to RNA secondary and tertiary structural barriers that were absent in the truncated versions of the RNA substrates used in the *in-vitro* binding experiments. This could explain the discrepancy between the *in-vitro* and *in-vivo* binding experiments results.

Chapter 6 - General Discussion

6.1 General Overview

The coding region determinant-binding protein (CRD-BP) is a member of the highly conserved family of VICKZ RNA-binding protein. CRD-BP has been shown to regulate the expression of multiple genes at the translational, mRNA localization and mRNA stability level (Git and Standart, 2002; Oberman et al., 2002; Farina et al, 2003; Chao et al., 2010). The expression of CRD-BP is present in early development, but then it declines rapidly (Doyle et al., 1998). CRD-BP, which is absent in normal adult tissues, is over-expressed in many adult human cancers. Accumulative evidence suggests that the oncogenic role of CRD-BP is through its physical association with mRNAs for genes implicated in cancers (Chao et al., 2010; Doyle et al., 1998; Farina et al., 2003; Sparanese and Lee., 2007; Vikesaa et al., 2006). Through truncation and deletion studies, it was shown that the KH domains, not the RRM domains, of CRD-BP are responsible for the physical interaction between CRD-BP and its various target RNAs (Git and Standart, 2002; Oberman et al., 2002; Farina et al, 2003; Chao et al., 2010). The CRD-BP KH domains contain a highly conserved Gly-rich linker (G-X-X-G motif) which has been shown to contribute to the binding of CRD-BP to its RNAs (Chao et al., 2010). The specific role of the amino acids within the G-X-X-G motif, mainly the first glycine (G-X-X-G), has not been determined. To date, there have been no studies employing site-directed mutagenesis to determine the specific role of the first glycine in the G-X-X-G motif within the KH domains of CRD-BP. One of the goals of this investigation was to analyze the specific role of the first glycine in the G-X-X-G motif in each of the KH domains of CRD-BP in binding its RNA substrates.

The first step in the investigation was to incorporate point mutations within CRD-BP using site-directed mutagenesis. Once the CRD-BP variants were generated, their ability to bind specific RNAs was analyzed both *in-vitro* and in cells. Following the analysis of CRD-BP variant's binding abilities, experiments were performed to assess specific oligonucleotides and antibiotics ability to break the CRD-BP-RNA interaction in cells, both of which that have shown the ability to break the CRD-BP-CD44 RNA interaction *in-vitro*.

6.2 Generation and Purification of the Coding Region Determinant-Binding Proteins

This goal of this step of the investigation was to generate CRD-BP mutant variants to be used in binding experiments both *in-vitro* and in cells. Mutations were incorporated into pet28b and pcDNA3 plasmids to be used in the *in-vitro* and *in-vivo* binding experiments, respectively. The pet28b plasmid was used to generate recombinant CRD-BP proteins to be used in the *in-vitro* binding experiments. The pcDNA3 plasmids is a eukaryotic expression vector, therefore, it was used in the binding experiment in cells. Site-directed mutagenesis was used to incorporate four mutations (Y5A, Q85A, E446D and D527E, Figure 5) into regions of CRD-BP not expected to be involved in binding; thus, they were meant to serve as negative controls. To determine the role of the first glycine of the G-X-X-G motifs, the first glycine of each motif was mutated to an aspartate (G214D, G284D, G422D and G506D, Figure 5). Single and double KH mutants were generated to determine if a single KH mutation alone had an effect on CRD-BP binding or if a double KH mutation at two KH domains are required to disrupt binding.

We mutated glycine to an aspartate because we hypothesized that aspartate can occlude binding through steric hindrance and electrostatic repulsion. Furthermore, such mutation in the KH domain of NusA has been shown to successfully interfere with the NusA-RNA interaction

(Zhou et al., 2002). The G-X-X-G motif is believed to aid in RNA binding due to its dynamic character. The R-group of glycine is a single proton (H^+), therefore, this linker region is flexible due to the free range of motion available to the peptide. Also, at physiological pH, the R-group of glycine is neutral. Therefore, the dynamic character of the linker and the favorable charge of glycine are thought to facilitate the flexibility of the region, thus enable the binding to RNAs. As described in Chapter 2, the generation of recombinant CRD-BP mutants and mutated pcDNA3 plasmid was successful.

6.3 Development of a fluorescent based method to study the CRD-BP-RNA interaction

Prior to investigating the CRD-BP mutant's ability to bind RNA *in-vitro* and in cells, a fluorescent-based method was developed to study the WT-CRD-BP-RNA interaction *in-vitro*. Fluorescence polarization methods have been previously used to study the protein-RNA interaction (Mao 2006; Meisner et al., 2007). After failed attempts at generating the RNA substrate using RNA ligation and 3'-fluorescein labeling methods, *in-vitro* transcription using fluorescein-labeled UTP was performed. Once the labeled RNA was generated, electrophoretic mobility shift assays was done to analyze the ability of WT-CRD-BP to bind the 181 nt c-myc CRD RNA substrate. Qualitative analysis of the binding profile revealed that WT-CRD-BP was able to bind the 181 nt c-myc CRD RNA substrate, but the CRD-BP-RNA complex was partially immobilized in the wells of the 8% native PAGE gel. This result prompted the switch to using radioactive EMSAs to further investigate the CRD-BP-RNA interaction *in-vitro*. However, because the initial rationale for choosing this method was to establish a system that could be translated to the fluorescent polarization method in the future, the qualitative result that yielded was deemed successful (i.e. WT-CRD-BP was able to bind the 181 nt c-myc CRD RNA fluorescently-labeled substrate).

6.4 Assessing the ability of CRD-BP mutant variants to bind c-myc CRD RNA *in-vitro*

The goal of this part of the investigation was to determine if mutations within CRD-BP affected the ability to bind the [^{32}P]-labeled c-myc CRD (nts 1705-1886) RNA substrate using EMSA. Confirming the WT-CRD-BP-c-myc CRD RNA interaction as previously shown was the initial step (Sparanese and Lee., 2007). Using EMSA, this binding study indicated that purified recombinant WT-CRD-BP was able to bind the c-myc CRD (nts 1705-1886) RNA substrate with a K_d of ~ 500 nM (Sparanese and Lee., 2007). Because the binding profiles of CRD-BP mutant proteins were going to be compared to that of the WT-CRD-BP, it was key to establish an accurate positive control; this step was completed and it was determined that WT-CRD-BP was able to bind the [^{32}P]-labeled c-myc CRD (nts 1705-1886) RNA substrate with a K_d of 398 ± 52.79 nM. This result was considered to be a suitable positive control based on comparable results from the previous CRD-BP binding study (Sparanese and Lee., 2007).

To investigate the role of the first glycine in the G-X-X-G motif of each CRD-BP KH domains, the glycine was mutated to an aspartate as previously mentioned. Previous work in Dr. Lee's lab has shown that single mutations in each of the four KH domains [G284D (KH2), G422D (KH3), G506D (KH4)] did not significantly disrupt the CRD-BP-c-myc CRD RNA interaction *in-vitro*; with the exception of the G214D (KH1) mutation which exhibited modest reduction in binding with a K_d of 723.07 ± 75.13 nM (van Rensburg, unpublished data). These results indicate that the first glycine of the G-X-X-G of KH2, KH3 and KH4 domains do not seem to play a significant role in binding the c-myc CRD (nts 1705-1886) RNA substrate *in-vitro*. However, due to the modest reduction in the binding ability of the KH1 single mutation, it can be suggested that the first glycine in the G-X-X-G motif of the KH1 domain may participate in binding the c-myc CRD (nts 1705-1886) RNA substrate *in-vitro*. It was also revealed that the

negative controls (E446D, D527E, Y5A, and Q85A) did not disrupt the CRD-BP-c-myc RNA interaction *in-vitro* (Figure 21, Chapter 4).

To further investigate the role of the first glycine of the G-X-X-G motifs of CRD-BP KH domains, KH double mutants were generated and their binding ability was tested using EMSA. It has been previously shown that CRD-BP orthologs KH domains work in tandem to bind their RNA substrates (Chao et al., 2010). Using x-ray crystallography, it was revealed that KH3 and KH4 form a pseudodimer which orients the G-X-X-G motifs of each domain in an antiparallel configuration to form the RNA binding surface (KH3 and KH4 RNA binding surfaces are red and blue, respectively. Figure 32) (Chao et al., 2010).

Figure 32. IMP1 KH34 pseudodimer configuration of reveals β-actin RNA-binding surfaces. Upon dimer formation, the KH domains adopt an antiparallel configuration that requires the RNA to undergo a ~180° bend in order to contact both RNA-binding surfaces; indicating that both G-X-X-G motifs of KH3 and KH4 are required for β-actin RNA binding (Chao et al., 2010).

Because it has been shown that the KH domains of CRD-BP orthologs work in tandem to bind its RNA substrates, it is reasonable to think that CRD-BP adopts the same binding mechanism when binding the c-myc CRD (nts 1705-1886) RNA substrate. Therefore, a single mutation within only one of the KH domains G-X-X-G motifs may not result in reduced binding

because the RNA is still in contact with the other KH domain of the dimer. This hypothesis was tested by generating the mutations in multiple KH domain G-X-X-G motifs simultaneously.

The results from KH double mutants' studies have revealed important insights to how CRD-BP KH domains interact with the c-myc CRD (nts 1705-1886) RNA substrate *in-vitro*. The KH1-2 (G214D-G284D) double mutant showed a complete loss in the ability to bind the c-myc CRD RNA substrate (Figure 21,D). This may be explained by KH1 and KH2 RNA-binding surfaces relative position to each other and if in fact the KH domains of VICKZ proteins work in tandem as previously described (Chao et al., 2010). It has been shown that the KH1 and KH2 domains of fragile X mental retardation protein (FMRP) and the KH1 and KH2 domains poly-C-binding protein 2 (PCBP1) function in tandem (Chao et al., 2010). Therefore, the result of the CRD-BP KH1-2 double mutant lacking the ability to bind the c-myc CRD RNA substrate can be explained by the mechanism it employs when binding. The KH1 and KH2 domains of CRD-BP most likely adopt a similar overall structure to that of IMP1 KH34 facilitating the interaction of the RNA substrate with both the G-X-X-G motif in KH1 and the G-X-X-G motif of KH2 (Chao et al., 2010). This explains why single KH mutants are not affected by the G>D mutation compared to multiple G>D mutations. The double mutation occludes binding by disrupting the protein-RNA interaction at multiple sites within CRD-BP, thus completely preventing binding.

Due to the many truncation studies performed on CRD-BP orthologs that focus on the function of single KH domains and mainly the function of the KH34 didomain, it may be more difficult to explain the result of the KH1-3 and KH2-4 double mutants (Git and Standart, 2002; Oberma et al., 2002; Farina et al., 2003; Chao et al., 2010). Models of the KH3-4 RNA complex showed that the 5' and 3' ends of the RNA substrates are at a distance of ~35 Å apart which can be spanned by as few as 5-6 nucleotides, but due to the energetically unfavorable conformation

of the RNA (strain on the RNA molecule), it is more likely that the RNA spans a farther distance than that of β-actin RNA (Chao et al., 2010). It is possible that the c-myc CRD RNA substrate is able to adopt many protein-RNA conformations *in-vitro* due to the limited constraints on the system (i.e. only protein and RNA are present in the EMSA reactions). Because the c-myc CRD RNA substrate is 181 nucleotides in length, it is possible that it contacts any combination of the KH domain binding surfaces (KH1-2, KH1-3, KH1-4, KH2-3, KH2-4...). This could explain why the KH1-3 and KH2-4 double mutants showed complete abolishment in the binding ability to the c-myc CRD RNA substrate. The antiparallel confirmation of the KH domains would support this theory. The KH3-4 didomain adopts the antiparallel orientation that places the RNA binding surfaces in a position that favors the 180° change in direction of the RNA substrate and it can be assumed that the KH1-3 and KH2-4 domains adopt a similar configuration (Chao et al., 2010). This would allow the c-myc CRD RNA substrate to interact with RNA binding surfaces that are antiparallel to each other. Further mutagenesis studies must be done to test this hypothesis. These studies would include mutating the G-X-X-G motifs in KH domains that have combinations of mutations that are spatially both antiparallel and parallel from each other.

Interestingly, the KH3-4 double mutant only showed a reduction, and not a complete abrogation, in the ability to bind the c-myc CRD RNA substrate *in-vitro*. The functional importance of both KH3 and KH4 domains in VICKZ RNA-binding proteins has been well documented (Git and Standart, 2002; Oberman et al., 2002; Farina et al., 2003; Chao et al., 2010). It was shown that KH3 and KH4 domains are required for Vg1RBP to bind both VLE and TGFβ RNAs *in-vitro* (Oberman et al., 2002). KH3 and KH4 domains of ZBP1 were shown to be responsible for the binding of β-actin mRNA *in-vitro* (Farina et al., 2003). The crystal structure of IMP1 KH3-4 didomain revealed the necessary RNA-binding surfaces for β-actin zipcode

mRNA (Chao et al., 2010). KH3 and KH4 form a didomain that facilitates the binding of its RNAs by positioning the RNA-binding surfaces in an energetically favorable orientation (Chao et al., 2010). Furthermore, both KH3 and KH4 domains possess the typical Type 1 eukaryotic $\beta_1\alpha_1\alpha_2\beta_2\beta_3\alpha_3$ topology with the G-X-X-G flexible loop connecting $\alpha_1\alpha_2$ and the variable loop that connects $\beta_2\beta_3$ (Chao, J.A., et al., 2010; Valverde, R., et al, 2008). The center of the groove between $\alpha_1\alpha_2$ is hydrophobic which is lined by hydrophilic, charged amino acid residues; furthermore, the pocket formed is 10Å in size which favors the binding of pyrimidines (Chao et al., 2010; Valverde et al, 2008). Therefore, RNA interacts with multiple regions within KH3 and KH4 domains. This explains why the mutation of G>D in KH3 and KH4 does not completely occlude binding. Although glycine seems to play some role in the interaction between CRD-BP and c-myc CRD RNA (based on the reduced binding observed), it does not seem to be completely responsible for CRD-BP-c-myc CRD RNA interaction. Due to the high level of conservation in the structure and function of CRD-BP KH3 and KH4 domains, the lack in the complete loss of binding when mutating the first glycine of the G-X-X-G motif of KH3 and KH4 is not surprising. We propose that interactions other than the c-myc RNA-(G-X-X-G) motif interactions are taking place to facilitate the binding to the KH3 and KH4 domains. Thus, the first glycine of KH3 and KH4 G-X-X-G motifs are not as important when binding the c-myc CRD RNA substrate.

6.5 Characterization of CRD-BP-RNA Interaction in Cells

The goal of this part of the investigation was to determine if mutations within CRD-BP affected the way that CRD-BP interacted with targets mRNA transcripts within HeLa cells. Using *in-vitro* electrophoretic mobility shift assays, it was previously shown that mutating the first glycine of the G-X-X-G motif to an aspartate in each of the KH domains individually, did

not disrupt the CRD-BP-c-myc CRD RNA interaction *in-vitro* (van Rensburg, unpublished data). To test the effect of the mutations on CRD-BP binding in cells, immuno-precipitation coupled with qRT-PCR was performed.

The negative controls generated (Y5A and Q85A) did not disrupt the CRD-BP-RNA interaction with β-actin, c-myc or CD44 mRNAs in cells (Figure 28, C,E,G). However, the KH domain single mutants [G214D (KH1), G284D (KH2), G422D (KH3), and G506D (KH4)] exhibited different preferences when it came to binding mRNAs in cells. KH1, KH2 and KH4 single mutants did not significantly disrupt the CRD-BP-β-actin mRNA interaction (Figure 28, D). However, KH3 single mutant show a significant reduction in the ability to bind β-actin mRNA in cells (Figure 28, D). It has been shown that KH3 and KH4 domains are required to bind β-actin mRNAs in cells (Farina et al., 2003). Truncations of ZBP1 were found to preferentially bind β-actin zipcode mRNA, with truncations that contains KH3 and KH4 showing the ability to bind the β-actin zipcode mRNA; whereas truncations that contained only the RRM domains or only KH1 and KH2 domains were unable to bind the β-actin zipcode mRNA (Farina et al., 2003). Therefore, the first glycine of the G-X-X-G motif in KH3 seems to play an important role in binding β-actin mRNA in cells. The fact that no significant reduction in the CRD-BP-KH4 mutant binding to β-actin mRNA, suggests that other regions of KH4 are responsible for binding with β-actin mRNA in cells. Furthermore, KH1 and KH2 mutants showed no effect on the binding of β-actin mRNA, which agrees with data on ZBP1 that indicated KH1 and KH2 are not involved in binding β-actin mRNA in cells (Farina et al., 2003). Therefore, introducing the mutations in KH1 and KH2 G-X-X-G motif did not cause a change in CRD-BP's ability to bind β-actin mRNA because these domains are not required for binding β-actin mRNA.

There was a significant drop in CRD-BP's ability to bind c-myc mRNA in KH1, KH2 and KH3 single mutants, however, KH4 single mutant did not have an effect on c-myc binding in cells (Figure 28, F). To date, there have not been deletion or truncation studies that analyzed the specific KH domains that are responsible for binding c-myc mRNA in cells. The results indicate that the G-X-X-G motifs, specifically the first glycine in each KH1, KH2 and KH3 domains play a role in binding c-myc mRNA in cells, whereas the first glycine of the G-X-X-G motif in KH4 does not seem to play a role in binding c-myc mRNA in cells. The importance of KH3 and KH4 domains of VICKZ proteins in binding RNAs is well established. There is, however, evidence that implicates Vg1RBP KH1 and KH2 domains as playing a role in the binding of RNAs (Git and Standart, 2002). Truncation studies have shown that Vg1RBP KH1 and KH2 domains are able to bind 3' UTR of the vegetal localization element, although affinity is reduced compared to KH3 and KH4 domain binding (Git and Standart, 2002). It is possible that CRD-BP's KH1 and KH2 domains are involved in binding c-myc mRNA in cells, thus introducing the mutations in the first glycine of the G-X-X-G motifs of KH1 and KH2 domains reduces CRD-BP's ability to bind c-myc mRNA in cells. KH4 single mutant did not show any effect on the c-myc mRNA binding ability of CRD-BP which agrees with previous *in-vitro* data (van Rensburg, unpublished data). There is, however, a significant difference in KH1, KH2 and KH3 mutants' ability to bind c-myc mRNA *in-vitro* versus in cells. This will be discussed below in section 6.6.

Finally, the ability for CRD-BP to bind CD44 mRNA in cells was significantly affected by the single mutation in the G-X-X-G motif of all four KH domains (Figure 28, H). To date, truncation studies to determine the specific KH domains of CRD-BP responsible for binding CD44 mRNA in cells has not been done. However, it is known that the CRD-BP ortholog, IMP1, binds to the 3' UTR of CD44 mRNA (Vikesaa et al., 2006). It is also known that ZBP1 relies on

both KH3 and KH4 domains to bind the zipcode region located in the 3' UTR of β-actin mRNA (Chao et al., 2010). It can be suggested that CRD-BP interacts with the 3' UTR region of CD44 mRNA in a similar, but not identical way. Mutations in G-X-X-G motif mutations in KH3 affected CRD-BP binding to both β-actin and CD44 mRNAs; suggesting the first glycine of the G-X-X-G motif in KH3 is important for CRD-BP binding to both β-actin and CD44 mRNAs in cells. Furthermore, the first glycine in the G-X-X-G motif of KH1, KH2 and KH4 are important for the CRD-BP-CD44 mRNA interaction in cells.

It is evident that CRD-BP interacts with its different target mRNAs quite differently in cells. A likely explanation for this observation is differences in the species of RNAs that are present in the cell. CRD-BP has been shown to bind *cis*-regulatory sequences, which some RNAs share in common. However, in the cell, the RNA species are full length transcripts with extensive secondary structures (Li et al., 2010). Therefore, differences in the secondary structures possessed by the RNAs can cause differences in the way the RNA interact with the protein. Because CRD-BP contains KH domains that have a 10Å binding cleft that favors pyrimidines over purines, the primary structure that dictates the secondary structure, may bind one RNA species more preferentially than another (Valverde et al., 2008). This may shed some light on the discrepancy between how CRD-BP binds β-actin, c-myc, and CD44 mRNAs in cells. Furthermore, the sizes of the mRNA transcripts within the cell differ significantly. CD44 mRNA has been shown to be 5.0 kb in size; c-myc mRNA has been shown to be 2.7 kb in size; and β-actin mRNA has been shown to be 2.8 kb in size (Ray et al, 2011; Joseph et al, 2012; Vikesaa et al., 2006). Therefore, the different lengths of the mRNAs within the cell may result in interactions with different regions of CRD-BP experienced by each mRNA transcript. This is not

a concern when looking at binding *in-vitro* because RNA substrates are truncated; therefore, the secondary structures are less complex.

6.6 Explaining the discrepancy between the CRD-BP-RNA interaction *in-vitro* and in cells

This section discusses the possible reasons that explain why there is a difference in the way the CRD-BP mutants interact with its RNAs *in-vitro* versus in cells. The complexity of a binding event *in-vitro* versus in cells is quite different. One might even describe a binding event *in-vitro* not complex at all, which is why *in-vitro* work is used when attempting to understanding complex mechanisms that occur in cells. *In-vitro* experiments are highly controlled and they usually contain only the essential components that allow for an interaction to be observed (i.e. protein and RNA to observe binding). *In-vitro* binding experiments were used to analyze the CRD-BP-RNA interaction. The results differed from the binding experiments performed in cells. Mutating the first glycine of the G-X-X-G motifs in CRD-BP KH2, KH3 and KH4 domains showed no significant effect on its ability to bind the c-myc CRD RNA (nts 1705-1886) substrate *in-vitro*. The single mutation in KH1 did show a modest reduction in binding to the c-myc CRD RNA (nts 1705-1886) substrate *in-vitro*. However, in cells, there was a significant reduction in the binding of CRD-BP KH1, KH2 and KH3 mutants to c-myc mRNA. Furthermore, the reduction in KH1 mutant binding in cells was more significant than that *in-vitro*. The explanation for the more significantly reduced binding in cells is due to the mechanism that CRD-BP employs when binding its RNAs in cells.

IMP1 has been shown to cooperatively and sequentially form dimers on RNA substrates leading to cytoplasmic RNP formation; $R + 2P \leftrightarrow RP_1 + P \leftrightarrow RP_2$ (Nielsen et al., 2004). The assembly of the RNP stabilizes the interaction between IMP1 and its RNA substrates and allows

for cellular localization. IMP1 has been shown to bind to the c-myc CRD and stabilize the mRNA and protect it from endonucleolytic attack by forming cytoplasmic mRNPs (Weidensdorfer et al., 2009; Sparanese and Lee., 2007; Jonson et al., 2007). Protein factors have been identified to associate directly with the IMP1, further stabilizing c-myc mRNA (Weidensdorfer et al., 2009). Using immuno-precipitation, FLAG-tagged IMP1-complexes were isolated from RNPs in U2OS cell lysate and through MS-analysis, HNRNPU, DHX9, SYNCRIP, YBX1, and ELAVL1 proteins were identified (Weidensdorfer et al., 2009). Furthermore, upon knockdown of HNRNPU, DHX9, SYNCRIP, YBX1, and ELAVL1 proteins, both c-myc protein and mRNA levels significantly decreased; suggesting that the regulation of c-myc protein and mRNA levels is dependent on more than IMP1 binding to the CRD regulatory element of c-myc (Weidensdorfer et al., 2009). To validate how the HNRNPU, DHX9, SYNCRIP, YBX1, and ELAVL1 proteins interacted with the IMP-RNA complex, micrococcal nuclease-treated cytoplasmic extracts of HEK293 cells was performed and it revealed that IMP1, HNRNPU, DHX9, SYNCRIP and YBX1 form a multi-protein complex that is dependent on IMP1 binding the c-myc CRD, which blocks the degradation of c-myc mRNA by cellular machinery (Weidensdorfer et al., 2009). This mechanism is believed to prevent the accelerated decay of c-myc mRNA and any disruption in the binding of IMP1 to the c-myc CRD or any disruption in the protein-protein interactions would lead to a dramatic decrease in the c-myc mRNA levels in cells (Weidensdorfer et al., 2009). Hence, the more significant decrease in the binding of CRD-BP to c-myc RNA observed is likely due to a compounded effect that is caused by the large number of factors involved in binding c-myc mRNA in cells. RNP formation is essential to the stabilization of the mRNAs within the cell; hence, the expression of c-myc may have been

altered. Therefore, any defect in the interaction, even minor, may lead to a dramatic change in the mRNA levels within the cell.

Vigilin, the KH domain RNA-binding protein, also forms a complex with other proteins within the cell. This further supports the theory that the significantly reduced levels of c-myc mRNA physically associated with the KH mutants are due to the associations that are occurring with CRD-BP and other protein factors in the cell (Dodson and Shapiro, 1997). It was shown that vigilin was localized immunocytochemically to polysomes which was identified to be cytoplasmic mRNPs that regulated the mRNA levels of vitellogenin (Dodson and Shapiro, 1997). There is evidence that RNA-binding proteins form a complex with other protein and RNA factors within the cell. The mutations within the G-X-X-G motifs of the CRD-BP KH domains caused a significantly larger decrease in the ability to bind c-myc RNA in cells as compared to that *in-vitro*. This finding is likely due to the necessary associations that CRD-BP has with other proteins in the cell to form RNPs, which greatly affect the stability of c-myc mRNA in cells.

6.7 Assessing the ability of oligonucleotides and small molecule antibiotics to disrupt CRD-BP-RNA interaction in cells.

This section of the investigation discusses the assessment of the ability of specific oligonucleotides and antibiotics to disrupt the CRD-BP-RNA interaction in cells. Using oligonucleotides and antibiotics to break protein-RNA interactions has been previously reported (Meisner et al., 2007; Zapp et al., 1993; Zellweger et al., 2002). Neomycin B was also shown to inhibit the Rev-RRE interaction in cells (Van Ryk and Venkatesan, 1999). Furthermore, low-molecular-weight inhibitors have been used to disrupt HuR dimerization and prevent HuR from binding its mRNA target (Meisner et al., 2007). Due to the ability of neomycin and small-molecular-weight inhibitors to break protein-RNA interactions, the antibiotic neomycin was used

in the attempt to break the CRD-BP-RNA interaction in cells. Furthermore, previous research in Dr. Lee's lab confirmed the ability of neomycin to break the CRD-BP-CD44 RNA interaction *in-vitro.*

Our results showed that the antibiotics did not exhibit the ability to disrupt the CRD-BP-RNA interaction in cells. It is believed that the lack of specificity of the antibiotics may contribute to the lack of effect on the CRD-BP-RNA interaction in cells. The chemical nature of neomycin at physiological pH promotes the interaction with any negatively charged particle within the cell, which explains why the specific CRD-BP-RNA interaction was unaffected by the treatment of the antibiotics. Moreover, the negative control antibiotic, lincomycin, also did not show any effect in disrupting the CRD-BP-RNA interaction in cells. This was expected due to its lack of ability to break the CRD-BP-CD44 mRNA interaction *in-vitro.*

The specific DD oligonucleotides that were designed to target the 3' UTR of CD44 mRNA did not disrupt the CRD-BP-RNA interaction in cells. The complexity of cellular mechanisms and the possible RNA secondary structures may have contributed to the results observed.

6.8 Future Studies

To further confirm the role of the G-X-X-G motifs in the CRD-BP KH domains, further investigations must be made. Firstly, a conformation of the integrity of the secondary and tertiary structure of CRD-BP when the mutations are introduced would confirm that any changes in the binding ability is due to the interaction at the first glycine of the G-X-X-G motif in the KH domains and not due to a loss in the native confirmation of CRD-BP. This investigation could be performed using circular dichroism (CD) spectroscopy. CD spectroscopy is a method that uses

circular polarized light to detect changes in protein structure (Adler et al., 1973). Therefore, any confirmation of the comparable protein secondary structure between CRD-BP mutants and WT-CRD-BP would confirm that the mutation did not drastically change the protein structure. This means that any difference in the binding ability of CRD-BP mutants is a result of the amino acid change introduced.

Furthermore, as mentioned earlier, the experiments with CRD-BP KH double mutants were only analysed *in-vitro*. Therefore, due to the difference in the binding profile observed for the KH single mutants *in-vitro* versus in cells, a complete analysis of the CRD-BP KH double mutants' ability to bind mRNAs in cells must be performed.

6.9 Concluding Remarks

The function of VICKZ proteins has been studied extensively and the structural features within the VICKZ proteins that are responsible for binding target RNAs and stabilizing the transcripts have be determined (Git and Standart, 2002; Oberman et al., 2002; Farina et al, 2003; Chao et al., 2010). Through association with many cellular factors and mRNA transcripts, VICKZ proteins play a significant role in cytoplasmic RNP granule formation which leads to the stabilization and localization of mNRAs within the cell (Vikesaa et al., 2006). These interactions are made through direct contact with VICKZ proteins structural domains, mainly the KH domains, and regulatory sequences within the mRNA transcripts, such as the coding region determinant (CRD) of c-myc mRNA (Nelson and Cox., 2008; Li et al., 2010).

Messenger RNA knockdown experiments and protein truncation studies have been used to determine the role VICKZ proteins play in mRNA regulation and also the structural domains within VICKZ proteins that are responsible for the physical interaction with their target RNAs.

For example, it was shown that KH3 and KH4 domains of Vg1RBP, not the RRM domains, were required for binding the vegetal localization element (VLE) and TGFβ mRNA (Git and Standart, 2000; Oberman et al., 2002). Farina *et al.* (2003) showed that KH3 and KH4 domains of ZBP1 facilitate β-actin mRNA binding which leads to the cytoplasmic granule formation and mRNA localization. Furthermore, ZBP1 KH3 and KH4 domain induces β-actin mRNA looping upon recognition by the zipcode region within the mRNA (Chao et al., 2010). CD44 mRNA has been shown to tightly regulate by the VICKZ ortholog IMP1; which has shown, upon knockdown of IMP1, that CD44 mRNA and protein levels significantly decrease, proving the importance of the IMP-CD44 interaction in cells (Vikesaa et al., 2006). The VICKZ protein CRD-BP has been shown to bind to c-myc mRNA and shield it from endonucleolytic attack (Sparanese and Lee., 2007). Although it is well documented of the role VICKZ proteins play in gene regulation and the structural features they use to facilitate these interactions, the specific amino acids within known binding regions of VICKZ proteins, mainly CRD-BP, has not been determined.

The data presented within this thesis show that the first glycine of the G-X-X-G motif within the CRD-BP KH domains is important for CRD-BP to interact with its target mRNAs and to different degrees depending on which KH domain and the RNA substrate in question. Upon introducing the aspartate mutation, CRD-BP altered binding varied from mRNA to mRNA. It was revealed that when it comes to CRD-BP binding its target RNAs, "one size, truly does not fit all".

Furthermore, in an attempt to disrupt the CRD-BP-RNA interaction in cells using oligonucleotides and antibiotics was unsuccessful. The nonspecific nature of the aminoglycoside antibiotics used proved to be ineffective in disrupting the highly specific CRD-BP-RNA interaction in cells. Also, when using oligonucleotides designed to target the 3' UTR of CD44

mRNA, it was shown to be ineffective at breaking the CRD-BP-RNA interactions in cells. RNA secondary structure complexity within cells was deemed to be responsible for the lack of ability for the oligonucleotides to break the interaction.

Nonetheless, important structural information on how on CRD-BP interacts with its RNAs was determined. These studies are critical for potential future investigations into the development of new classes of anti-cancer drugs that act by breaking protein-RNA interactions.

References

Adler et al. (1973). Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods in Enzymology*, 675-735.

Alberts et al. (2008). Molecular Biology of The Cell, 5th ed.

Barbieri et al. (2007). Defining the Molecular Forces That Determine the Impact of Neomycin on Bacterial Protein Synthesis: Importance of the 2'-Amino Functionality. *Antimicrob Agents Chemotherapy*, 51(5): 1760-1769.

Barnes et al. (2009). Identification of Apurinic/apyrimidinic endonuclease 1 (APE1) as the endoribonuclease that cleaves c-myc mRNA. *Nucleic Acids Res.*, 37(12):3946-58.

Benoit et al. (2010). The X-ray Crystal Structure of the First RNA Recognition Motif and Site-Directed Mutagenesis Suggest a Possible HuR Redox Sensing Mechanism. *Journal of Molecular Biology*, 1231-1244.

Bustin and Nolan. (2004). Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction. *Journal of Biomolecular Techniques*, 15(3): 155-167.

Chao et al. (2010). ZBP1 recognition of b-actin zipcode induces RNA looping. *Genes Dev.*, 24: 148-158.

Christiansen et al. (2009). IGF2 mRNA-binding protein 2: biological function and putative role in type 2 diabetes. *Journal of Molecular Endocrinology*, 43: 187-195.

Diaz-Cano. (2012). Tumor Heterogeneity: Mechanisms and Bases for a Reliable Application of Molecular Marker Design. *Int. J. Mol. Sci.*, 13: 1951-2011.

Dodson and Shapiro. (1997). Vigilin, a Ubiquitous Protein with 14 K Homology Domains, is the estrogen-inducible vitellogenin mRNA 3'untranslated region-binding protein. *Journal of Biological Chemistry*, 12249-12252.

Doyle et al. (1998). The c-myc coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins. *Nuc. Acid. Res.*, 26: 5036-5044.

Dykxhoorn et al. (2006). The silent treatment: siRNAa as small molecule drugs. *Gene Therapy*, 13: 541-552.

Elcheva et al. (2008). Overexpression of mRNA-binding protein CRD-BP in malignant melinomas. *Oncogene*, 27: 5069-5074.

Elcheva et al. (2009). CRD-BP protects the coding region of betaTrCP1 mRNA from miR-183-mediated degradation. *Mol. Cell*, 35: 240-246.

Farina et al. (2003). Two ZBP1 KH domains facilitate b-actin mRNA localization, granule formation, and cytoskeletal attachment. *The Journal of Cell Biology*, 77-87.

Farina et al. (2008). Embedding mRNA Stability in Correlation Analysis of Time-Series Gene Expression Data. *PLoS Comput Biol.*, 4(8): e1000141.

Git and Standart. (2002). The KH Domains of Xenopus Vg1RBP mediate RNA binding and self-association. *RNA*, 8: 1319-1333.

Goodison et al. (1999). CD44 cell adhesion molecules. *J Clin Pathol: Mol Pathol*, 52: 189-196.

Goswami et al. (2010). MicroRNA-340-mediated degradation of microphthalmia-associated transcription factor mRNA is inhibited by the coding region determinant-binding protein. *Journal of Bological Chemistry*, 20532-20540.

Hanahan and Weinberg. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, 646-674.

Herr and Slomp. (1967). Lincomycin. II. Characterization and gross structure. *J. Am. Chem. Soc.,*, 89 (10): 2444–2447.

Hogan et al. (2008). Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System. *PLoS Biol.*, 6(10): e255.

Jonson et al. (2007). Molecular composition of IMP1 ribonucleoprotein granules. *molecular and cellular proteomics*, 798-811.

Levy et al. (1998). Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR. *Journal of Biological Chemistry*, 273: 6417-6423.

Lewis et al. (1999). Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure*, 191-203.

Li et al. (2008). Expression of RNA-binding protein IMP3 (KOC) in benign urothelium and urothelial tumors. *Human Pathology*, 1205-1211.

Li et al. (2010). Endoribonucleases – enzymes gaining spotlight in mRNA metabolism. *FEBS journal*, 627-641.

Li et al. (2010). Predictin in vivo binding sites of RNA-binding proteins using mRNA secondary structure . *RNA*, 16: 1096-1107.

Liao et al. (2005). The RNA-binding protein IMP-3 is a translational activator of insulin-growth factor II leader-3 mRNA during proliferation of human K562 leukemia cells. *The journal of biological chemistry*, 280: 18517-18524.

Life Technologies Corporation . (2011). Real-time PCR: Understanding Ct. *online protocol*, 1-6.

Mao et al. (2006). Analysis of RNA-protein interactions by a microplate-based fluorescence anisotropy assay. *Analytical Biochemistry*, 222-231.

Mayer. (2013, April 15). *Antibiotics- Protein synthesis, nucleic acid synthesis and metabolism*. Retrieved from Microbiology and Immunology On-line: http://pathmicro.med.sc.edu/mayer/antibiot.htm

Meisner et al. (2007). Identification and mechanistic characterization of low-molecular-weight inhibitors of HuR. *Nature Chemical Biology*, 508-515.

Mongroo et al. (2011). IMP-1 displays crosstalk with K-Ras and modulates colon cancer cell survival through the novel pro-apoptotic protein CYFIP2. *Cancer Research*, 71(6): 2172-2182.

Mullen and Marzluff. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degration of the mRNA both 5' to 3' and 3' to 5'. *Genes and Development*, 22(1): 50-65.

Nakayama and Nakayama. (2006). Ubiquitin ligases: cell-cycle control and cancer. *Nature Reviews*, 369-381.

Nelson and Cox. (2008). *Lehninger Principles of Biochemistry, 5th ed.* New York, NY: W.H. Freeman and Company.

Nielsen et al. (1999). A Family of Insulin-Like Growth Factor II mRNA-Binding Proteins Represses Translation in Late Development. *Molecular and Cellular Biology*, 19(2): 1262-1270.

Nielsen et al. (2004). Sequential dimerization of human zipcode-binding protein IMP1 on RNA: a cooperative mechanism providing RNP stability. *Nucleic Acids Research*, 32(14): 4368-4376.

Noubissi et al. (2006). CRD-BP mediates the stabilization of βTrCP1 and c-myc mRNA in response to beta-catenin signalling. *Nature*, 441: 898-901.

Noubissi et al. (2009). Wnt signalling stimulates transcriptional outcome of the Hedgehog pathway by stabilizing GLI1 mRNA. *Cancer Res.*, 69: 8572-8578.

Oberman et al. (2002). VICKZ proteins mediate cell migration via their RNA binding activity. *RNA*, 13: 1558-1569.

Pagano et al. (2011). Quantitative appraoches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA*, 17: 14-20.

Pfaffl et al. (2002). Real-time RT-PCR quantification of insulin-like growth factor (IGF)-1, IGF-1 receptor, IGF-2, IGF-2 receptor, insulin reeceptor, growth hormone receptor, IGF-binding proteins 1, 2 and 3 in the bovine species. *Domestic Animal Endocrinology*, 22: 91-102.

Quellhorst and Rulli. (2008). A Systematic Guideline for Developing the Best Real-Time PCR Primers. *SABiosciences TECHNICAL ARTICLE*, 1-8.

Reines and Cantor. (1974). New fluorescent hydrazide reagnets for the oxidized 3'-terminus of RNA. *Nucleic Acids Research*, 767-786.

Richards. (1972). The 1972 Nobel Prize for Chemistry. *Science*, 178 (4060): 492–493.

Ryder, et al. . (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods Mol Biol.*, 488: 99-115.

Sanford et al. (2008). Identification of Nuclear and Cytoplasmic mRNA Targets for the Shuttling Protein SF2/ASF. *PLoS ONE*, 3(10): e3369.

Sparanese and Lee. (2007). CRD-BP shields c-myc and MDR-1 RNA from endonucleolytic attack by a mammalian endoribonuclease. *Nucleic Acids Research*, 35: 1209-1221.

Stark et al. (2006). An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs. *RNA*, 12:2014-2019.

Sun et al. (1997). Synthesis of Fluorinated Fluoresceins. *Journal of Organic Chemistry*, 62(19): 6469-6475.

Tessier et al. (2004). Mammary tumor induction in transgenis mice expressing an RNA-binding protein. *Cancer Research*, 64: 209-214.

Valverde et al. (2008). Structure and Function of KH domains. *FEBS*, 2712-2726.

van Rensburg, G. (unpublished data). Investigating the Roles of the Four KH domains of CRD-BP and their interaction with mRNAs. *UNBC Undergraduate Thesis*, unpublished data.

Van Ryk and Venkatesan. (1999). Real-time Kinetics of HIV-1 Rev-Rev Response Element Interactions: Definition of minimal binding site on RNA and protein and stoichiometric analysis. *Journal of Biological Chemistry*, 274: 17452–17463.

Vikesaa et al. (2006). RNA-binding IMPs promote cell adhesion and invadopodia formation. *EMBO*, 25: 1456-1468.

Weidensdorfer et al. (2009). Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA*, 15: 104-115.

Wetlaufer and Xie. (1995). Control of aggregation in protein refolding: A variety of surfactants promote renaturation of carbonic anhydrase I1. *Protein Science*, 4:1535-1543.

Whelan et al. (2012). Post-transcriptional regulation of the Ras-ERK/MAPK signaling pathway. *Journal of Cellualr Physiology*, 1235-1241.

Yisraeli. (2005). VICKZ proteins: a multi-talented family of regulatory RNA-binding proteins. *Biol. Cell*, 97: 87-96.

Zapp et al. (1993). Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell*, 74: 969-978.

Zellweger, T., et al. (2002). Enhanced Radiation Sensitivity in Prostate Cancer by Inhibition of the Cell Survival Protein Clusterin1. *Clinical Cancer Research*, 3276-3284.

Zhou et al. (2002). Evidence that the KH RNA-binding Domains Influence the Action of the E. coli NusA Protein. *J. Mol. Biol.*, 318: 1175-1188.

Zhou et al. (2009). Insulin-like growth factor II mRNA binding protein 1 modulates Rev-dependent human immunodeficieny virus type 1 RNA expression. *Virology*, 210-220.