MOLECULAR CHARACTERIZATION OF THE CODING REGION DETERMINANT-BINDING PROTEIN INTERACTION WITH KRAS ONCOGENE MRNA

by

Sebastian J. Mackedenski

B.Sc., University of Northern British Columbia, 2012

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
BIOCHEMISTRY

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

July 2015

Abstract

The coding region determinant binding protein (CRD-BP) plays a role in some cancers, regulating gene expression through mRNA trafficking, turnover, and translational repression. Recent evidence reveals CRD-BP binding to KRAS mRNA – a proto-oncogene frequently mutated and/or overexpressed in cancer, leading to up-regulation. A CRD-BP-KRAS mRNA interaction therefore presents a novel therapeutic opportunity to target KRAS-driven cancers and is the aim of this thesis. Electrophoretic mobility and immunoprecipitation experiments presented here using GXXG-mutant CRD-BP variants show four K-homology domains are required to bind a 57-nucleotide minimum binding sequence in the coding region of KRAS RNA; these mutations do not compromise protein folding as evidenced by circular dichroism spectroscopy. DNA-based anti-sense oligonucleotide inhibitors were developed based on the minimum binding sequence that successfully abrogate the CRD-BP-KRAS mRNA interaction in *vitro*. Finally, a CRD-BP KH1toKH4 construct purified using a novel "on-column"-refolding scheme suggests a possible role for RRM-domains as negative regulators of RNA-binding.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank first and foremost my supervisor, Dr. Chow Lee for giving me the opportunity to work in cancer research. My time spent in his lab has helped me grow as an academic and allowed me to participate in meaningful research projects while developing the skills required to succeed in a scientific environment. For the much appreciated technical assistance and often insightful conversations I acknowledge Maggie Lee, Andrea Gorrell, Martha Stark as well as Gerrit van Rensberg and Mark Barnes for preparing the IP-RNA for this project. I would also like to extend my thanks to Dr. Stephen Rader, Dr. Brent Murray, and Dr. Jane Young for providing my recommendations for admission into the M.Sc. Biochemistry program at UNBC, without whom this excellent experience would not have been possible. Finally, my family and close friends – Barry, Laurie, Kris, Lisa, Brahm, Alex and Corbin, deserve my utmost gratitude for their support through the rough times, and company during the good.

Chapter 1

Introduction

Cellular responses to the wide variety of physiological conditions that organisms are regularly exposed to depends heavily on the appropriate regulation of gene expression. Such precise control of gene expression stems from a complex system of signalling molecules capable of exerting influence on the various biochemical pathways involved in cell activities. From homeostasis to blood clotting to induction of wound healing, the ability to selectively modulate various protein levels within the cell is critical for appropriate response to stimuli and optimal functioning. First defined in the late 1950s, the central dogma of molecular biology states that DNA is transcribed into an RNA messenger molecule (mRNA) and subsequently translated into a string of amino acids with defined structure known as protein. Much work has focussed on the fine details of this process and those efforts have delivered an ever-growing body of knowledge regarding the precise workings of how genetic sequences stored in DNA turn into the fully functioning molecular machines and building blocks that cells are largely made of. As astonishing as this process of protein production may be, it is not enough to simply have proteins being made in the cell; to avoid diseases like cancer, they must also be regulated.

1.1 Importance of gene expression in cancer

Cancer is a particularly devastating disease among a host of others, and is the result of genetic instability fostering the deregulation of biochemical pathways effecting cell growth, cell death (apoptosis), proliferation, replicative immortality, angiogenesis, and metastasis (reviewed in Hanahan and Weinberg, 2011). Genetic mutations affecting transcription or altering the functioning of important enzymes in terms of reaction rate can lead to changes in pathway activity. When these altered pathways affect cell cycle events, notable abnormalities can arise

that impact the health and well-being of an afflicted individual. For example approximately 40% of melanomas contain mutations in the signalling molecule B-Raf resulting in constitutive induction of the MAP-Kinase pathway. This leads to enhanced transcription and expression of many genes including C-Fos, which in turn causes changes in cell division (Davies and Samuels, 2010). Abnormal expression of the vascular endothelial growth factor (VEGF) gene is often associated with hypoxic disorders and oncogenic signalling, and potentiates heightened cell receptor signalling involved in the un-governed blood vessel formation required by aggressive solid tumours (Ferrara 2010). The canonical epithelial-mesenchymal transition (EMT) is a key step in the initiation of carcinoma metastasis and can result from the ectopic induction and chronic overexpression of several transcription factors including Zeb1/2 and Twist as reviewed by Yang and Weinberg (2008). Gene expression is influenced by many cellular processes and between transcription and the final protein product, post-transcriptional regulation mechanisms provides a major source of gene control.

1.2 Post transcriptional gene regulation and mRNA stability

Control of cell activities, stability and even the complex string of events that guide embryo development is dependent on mRNA encoded information, the strength of these signals affected by steady-state mRNA levels - the balance between mRNA production and destruction (Brewer 2003). Initial transcription rates, RNA splicing, RNA transport, and mRNA degradation all play major roles in determining steady-state mRNA levels and in particular the modulation of the molecule's longevity, or half-life, within the cell is proving to be increasingly important (reviewed in Filipowicz *et al.*, 2008). Evidence for the importance of stabilizing/destabilizing effects on steady state mRNA level come from global cDNA array experiments comparing stress-induced transcription rates and steady-state mRNA levels of many genes, as well as kinetic

studies of mRNA decay. Stabilization of mRNA was found to enhance the gene expression of many gene groups under various stresses, and similarly destabilization was found to account for reduce gene expression in the examined mammalian cells (Fan *et al.* 2002). Critically, it has been found that large-scale mRNA half-life changes are not a prerequisite for significant effects, as even small changes in mRNA stability can dramatically alter gene expression (Ross *et al.* 1995).

Factors affecting mRNA longevity in the cell are diverse, and our understanding of these processes has been a rapidly evolving process. With the discovery of RNA interference by Fire and Mello (1998), it has become clear that post-transcriptional level regulation is a core means of tightly regulating mRNA in the cell. In general a dynamic combination of cis-elements and transacting factors produce the variations in mRNA stability. Cis-elements are intrinsic sequences that exist throughout the length of a transcript, found in the coding regions, 5` UTR, and the 3`UTR, that are associated with mRNA (in)stability, with the latter being the most common region of destabilizing cis-elements in the form of AU-rich sequence elements, or AREs (Shaw *et al.* 1986). Cis-elements can exist singly or together with other regulatory elements on a single transcript and collectively provide rapid, adaptive responses to a number of different cell signals (Chen *et al.* 1998). Cis-elements often act in concert with trans-acting factors to achieve their effect, a group consisting of proteins and non-coding nucleic acids, such as ribonucleases, RNA-binding proteins (RPBs) and microRNAs capable of recognizing the mature RNA transcript and modulating its physiological survivability (Neugarbauer *et al.* 2002).

1.3 RNA binding proteins and mRNA turnover

Mechanisms of gene regulation are multiple, and the specific interplay between regulating factors increase in complexity in eukaryotic cells (Haynes 1999). Control of gene expression by mRNA regulation begins immediately following transcription in the nucleus, and even export

into the cytoplasm as heterogenous nuclear ribonucleoprotein (HnRNP) particles requires an important class of proteins called RNA-binding proteins (Brewer 2003). As proteins larger than 50 kDa cannot freely diffuse across the nuclear membrane, specific nuclear export signals that regulate mRNA export are required of the RBPs for recognition and transit through the nuclear pore complex (Derrigo *et al.* 2000). RBPs also play an instrumental role in RNA metabolism, and are of particular interest because they serve as an adaptor between mRNA regulatory sequences and modulated expression of the gene. RBPs are a large family of proteins and hundreds exist (mostly uncharacterized) within the human genome. Together with all the RBP-binding sites within the transcriptome, these RPBs form an important tier of gene regulation based on sequence selectivity and specific affinity, collectively termed the "posttranscriptional regulatory code" (Keene 2007). Several types of RNA binding proteins have been identified and each function in different ways by identifying unique cis-elements, binding different mRNA regions and altering translation rates or stabilizing/destabilizing the mRNA transcript. P27Kip1 for example is a regulator of cyclins during G1 phase, and its activity is down-regulated by the association of its 5'UTR mRNA with AU-rich binding proteins, HuR and HuD. Such physical interaction restricts ribosome access to the transcript and therein reduces translation (Yeh *et al.*, 2008). BRF-1 is another RNA binding protein which functions not by regulating translation, instead achieving its effect by destabilizing AU-rich mRNA target transcripts (interleukin-3, interleukin-6 and TNF-alpha) through deadenylation (Raineri *et al.*, 2004). RBPs can also serve to stabilize mRNAs and increase relative gene expression, a phenomenon that when kept in-check serves to appropriately regulate various cell events. Erroneous over-expression of RBPs however can bolster multiple downstream mRNA targets simultaneously, leading to global cell

molecular profile shifts implicated in the development of inflammatory disease, immune dysfunction, and even cancer (Lin *et al.*, 2006; Dean *et al.*, 2004; Lebedeva *et al.*, 2011).

1.4 A brief history of CRD-BP

The Coding Region Stability Determinant Binding Protein (CRD-BP) is an RNA-binding protein that was first characterized after UV-crosslinking experiments revealed an RBP of approximately 70 kDa that selectively binds and stabilizes the coding region determinant sequence of c-myc mRNA (Bernstein et al., 1992). CRD-BP is a member of the "VICKZ" protein family (also known as zipcode-binding proteins), a small group of highly conserved (77-98% identity; Git and Standart, 2002) orthologous RBPs with similar function that include Vg1RBP/VERA (*Xenopus laevis),* IMP-1,2,3 (*Homo sapiens*), CRDBP (*Mus musculus*), KOC (*Homo sapiens*) and ZBP1 (*Gallus gallus domesticus*) (reviewed in Yisraeli 2005). There is however no established standard for the naming, and the designations are often used interchangeably depending on the laboratory. For example, Patel *et al.*, (2011) refers to ZBP1 as IMP1, with IGF2BP1 (another name for CRD-BP) listed in brackets. Members of this family have been identified as important post-transcriptional mRNA regulators with roles in mRNA stability, as well as translational control and mRNA localization involved in lamellipodia formation and bleb extensions in motile cell types (Bernstein *et al.*,1992; Condeelis and Singer, 2005; Poincloix *et al.*, 2011). VICKZ proteins are therefore each multifunctional and have been further implicated in early developmental stages where they are involved in critical events such as oogenesis, synaptogenesis, cell motility, and embryo polarization (Boylan *et al.*, 2008).

Of particular interest is CRD-BP's ability regulate gene expression by affecting the half-life of mRNA targets which influences protein levels. Demonstrating the potency of CRD-BP's effect on steady-state mRNA levels, an experiment utilizing transgenic mice over-expressing

CRD-BP reported a dramatic 100-fold increase in IGFII mRNA, an established target of the protein. Similarly, Bernstein *et al.* (1992), showed that c-myc mRNA levels were stabilized by at least 8-fold by CRD-BP in a cell-free mRNA decay assay. There are numerous examples of this protein's effect on target mRNA levels and pathway signalling, yet how it fits into an overall picture and its contribution to disease is not yet fully understood. However, expression patterns and involvement in certain pathways are beginning to reveal the importance of CRD-BP.

1.5 CRD-BP in early development and carcinogenesis

CRD-BP is normally under strict spatiotemporal control with expression peaking between the zygote and early embryo stages on embryonic day 12.5 in mice. In the days leading up to birth, CRD-BP becomes undetectable or is expressed at very low levels in all rodent tissues (Hansen *et al.* 2004). Similarly, adult human tissue shows little IMP-1 expression compared to the robust levels found in various embryo stages (Leeds *et al.* 1997; Nielsen *et al.* 1999; Le *et al.* 2012). Importantly, CRD-BP expression in rats does not re-emerge during liver tissue regeneration (partial hepatectomy) agreeing with the standing theory that CRD-BP is normally limited to developmental stages only and not just tissue growth/regeneration (Leeds *et al.* 1997). Also consistent with the notion that CRD-BP is an important player in the earliest life stages in mammals, CRD-BP-deficient mice generated by gene trap insertion show severe abnormalities such as dwarfism, improper gut formation and higher perinatal mortality rates (Hansen *et al.*, 2004). On the other hand, over-expression of CRD-BP in a mouse model allowed for normal development, but produced populations with a high incidence of mammary adenocarcinomas (95% occurrence compared to no tumours in the non-transgenic mice) demonstrating a potent proto-oncogene feature of CRD-BP (Tessier *et al.* 2004).

Re-expression of CRD-BP (IMP-1) in later life stages can occur and is strongly associated with several types of cancer. Indeed, a study examining breast cancer patient tumours revealed elevated expression of CRD-BP in 58.5% of human primary breast carcinomas, with ~18.3% as a result of gene duplication in the 17q21 band region where the CRD-BP gene is located (Ioannidis *et al.* 2003). In similar studies, half of neuroepithelial (brain) tumours, nearly a third of non-small cell lung carcinomas, 73% of malignant mesenchymal tumours and 69% of ovarian cancers were found to test positive for CRD-BP (Ioannidis *et al.* 2001, 2003; Kobel *et al.* 2007). Even more incriminating evidence regarding this protein's involvement in neoplasias is a study that found 81% (17 out of 21) of tested colorectal tumours were positive for significant CRD-BP expression, while normal colon tissue and inflammatory bowel diseased tissue showed no detectable levels (Ross *et al.* 2001). Reasons for this re-expression are mostly unknown, however recent study of CRD-BP in melanoma cells cultured under hypoxic conditions that mimic the tumour microenvironment, has shown CRD-BP modulation by the HIF1α (hypoxia induced factor 1 alpha) transcription factor through acetylation of the CRD-BP gene. MTS cell proliferation assays in the same paper also uncovered a strong link between CRD-BP and the melanoma cell's proliferative capacity, as sh-RNA (short hairpin RNA) directed against CRD-BP was able to abrogate the cancer's growth (Craig *et al.* 2012). Interestingly, HIF1α induction is also seen in other cancers where CRD-BP appears to play a role, such as breast and colorectal tumours (Kieth *et al.* 2012). Hypoxia-independent induction has also been observed by Noubissi *et al.* (2006), where the researchers demonstrated that CRD-BP can be induced by enhanced β-catenin/Tcf signalling, a common feature in many colorectal cancers (Morin *et al.* 1997).

The clinical value of CRD-BP is evident in prognostic statistics, and in general, human CRD-BP expression levels correlate with the diagnosed grade of malignancy for several cancers

(Yaniv and Yisraeli 2002). Additionally, it has been found that CRD-BP levels correlate with reduced recurrence-free and overall survival rates in ovarian cancer. So far based on singlevariate and multivariate analysis, it is unclear whether actual prognostic value exists for ovarian cancer patients with research requiring larger sample sizes (Kobel *et al.* 2007). In colorectal carcinomas, where CRD-BP expression is most prevalent, positive testing correlated strongly with overall aggression and metastasis and offers a molecular factor in prognosis (Dimitriadis *et al.* 2007). While the role of CRD-BP in some cancers is more significant than in others, it is clear that CRD-BP is generally an important ingredient for carcinogenesis. Although its precise contributions and mode of action appear to be complex, knowledge of its targets and how precisely it interacts with those targets will allow scientists to take aim at CRD-BP in a prognostic and even therapeutic sense. Unfortunately there is much ground yet to tread in this realm, but how this RBP does its job is progressively coming to light.

1.6 VICKZ protein family structure and function

VICKZ proteins have been shown to bind multiple targets including the mRNAs for c-myc, IGF-II, beta-actin, H19, tau and others (reviewed in Yisraeli 2005). How they accomplish this is through the multiple RNA-binding domains present in these proteins. CRD-BP and other VICKZ family members are composed of six RNA-binding domains; 4 K-homology domains (KH1-4) with a βααββα topology, and 2 RNA recognition motifs (RRM1-2) with a βαββαβ topology that make up the majority of the protein's structure. High sequence conservation of the VICKZ family RNA-binding domains, as well as the linker regions between KH1-KH2 and KH3-KH4, suggest that the domains may function together as didomains (KH12 and KH34) (Git and Standart, 2002). Indeed, studies of the ZBP1 homolog show that linker region amino acids are critical for proper protein function. In this experiment, a truncated KH34 ZBP1 construct was

synthesized and the linker amino acids joining ZBP1 KH3 and KH4 pseudo-domains were replaced with those from the adjacent KH1 – KH2 linker region. Only the identity of the amino acids was changed, maintaining the overall length of the linker. Change in the amino acid identities alone produced a marked negative affect on the observed Kd for zipcode binding as determined through EMSA, indicating the important role these linker regions play in RNA binding (Chao *et al.* 2010). Physiological significance was determined through cell-based experimentation, where a GFP-fused ZBP1 linker mutant (full length protein containing the same linker mutation) was expressed in mouse embryonic fibroblasts to assess any functional impact. When analyzed by confocal microscopy, the linker mutant no longer retained the wild-type sub-cellular localization and was found to with those in-between amino acids serving as more than just a hinge (Chao *et al.* 2010).

Part of CRD-BP's function is to bind and localize mRNA within the cell. How it accomplishes this task and the question of which, if any, individual domains are primarily responsible for binding has been answered with varying results by studying CRD-BP and its closely related orthologs. Granular RNP formation assists in localizing various mRNA transcripts within the cell, and systematic deletion of the individual domains of IMP1 revealed nuclear export signals (NES) in KH2 and KH4 domains (Nielsen *et al.* 2004). To address the question of domain redundancy in target-binding, Nielsen *et al.* (2002) designed an experiment using GFP-tagged deletion constructs of IMP1 that had the RRM and KH domains systematically removed, and subsequently analyzed for presence of IMP1-containing RNP granule formation, sub-cellular localization and mRNA binding capacity. It was found that in terms of cytoplasmic trafficking, loss of both RRM domains had little impact on subcytoplasmic localization, with all four KH domains proving necessary and sufficient for normal IMP-1

distribution and granular appearance. Similar results were obtained with regard to RNA-binding to known target H19 mRNA, where the various deletion constructs were assayed using electrophoretic mobility shift assay (EMSA) and KH domains 1-4 were all found to be required for effective binding (Nielsen *et al.* 2002). Conversely, ZBP1 - the founding member of the VICKZ protein family, was found through similar experiments to only require KH 3-4 domains to bind its target beta-actin mRNA, with the other four domains (RRM1-2 and KH1-2 constructs) unable to bind with an appreciable Kd (Chao *et al.* 2010). Despite high similarity, and otherwise identical behaviour in same-substrate binding assays, the different domain requirements demonstrate that VICKZ proteins (including CRD-BP) may interact differently on an individual target basis (VICKZ protein truncation studies summarized in Table 1.1).

Further insight into the structural details of VICKZ family proteins and how they bind RNA was addressed by means of X-ray crystallography. Studying ZBP1 in complex with beta-actin mRNA, it was found that binding of RNA was mediated by interaction with a minimum 29 nt RNA fragment in 2 spots, with both the KH3 and KH4 domains arranged in an anti-parallel fashion such that the RNA binding surfaces ran in opposite directions (revealed by a IMP-1 KH3-KH4 crystal structure). This didomain configuration was found to induce a unique 180° looping of the RNA strand with two non-sequential, spaced RNA sequences bound. It is thought that this is likely the *in vivo* conformation and mechanism of binding for VICKZ proteins because the interaction of the KH domains is stabilized by sequestering a large surface area and several hydrophobic residues present in all VICKZ homologs (Chao *et al.* 2010). Adding another level of complexity, surface plasmon resonance and NMR experiments have demonstrated that the KH3-4 didomain contains a dimerization motif. Weak protein-protein interaction between the separate KH-3-4 didomains was observed and this interaction is stabilized by the presence of

RNA, leading to alternative stoichiometry possibilities in VICKZ RNP complexes and granule composition (Git & Standart, 2002; Lewis *et al.* 2000). In further support of this notion, kinetic analysis and advanced atomic force microscopy of IMP1 suggests positive cooperativity, with multiple IMP-1 molecules binding to a single mRNA target (Nielsen *et al.* 2004; Jonsen *et al.* 2007). Protein-RNA interaction in this model occurs twice, with the first binding event characterized by a highly transient intermediate of low stability, followed by a second IMP1 binding. Protein-protein interaction via a KH3-4 dimerization motif then further stabilizes the RNP complex, lending explanation to the hyper-stable RNP granules responsible for transporting target mRNAs across the cell (Nielsen *et al.* 2004). Little information regarding the actual transport of these protective RNP complexes exists. However, immuno-staining and confocal microscopy revealed that IMP-1 does co-localize with microtubules and f-actin, suggesting cytoskeleton association with ATP-dependent movement (Nielsen *et al.* 2002).

Despite the numerous studies displaying tight association with whittled-down target mRNA regions, and even X-ray structural models, no consensus binding sequence for VICKZ proteins as a family, or individually, has been identified with only case-specific generalizations having been made that do not fully account for observed binding. For example, CRD-BP interaction with c-myc mRNA has been shown to predominantly occur in the region corresponding to nucleotides 1763 to 1777 (5'-AGCCACAGCAUACAU-3'), but it is unknown which nucleotides comprise the actual binding motif (Coulis *et al.* 1999). Even between experiments using identical proteins in similar conditions, different conclusions have been drawn. Jonsen *et al.* (2007), concluded that ZBP1 has a high affinity for guanosine-rich and cytosine-poor RNA regions, and putatively identified a 5'-CCYHHCC-3' binding motif (Y=C or U, H = A or C or U). Using a different experimental approach, other researchers later concluded

that the ZBP1 consensus binding motif was an enrichment of 5'-CAUH-3' repeats in the target transcript (Hafner *et al.*, 2010). VICKZ protein RNA-binding requirements for mRNA transcripts are still largely tentative, however Patel *et al.* (2011), demonstrated a strict binding requirement for just the ZBP1 KH34 didomain, with the KH4 domain recognizing a 5'-CGGAC sequence, and the KH3 domain recognizing a 3'-C/A-CA-C/U sequence, in line with the aforementioned "RNA-looping" model (Patel *et al.* 2011; Chao *et al.* 2010). These discrepancies may say something about the methods used, or it may be that the different identified sequences stem from different KH domain preferences. It may also be that these multi-RNA-binding-domain proteins are targeting a common three-dimensional structure as opposed to just a linear sequence.

Table 1.1. Summary of mentioned VICKZ protein mutation studies

CRD-BP Ortholog	Type of Mutation	Type of Study	Conclusion	Ref.
IMP1	Deletion construct	Cell-based, GFP-tagged	NES present in KH2 & KH4	Nielsen *et al.* 2004
ZBP1	KH3-4 linker	In vitro EMSA	Increased zipcode RNA binding Kd.	Chao *et al.* 2010
ZBP1	KH3-4 linker	Cell-based, GFP-tagged	Reduced granulation and cytoplasmic sub-localization	Chao *et al.* 2010
IMP1	Deletion construct	In vitro, EMSA Cell-based, GFP-tagged	KH domains 1-4 all required for effective binding of H-segment (173 nt fragment) H19 RNA	Nielsen *et al.* 2002
ZBP1	Deletion construct	In Vitro, EMSA	KH 3-4 didomain important in binding beta-actin mRNA (54 nt binding region)	Chao *et al.* 2010

1.7 Oncogenic KRAS gene expression and CRD-BP

Deregulation of critical pathway elements involved in cell signalling and cell-cycle control are at the core of cancer. Many of these oncogenic pathways are well characterized and are particularly common offenders in a variety of cancers. Colorectal, pancreatic, endometrial, prostate, and lung cancers are examples that consistently show up-regulation of specific pathways including the Wnt/beta-catenin, PI3-K, VEGF and EGFR pathways (Takahashi *et al.* 1995; Rodriguez *et al.* 1996; Hennessy *et al.* 2005). Common to each of these is a constituent GTPase signalling protein called the Kirsten rat sarcoma viral oncogene homolog (KRAS). Abnormalities in this signalling molecule are implicated in 90% of pancreatic cancers, 60% of colorectal cancers (CRC), and 50% of lung cancers (Dergham *et al.* 1997; Bos *et al.* 1987; Mascaux *et al.* 2005). A common mutation in the glycine-12 position to aspartic acid strongly reduces the enzyme's GTPase activity, effectively locking the molecule in the "on" state by preventing hydrolysis of the activating GTP molecule once it has bound, and is considered an important biological factor in many cancers (Georges *et al.* 1993; Pylayeva-Gupta *et al.* 2011). This in effect bypasses externally regulated signalling and provides the necessary biochemical environment to induce the cancer phenotype and is in fact the single most common activating mutation in human cancers (Shaw *et al.* 2011). RNA interference experiments have further shown that changes in KRAS expression can independently regulate the downstream effectors of all the aforementioned pathways such as the mitotic factor cyclin D1, translation enhancing S6K, VEGF, and apoptotic factor FOXO making Kras a lucrative target for future antineoplastic therapies (Miura *et al.* 2005).

Deregulated KRAS activity in many cancers however can also arise from over expression of the wild-type form. Indeed, a comprehensive study of KRAS lesions in endometrial carcinoma patients found that high wild-type KRAS expression correlated more strongly with poor survival rates and unfavourable treatment outcomes than did any mutant variations. Elevated KRAS mRNA expression and KRAS gene amplification specifically correlated with an aggressive phenotype, as well as increased from primary to metastatic lesions (Birkeland *et al.* 2012). A role for wild-type KRAS in colorectal cancer is echoed by Valtorta *et al.* (2013) who demonstrated KRAS gene amplification correlates with anti-EGFR antibody therapy resistance with cetuximab and panitumumab in some CRC patients. KRAS induction is independently able to induce a cancer phenotype in rat pancreas, and knockdown has demonstrated that many pancreatic cancer cell lines are dependent on KRAS over-expression (Tuveson *et al.* 2004; Zhang *et al.* 2006) IMP1 has been found to be a major player in the regulation of KRAS levels by binding directly to the mRNA transcript (Mongroo *et al.* 2011). UV-crosslinking has demonstrated that IMP-1 directly interacts with KRAS mRNA in the coding region (CDS) and the 3' UTR and that down-regulation of IMP1 directly results in the reduced mRNA level and protein level of Kras in CRC cell lines (Mongroo *et al.* 2011). Positive regulation of KRAS expression by IMP1/CRD-BP therefore presents a novel opportunity to target KRAS in cancers that co-express CRD-BP.

1.8 Targeting KRAS in KRAS-driven cancers

KRAS driven malignancies are very common, effecting up to a third of all cancers but most prominent in leukemias, colorectal, pancreatic and lung cancers and as such, a diverse range of anti-KRAS therapeutic strategies have been pursed. There is still no prescribable go-to drug for targeting KRAS in cancer, with many experimental treatments falling short due to systemic toxicity or poor inhibition, which has limited the degree to which KRAS cancers can be treated

(Baines et al. 2011). However, there are several promising avenues of research that have produced hopeful candidate proto-drugs recently.

Directly targeting the KRAS protein with small molecules has proven an ill-fated strategy thus far. The protein's shallow grooves and lack of binding pockets have limited the success in this area to molecules incapable substantially modulating KRAS activity due to low-affinity binding interactions (Maurer *et al.* 2012; Sun *et al.* 2012; Shima *et al.* 2013). However, Ostrem *et al.* (2013) made a recent KRAS targeting breakthrough when experiments utilizing thiol-based inhibitors showed strong protein binding character. These inhibitors form a disulphide bond with a nascent cysteine on the KRAS protein, and are able to coax KRAS into its inactive form by stabilizing the GDP-bound conformation. Furthermore, the top candidate molecule displays an inherent preference for cells expressing the G12D mutant version of KRAS, commonly found in many cancers and achieve some specific anti-proliferation activity. However, despite the promising binding data, it is emphasized that these thiol-based inhibitors are unsuitable as cancer drugs in the current form, as Ras signalling was only marginally reduced. Development of this class of inhibitors remains an active area of research pursing chemical "warheads" more capable of inhibiting KRAS signalling.

Farnesyl-transferase inhibitors (FTI's) have been thoroughly explored in the laboratory and in the clininc (reviewed in Appels et al. 2005). This class of KRAS inhibitor disrupts a post-translational farnesylation modification of the KRAS protein, whereby normally a large 15-carbon hydrophobic tail is added to allow effective membrane association. As KRAS achieves its signalling function through interactions with other proteins near the cell membrane, failure to attach a lipid anchor results in a significant inhibition of KRAS signalling. In practice, FTIs such as Lonafarnib and Tipifarnib only proved effective against the less common H-RAS and N-RAS

members of the RAS superfamily. Unlike the other RAS proteins, KRAS as it turns out can be alternatively post-translationally modified by geranyl transferases upon inhibition of farnesyl transferase enzymes, providing a compensatory redundancy largely limiting the success of FTIs against KRAS driven cancers. Targeting both the farnesyl and geranyl transferase enzymes simultaneously proved lethal to non-cancerous and cancerous cells alike, as many proteins necessary for basic cell functions are substrates of these same enzymes.

Preventing KRAS from functioning at the cell membrane remained an enduring therapeutic strategy with the introduction of PDEδ inhibitors. PDEδ is the delta subunit of cGMP phosphodiesterase and is a KRAS binding partner that preferentially interacts with farnesylated KRAS and aids in the translocation of KRAS to the membrane (Chandra *et al.* 2012). Small molecules developed to inhibit this critical step, by specifically interrupting PDEδ's association with the farnesylated tail of KRAS but still allowing the actual fanesylation step to occur, culminated in the introduction and pre-clinical trial of Deltarasin. Dose-dependent delocalization of farnesylated KRAS using deltrarasin was recently demonstrated in human pancreatic cancer cell lines and in a mouse model of pancreatic ductal adenocarcinoma where tumour growth was noticeably reduced, providing promising evidence that this strategy is a valid method to target KRAS. However, the drug's lack of specificity produces high systemic toxicity and future work still needs to be done before pushing this proto-drug to human trials (Zimmerman *et al.* 2013).

Salirasib (S-trans,trans-farnesylthiosalicylic acid) is another small molecule that has been shown to interfere with proper membrane association of KRAS. Salirasib however does not directly target any protein or enzyme important in the post-translational modification or intracellular transport of KRAS, instead binding the hydrophobic farnesyl tail directly, reducing the overall hydrophobicity of the tail and impairing its effectiveness to interact with the cell

membrane (Weisz et al. 1999). Recent pre-clinical and clinical development on Salirasib has further shown that the molecule is well-tolerated in patients with metastatic pancreatic ductal adenocarcinoma (PDA) and enhances the effectiveness of gemcitabine chemotherapy when used in combination (Laheru et al. 2012). Similar problems limit the effectiveness of salirasib, as it universally affects all RAS isoforms with no cancer cell specificity.

Inhibitors designed to block KRAS from reaching, or interacting with the membrane are only strategy, but another promising avenue is preventing the activation of KRAS by inhibiting the RAS guanine nucleotide exchange factor: Ras-GEF SOS, required to induce GTP-binding and turn KRAS "on". The exchange of GDP for GTP is a key step in downstream signalling of KRAS and has been the specific target of Patgiri *et al.* (2011) who have produced a small molecule, a synthetic α-helix that mimics the binding motif of the SOS protein, that competes with the KRAS-Ras-GEF SOS protein-protein interaction, blocking the exchange of GDP for GTP and hindering KRAS activation through receptor tyrosine kinase stimulation. While this work has been limited thus far to cell lines only, it is undoubtedly a promising avenue of research.

Attacking KRAS at a genetic level by altering the rate of transcription is a more recent approach pioneered by Cogoi *et al.* (2013). Within the promoter region of the KRAS gene, it was determined that stretches of the DNA form G4-quandruplex structures which are recognized by a myc-associated zinc finger transcription factor (MAZ). MAZ protein binding promotes KRAS transcription, and by using a synthetic DNA oligonucleotide G4-quadruplex mimic that acts as a MAZ decoy, KRAS expression could be effectively reduced in human pancreatic cell lines (Panc-1) and reduce KRAS-driven tumour growth in mice by 64%. The pancreatic cancer cells were shown to display increased apoptosis upon administration of the MAZ-decoy and validate

the concept of MAZ-specific G4 quadruplex competitor molecules. Further work in this area will focus on increasing physiological stability of the molecules, as a large dose was required to achieve the effect presumably due to degradation of the DNA-based molecule.

It is clear that KRAS is a current topic, and a strategic therapeutic target in the war on cancer, however despite literally decades of research, KRAS remains an elusive target in the clinic. Complexity in targeting this GTPase signalling molecule arise from the structurally similar, but functionally distinct members of the Ras family, whereby targeting KRAS usually inadvertently targets the other related proteins unintentionally. Approaches whereby KRAS is directly targeted have only recently shown moderate success, using a thiol-based binding scheme, but these still are limited in ability to actually knockdown KRAS activity. FTIs showed great early promise by preventing a critical post-translational lipid modification necessary for KRAS membrane association, but the co-activity of both farnesyl- and geranyl-transferase inhibitors proved toxic to cells. Similar toxicity plagues the Salirasib approach, where obstruction of the hydrophobic tail of KRAS by a masking molecule delocalizes the protein. Interfering with the proper transport of farnesylated KRAS protein to the inner cell membrane using a PDEδ inhibitor has shown some promise in mouse models, but again systemic toxicity and poor drug tolerance has limited the effectiveness of this molecule and requires further optimization still. Inhibiting the activation of KRAS showed moderate success. Ras-GEF SOS is required to induce KRAS to bind GTP, a necessary step in KRAS signalling, and with the use of synthetic α-helix mimics designed to mimic the KRAS binding site, KRAS docking with the guanine nucleotide exchange factor could be inhibited and overall activity reduced. Perhaps the single most promising work, including cell and mouse-model data was the G4-quadruplex MAZ-decoy approach, as it was able to specifically target KRAS at the genetic level, and not the other RAS isoforms, but has not yet

progressed to human clinical trials. Unfortunately, despite the diverse array of strategies employed to target KRAS, no silver bullet has been proven although there are some promising leads, necessitating the need for further research into the existing aforementioned drug platforms as well as new novel approaches.

1.9 Research Goals

CRD-BP has a prime role in the initiation and progression of several cancers. Of particular interest is the recent discovery that CRD-BP binds the KRAS mRNA transcript, resulting in overexpression of this well-established oncogene prevalent in colorectal, pancreatic, and lung cancers. Targeting KRAS-driven cancers has proven extremely difficult and despite thorough investigation of a wide variety of anti-KRAS strategies, there remains still no effective, targeted therapy. Importantly, CRD-BP is an oncofetal protein, and is normally only expressed during early zygotic and embryonic stages of development. Adult tissue normally has virtually no detectable levels of CRD-BP; however robust re-expression occurs in many cancer cells and therein lies a novel therapeutic opportunity that targets the core of some cancers and also is inherently specific. Abrogating the KRAS mRNA interaction with CRD-BP therefore may be an effective means of attacking some KRAS-driven cancers, and is the focus of this research.

The first major research aim was to confirm that CRD-BP binds KRAS mRNA in the coding region and 3'UTR using the EMSA assay and to further this work by mapping CRD-BP binding sites on the KRAS mRNA transcript using purified recombinant CRD-BP protein and IVT generated KRAS sub-regions. Determination of a minimum binding sequence was the end goal for these experiments, contributing elements towards a final CRD-BP consensus sequence and serving as a focus for further study of the molecular nature of the CRD-BP-KRAS mRNA interaction. The second major goal was the development of binding inhibitors, spearheaded using

DNA-based oligonucleotides that complementarily bind through Watson-crick base pairing to contiguous stretches of a minimum binding sequence. This would allow site-specific assessment of their ability to block KRAS RNA recognition by CRD-BP and have possible use as proto-drugs in cell assays and mouse models targeting KRAS-driven cancers. The third major goal was to use a minimum binding sequence and CRD-BP mutational analysis to gain insight into the KH domain requirements for KRAS binding. As abrogating CRD-BP interactions are an ultimate goal of VICKZ protein research, knowledge of the unique KH domain requirements for each mRNA target sheds light onto some necessary properties small molecule inhibitors must have - which sub-set of KH domains need be blocked. The final objective was to elucidate structural properties or CRD-BP. Protein crystallization for use in X-ray diffraction was to be explored to reveal the precise mechanism of CRD-BP/VICKZ protein RNA-binding. Additionally, circular dichroism was to be employed in order to assess the possibility of a dynamic CRD-BP protein that undergoes conformational shift in secondary structure upon target-specific mRNA-binding. Determining structural changes that occur upon binding target RNAs would offer an additional means of screening potential RNA target molecules using CD spectroscopy.

Chapter 2

In vitro characterization of KRAS mRNA binding to CRD-BP and development of antisense oligonucleotide inhibitors.

The following chapter covers experiments carried out to determine a minimum binding sequence for KRAS mRNA required to bind CRD-BP. Additionally, the development of anti-sense DNA-based oligonucleotide inhibitors capable of abrogating the CRD-BP-KRAS mRNA interaction will be presented and discussed.

2.1 Methodology – Mapping of CRD-BP binding sites on KRAS mRNA

2.1.1 Transformation of BL21(DE3) cells with pET28b-CRD-BP plasmid.

Plasmid containing wild-type CRD-BP was previously acquired from the Dr. J. Ross lab. Full-length mouse CRD-BP coding sequence cut with NcoI and XhoI was subcloned into pET28b(+) (Novagen), flanked by an amino terminus FLAG epitope and a C-terminus 6x His-tag. The plasmid contains a kanamycin resistance gene and expression is driven by a T7 promoter as depicted in Figure 2.1. Approximately 100 µl of competent BL21(DE3) cells were thawed from -80°C storage on ice for one half hour before adding 10 ng of pET28b-CRD-BP plasmid DNA. Cells were subsequently placed on ice for an additional 25 minutes followed by 90 seconds at 42°C. Cells were immediately supplemented with 300 µl of

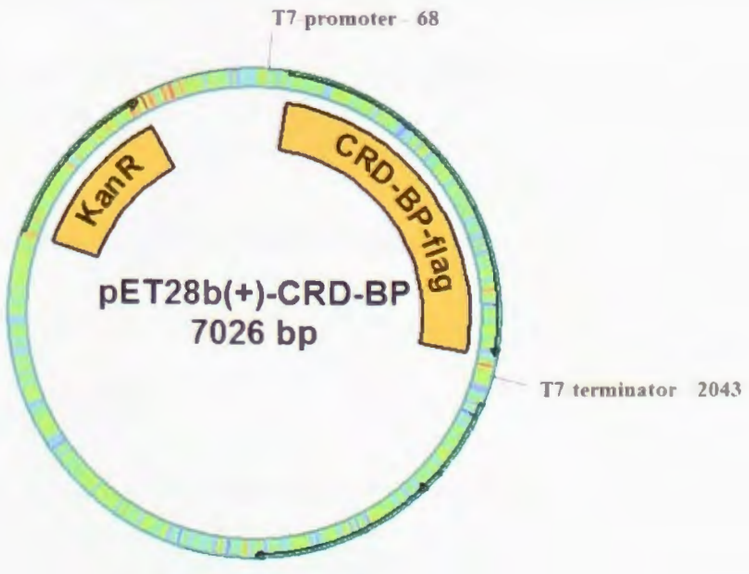


Figure 2.1. Plasmid map of pET28b(+)-CRD-BP

L.B. broth (Invitrogen) and incubated at 37°C for one half hour before plating on L.B. agar (Invitrogen) plates containing 25 ug/ml kanamycin sulfate (Fisher Scientific) antibiotic. Plates were then allowed to incubate overnight for approximately 16 hours, yielding colonies of transformants harbouring the pET28b-CRD-BP plasmid.

2.1.2 Induction and Growth of pET28b-CRD-BP E.coli cells

Transformant colonies (~20) containing the CRD-BP expression vector were used to inoculate 100 ml of L.B. broth (Invitrogen) containing 25 ug/ml kanamycin sulfate (Fisher Scientific) in a 250 ml Erlenmeyer flask. Cell suspension was placed on an incubator-shaker (New Brunswick Scientific, model Innova 40) set at 200 rpm, 1 inch stroke, at 37°C for 3 hours before being transferred to a 1L Fernbach flask containing 900 ml of L.B. broth with 25 µg/ml kanamycin. Culture was grown until cell $O.D._{600}$ equaled 0.5 at which point the cells were induced with 1mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG, Bio Basic Inc.) and incubated under the same conditions for another 6 hours. Centrifugation of 1L culture at 4 °C, 2,100 xg, for 15 minutes (sub-divided into 4 volumes of 250 ml) produced 4 cell pellets which were stored at -80 °C.

2.1.3 Purification & quantification of recombinant CRD-BP protein

Recombinant CRD-BP purification was achieved using the following buffers and scheme as outlined in table 2.1 below.

Table 2.1. Composition of buffers used in denaturing purification of recombinant CRD-BP protein from transformed BL-21(DE3) E.coli cells.

Buffer	NaH_2PO_4	Tris-Cl	Urea	pH
B	100 mM	10 mM	8 M	8
C	100 mM	10 mM	8M	6.5
D	100 mM	10 mM	8M	6.3
E	100 mM	10 mM	8M	5.9
F	100 mM	10 mM	8M	4.5

E.coli BL21(DE3) Cell pellets were thawed on ice for 15 minutes before being re-suspended into 12 ml of lysis buffer B. Cells in lysis buffer were incubated for 1 hour on ice while being agitated using a waver on medium setting (VWM Scientific Products). Lysate was cleared by centrifugation at 21,000 xg for 30 minutes at 4 °C, producing a pellet of cell debris. Supernatant was then syringe filtered (0.45 µm) and prepared for batch-binding by mixing in 1 ml of pre-equilibrated (buffer B) nickel-nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen) and incubating the slurry on ice for 1 hour. Resin-lysate mix was applied to a mini-column (Qiagen) and washed using 5 ml of buffer B, followed by 5 ml of buffer C and finally 5 ml of buffer D. CRD-BP protein was eluted by passing 12 ml of buffer E over the column followed by 12 ml of buffer F, collecting 0.5 ml fractions of protein solution. Denatured protein samples were immediately stored at -80 °C.

Purified recombinant mouse CRD-BP samples were re-natured using 3,500 MWCO mini dialysis units (Thermo Scientific) in a multi-step buffer exchange. Denatured protein samples containing 8M urea were dialyzed against a 500-fold volume of refolding buffer for 24 hours. Following the protein refolding stage, sample was then dialyzed against a 500-fold volume of final buffer (buffer recipe below in table 2.2) for 2 hours, followed by another dialysis against a 2,500-fold volume of final buffer to completely remove all traces of urea. Finally concentration of protein was determined using a BCA Protein Assay Kit (Thermo Scientific).

Table 2.2. Buffers used for refolding and storage of purified recombinant CRD-BP protein.

Refolding buffer	Storage (final) buffer
200 mM NaCl	200 mM NaCl
20 mM Tris-Cl	20 mM Tris-Cl
2 M urea	10% glycerol
10% glycerol	0.01% triton-X
0.01% triton-X	pH 7.4
1 mM Glutathione - reduced	-
0.1 mM Glutathione - oxidized	-
pH 7.4	-

2.1.4 Generation of KRAS mRNA sub-region cDNA templates

In order to assess KRAS mRNA sub-region affinity for CRD-BP, smaller fragments were first generated by polymerase chain reaction (PCR) using primers designed to amplify contiguous stretches of a pCMV6-entry Human KRAS cDNA clone which includes the full coding sequence as well as a partial 3' untranslated region (UTR) (Origene, NM_033360). Primer pairs were designed to amplify specific sub-regions of the KRAS mRNA cDNA with T7 promoter sequence incorporated at the 5' end of all forward primers to drive in vitro transcription of cDNA templates for subsequent use in binding assays (table 3, 4 and 5) and PCR reactions carried out according to the template below.

Template	100 ng
10 x PCR buffer	3.5 µl
dNTPs (2.5 mM)	3.5 µl
forward primer	100 ng
reverse primer	100 ng
Taq. Polymerase (5,000 U/ml)	0.5 µl
Water	to 35 µl

Thermocycler program

94°C	30 sec	
50°C	30 sec	30 cycles
72°C	45 sec	

Table 2.3. PCR primers used for the generation of KRAS mRNA sub-region cDNA templates A-F for in vitro transcription. KRAS mRNA sub-region nucleotide positions in brackets are relative to the KRAS AUG start codon.

Region	Primer	DNA sequence
A(1-185)	forward	**5'**GGATCCTAATACGACTCACTATAGGATGACTGAATATAAACTT **3'**
	reverse	**5'**-TCATGACCTGCTGTGTCG-**3'**
B(175-401)	forward	**5'**-GGATCCTAATACGACTCACTATAGG GCAGGTCATGAGGAGTAC-**3'**
	reverse	**5'**-GCTAAGTCCTGAGCCTGT-**3'**
C(388-610)	forward	5'-GGATCCTAATACGACTCACTATAGGGCTCAGGACTTAGCAAGA-3'
	reverse	5'-CCACTTGTACTAGTATGC-3'
D(568-793)	forward	5'-GGATCCTAATACGACTCACTATAGGATACAATTTGTACTTTTT-3'
	reverse	5'-CACAGGCATTGCTAGTTC-3'
E(772-988)	forward	5'-GGATCCTAATACGACTCACTATAGGTTTTGAACTAGCAATGCC-3'
	reverse	5'-CCAATTAGAAGGTCTCAA-3'
F(971-1155)	forward	5'-GGATCCTAATACGACTCACTATAGGTTGAGACCTTCTAATTGG-3'
	reverse	5-CATCATCAGGAAGCCCAT-3'

Table 2.4. PCR primers used for the generation of KRAS mRNA sub-region cDNA templates KRAS-A(1-4) for in vitro transcription. KRAS mRNA sub-region nucleotide positions in brackets are relative to the KRAS AUG start codon.

Region	Primer	DNA Sequence
A1-3(1-141)	forward	5' GGATCCTAATACGACTCACTATAGGATGACTGAATATAAACTT 3'
	reverse	5'AATTACTACTTGCTTCCT 3'
A1-2(1-92)	forward	5' GGATCCTAATACGACTCACTATAGGATGACTGAATATAAACTT 3'
	reverse	5' TCGTCCACAAAATGATTC 3'
A1(1-46)	forward	5' GGATCCTAATACGACTCACTATAGGATGACTGAATATAAACTT 3'
	reverse	5' TGCCTACGCCACCAGCTC 3'
A2-4(47-185)	forward	5' GGATCCTAATACGACTCACTATAGGAGAGTGCCTTGACGATAC 3'
	reverse	5' TCTTGACCTGCTGTGTCG 3'
A3-4(93-185)	forward	5' GGATCCTAATACGACTCACTATAGGATATGATCCAACAATAGA 3'
	reverse	5' TCTTGACCTGCTGTGTCG 3'
A4(139-185)	forward	5' GGATCCTAATACGACTCACTATAGGGATGGAGAAACCTGTCTC 3'
	reverse	5' TCTTGACCTGCTGTGTCG 3'
A3(93-138)	forward	5' GGATCCTAATACGACTCACTATAGGATATGATCCAACAATAGA 3'
	reverse	5' AATTACTACTTGCTTCCT 3'

Table 2.5. PCR primers used for the generation of KRAS mRNA sub-region cDNA templates KRAS-A34(*a-f*) for in vitro transcription. KRAS mRNA sub-region nucleotide positions in brackets are relative to the KRAS AUG start codon.

Region	Primer	DNA sequence
A3-4*a* (111-185)	forward	5' GGATCCTAATACGACTCACTATAGGGGATTCCTACAGGAAGCA 3'
	reverse	5' TCATGACCTGCTGTGTCG 3'
A3-4*b*(129-185)	forward	5' GGATCCTAATACGACTCACTATAGGAGTAGTAATTGATGGAGA 3'
	reverse	5' TCATGACCTGCTGTGTCG 3'
A3-4*c*(147-185)	forward	5' GGATCCTAATACGACTCACTATAGGAACCTGTCTCTTGGATAT 3'
	reverse	5' TCATGACCTGCTGTGTCG 3'
A3-4*d*(93-167)	forward	5' GGATCCTAATACGACTCACTATAGGATATGATCCAACAATAGA 3'
	reverse	5' AGAATATCCAAGAGACAG 3'
A3-4*e*(93-149)	forward	5' GGATCCTAATACGACTCACTATAGGATATGATCCAACAATAGA 3'
	reverse	5' GTTTCTCCATCAATTACT 3'
A3-4*f*(93-131)	forward	5' GGATCCTAATACGACTCACTATAGGATATGATCCAACAATAGA 3'
	reverse	5' ACTTGCTTCCTGTAGGAA 3'

2.1.5 Generation of [^{32}P]-labelled KRAS RNA substrates by IVT

In vitro transcription (IVT) of KRAS sub-region cDNA templates in the presence of [^{32}P]-labelled UTP was used to produce internally labelled RNA substrates. IVT reaction mixtures were prepared for each cDNA template according to table 2.6.

Table 2.6. IVT reagents used for KRAS mRNA sub-region RNA synthesis.

IVT reagent	Amount in 20 ul reaction volume
5x transcription buffer	4 µl
100 mM DTT	2 µl
20 U/ul RNasin	1 µl
10 mM ATP	1 µl
10 mM CTP	1 µl
10 mM GTP	1 µl
100 µM UTP	2.5 µl
15 U/ul T7 RNA polymerase	1 µl
DEPC H_2O	To 20 µl
[^{32}P]-UTP	2.5 µl

IVT reactions were mixed and incubated at 37°C for 1 hour prior to a 10 min incubation with 10 µl of RNase-free DNase (1U/ µl) at 37°C. IVT reactions were stopped with a urea loading dye (9M urea, 0.01% bromophenol Blue, 0.01% xylene cyanol, 0.01% phenol) and RNA transcripts gel purified using an 8% polyacrylamide (29:1 bis:acrylamide) denaturing gel (7M urea) ran at 25 mA for 1 hour with 0.5x TBE buffer. Excised gel-slices were crushed and RNA extracted in 400 µl of H_2O at 70 °C for 5 minutes prior to further purification by Performa DTR gel-filtration cartridges (Edge Bio). RNA was then subjected to phenol-chloroform extraction and ethanol precipitation before being resuspended in 25 µl of DEPC H_2O. Purified [^{32}P]-labelled IVT products were quantified using a liquid scintillation counter (Hidex) and stored at -20 °C.

2.1.6 EMSA mapping of CRD-BP binding sites on KRAS mRNA

Binding reactions were prepared separately for each generated [^{32}P]-labelled KRAS sub-region RNA according to table 2.7 and CRD-BP concentration varied from 0 to 540 nM. EMSA reactions used for mapping high affinity KRAS-CRD-BP binding sites contained 20,000 cpm RNA probe.

Table 2.7. EMSA reagents used for KRAS-CRD-BP affinity mapping and screening of antisense oligonucleotide inhibitors.

EMSA binding reaction reagents (19 ul)	**Volume**	**Final Concentration**
Binding buffer (see below)	4 µl	-
10 mg/ml baker's yeast tRNA (Sigma Aldrich)	1 µl	0.53 mg/ml
10 mg/ml BSA (New England Biolabs)	1 µl	0.53 mg/ml
40 U/µl rRNasin (Promega)	1 ul	2 U/µl
CRD-BP	x	0 to 540 nM
[^{32}P]-labelled RNA probe	2 µl	20,000 cpm/reaction
H_2O	to 19 µl	-
Binding buffer reagents (per 1 ml)		
1M tris-Cl pH 7.4	50 µl	50 mM
0.5 M EDTA (pH 8.0)	25 µl	12.5 mM
100 mM DTT	100 µl	10 mM
50% glycerol	500 µl	25%
10% triton X-100	1 µl	0.01%
H_2O	324 µl	-

RNA probes prior to reagent mixing were first heated to 55 °C for 5 minutes and allowed to cool to room temperature for 7 minutes in order to facilitate proper folding. Reagents were added separately in a 1.5 ml eppendorf tube and centrifuged at 3000 xg for 10 seconds to begin binding reaction; mixtures were then incubated at 37 °C for 10 minutes, followed by a 5 minute

incubation on ice, another 10 minutes at 37 °C and a final 5 minute step on ice totalling 30 minutes. 2 µl of EMSA loading dye consisting of 250 mM Tris-HCl, 0.2% Bromophenol blue, 0.2% xylene cyanol, and 40% sucrose was added to each sample then resolved by loading 15 µl onto an 8% polyacrylamide (29:1 Acrylamide/Bis, Bio-rad) non-denaturing gel in 0.5 x TBE running buffer and subjected to electrophoresis for 90 minutes at 25 mA. EMSA gels were exposed and visualized using a cyclone storage phosphor imager system (Packard Instrument Company, Inc.) for 18 hours at -80°C with final gel image analysis completed on Optiquant software version 4.0 (Packard Instrument Company, Inc.)

2.2 – Methodology – Antisense oligonucleotide inhibitors

2.2.1 EMSA Competition assay with specific anti-sense oligonucleotides

Radiolabelled KRAS-A (nts 1-185) was synthesized using *in vitro* transcription in the presence of [^{32}P]-UTP (method described in section 2.1.5) and used in EMSA with the parallel addition of competitor molecules. EMSA experimental conditions were identical to previously described experiments and can be reviewed in section 2.1.6. Anti-sense oligonucleotides (AONs) of 23-24-mer were designed to cover the entirety of the KRAS-A region in a complementary fashion end-to-end, and were synthesized by Integrated DNA Technologies Inc. (table 2.8). Screening of the anti-sense oligonucleotides, labelled AON-SM(1

Table 2.8. DNA sequences (5' to 3') of candidate inhibition antisense oligo nucleotides used in EMSA inhibition assays.

AON	DNA sequence
SM1	5' ACCACAAGTTTATATTCAGTCAT 3'
SM2	5' TGCCTACGCCACCAGCTCCAACT 3'
SM3	5' TAGCTGTATCGTCAAGGCACTCT 3'
SM4	5' TCGTCCACAAAATGATTCTGAAT 3'
SM5	5' AATCCTCTATTGTTGGATCATAT 3'
SM6	5' AATTACTACTTGCTTCCTGTAGG 3'
SM7	5' TCCAAGAGACAGGTTTCTCCATC 3'
SM8	5' TCATGACCTGCTGTGTCGAGAATA 3'

through 8), proceeded initially using two concentrations: 10-fold and 50-fold relative to the KRAS-A RNA probe concentration. For example, for a probe concentration of 1, the inhibitors would be used at 10 and 50. Effectiveness of the individual AONs were gauged by their ability to abrogate the interaction between KRAS-A RNA and CRD-BP, as indicated by the loss of an upper higher molecular weight EMSA gel band. Binding data was analyzed by densitometry of the EMSA autoradiograph using the Cyclone Storage Phosphor-System and Optiquant software to quantify the degree of complex inhibition and quantitatively assess the effectiveness of candidate AON molecules. Further characterization AONs SM-6 and SM-7 was pursued using a greater concentration range from 1X to 4000X molar-fold concentration (equivalent to 118 pM – 472 nM, see table 2.9). Additionally, AON SM-2 was further characterized as a negative control for inhibition with concentration tested from 1X to 1000X molar-fold (118 pM to 11.8 nM). Densitometry analysis using Optiquant software was utilized to produce inhibition plots using KaleidaGraph™ software.

Table 2.9. Concentration of antisense oligo nucleotides (AONs) used in EMSA inhibition assays. Molar-fold concentrations are based on a calculated 118 pM KRAS-A probe.

AON molar-fold	Concentration (nM)
1X	0.118
5X	0.590
10X	1.18
25X	2.95
50X	5.9
100X	11.8
250X	29.5
500X	59
750X	88.5
1000X	118
2000X	236
4000X	472

2.3 Results and Discussion

2.3.1 Synthesis of KRAS mRNA sub-regions by IVT

Generation of KRAS mRNA sub-region cDNA templates for use in *in vitro* transcription by means of PCR was successful with all six fragments, labelled KRAS A through F and are represented graphically in figure 2.2.

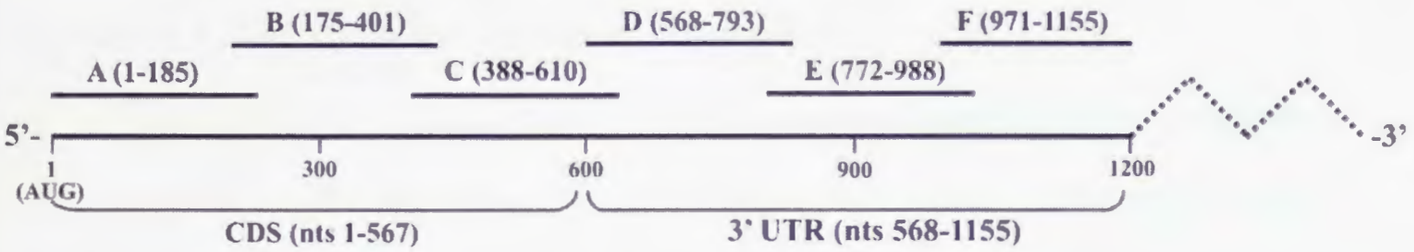


Figure 2.2. Graphical representation of KRAS mRNA sub-regions labelled KRAS-(A through F) covering the entire coding sequence and partial 3' UTR. Nucleotide positions in brackets are relative to the KRAS mRNA start codon (AUG) and the remaining unstudied 3' UTR region is represented by the double-peak dotted lines.

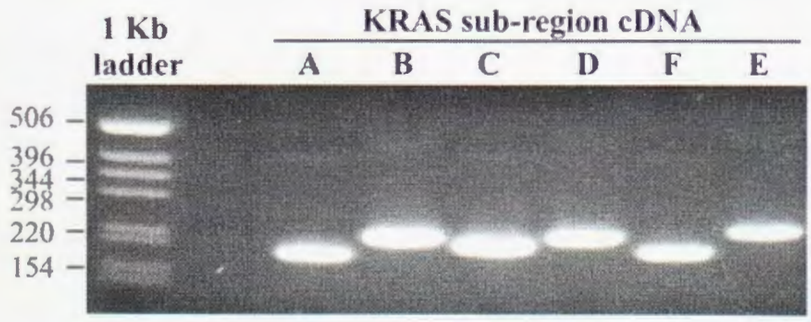


Figure 2.3. IVT cDNA templates for KRAS sub-regions A through F. DNA samples were run on a 1% agarose gel with invitrogen 1 Kb ladder and visualized with ethidium bromide staining.

DNA bands were of the expected size as determined by ethidium bromide UV visualization on a 1% agarose gel analysis (figure 2.3). Subsequent T7-driven in vitro transcription reactions produced [^{32}P]-internally labelled radioactive RNA probes of the expected relative size with the gel extraction purification method typically yielding 30 µl of 0.5-2 million cpm/µl. RNA sub-regions were designed to contain a small amount of sequence overlap with the neighboring fragments (approximately 20 nucleotides) to avoid splitting a potential binding-sequence or disrupting a specific RNA structure such as a hairpin loop. While the full

KRAS 3' UTR is approximately 5 Kb in length, only the first 588 nucleotides were covered. The rationale for this choice resides in the fact that CRD-BP typically binds target RNA either in the coding region as is the case for βTrCP1, GLI1, c-myc and N-myc mRNA, or within the first few hundred nucleotides of the 3' UTR as with β-actin, and MITF (Noubissi *et al.* 2006; Noubissi *et al.* 2009; Prokipcak *et al.* 1994; Chao *et al.* 2010; Goswami *et al.* 2010).

2.3.2 EMSA CRD-BP affinity mapping of KRAS mRNA

CRD-BP affinity for KRAS mRNA sub-regions was assessed using electrophoretic mobility shift assay (EMSA). Radioactive [^{32}P]-labelled RNA probes (20,000 cpm/reaction) covering the entire coding region and partial 3' UTR were titrated with CRD-BP and their ability to form a complex assessed by the formation of a band with reduced migration distance, indicative of a higher molecular weight complex formed by binding of CRD-BP with the RNA fragment. Firstly KRAS coding region fragments 'A'(nts 1-185), 'B'(175-401) and 'C'(388-610) were tested (figure 2.4.)

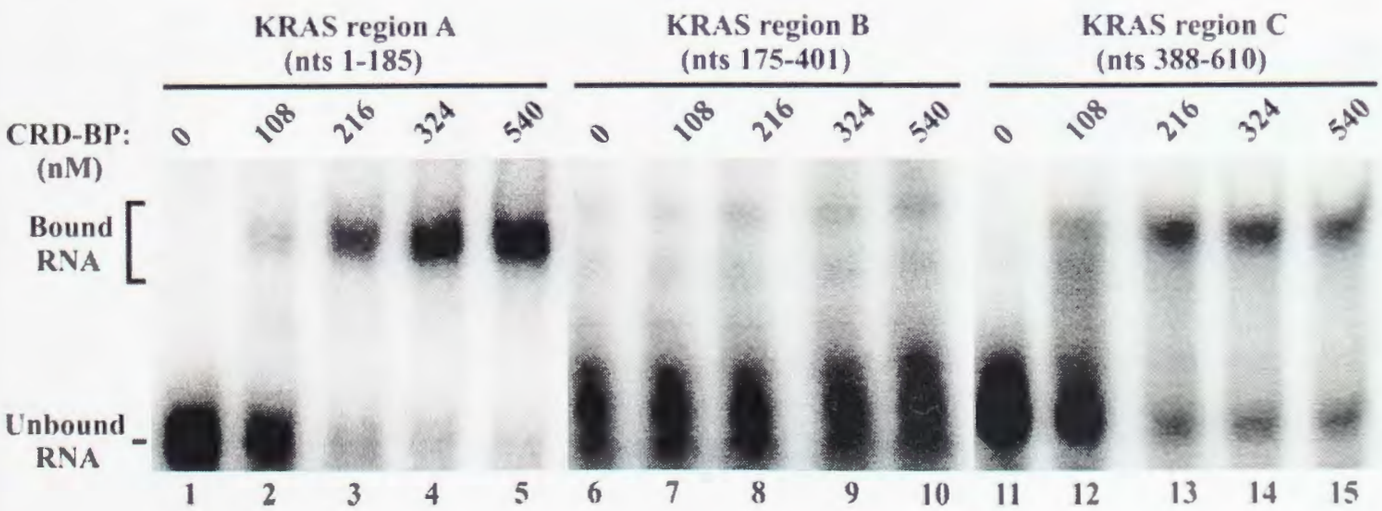


Figure 2.4. EMSA assessing KRAS coding sequence RNA sub-regions A to C for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lanes 1,6 and 11) forms a larger molecular weight complex upon titration with CRD-BP, as indicated by the bound RNA fraction (lanes 2-5, 7-10, 12-15). Binding reactions were incubated for a total of 20 minutes at 37°C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

Relatively strong CRD-BP affinity was found for sub-regions KRAS-A and KRAS-C, both being able to form a higher-order band suggesting the formation of an RNA-protein complex. When comparing lanes KRAS-A (lanes 1-5) to lanes KRAS-C (lanes 11-15), the two high affinity probes, it is clear that the region consisting of the first 185 nucleotides of the coding sequence forms a more distinct complex and has a larger bound fraction of RNA at 540 nM of CRD-BP. In contrast, KRAS-B (lanes 6-10) consistently showed a weak affinity with the gradual appearance of two faint higher bands but with a negligible change in the unbound RNA fraction up to 540 nM concentration, and thus was considered to be a relatively low affinity region and investigated no further. RNA fragments representing sub-regions of the partial KRAS 3' UTR mRNA were next tested for affinity in the same manner.

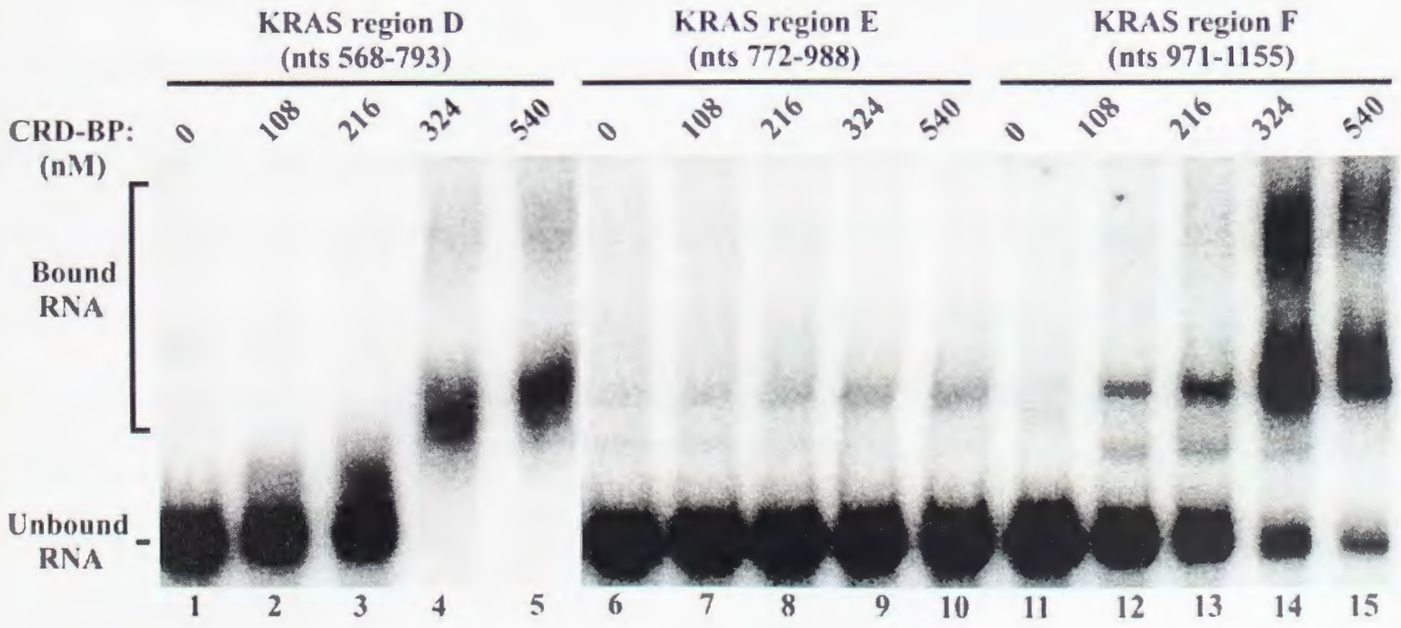


Figure 2.5. EMSA assessing KRAS 3' UTR RNA sub-regions D to F for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lanes 1,6 and 11) forms a larger molecular weight complex upon titration with CRD-BP, as indicated by the bound RNA fraction (lanes 2-5, 7-10, 12-15). Binding reactions were incubated for a total of 20 minutes at 37°C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

RNA KRAS-D and KRAS-F both displayed the ability to bind CRD-BP and a significant change in the unbound fraction of RNA occurred with increasing concentrations of CRD-BP. Interestingly, the KRAS-F region and to a lesser extent the KRAS-E region, forms multiple higher order complexes with a faint band forming just above the unbound RNA fraction in tandem with a more prominent band above that at twice the distance. This observation suggests that for only these regions (nts 971-1155 and 772-988), two binding events occur separately; however the phenomenon does not appear to correlate with overall affinity for CRD-BP. Both binding modes have been reported previously with various targets of CRD-BP/VICKZ proteins. IMP-1, which has a 99% sequence identity to CRD-BP, and is capable of binding as a single species to H19 mRNA and appears very similar to the EMSA shift pattern of KRAS-A. The IGF-II mRNA "zipcode" region binds IMP-1 in a sequential dimerization mode as determined by EMSA, much like KRAS-F and KRAS-E (Nielsen *et al.* 2002; Chao *et al.* 2010). Binding data for all six fragments was analyzed by densitometry of the autoradiograph using the Cyclone Storage Phosphor-System and Optiquant software and Kd plots generated by fitting to the Hill equation for quantitative comparison (figure 2.6). KRAS-C region was not included in the Kd calculation due to repeated inability to achieve greater than 70% bound fraction in the EMSAs. KRAS mRNA regions D (nts 568-793) and F (nts 971-1155), representing the first 588 nucleotides of the 3' UTR, also revealed multiple binding sites for CRD-BP (figure 2.5, lanes 1-5, and 11-15 respectively).

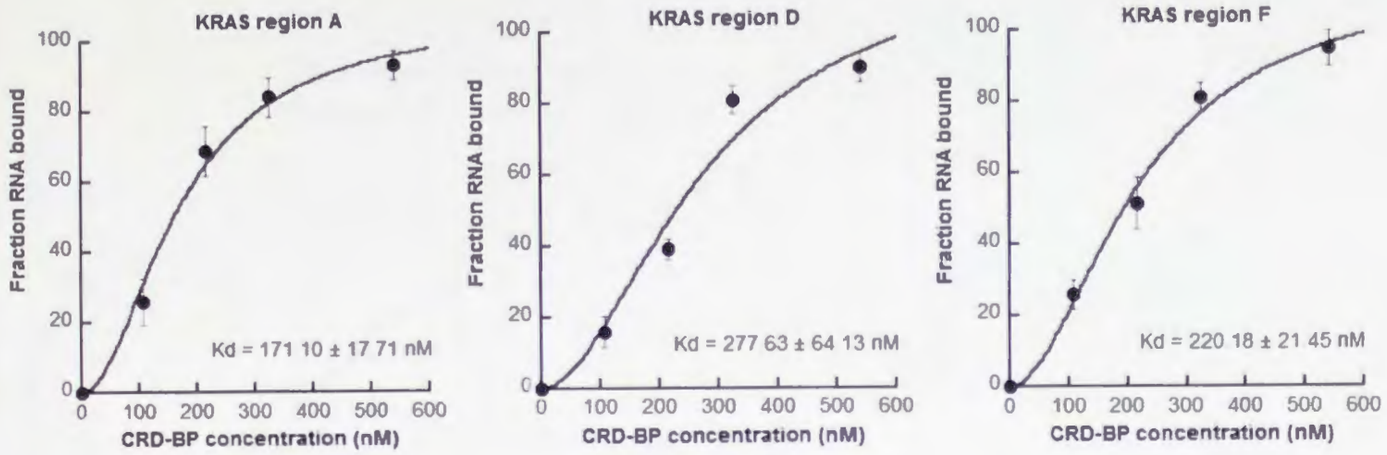


Figure 2.6 Kd plots for KRAS mRNA sub-regions A, D and F. Data points are averaged from two biological sets of triplicate experiments and fit to the Hill equation for Kd determination (Hill coefficient = 2); error bars represent standard deviation.

Comparing the calculated Kd values doubly confirmed KRAS-A as the KRAS mRNA sub-region possessing the highest affinity binding site of the examined fragments, the calculated Kd value for KRAS-A (171.10 ± 17.71 nM) being lower than KRAS-F and almost 1.6-fold lower than KRAS-D (220.18 ± 21.45 nM and 277.63 ± 64.13 nM, respectively). Further binding characterization of the KRAS-A fragment was done with additional CRD-BP concentrations to more accurately determine the binding character and similarly a saturation binding curve plotted to determine the Kd and Hill Coefficient (figure 2.7 and 2.8, respectively).

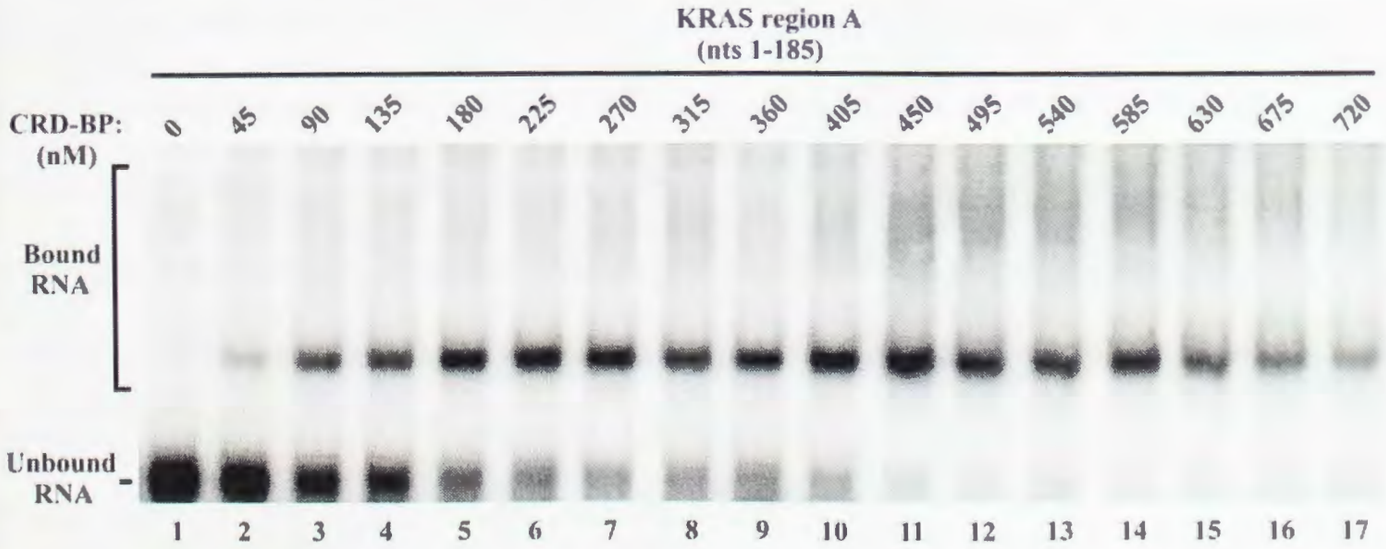


Figure 2.7. EMSA saturation binding experiment between CRD-BP and KRAS-A (nts 1-185). RNA (20,000 cpm) with no protein added (lane 1) forms a larger molecular weight complex upon titration with CRD-BP, as indicated by the bound RNA fraction (lanes 2-17).

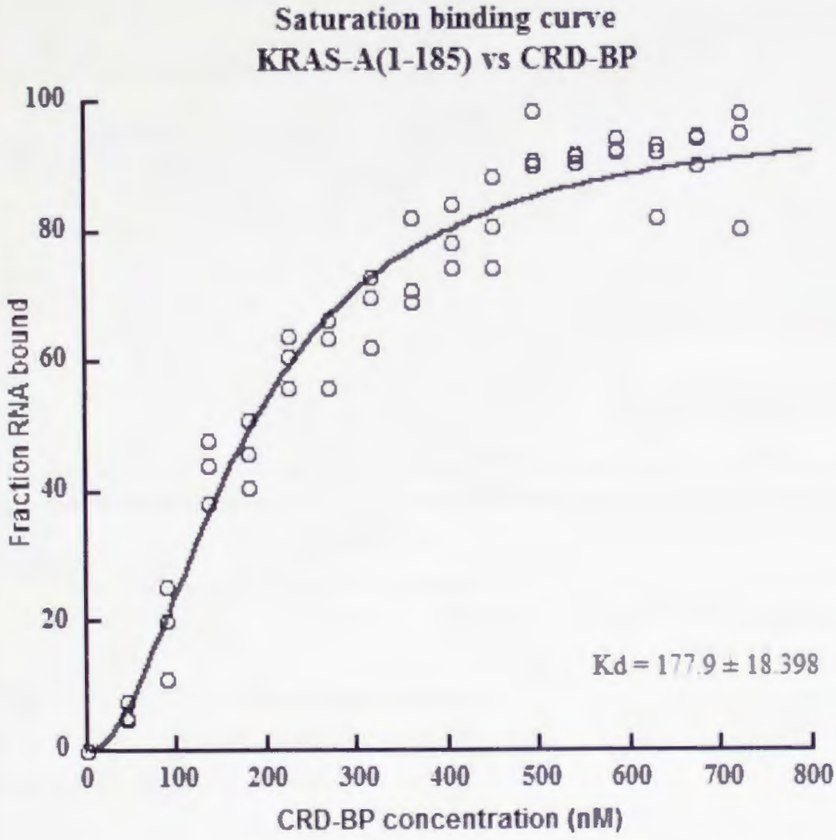


Figure 2.8. EMSA Saturation binding curve for CRD-BP vs KRAS-A (nts 1-185). Data was derived in triplicate from saturation EMSA experiments where CRD-BP was incubated with [^{32}P]-labelled KRAS-A RNA(20,000 cpm) and fit to the Hill equation.

Further refinement of the KRAS binding sequence was pursued in a similar manner by producing 5' and 3' truncations of the high affinity KRAS-A region. Small cDNA templates covering arbitrarily assigned sub-regions of KRAS-A (1 through 4) were successfully generated using PCR (figure 2.10). If upon retention or removal of one of the four approximately 50 nucleotide regions resulted in drastic changes in CRD-BP affinity, it could be deduced by combining data where the high affinity site(s) is located. IVT reactions carried out in the presence of radio-labelled UTP were used to synthesize RNA strands corresponding to the representative mRNA regions as depicted in figure 2.9.

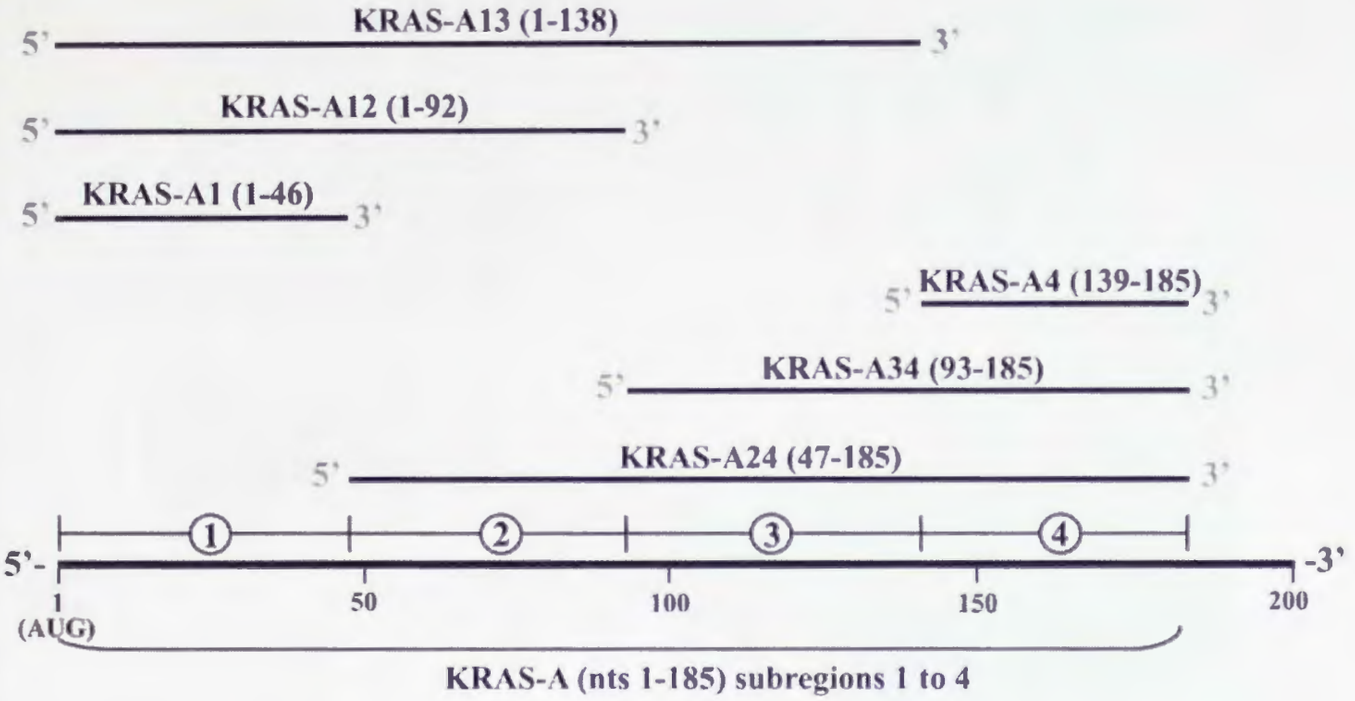


Figure 2.9. Graphical representation of KRAS-A(1 through 4) RNAs. Fragments cover the first 185 nucleotides of the coding sequence and nucleotide positions in brackets are relative to the KRAS mRNA start codon (AUG).

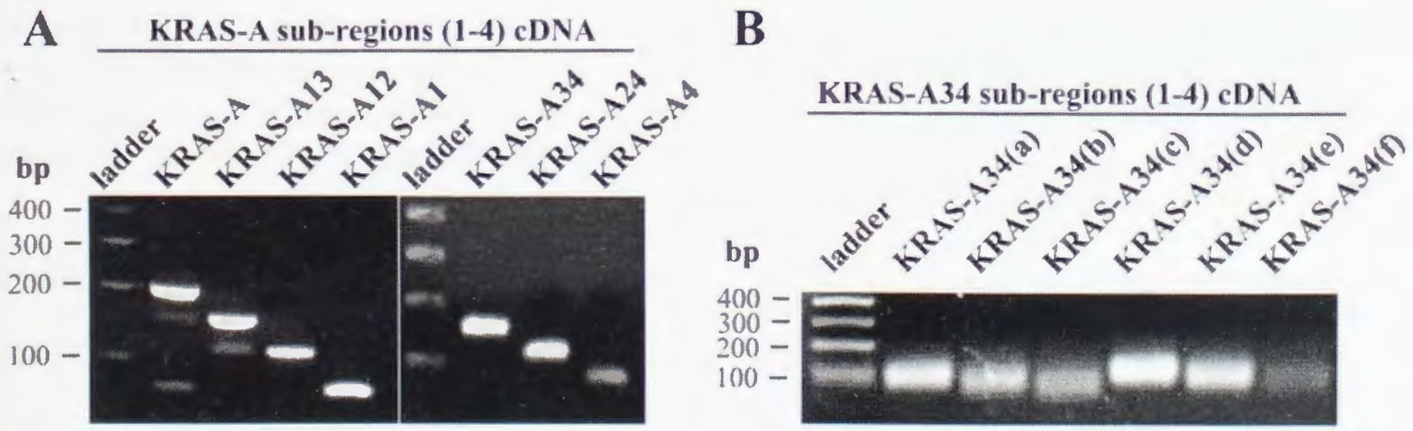


Figure 2.10. KRAS sub-region cDNA templates for in vitro transcription. (A) 5' and 3' truncations of KRAS region A and (B) 5' and 3' truncations of KRAS region A34 resolved on a 1% agarose gel visualized by UV-ethidium bromide staining.

Synthesized RNA fragments were again assessed for relative affinity using EMSA and [^{32}P]-labelled RNA probes (20,000 cpm/reaction) titrated with CRD-BP up to 540 nM. Results of the binding assays for the 3' truncated fragments A13, A12 and A1 can be seen in figure 2.11 in lanes 1-5, 6-10, and 11-15 respectively.

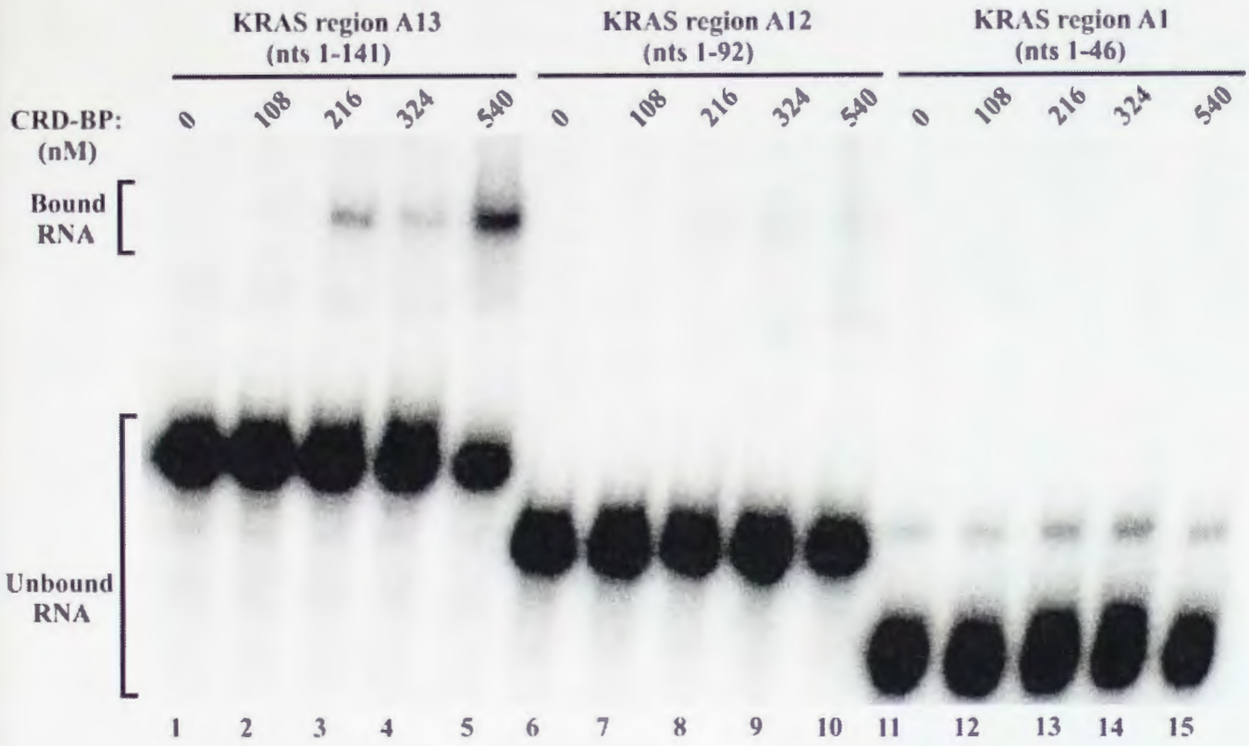


Figure 2.11. EMSA assessing KRAS-A 3' truncated sub-regions A13, A12 and A1 for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lanes 1,6 and 11) forms a larger molecular weight complex with CRD-BP (lanes 2-5, 7-10, 12-15). Binding reactions were incubated for a total of 20 minutes at 37°C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

It is immediately apparent upon visual inspection of lanes 1 to 5 (figure 2.11) that only the KRAS-A13 region corresponding, to nucleotide positions 1-141 of KRAS mRNA retained the ability to bind CRD-BP with measurable affinity. From the systematic 3' truncation of KRAS-A, it was determined that KRAS-A12, one full half of KRAS-A consisting of 92 nucleotides, did not likely possess any binding information. Further EMSA binding experiments were carried out using 5' truncated fragments of KRAS-A to examine the possibility of a binding sequence present in the 3' end of KRAS-A (figure 2.12).

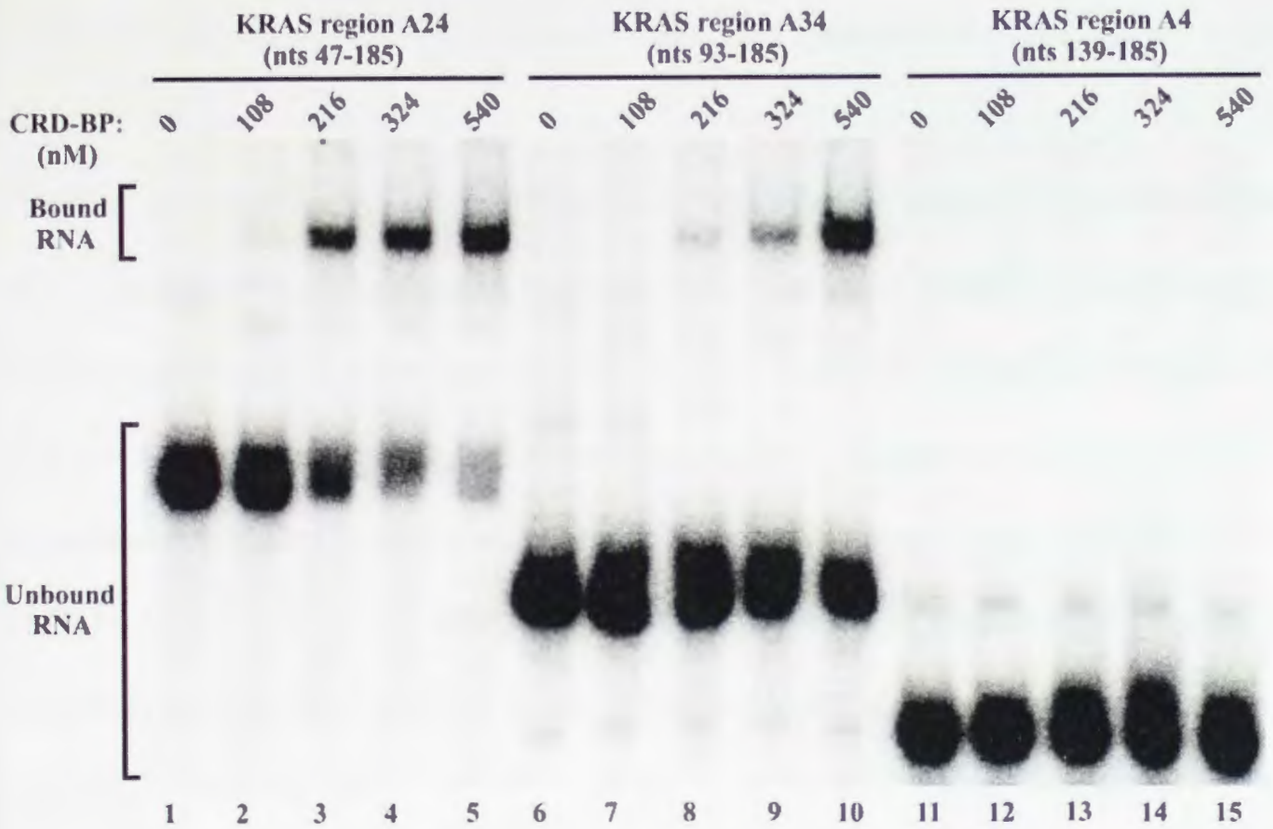


Figure 2.12. EMSA assessing KRAS-A 5' truncated sub-regions A24, A34 and A4 for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lanes 1,6 and 11) forms a larger molecular weight complex with CRD-BP (lanes 2-5, 7-10, 12-15). Binding reactions were incubated for a total of 20 minutes at 37°C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

From the EMSA binding assays with 5' truncations of KRAS-A in figure 2.12, It can be seen that both KRAS-A24 and KRAS-A34 are both capable of forming a protein-RNA complex while KRAS-A4, the tailing 47 nucleotide 3'end, shows no interaction. A faint band in KRAS-A4 can be seen directly above the unbound RNA, however it is present in the absence of any protein and its magnitude does not change with increasing CRD-BP concentration and is thus considered a non-binding region. To further confirm the binding disparity between KRAS-A12 and KRAS-A34, another experiment was completed pitting the two fragments side-by-side on the same gel and effectively doubling the protein concentration to achieve binding saturation for Kd determination (figures 2.13 and 2.14). Direct comparison of the 2 regions show that KRAS-A34 binds to CRD-BP more aggressively than KRAS-A12, despite a size difference of only one

nucleotide, showing that there is a specific recognized element in the A34 region and that the affinity is not due to non-specific binding. Returning to KRAS-A24 and KRAS-A34, it can be observed that while both clearly show complex formation, the larger KRAS-A24 (139 nt) has an affinity similar to the full KRAS-A fragment, with nearly all of the RNA present in the bound fraction by 540 nM CRD-BP concentration. The 139 nucleotide size of KRAS-A24 is consistent with literature, in that most VICKZ protein binding RNA elements are greater than 100 nucleotides, with the β-actin zipcode being an exception (Yisraeli *et al.* 1990; Nielsen *et al.* 2003; Prokipcak *et al.* 1994). KRAS-A12 by itself proved to be a poor binding region, suggesting that perhaps it was region KRAS-A3 that contained the binding sequence as only fragments containing region-3 of KRAS-A displayed binding. Hence, a small [^{32}P]-labelled RNA probe (20,000 cpm/reaction) with sequence representing region KRAS-A3 was synthesized and tested for CRD-BP affinity by the same method (figure 2.15).

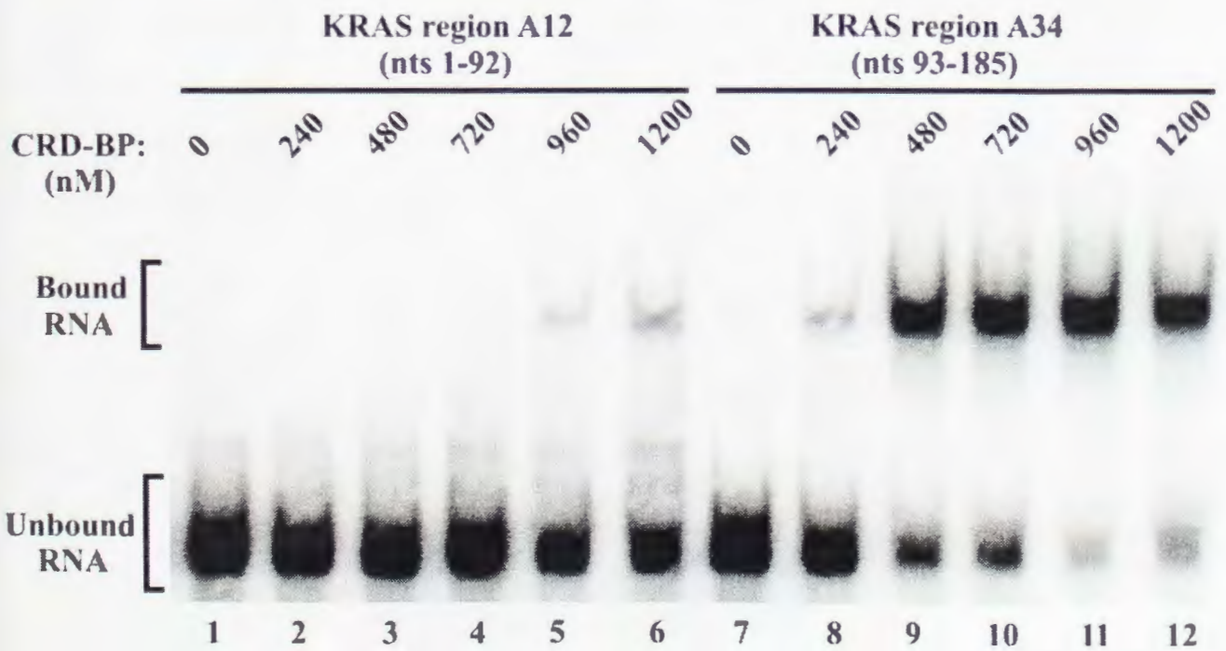


Figure 2.13. EMSA assessing RNAs KRAS-A12 and KRAS-A34 for CRD-BP binding. [^{32}P]-labelled unbound KRAS-A12 and KRAS-A34 (20,000 cpm/reaction) with no protein added (lane 1 and 7, respectively) forms a larger molecular weight complex upon titration with CRD-BP for the KRAS-A34 RNA fragment but only to a limited extent with KRAS-A12 up to 1200 nM, as indicated by the bound RNA fraction (KRAS-A12: lane 2-6; KRAS-A34: lanes 8-12).

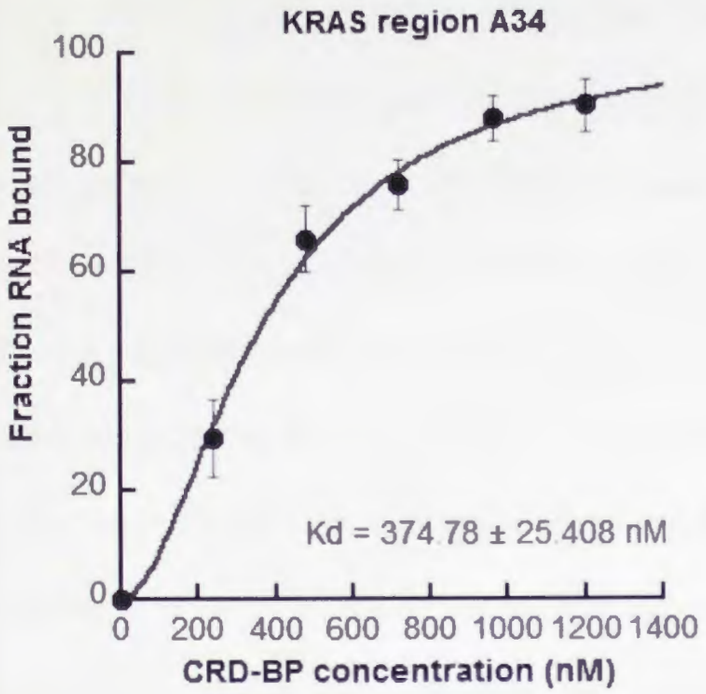


Figure 2.14. Saturation binding curve for CRD-BP vs KRAS-A34 RNA (nts 93-185). Data was produced from saturation EMSA experiments shown in figure 13 where CRD-BP was incubated with [^{32}P]-labelled KRAS-A34 and KRAS-A12 RNA(20,000 cpm) and fit to the Hill equation.

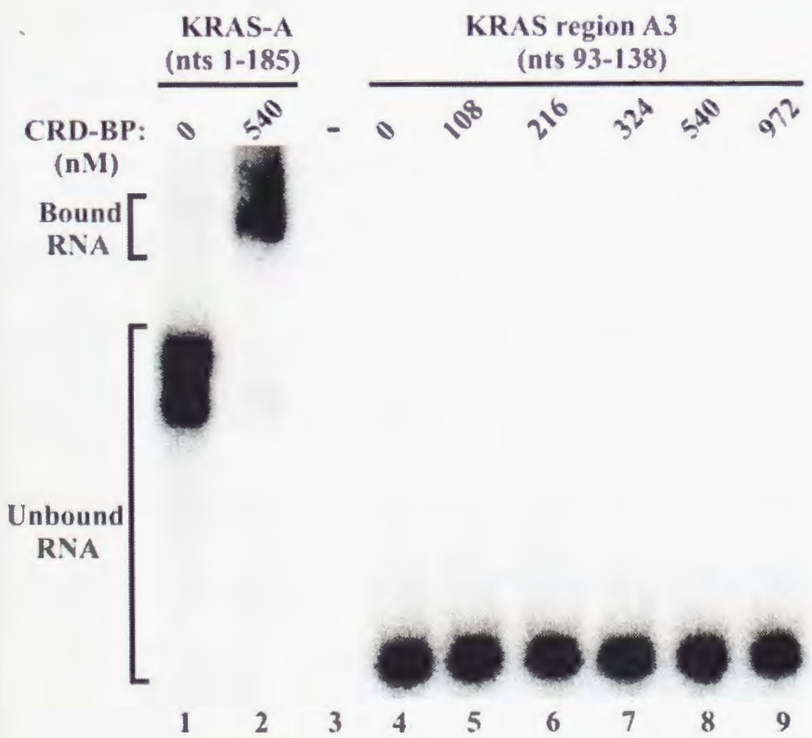


Figure 2.15. EMSA assessing KRAS-A3 RNA for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lane 1 and 4) forms a larger molecular weight complex upon titration with CRD-BP only for the control KRAS-A RNA fragment but not with A3, as indicated by the bound RNA fraction (lane 2).

KRAS-A3 did not show any significant binding, even with CRD-BP concentration up to 972 nM. This unexpected result may have been due to an unforeseen requirement of nearby sequence regions (specific or non-specific) outside region A3 that are instrumental in proper RNA folding for recognition by CRD-BP. Also, it could be that binding of KRAS-A is achieved through a bipartite mode, whereby multiple binding sites exist along the RNA strand to facilitate full binding. In this scenario, a single site may fall within region KRAS-A3, while another site(s) may exist in the flanking regions KRAS-A2 and KRAS-A4 such that only a combination of two binding events are required for stable binding. Evidence for this scenario comes from SELEX and RIP-chip experiments identifying conserved 5' and 3' sequences in many amplified RNA targets of VICKZ proteins, suggesting multiple binding sites are important for any given RNA target. However, among the cohort of amplified RNA targets, some sequences were present that did not share any obvious sequence similarity with the others identified (Farina *et al.* 2003; Patel *et al.* 2012).

Further mapping of the CRD-BP binding sequence identified in KRAS-A34 was carried out in one more set of 5' and 3' truncations of KRAS-A34 according to the graphic below (figure 2.16). While the previous systematic truncation of KRAS-A into four roughly equally sized segments (regions 1 through 4), this series of truncations aimed instead to simply shave off terminal nucleotides from both the 5' and 3' end of KRAS-A34, to find the smallest fragment that can still produce a complex as determined by EMSA.

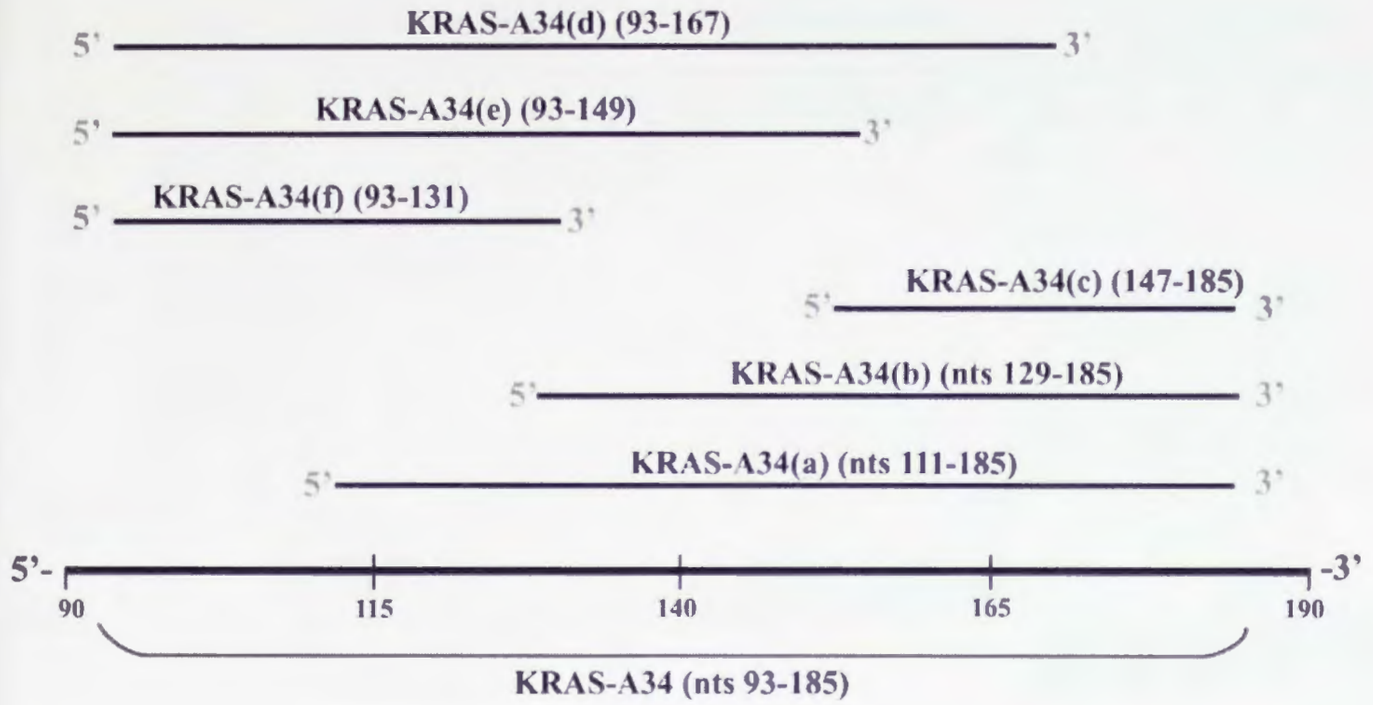


Figure 2.16. Graphical representation of KRAS-A34(a-f) RNAs. Fragments represent sub-divisions of the KRAS-A34 region, covering nucleotides 93-185 of the coding sequence. Nucleotide positions in brackets are relative to the KRAS mRNA start codon (AUG).

To further narrow down a minimum CRD-BP binding sequence, an additional set of RNA fragments representing sub-division of the parent KRAS-A34 positive binding region were synthesized. The ability of the 3' truncated fragments to bind CRD-BP was assessed by EMSA and the results for RNAs labelled KRAS-A34(d, e or f) can be seen in figure 2.17. Analyzing the results of the 3' truncated fragments of binding region KRAS-A34 showed virtually no binding. As with the previously discussed fragment KRAS-A4 which presented with a faint upper band but did not correlated with CRD-BP concentration, it can be seen that KRAS-A34(d) - covering nucleotides 93-167, also presents with a similar pattern, and is concluded to be a non-binding fragment.

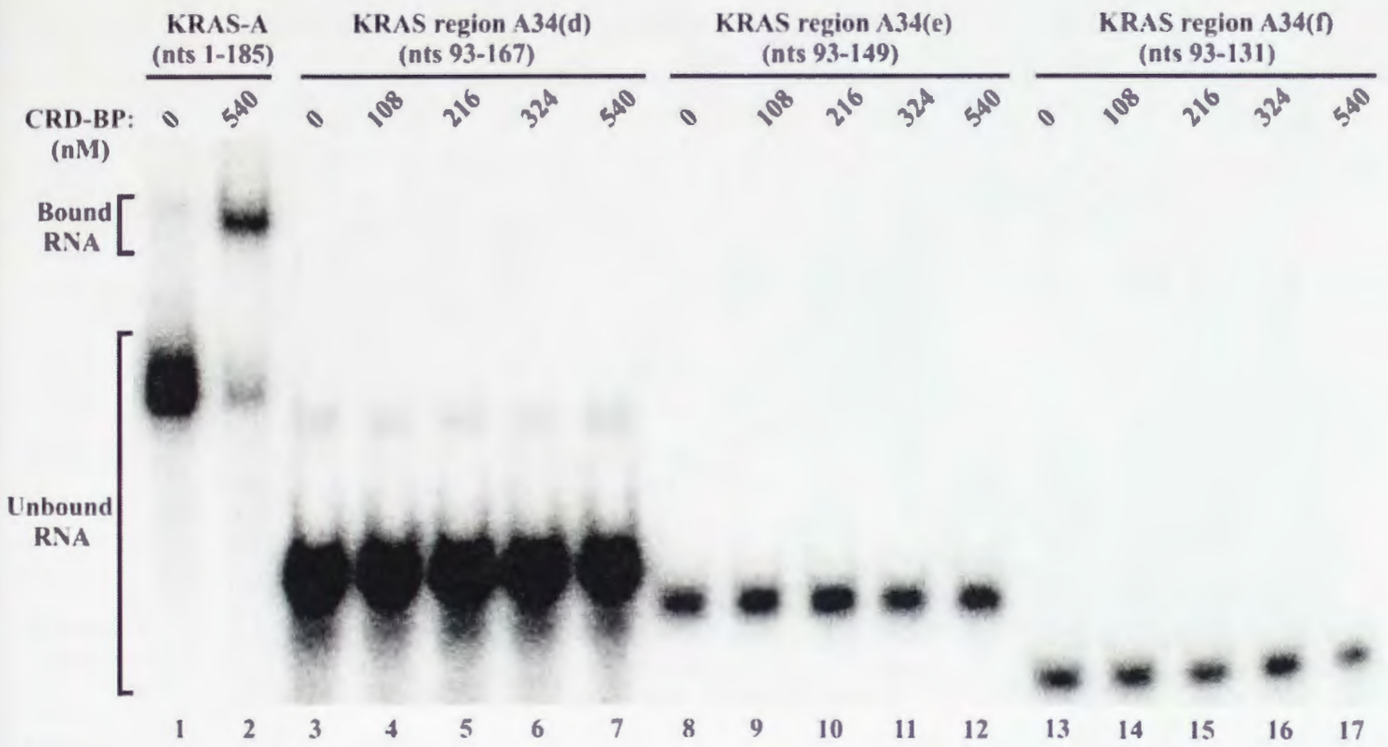


Figure 2.17. EMSA assessing the KRAS-A34(d, e and f) RNAs, for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lanes 1, 3, 8 and 13) forms a larger molecular weight complex upon titration with CRD-BP only in the case of the KRAS-A probe, which was used as positive control and included to demonstrate binding contrast for these non-binding fragments. Binding reactions were incubated for a total of 20 minutes at 37°C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

To ensure that the lack of complex formation in any of the 3' truncated fragments was not due to experimental error, a positive control previously determined to bind (KRAS-A) was included. From this, it was concluded that the initial 3' truncation, the largest of the three which removed 19 nucleotides to produce region KRAS-A34(d), likely contained binding elements. The complementary 5' truncated fragments, labelled as KRAS-A34(a, b and c) which retained the aforementioned 19 nucleotide sequence, were also tested for binding using EMSA (figure 2.18.) Complex formation can be readily observed as a distinct band above the unbound fraction of RNA in fragments KRAS-A34(a) and KRAS-A34(b) with binding almost entirely lost in KRAS-A34(c).

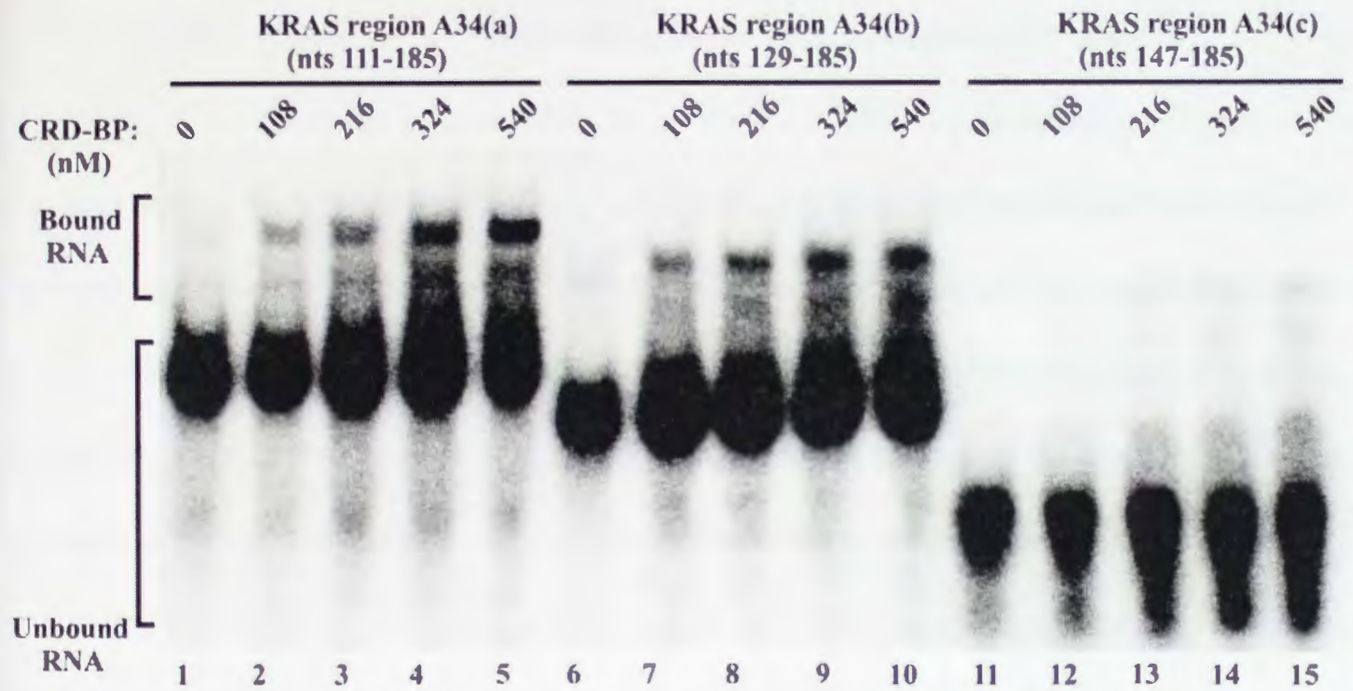


Figure 2.18. EMSA assessing KRAS-A34(a, b and c) for CRD-BP binding. [^{32}P]-labelled unbound RNA (20,000 cpm/reaction) with no protein added (lanes 1, 6 and 11) forms a larger molecular weight complex upon titration with CRD-BP as indicated by formation of a band in the area labelled "bound RNA". Binding reactions were incubated for a total of 20 minutes at 37°C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

Oddly, the addition of CRD-BP appears to shift the unbound KRAS-A34(c) RNA further down the gel, increasing migration distance as indicated by the increasingly dense autoradiographic exposure appearing below the free RNA in lanes 11 to 15. While this may be considered binding, a lack of a distinct band suggests at most non-specific interaction and was left as an anomaly. Interestingly it is possible to distinguish an intermediate band between the upper shifted complexes, and the unbound RNA which was not readily apparent in the non-truncated KRAS-A34 fragment. Hence, the smallest region of KRAS mRNA presented here capable of binding CRD-BP as determined by EMSA is the 57 nucleotide region labelled KRAS-A34(b); nucleotides 129-185.

2.3.3 Development of specific antisense oligonucleotide inhibitors

Development of molecular inhibitors of the KRAS-CRD-BP interaction is paramount to the long term goals of studying CRD-BP and its mRNA targets. With CRD-BP under strict spatiotemporal control during the earliest weeks of embryogenesis and tapering off to extremely low or undetectable levels in later life, its presence and activity in tumor tissue is a clear anti-proliferative target given its involvement in several growth pathways. Elevated KRAS signalling has been demonstrated to drive certain cancers and it has been shown that CRD-BP is capable of enhancing expression of this oncogene (Mongroo et al. 2011) singling out the KRAS-CRD-BP interaction as an ideal target for molecular intervention. To approach this task, 23-24 mer DNA-based antisense oligonucleotides were designed to contiguously cover the entire region of KRAS-A in a complementary fashion (nts 1-185) with the rationale being that important RNA sequence or structural elements within KRAS-A, once bound through Watson-Crick base pairing, will impede CRD-BP binding to KRAS mRNA. Any discovered DNA-based oligonucleotide inhibitors could subsequently be synthesized as an O-methyl RNA oligonucleotide derivative and used in the context of cell-based anti-proliferative assays for further development. Under binding conditions identical to the previously outlined EMSA binding experiments, [^{32}P]-labelled KRAS-A (nts 1-185) RNA was incubated with 216 nM CRD-BP in the presence of eight potential inhibitor oligos at 10X and 50X molar-fold concentrations (figure 2.19). Protein concentration was chosen to only fractionally bind the RNA probe (as opposed to full shift at 540 nM) to avoid over saturation of the RNA and increase the protein-RNA complex sensitivity to the inhibitory effect of the oligonucleotides for screening. *In vitro* transcribed unlabelled KRAS-A RNA was also incubated with CRD-BP at 10X and 50X concentrations to serve as a positive control for inhibition (figure 2.19, lanes 3-4). Concentration

of radio-labelled KRAS-A probe was determined to be 118 pM and thus competitor molecules were made to 1.18 nM and 5.9 nM respectively (10X and 50X). Successful inhibitors were then selected based on ability to reduce the bound fraction of RNA in the EMSA.

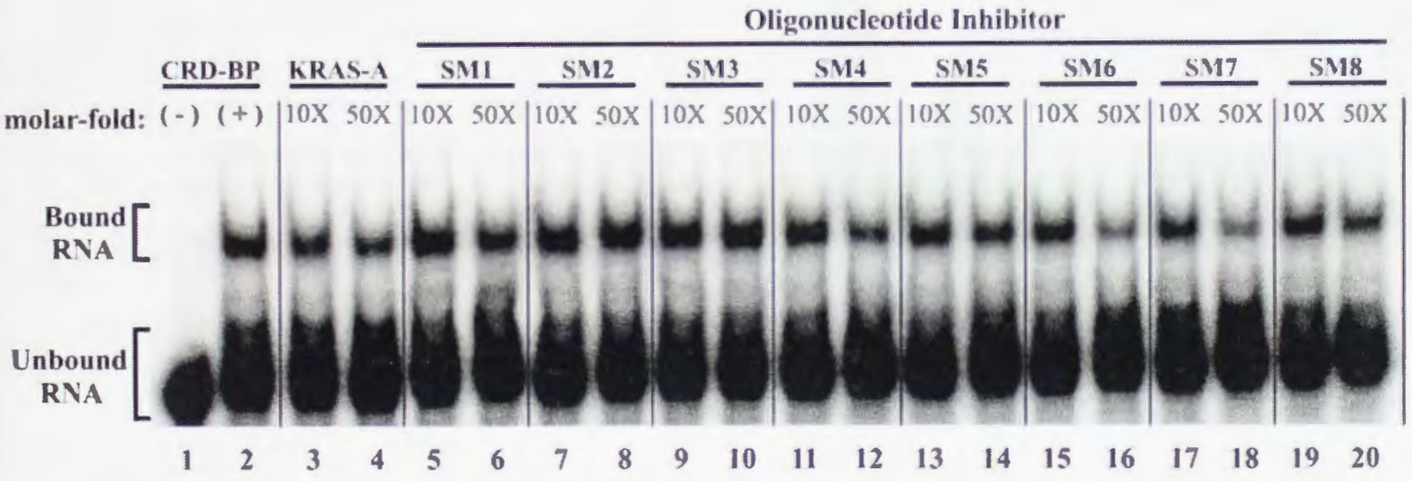


Figure 2.19. EMSA screen of antisense oligonucleotide AON-SM(1-8) inhibitors. [^{32}P]-labelled KRAS-A RNA probe (-/lane 1) was partially shifted by 216 nM CRD-BP (+/lane 2) in the presence of 23-24-mer antisense oligonucleotides in 2 concentrations: 10X molar excess relative to 118 pM radioactive probe (1.18 nM), and 50X (5.9 nM) in lanes 3-20. Effectiveness of potential inhibitors is evidenced by a reduction in the bound RNA fraction indicated on the figure left. Error bars are standard deviation.

Screening of SM1 to SM8 inhibitor oligonucleotides was completed in triplicate and when the bound RNA in the protein-only reaction (lane 2) is compared to those with inhibitors added, it can be seen in the representative figure that a reduction in complex formation occurred for inhibitors SM6 (lanes 15-16), SM7 (lanes17-18) and to a lesser extent SM4 (lanes 11-12). Cold, unlabelled KRAS-A which was used as a positive control for inhibition did display a reduction of the complex as expected, however the extent of inhibition was somewhat less than predicted possibly due to error in RNA quantification of either the radioactive probe, or the unlabelled competitor RNA. Densitometry of the autoradiographs using the Cyclone Storage Phosphor-System and Optiquant software was used to construct a graphical representation of the EMSA

(figure 2.21, lanes 13-17). As stated earlier in the introduction, KRAS is a bit of a mixed bag and in addition to acting as an oncogene, in some circumstances it can act as a tumor suppressor (Zhang *et al.* 2001). In such cases where CRD-BP correlated KRAS expression may be acting as a tumor suppressor, enhancing CRD-BP binding using oligonucleotides (or small molecules with the same effect) may have beneficial uses.

Further quantification of SM6 and SM7 was carried out by expanding the concentration range of the inhibitor oligonucleotides from 1X to 1000X, or 118 pM to 118 nM (figure 2.21). EMSA conditions remained the same except for CRD-BP concentration was set at 540 nM to achieve maximal KRAS-A shift, and allow characterization of SM6 and SM7 over the entire KRAS-CRD-BP binding range. Here, SM2 was also included for comparison based on the initial anti-sense oligonucleotide screening that suggested it may function as a good negative control.

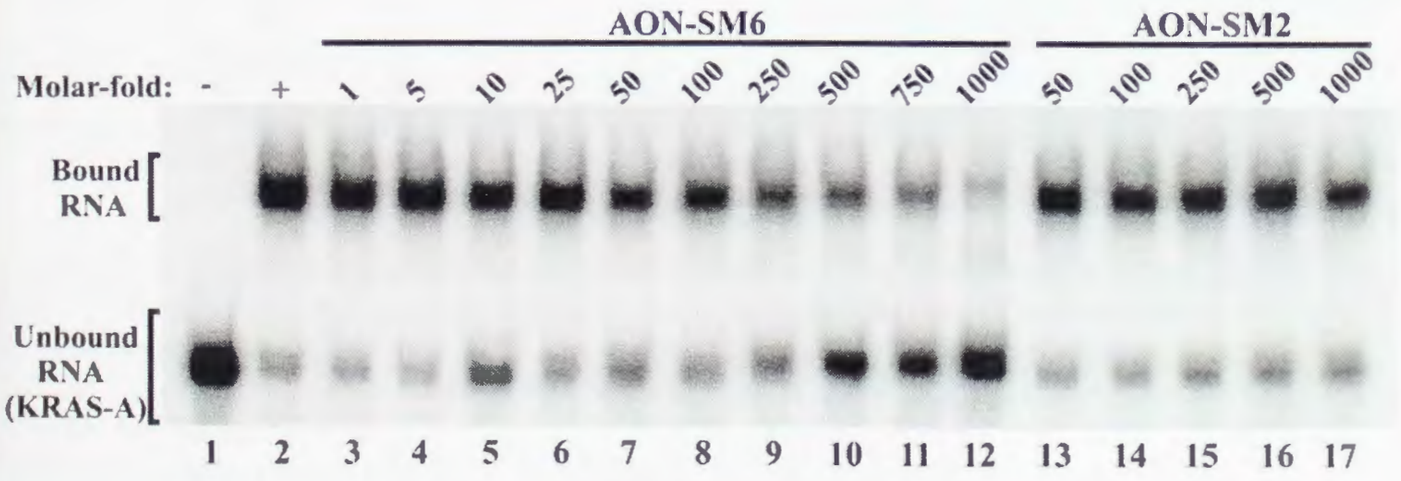


Figure 2.21. EMSA characterizing CRD-BP-KRAS-A RNA complex inhibition by AON-SM6. [^{32}P]-labelled RNA - region KRAS-A probe (unbound in lane 1) was incubated with 540 nM CRD-BP (lane 2). Inhibitor SM6 was present (lanes 3-12) in otherwise identical conditions to lane 2, ranging in concentration from 1X molar-fold relative to probe concentration (118 pM) to 1000 molar-fold (118 nM) which correlated with a loss of CRD-BP- bound RNA. In contrast AON-SM2 displayed drastically reduced ability to interrupt the CRD-BP-KRAS-A RNA complex.

of radio-labelled KRAS-A probe was determined to be 118 pM and thus competitor molecules were made to 1.18 nM and 5.9 nM respectively (10X and 50X). Successful inhibitors were then selected based on ability to reduce the bound fraction of RNA in the EMSA.

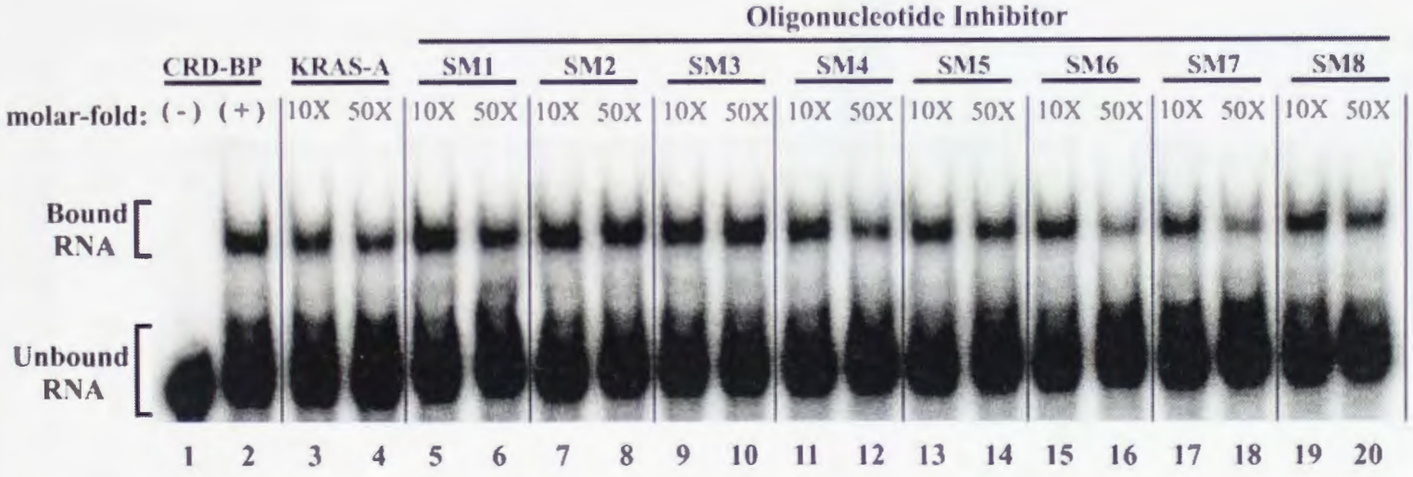


Figure 2.19. EMSA screen of antisense oligonucleotide AON-SM(1-8) inhibitors. [^{32}P]-labelled KRAS-A RNA probe (-/lane 1) was partially shifted by 216 nM CRD-BP (+/lane 2) in the presence of 23-24-mer antisense oligonucleotides in 2 concentrations: 10X molar excess relative to 118 pM radioactive probe (1.18 nM), and 50X (5.9 nM) in lanes 3-20. Effectiveness of potential inhibitors is evidenced by a reduction in the bound RNA fraction indicated on the figure left. Error bars are standard deviation.

Screening of SM1 to SM8 inhibitor oligonucleotides was completed in triplicate and when the bound RNA in the protein-only reaction (lane 2) is compared to those with inhibitors added, it can be seen in the representative figure that a reduction in complex formation occurred for inhibitors SM6 (lanes 15-16), SM7 (lanes17-18) and to a lesser extent SM4 (lanes 11-12). Cold, unlabelled KRAS-A which was used as a positive control for inhibition did display a reduction of the complex as expected, however the extent of inhibition was somewhat less than predicted possibly due to error in RNA quantification of either the radioactive probe, or the unlabelled competitor RNA. Densitometry of the autoradiographs using the Cyclone Storage Phosphor-System and Optiquant software was used to construct a graphical representation of the EMSA

data, where any inhibitory effect was expressed as a reduction in the complex intensity, with 0% being no change and 100% being complete loss of the bound RNA band (figure 2.20).

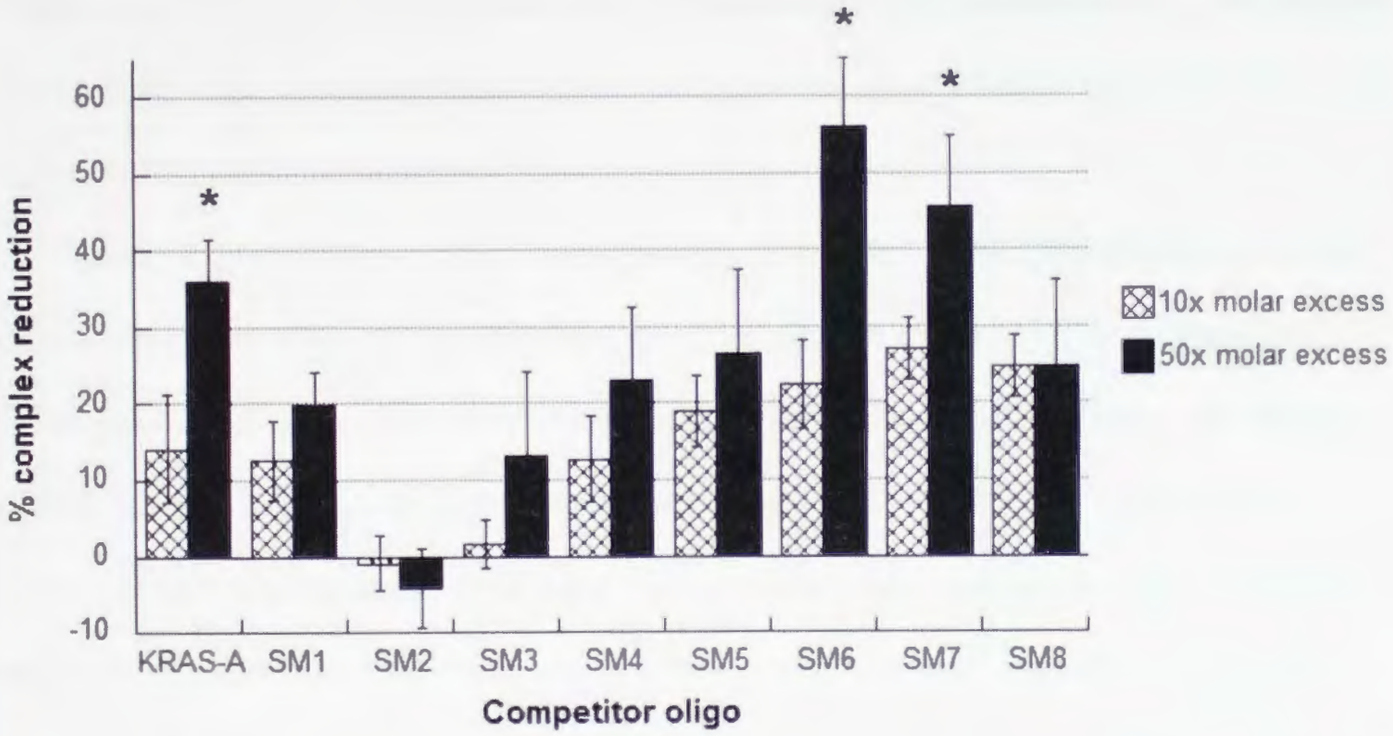


Figure 2.20. Bar graph of antisense oligonucleotide molecules SM(1-8) inhibitory effect on CRD-BP-KRAS RNA complex formation. Data is presented as the average percent reduction of bound RNA compared to a control with no inhibitor present in an electrophoretic mobility shift assay. Data was subjected to statistical analysis using a one-way ANOVA; $F=3.47$, $p<0.02$, $n=3$ and student's t-test ($p<0.05$). Antisense oligonucleotides (23-24 mer) were present in 2 concentrations: 10X molar excess relative to 118 pM radioactive probe (1.18 nM), and 50X (5.9 nM).

Plotting the antisense inhibitor data as a bar graph in figure 2.20 more clearly depicts the relatively strong ability for SM6 and SM7 to reduce the CRD-BP-KRAS RNA complex formation. At 50X molar excess (5.9 nM), SM6 and SM7 both achieved approximately half reduction of the bound RNA band, twice the effectiveness of the next most effective inhibitor. Interestingly, SM2 and SM3 appeared to be capable of strengthening the KRAS-A-CRD-BP complex formation in these experimental conditions. This phenomenon of increasing the bound RNA fraction was repeatable in the initial screen, but was not evident at higher concentration

(figure 2.21, lanes 13-17). As stated earlier in the introduction, KRAS is a bit of a mixed bag and in addition to acting as an oncogene, in some circumstances it can act as a tumor suppressor (Zhang *et al.* 2001). In such cases where CRD-BP correlated KRAS expression may be acting as a tumor suppressor, enhancing CRD-BP binding using oligonucleotides (or small molecules with the same effect) may have beneficial uses.

Further quantification of SM6 and SM7 was carried out by expanding the concentration range of the inhibitor oligonucleotides from 1X to 1000X, or 118 pM to 118 nM (figure 2.21). EMSA conditions remained the same except for CRD-BP concentration was set at 540 nM to achieve maximal KRAS-A shift, and allow characterization of SM6 and SM7 over the entire KRAS-CRD-BP binding range. Here, SM2 was also included for comparison based on the initial anti-sense oligonucleotide screening that suggested it may function as a good negative control.

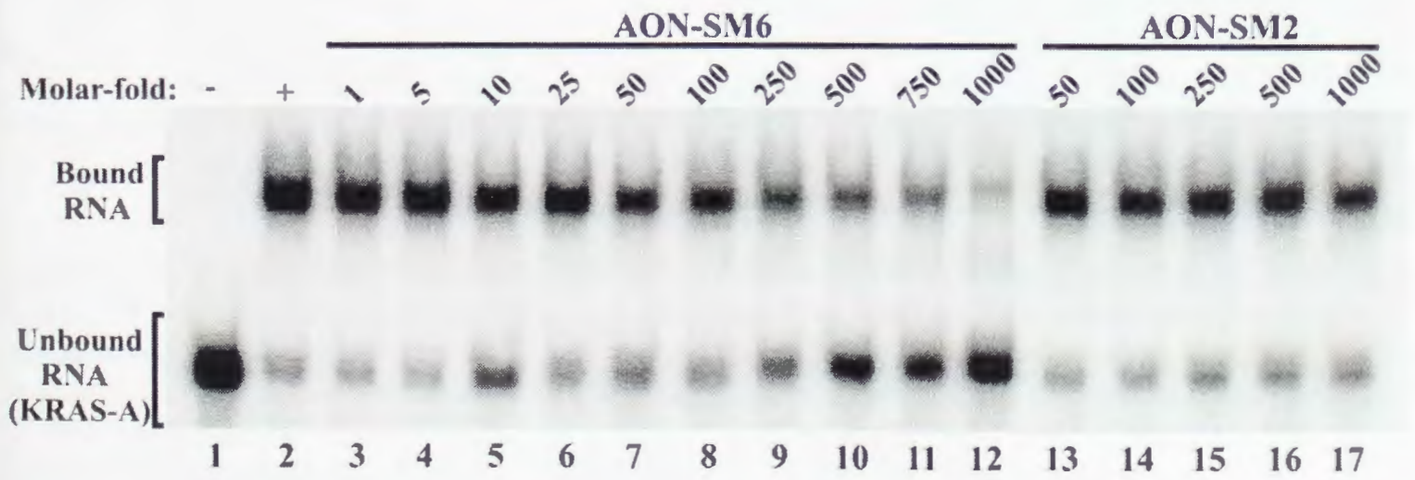


Figure 2.21. EMSA characterizing CRD-BP-KRAS-A RNA complex inhibition by AON-SM6. [^{32}P]-labelled RNA - region KRAS-A probe (unbound in lane 1) was incubated with 540 nM CRD-BP (lane 2). Inhibitor SM6 was present (lanes 3-12) in otherwise identical conditions to lane 2, ranging in concentration from 1X molar-fold relative to probe concentration (118 pM) to 1000 molar-fold (118 nM) which correlated with a loss of CRD-BP- bound RNA. In contrast AON-SM2 displayed drastically reduced ability to interrupt the CRD-BP-KRAS-A RNA complex.

Inspection of figure 2.21 reveals SM6 as a capable KRAS-CRD-BP inhibitor, as evidenced by the gradual disappearance of the bound fraction of RNA over the oligo concentration range. While this effect is mirrored to some extent by SM2, with a visible increase in the free RNA fraction over the same concentration range, it is clear SM6 is far more active and suggests a degree of specificity for the observed effect beyond non-specific interaction. Inhibitor AON-SM7 characterization was also expanded in parallel under identical conditions. However, in an attempt to remove the CRD-BP-KRAS complex entirely, the concentration range of SM7 was extended to 4000-fold (472 nM) as can be seen in figure 2.22.

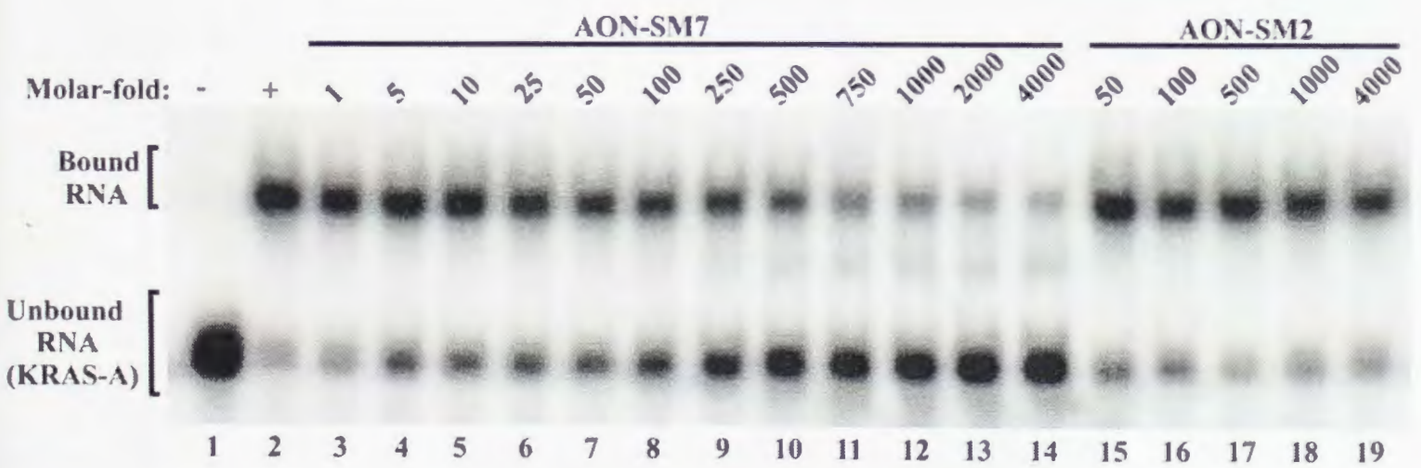


Figure 2.22. EMSA characterizing CRD-BP-KRAS-A RNA complex inhibition by AON-SM7. [^{32}P]-labelled RNA - region KRAS-A probe (unbound in lane 1) was incubated with 540 nM CRD-BP (lane 2). Inhibitor SM7 was present (lanes 3-14) in otherwise identical conditions to lane 2 ranging in concentration from 1X molar-fold relative to probe concentration (118 pM) to 4000 molar-fold (472 nM) which correlated with a loss of CRD-BP complexed bound RNA. AON SM2, in contrast displayed reduced ability to interrupt the CRD-BP-KRAS-A RNA complex.

As can be seen in the SM7 inhibition assay (figure 2.22), a similar trend occurred with SM7 as with SM6. Both inhibitors show a reduction of the bound RNA fraction and a proportional increase in unbound RNA, indicative of a loss of stability of the CRD-BP-KRAS RNA complex. Once again the SM2 oligonucleotide displayed a similar pattern as in figure 2.21,

with no activity against RNP formation. Pushing the upper limit of the inhibitor concentration range to 4000-fold produced only a marginal further reduction in the bound RNA fraction and can be seen to be effective far before this concentration (lanes 11-14). Densitometry analysis of the inhibitory capacity against the CRD-BP-KRAS RNA complex of both AON SM6 and SM7 was completed using the phosphoimager system and optiquant software. Numerical data for both oligonucleotides up to 1000X concentration was plotted graphically using KaleidaGraph™ software as seen below (figure 2.23).

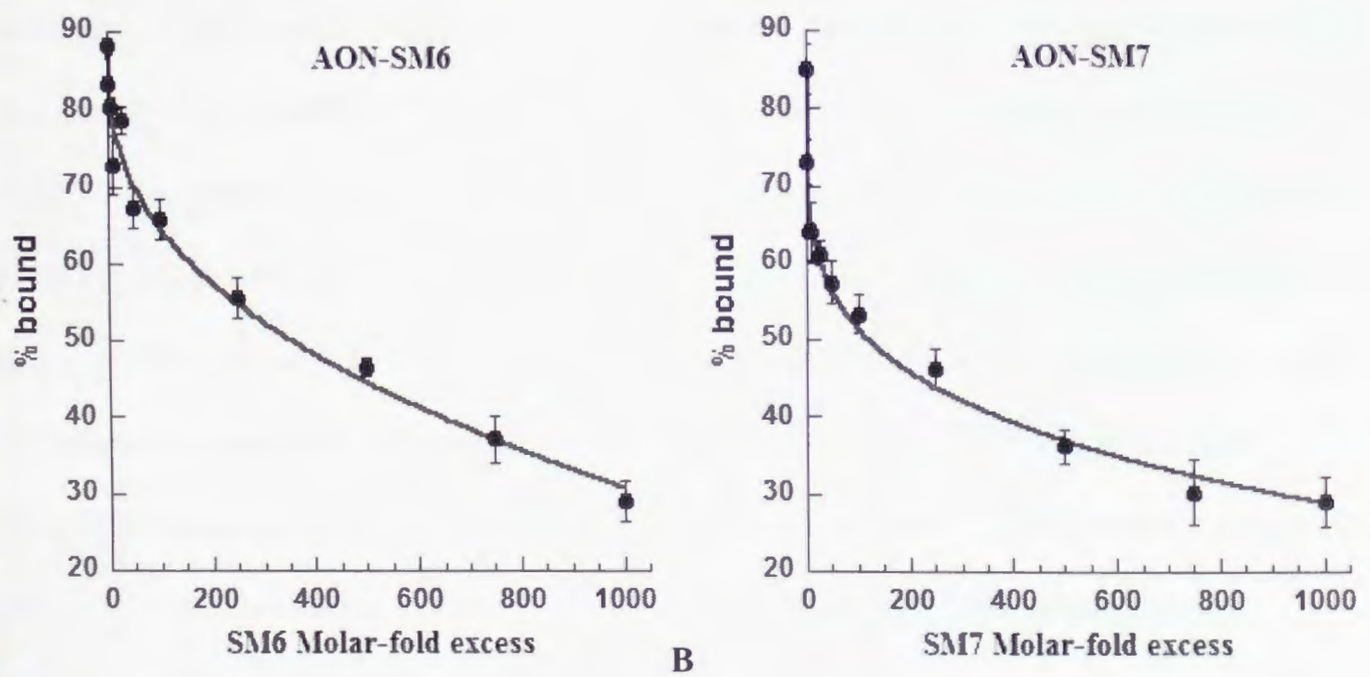


Figure 2.23. Inhibition plot of AON SM6 (A) and SM7 (B) against CRD-BP-KRAS-A RNA complex formation. Graphs represent quantified data from densitometry analysis of electrophoretic mobility shift assays with inhibitor oligonucleotide concentrations ranging from 118 pM at 1X molar-fold to 118 nM at 1000X molar-fold excess. [^{32}P]-labelled KRAS-A probe was 118 pM with CRD-BP protein maintained at 540 nM.

Graphical presentation more clearly depicts a trend whereby both AON SM6 and SM7 inhibitors show effective dose-dependent abrogation of the RNA-protein interaction; however SM7 is more active at lower concentrations (IC_{50} of approximately 100X or 1.18 nM for SM7,

compared to 400X or 4.72 nM for SM6) and is overall deemed the most effective single DNA oligonucleotide of the set. Another look at the inhibitors in parallel with computer simulated RNA folding of KRAS-A with *mfold* place SM6 (labelled in blue) and SM7 (labelled in black) as complementary to adjacent stem-loop structures, possibly implicating these loop structures as important for CRD-BP recognition (figure 2.24). Also in the same figure, the KRAS-A34(b) RNA region is indicated in red - the smallest CRD-BP-binding RNA fragment identified through EMSA mapping (section 2.3.2), and overlap can be seen with both the SM6 and SM7 inhibitors. This striking result complements the previous RNA affinity mapping data, and reinforces the conclusion that the KRAS-A34(b) region is important for KRAS mRNA binding and likely is in direct contact with the CRD-BP protein, at least in these in vitro assays. Mongroo *et al.* (2011) showed that the KRAS mRNA coding region directly interacts with IMP-1 (Human ortholog of CRD-BP) in vivo through UV cross-linking, and based on the described experiments here, that interaction likely occurs in the KRAS-A34(b) region between nucleotides 129-185 relative to the start codon. To explain the reduction in overall affinity of this smaller 57nucleotide region compared to the larger KRAS-A (185 nt), it must be considered that CRD-BP possess multiple domains that each contribute to an overall Kd for a given RNA substrate, as is the case for hnRNP K and FMR1 where multiple KH-domains are required for a stable interaction with RNA (Siomi *et al*, 1994). Further emphasis on the requirement for multiple KH-domain interactions comes from experiments involving IMP-1, the Human ortholog of CRD-BP, where all four KH-domains were shown to be necessary for proper cytoplasmic trafficking and RNA binding (Nielsen *et al.* 2002). As a result of this property, it is possible that the smallest sequence identified capable of forming a complex with the protein – KRAS-A34(b) may only be

nteracting with a sub-set of the RNA-binding KH domains, and that the truncation process may

ıave removed other lower-affinity sites that would normally collectively contribute to binding.

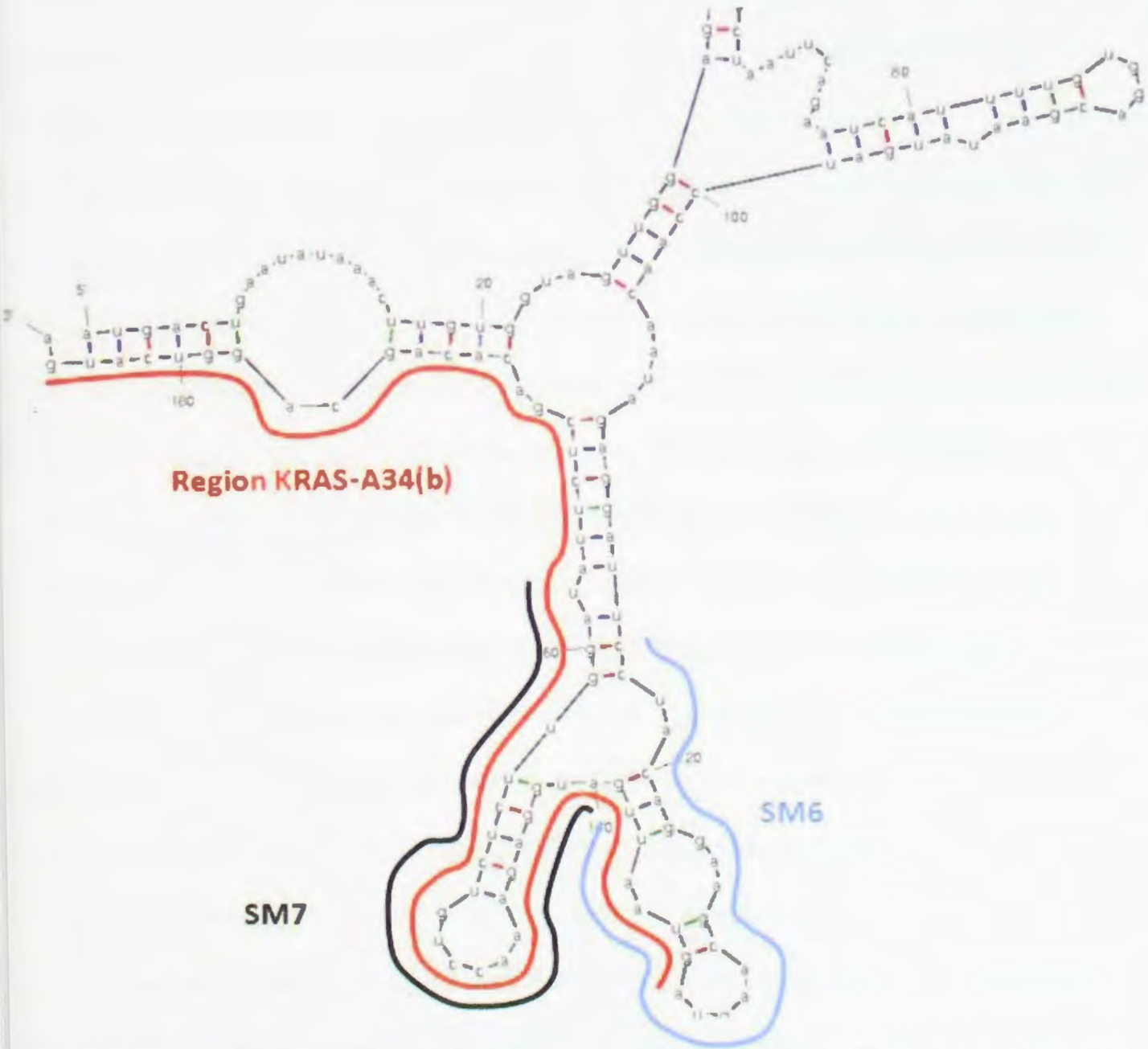


Figure 2.24. Computer simulated RNA folding of KRAS-A RNA with KRAS-A34(b) and AON inhibitors SM6 and SM7 target regions indicated. RNA folding of KRAS-A (nts 1-185) was completed using the mfold web server; the presented structure having a predicted free energy of dG = -38.63 [initially -46.9] Kcal/mol.

interacting with a sub-set of the RNA-binding KH domains, and that the truncation process may have removed other lower-affinity sites that would normally collectively contribute to binding.

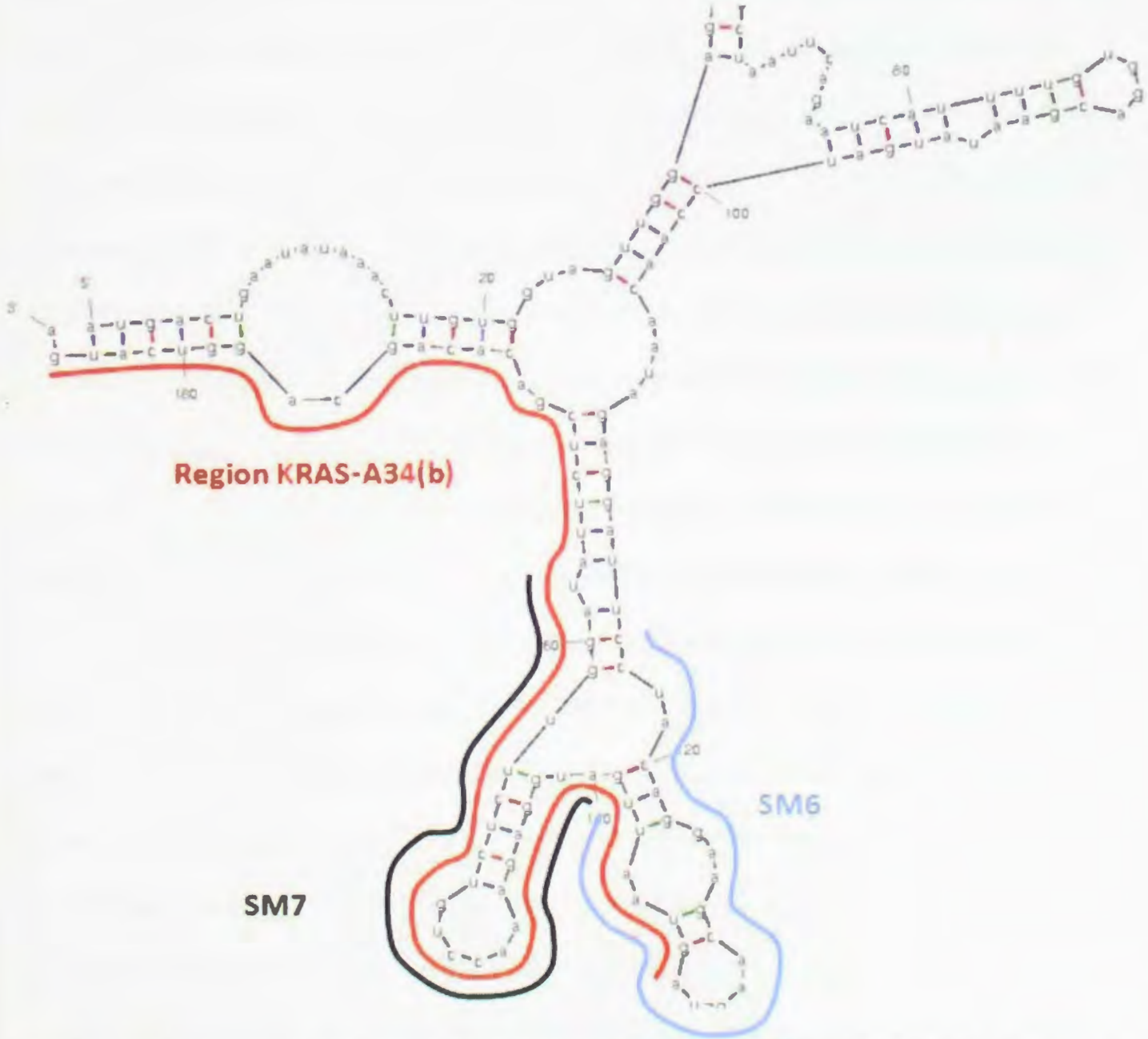


Figure 2.24. Computer simulated RNA folding of KRAS-A RNA with KRAS-A34(b) and AON inhibitors SM6 and SM7 target regions indicated. RNA folding of KRAS-A (nts 1-185) was completed using the mfold web server; the presented structure having a predicted free energy of dG = -38.63 [initially -46.9] Kcal/mol.

2.3.4 KRAS RNAminimum binding sequence analysis

Physical separation of the KH domains and their precise geometry relative to one another play a critical role in the selection of RNA targets, as well as nucleotide sequence. KH-domain architecture does not allow for recognition of long sequences, but instead direct interaction between an RNA substrate and an individual domain is limited to only to 4 or 5 nucleotides and is of weak affinity (Chao *et al.* 2010). To try and identify some of the sequences that may be important in facilitating this interaction, a search for putative RNA-binding sequences for the VICKZ protein family was undertaken and local sequence alignments were completed using previously identified consensus sequences and the putative KRAS mRNA binding region. While there is little agreement regarding a precise consensus binding sequence for this family of proteins, with different experiments returning largely different conclusions, a few interesting results are mentioned here that positions KRAS-A34(b) in relation to the available literature.

Vg1RBP is a closely related ortholog of CRD-BP whose known mRNA targets: VLE and Vg1, contain a sequencing clustering of 5'-UUUCUA-3' or 5'-UUCAC-3' (Git and Standart, 2002). In KRAS-A34(b) RNA, a sequence very similar to those aforementioned, 5'-UUCUC-3', appears in close proximity to a repeat which runs in the opposite direction, separated by only 5 nucleotides. Strikingly Du *et al.* (2008) suggest that a minimum of exactly 5 nucleotides between binding-sequences could fill the gap between KH-domains and is required for efficient binding for hnRNP and PCB1/2, which are related RNA-binding proteins that contain only three KH domains. Furthermore, the RNA targets reported for hnRNP and PCB1/2 require a minimum of 2

repeats of a 19 nucleotide sequence: 5'CCCCACC**CUCUU**CCCCAAG3' which contains the 5'-CUCUU-3' sequence found in KRAS-A34(b). Importantly, in developing a model for binding, the ambiguity of binding-sequence order within a transcript has been demonstrated by Chao *et al.* (2010) with only the overall polarity relative to the KH-domains being an important factor. This group also determined from 3 rounds of selection and amplification (SELEX) of a degenerate RNA library consisting of random mutations introduced into the target "zipcode" sequence, that ZBP1 recognized two conserved sub-sequences: 5'-GGACU-3' and 5'-ACA-3'. KRAS-A34(b) region does not possess the prior identified sequence within its 57 nucleotide length, however the latter sequence appeared twice in partial overlap as 5'-ACACA-3' which is not surprising given the simplicity of the 5'-ACA-3' sequence and its likelihood of appearing by chance. However, later work with ZBP1 completed by Patel *et al.* (2012), focussing on the ZBP1 KH3-4 didomain only, separately determined similar binding elements except for a slightly lengthier 3' bipartite sequence (5'-ACAC-3') and this still matches that found in the KRAS-A34(b) region. Additionally, using 1H-15N HSQC NMR it was also reported by Patel that the 5'-ACAC-3' motif interacts with the IMP-1 KH3 domain, as evidenced by resonance shifts within the residues of the KH3 domain when free protein was complexed with this small oligonucleotide. Pairing RNA sequence with a specific KH domain is useful information for constructing a model of KRAS RNA binding.

PAR-CLIP (photoactivateable-ribonucleoside-enhanced crosslinking and immunoprecipitation) experiments, where UV-reactive nucleosides are added to cell cultures to be incorporated into mRNA transcripts, as well as other RNA molecules, revealed that IGF2BP1 (aka IMP-1) co-precipitates with cross-linked RNA sequence: 5'-CAUH-3' (where H=A,C or U). This was in fact the only single consensus sequence found, from this study, however the

actual cross-linked residue occurred within this motif only 90% of the time, with the other 10% occurring within a 5 nt range implying some flexibility (Hafner *et al*, 2010). KRAS-A34(b) does not contain this 5'-CAUH-3' motif. However, the slightly larger KRAS-A24, mentioned previously to have affinity similar to the full length KRAS-A (section 2.3.2), does contain this motif in the form of 5'-CAUU-3' that occurs once at nucleotide position 79 thus offering a potential RRE sequence for an upstream binding sequence hypothesized in section 2.3.2 to work cooperatively with the others previously mentioned within KRAS-A34(b). Keeping in mind that KRAS-A12, which includes the A2 region that has the aforementioned 5'-CAUH-3' motif, does not seem to bind CRD-BP, it is thus not sufficient on its own as a binding motif and seems to require information in the KRAS-A34 region, which does bind CRD-BP, albeit with reduced affinity.

Nielsen et al. (1999) proposed a single putative binding sequence for IMP-1: 5'-UUCACGUUCAC-3' based on EMSA and UV-crosslinking methods with mRNA target IGF-II. By aligning DNA equivalents of KRAS-A34 with the proposed IGF-II binding sequence, a region of highest similarity can be seen and this falls within the KRAS-A34(b) region. The RNA equivalent is 5'-UUCUCGACAC-3' in the KRAS-A34(b) region. A proposed binding sequence for ZBP1 is 5'-ACACCC-3' (Ross *et al.* 1997) and a similar alignment placed the region of possible importance within the same range. Poly-C and poly-U pyrimidine homopolymers compete significantly better than poly-A and poly-G purine homopolymers against VLE mRNA with Vg1RBP (Git& Standart 2002). Transcripts with more C/U content should therefore bind better, however the total composition of A34(b) does not agree with high C and high U; 44% C/U and 56% AG. However binding assays with poly-C homopolymers do show that the binding mode with these oligonucleotides is different, being additive and competing away the actual

VLE target sequence in a linear, non-cooperative fashion, as opposed to the target VLE sequence which binds cooperatively much like KRAS-A. Thus, while KRAS-A34(b) does not fall under the C/U-rich targets, it may be a different mechanism at work explaining homopolymer binding. Perhaps it is not so much the raw overall C/U content, but instead a certain C/U-rich sub-sequence that aids in binding VICKZ proteins. Notably, the previously mentioned 5'-CUCUU-3' sequence repeat in KRAS-A34(b) as well as the 5'-UUCUCGACAC-3' may fall in-line with this idea.

Figure 2.25. Local DNA alignment of KRAS-A34 and putative binding sequences for IMP-1 and ZBP1 CRD-BP orthologs. (A) IMP-1proposed binding sequence and aligned with KRAS-A34 and (B) ZBP1 proposed binding sequence aligned with KRAS-A34. DNA sequences were converted from their RNA sequence origins for the purpose of the alignment, and both highlight an overlapping possible binding region.

Chapter 3

Assessing the role of CRD-BP KH domains involved in KRAS mRNA binding in vitro and in cells.

This chapter outlines the experiments undertaken to determine the role of each of the four individual CRD-BP KH domains in binding KRAS mRNA. The first section focuses on in vitro electrophoretic mobility shift assay (EMSA) experiments using point mutation CRD-BP protein variants. The second section focuses on determining the importance of KH domains in cells, using RNA immuno-precipitation coupled with quantitative reverse transcriptase PCR (RT-qPCR).

3.1 Methodology – CRD-BP KH domain requirements for binding KRAS mRNA *in vitro*.

3.1.1 Purification of CRD-BP KH domain variants

CRD-BP variants were previously generated by another member in Dr. Lee's lab using site-directed mutagenesis against the first glycine (to aspartate) in each of the four G-X-X-G motifs within each KH domain (G212D, G293D, G422D, G504D). These motifs are critical for the interaction of the KH domain with target mRNA, and by mutating each individually as well as in tandem as double-KH mutants, individual KH domain importance in binding KRAS mRNA could be assessed. Proteins harbouring a specific single mutation were labelled as "KH1" through "KH4" depending on the KH domain affected, with double mutants harbouring two affected domains labelled as "KH1-2" through "KH3-4"; information summarized in Table 3.1.

Table 3.1. Summary of generated CRD-BP KH domain variants. KH domains are individually affected as a result of site-directed mutation in the canonical G-X-X-G motif.

CRD-BP mutant	KH domain mutated			
	KH1	KH2	KH3	KH4
KH1	•			
KH2		•		
KH3			•	
KH4				•
KH1-2	•	•		
KH1-3	•		•	
KH1-4	•			•
KH2-3		•	•	
KH2-4		•		•
KH3-4			•	•

E. *coli* BL21 (DE3) cells were heat-shock transformed with pET28b-CRD-BP variant plasmids and purified using a Ni-NTA agarose column followed by dialysis into the final buffer: 200 mM NaCl, 20 mM Tris-Cl, 10% glycerol, 0.01% Triton-X, pH 7.4 as described in section 2.1.1. Quantification of each protein sample was completed using a BCA Protein Assay Kit (Thermo Scientific) with CRD-BP protein samples typically ranging between 3 and 7 μM.

3.1.2 Generation of [^{32}P]-labelled RNA substrates by IVT

Experiments with CRD-BP KH domain variants were completed using the KRAS-A RNA fragment (nts 1-185) determined to possess strong binding affinity (section 2.3.2), as well as the coding region determinant region of c-myc mRNA (nts 1705-1886), an established mRNA target of CRD-BP (Prokipcak *et al.* 1994). T7-driven *in vitro* transcription of both [^{32}P]-labelled RNA regions from cDNA templates was completed as described previously in section 2.1.5.

3.1.3 EMSA experiments with CRD-BP KH domain variants

EMSA binding reactions were completed for each generated [^{32}P]-labelled KRAS and c-myc sub-region and contained CRD-BP KH-mutants concentration varied from 0 to 540 nM. All EMSA reactions used for KH domain binding assessment contained KRAS or c-myc probe at

20,000 cpm. Protein-RNA mixtures were incubated at 37 °C for one half hour and complexes resolved on an 8% polyacrylamide (29:1 Acrylamide/Bis, Bio-rad) non-denaturing gel in 0.5 x TBE running buffer and subjected to electrophoresis for 90 minutes at 25 mA in a similar fashion to section 2.1.6. Dissociation constants were calculated based on data acquired from densitometry analysis of the EMSA autoradiograph using the Cyclone Storage Phosphor-System and OptiQuant software.

3.2 Methodology – CRD-BP KH domain requirements for KRAS mRNA binding in cells.

3.2.1 Preparation of IP-RNA from HeLa cells

Preparation of purified immunoprecipitated (IP) RNA was part of previous projects and was already on hand for the described experiments; a condensed description of methods used to acquire the IP-RNA follows. CRD-BP KH domain mutant variants as described in the previous section (3.1.1), were sub-cloned into a pcDNA mammalian expression vector (+FLAG) and transfected into HeLa cells using Lipofectamine 2000 reagent (Invitrogen) followed by a 48 hour incubation. Cells were lysed with total cell lysis (TCL) buffer in the presence of RNasin and protease inhibitor tablet (Roche diagnostics) and lysate incubated with anti-FLAG antibody overnight at 4 °C. Immunoprecipitation (IP) of the different CRD-BP variants containing point mutations in the GXXG motif (single and double mutants) using Protein G-agarose beads was performed and equal amounts of immune-precipitated CRD-BP was recovered as determined by SDS-PAGE, followed by treatment with Proteinase K and a final phenol-chloroform-isoamyl alcohol extraction. IP-RNA was quantified using a Nanodrop Spectrophotometer (Thermo Scientific). RT-qPCR to measure the relative amounts of KRAS mRNA associated with each of the variants as compared to a wild-type control.

3.2.2 cDNA synthesis from IP-RNA

To ensure no DNA contamination in the RT-qPCR reactions which may have co-precipitated with the CRD-BP protein granules, the "DNA-free" DNA removal kit (Ambion) was employed. Firstly to each 1 µg of IP-RNA co-precipitated with each CRD-BP KH variant sample (all single and double mutants in triplicate except KH2 and KH3, of which only 1 biological replicate was available), 1 µl of DNase1 (RNase free DNase) (Ambion), 1 µl 10x buffer and H_2O made up to 10 µl reaction volumes. Reactions were incubated at 37 °C for 30 min followed by the addition of 2 µl of rDNase deactivating reagent after which a 2 min room temperature incubation period followed. Deactivated reactions were spun at 10,000 xg for 1.5 min to sediment rDNase1 molecules and 10 µl of the supernatant containing the purified IP-RNA was drawn off for use in cDNA synthesis prior to qPCR.

Synthesis of complementary cDNA from purified IP-RNA samples was accomplished using a Biorad iScript cDNA synthesis kit; 20 µl reactions prepared with 4 µl 5x iScript buffer mix containing random hexamer primers, 5 µl nuclease-free H_2O, 10 µl of DNA-free IP-RNA, and 1 µl reverse transcriptase (RTase) assembled specifically in that order. Assembled reactions were placed in a thermocycler and subjected to incubation at 25 °C for 5 min (random hexamer primer annealing), 42 °C for one half hour (cDNA synthesis), 85 °C for 5 min (RTase inactivation) followed by 4 °C storage until use. As a control, each KH variant IP sample was also subjected to the same procedure without the addition of reverse transcriptase (- RT/no template) to account for the possibility of genomic DNA stability.

3.2.3 KRAS mRNA RT-qPCR

Quantitative PCR was used to measure relative amounts of KRAS mRNA associated with the CRD-BP variants to determine involvement/importance of individual KH domains. Firstly,

KRAS specific amplification primers designed to amplify a 126 nucleotide region of KRAS coding sequence were synthesized by Integrated DNA technologies Inc. (IDT), as well as a complementary TaqMan probe containing a 5' 6-FAM (fluorescein) fluorophore and 3' BHQ-1 Black Hole Quencher. Sequences for KRAS qPCR primers and TaqMan probe (Table 3.2) were based on described experiments with pancreatic cancer cell lines testing DNA oligonucleotide based modulators of KRAS gene transcription (Cogoi *et al.* 2013).

Table 3.2. KRAS RT-qPCR amplification primers and TaqMan probe.

Primer/probe	DNA Sequence
Forward primer:	5' – CGAATATGATCCAACAATAGAG - 3'
Reverse primer:	5' – AATGAGGGACCAGTACAT - 3'
KRAS TaqMan probe:	5' - /56-FAM/TACTCCTCTTGACCTGCTGTG/3BHQ-1/-3'

Quantitative PCR samples were setup as 25 µl reactions and completed in triplicate for each biological replicate totally 9 reactions per CRD-BP variant. As previously mentioned, only a single biological replicate was available for KH2 and KH3 CRD-BP variants and as such, only 3 values were obtained for these. Additionally, for each RIP sample a no RT control was measured in an otherwise identical fashion to ensure no genomic DNA contamination. In a 96 well PCR microplate (Axygen Scientific Inc.) reactions were assembled according to table 3.3 from an iQ™ Supermix kit (Bio-Rad). In an attempt to reduce pipetting error and inter-sample variability, all qPCR experiments were completed by first creating a master-mix that contained all reagents except the cDNA. To each sample well of the PCR microplate, 24 µl of master mix was added, with the final 1 µl cDNA volume added last. Following addition of all reagents, the used PCR microplate wells were sealed with sealing film suitable for real time PCR (Axygen Scientific Inc). As a final step, the PCR microplates were centrifuged at 200 xg for 1 minute at room temperature to ensure proper mixing. Following reaction assembly, qPCR was performed

on an iQ5 Multicolour Real-Time PCR Detection System and data analysed using iQ5 optical system software (Bio-Rad).

Table 3.3. Reagents used for KRAS qPCR reactions. IQ supermix is from the Bio-Rad iQ™ Supermix kit and contains all necessary buffer ingredients, dNTPs and enzyme. All reactions were assembled in triplicate for each CRD-BP mutant variant and mixed as a master mix prior to adding final cDNA.

qPCR reagent	Volume used per qPCR reaction
Forward primer (10µM)	0.5 µl
Reverse primer (10 µM)	0.5 µl
KRAS TaqMan probe (5 µM)	0.5 µl
IQ™ Supermix (2X)	12.5 µl
cDNA	1 µl
H_2O	10 µl
Total volume	25 µl

3.3 Results and Discussion – KH domain requirements for binding KRAS mRNA

3.3.1 CRD-BP single KH domain variant EMSA binding assays

To investigate the functional role of each individual KH domain in binding KRAS mRNA, CRD-BP variants harboring mutations in each GXXG motif of the four KH domains were used in EMSA binding assays with the high affinity KRAS sub-region A (section 2.3.2) [^{32}P]-labelled probe. EMSA containing the KH mutants and KRAS-A probe was carried out with all single and double mutants beginning with KH1 through KH4 single mutant variants and were performed in a manner identical to chapter 2 describing KRAS mRNA affinity mapping. Results from the single KH mutants displayed a stark contrast in binding capacity when compared to wild-type CRD-BP (figure 3.1). Beginning with figure 3.1A, it can be seen that the wild-type CRD-BP fully shifts the free KRAS-A RNA probe within the experimental protein concentration range of 0-540 nM (lanes 1-6). However, with a single GXXG motif point mutation in either KH1 or KH2 (lanes 7-11 and 12-16 respectively), the binding pattern changes considerably, both

mutants displaying limited mobility shift of the free RNA. Interestingly, KH1 and KH2 both have unique binding patterns; KH1 remains capable of forming a specific molecular weight complex of the same size as wild-type, albeit at a reduced quantity (figure 3.1A, lanes 7-11), while KH2 appears to have a complete absence of this band. A faint, much larger complex can be seen near the top of the gel for KH2 (figure 3.1A, lanes 12-16) as a smear that is absent in the KH1 lanes but still present in the wild-type shift pattern. CRD-BP variants KH3 and KH4 were assessed for KRAS-A RNA binding capacity in the same experiment, and similar results were found. As with KH1, a GXXG mutation in the KH3 domain yielded drastically reduced binding ability, with only an extremely faint trace of the specific complex that is so prominent in the wild-type version of CRD-BP, and to a lesser extent the KH1 mutant variant (figure 3.1B, lanes 7-11). Higher molecular weight smearing was also not as prominent when compared to wild-type. KH4 domain mutation in CRD-BP produced a very clear result, with no hint of any significant binding whatsoever, neither as a specific complex band, nor as a higher molecular weight smearing. This high molecular weight smearing observed may suggest partial binding of the protein to KRAS RNA.

Due to the striking reduction in KRAS-A probe binding affinity when any single CRD-BP KH domain GXXG motif is mutated, we wondered if perhaps these mutant protein samples had somehow mis-folded or were otherwise producing artifact data. To test this possibility, an additional set of EMSA experiments were completed utilizing a positive control where the binding pattern with CRD-BP KH mutants was already known. c-myc mRNA contains a region known as the coding-region instability determinant (CRD), and is a well established RNA target sequence of CRD-BP (Bernstein *et al.* 1992; Prokipcak *et al.* 1994; Leeds *et al.* 1997).

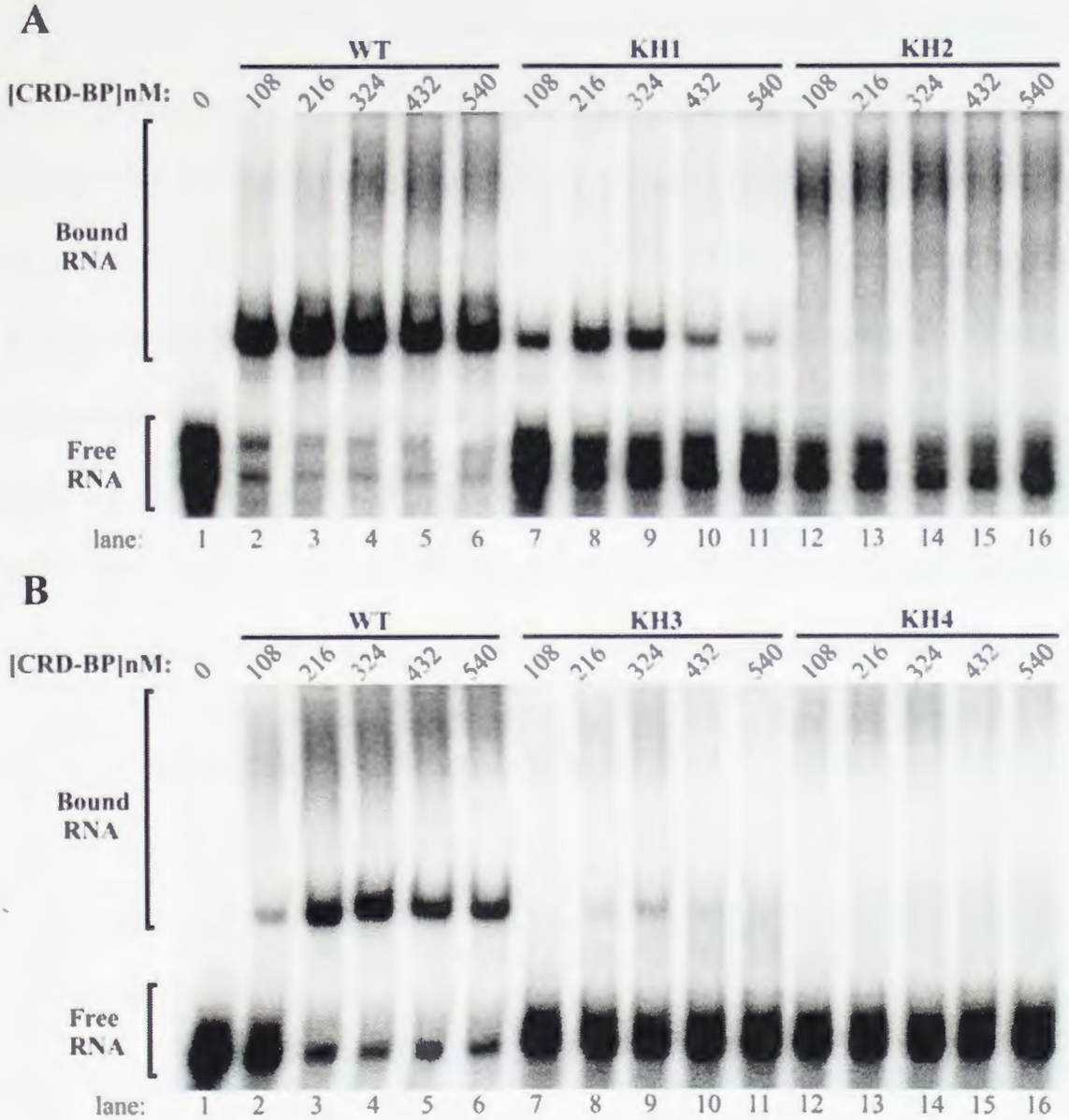


Figure 3.1. EMSA assessing CRD-BP KH variant with single point mutation for binding to sub-region KRAS-A RNA; (A) KH1 and KH2 variants, (B) KH3 and KH4 variants. [^{32}P]-labelled unbound KRAS-A (nts 1-185). RNA probe (20,000 cpm/reaction) with no protein added (lanes 1, A and B) forms a larger molecular weight complex upon titration with WT CRD-BP and single KH variants possessing GXXG motif mutation (first glycine to aspartate), as indicated by the bound RNA fraction (lanes 2-16, A and B).

This RNA probe consisting of 182 nucleotides, was synthesized by IVT and was used in additional EMSAs (figure 3.2). Finally, it is worth noting that higher radioactivity (80,000 cpm/reaction) was used in this latter set of EMSA experiments in order to increase sensitivity to complex formation, and the results can be seen as the presence of a weak KRAS-CRD-BP complex formation in all four KH mutant samples. This is in contrast to the previous set of

EMSAs (figure 3.1), where only KH1 formed a visible band of notable intensity and was able to reduce the amount of free RNA. However, the overall pattern of KH1 producing the strongest shifted RNA intensity still remains unchanged, also suggesting that the KH1 domain has a necessary (for full binding) but reduced role in binding the KRAS-A sub-region fragment. EMSA experiments incorporating the c-myc RNA as a positive control (figure 3.2, A-D) yielded data which was very similar to the previous set of single KH mutant CRD-BP variants, but also provided some comparisons regarding the binding mode of KRAS RNA as compared to c-myc RNA. Firstly, it can be seen that CRD-BP displays redundancy in binding c-myc mRNA with a single mutation in any one KH-domain failing to strongly affect overall binding.

Furthermore, CRD-BP appears to bind c-myc (and the KH mutant variants) cooperatively as a dimer with two separate Kd values as evidence by two distinct bands clearly visible at 108 nM (figure 3.1 A-D lane 8). First, CRD-BP binding likely occurs as a single protein complex, with the dimer forming subsequently. On the other hand, KRAS-A RNA does not appear to form a strong complex with a single CRD-BP molecule, with only an extremely faint intermediate band visible in the wild-type lanes (figure 3.1 A-D, lanes 2-6). The c-myc and KRAS RNA formed a dominant upper complex with similar migration distance, and the RNA molecules being of similar size (182 and 185 nucleotides respectively), hence it is also likely that KRAS-A RNA binds almost exclusively in the dimerized form with 2 CRD-BP molecules attached in one complex. It could be that an initial binding event with a single CRD-BP molecule induces an RNA conformation change that produces or reveals a second CRD-BP binding site of very high affinity, which is immediately bound. In this induced fit model, the initial binding event forming complex 1 (one molecule bound) is rate-limiting and the second CRD-BP molecule (complex 2) binds almost simultaneously. This idea is in partial contrast to a previously proposed binding

model by Nielsen *et al.* (2004) where sequential binding of IMP-1 produces dimerization on the IGF-II mRNA target, with a fast initial binding event (complex 1) and a much slower, clearly sequential formation of complex II. In this same model put forward by Nielsen *et al.* (2004), the formation of complex 1 is considered non-specific and a relatively low stability precursor intermediate, and it is the successful formation of complex 2 that "locks" the target mRNA into a stable RNP complex. If this model is indeed correct, then the KRAS-A sub-region is a bona fide target of CRD-BP with unique binding character in that the dimerization event occurs either simultaneously with the formation of complex 1 or at the very least, extremely rapidly. Another possibility is that upon an initial CRD-BP binding, not the RNA, but rather the protein experiences a dynamic conformational shift, a mechanism suggested by Wang *et al.* (1999) for the binding of a related KH domain containing protein: Poly-A binding protein (PABP) with target α-globin mRNA. This change in folding character may then lead to the creation of a protein-protein interaction site recognized by the second CRD-BP that can lead to the dimerization event apparent in the EMSA figures presented here. In support of either model (i) RNA or (ii) protein conformation change, experiments conducted to determine if IMP-1 can dimerize independent of an RNA substrate indicate that RNA-binding is a prerequisite for dimerization and that complex I precedes the formation of complex II (Nielsen *et al.* 2004).

3.3.2 CRD-BP double KH domain mutant EMSA binding assays

Further work in assessing the specific requirements of KH domains in binding KRAS mRNA was also completed using the EMSA. CRD-BP KH variants harbouring point mutations at two GXXG motif sites located in separate KH domains were used in a similar fashion to assess whether or not all traces of binding could be abolished. Indeed with the mutation of any two KH domains, absolutely no binding was observed (figure 3.3 A-D).

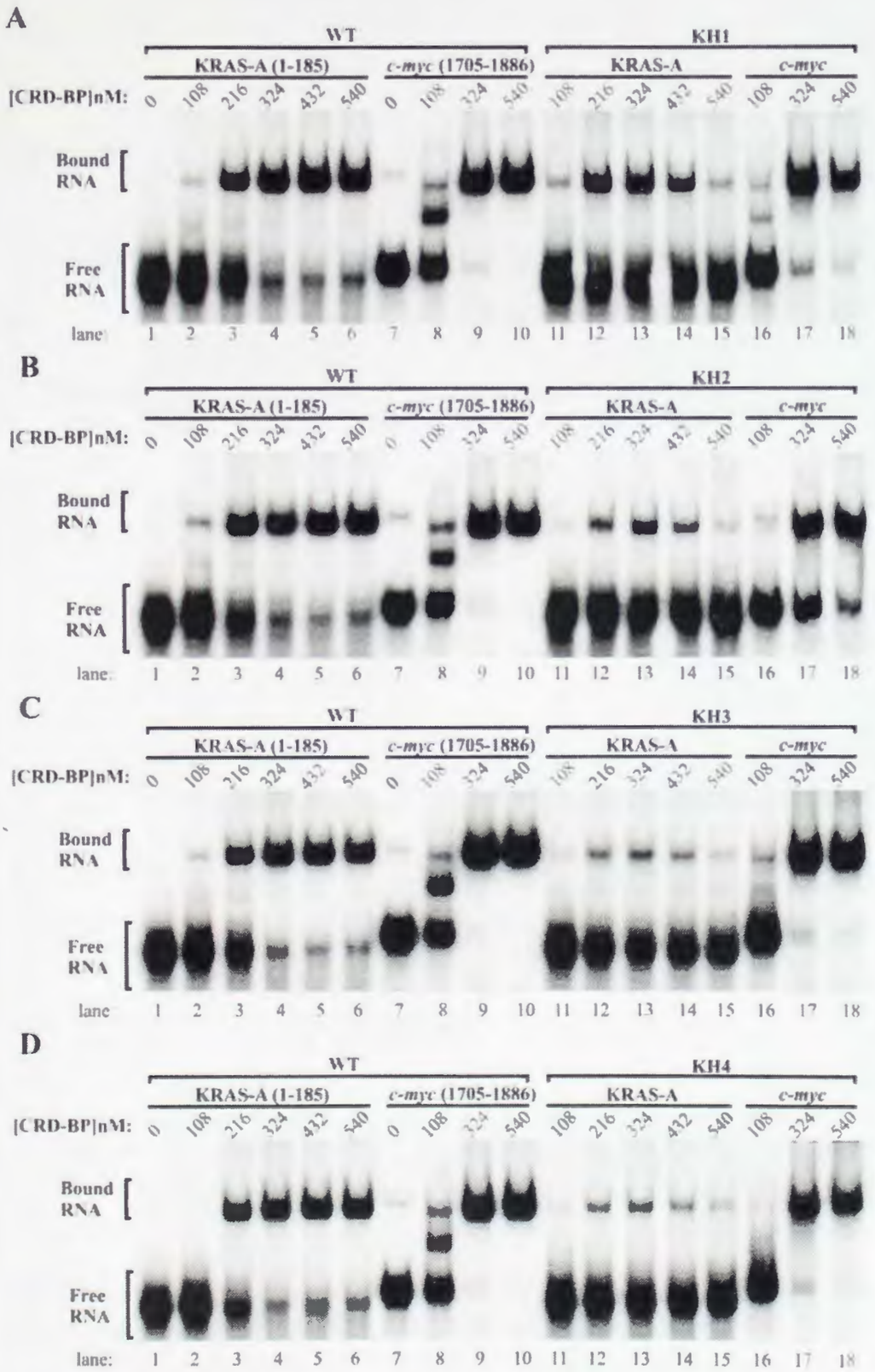


Figure 3.2. EMSA assessing CRD-BP KH variants with single point mutation for binding to KRAS-A and c-myc RNAs; (A) KH1 variant, (B) KH2 variant, (C) KH3 variant, (D) KH4 variant. [^{32}P]-labelled unbound KRAS-A (nts 1-185) and c-myc (nts 1705-1886) RNAs (80,000 cpm/reaction) with no protein added (lanes 1 and 7, A-D)

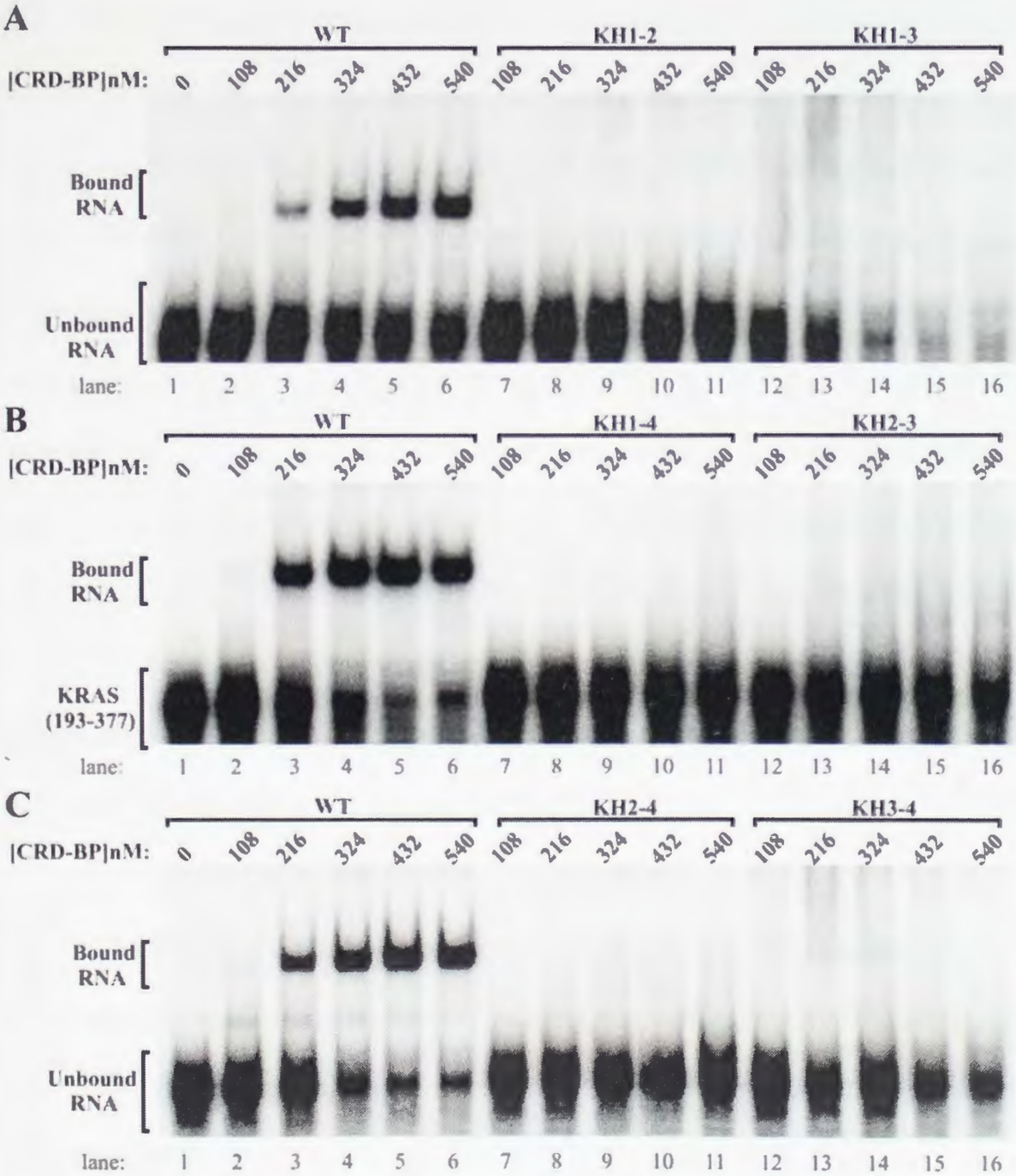


Figure 3.3. EMSA assessing the CRD-BP KH variant with two point mutations for binding to sub-region KRAS-A RNA; (A) KH1-2, KH1-3 (B) KH1-4, KH2-3 (C) KH2-4 and KH3-4. [^{32}P]-labelled unbound KRAS-A (nts 1-185) RNA probe (80,000 cpm/reaction) with no protein added (lanes 1 in all gel images) forms a larger molecular weight complex upon titration with WT CRD-BP and double KH variants possessing GXXG motif mutations (first glycine to aspartate), as indicated by the bound RNA fraction (lanes 2-16, A and B). Binding reactions were incubated for a total of 20 minutes at 37 ^{o}C and 10 minutes on ice before being resolved on an 8% polyacrylamide gel.

A complete elimination of binding in the double KH mutant EMSA experiments provides compelling evidence that KRAS mRNA interacts primarily and likely entirely through the GXXG flexible loop regions, and not by any other non-canonical mechanism.

3.3.3. KH domain requirements for KRAS mRNA binding in HeLa cells

CRD-BP KH domain binding requirements for KRAS mRNA was also assessed in cells through the use of RNA immuno-precipitation (RIP) method coupled with RT-qPCR. HeLa cells were transfected with pcDNA plasmids containing CRD-BP-FLAG KH mutant variants as well as a wild-type version. Anti-FLAG antibody was then used to immune-precipitate the ribonucleoprotein complexes. Co-precipitated KRAS mRNA associated with each of the CRD-BP KH mutants was then analyzed by RT-qPCR. As a control for each KH variant sample, a no reverse transcriptase (- RT/no template) sample consisting only of the purified mRNAs co-precipitated with CRD-BP was run under identical qPCR conditions. This control produced no detectable fluorescence within the 40 cycle limit of the system and confirmed that no genomic DNA contamination was present in any of the samples tested.

The KRAS qPCR primers were tested for efficiency using plasmid dilutions containing KRAS cDNA template and determined to be 97.8% with an R^2 of 0.998. Data group Ct values (figure 3.4) from RT-qPCR samples were then analyzed using a One Way Analysis of Variance (one-way ANOVA) test and a <0.0001 p-value (assuming null hypothesis of all groups being of the same population) was calculated with an F-stat value of 198.1, indicating a significant difference between one or more of the analyzed groups. From the rejection of null hypothesis that all samples were equal or from the same population, a post-hoc Tukey's HSD multiple comparison test was calculated for each CRD-BP KH mutant pair, and results revealed that levels of KRAS mRNA associated with wild-type CRD-BP were significantly different than the

KH-mutants. Furthermore, KH1 KRAS mRNA levels were also found to be significantly higher than the other remaining CRD-BP mutants (p=0.01). Finally an unpaired t-test was conducted to compare the Ct-values between wild-type and each individual CRD-BP mutant samples. This analysis also showed that all of the qPCR data groups (KH single and double mutants) were significantly different from wild-type at p=0.01. Cycle threshold values were normalized to wild-type CRD-BP and plotted on a logarithmic scale which was required considering the large relative differences in mRNA levels between wild-type and the double KH mutants (figure 3.5).

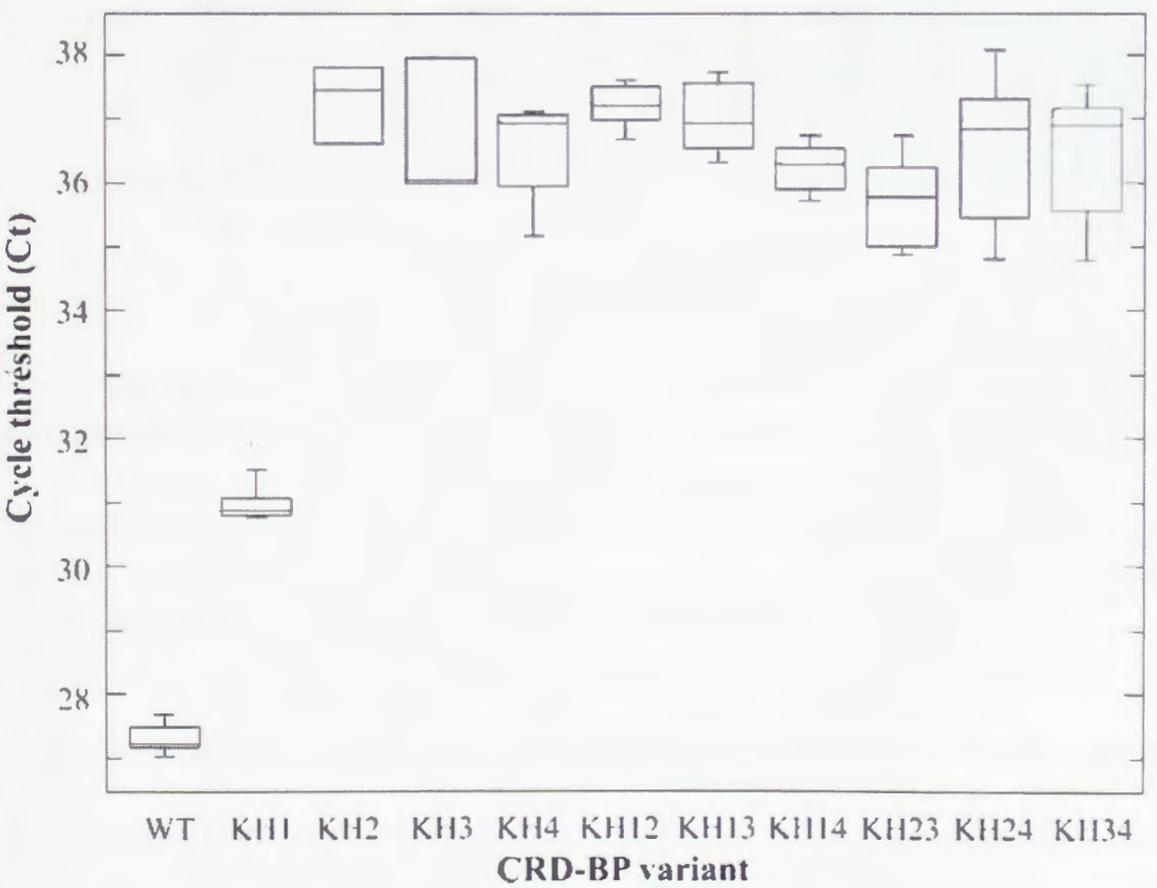


Figure 3.4. Box plot pf the RT-qPCR cycle threshold values from immune-precipitated KRAS mRNA associated with CRD-BP KH variants. Horizontal middle line in the encompassing box represents the overall mean, the top box represents the third quartile and bottom box the first quartile. Bars are the absolute minimum and maximum Ct values obtained.

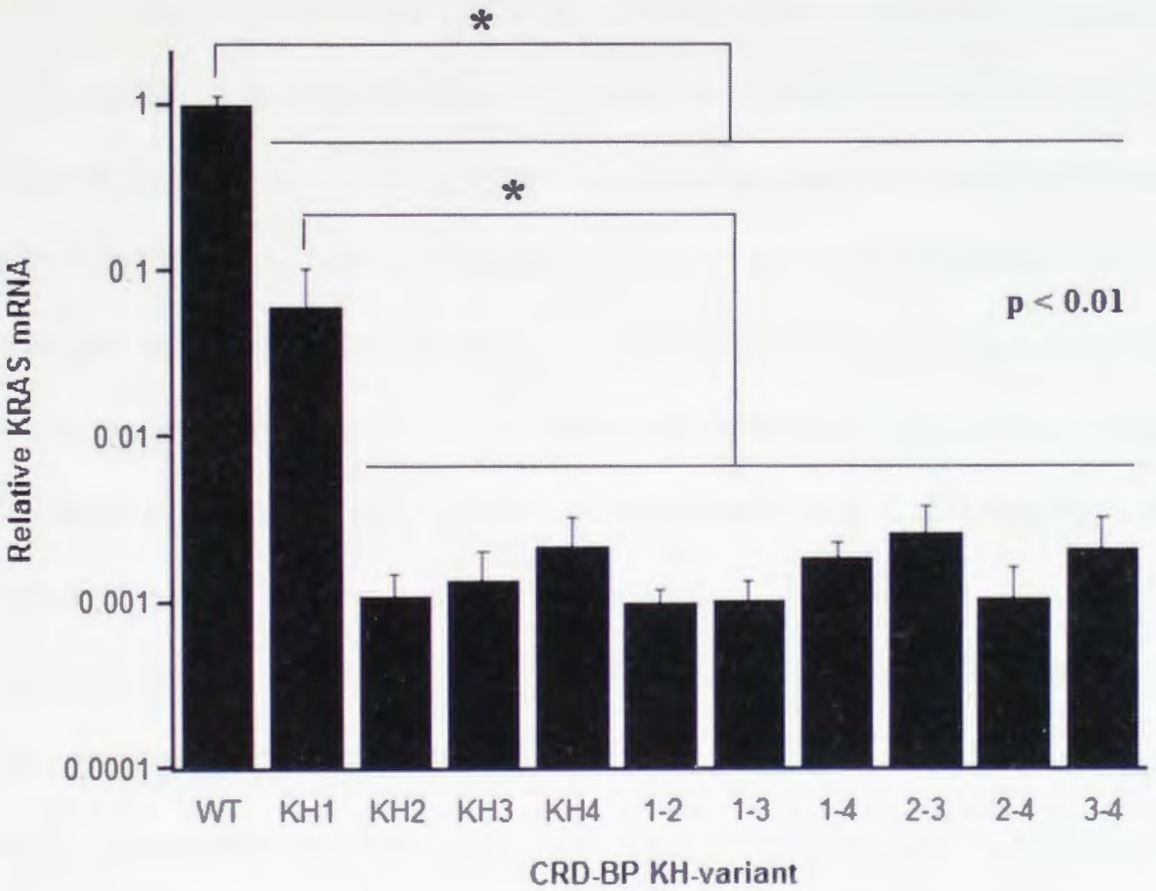


Figure 3.5. Relative KRAS mRNA associated with KH domain variants from immune precipitation of CRD-BP in HeLa cells as determined by RIP RT-qPCR. Each cycle threshold value is from averaging 3 biological samples in triplicate and error bars here represent the standard deviation, with the exception of KH2 and KH3 of which only a single biological replicate was available for RT-qPCR. One-way ANOVA determined $P<0.0001$ and post-hoc Tukey HSD and Student's t-test revealed significance between pairs.

To understand the molecular nature of how CRD-BP binds KRAS mRNA, experiments in the form of EMSA gel shift assays and also RT-qPCR analysis were undertaken. Data from in vitro EMSA experiments correlated well with the result from IP experiments using HeLa cells. Together, the data presented in this chapter strongly support a model where all four KH domains of CRD-BP are required for KRAS mRNA to maximally bind CRD-BP. Mutation of the first GXXG motif glycine to aspartate in of any one of the four KH domains resulted in a near 1000-fold reduction in pull-down KRAS mRNA, save for KH1. The conclusion of a reduced role for KH1, that was initially suggested by the EMSA data (figure 3.1A and 3.2A), is further echoed by

the RT-qPCR data. Comparing KH1 (Ct – 30.008) it can be seen that KRAS mRNA levels are about 13-fold lower than wild-type (Ct-27.310) levels, while the other CRD-BP KH mutants are approximately 1000-fold lower (figure 3.5). Together, the EMSA and IP-qPCR data support the notion that indeed all four KH domains are required for CRD-BP to bind KRAS mRNA, with perhaps a mitigated, but still necessary role for the KH1 domain. This conclusion infers that the interaction is highly specific; four RNA sequence elements evidently required to contact each KH domain for binding, and is in-line with the literature notion that many RNA-binding proteins achieve their specificity through interactions with multiple domains.

Results by Barnes *et al.* (2015) found similar RT-qPCR results with c-myc and CD44 mRNAs. Both mRNAs, like KRAS, were found to display reduced association with KH single and double mutant variants of CRD-BP (except KH4, which retained affinity for c-myc). However in stark contrast to KRAS data, EMSA CRD-BP binding results for these same mRNAs (c-myc and CD44) display an opposite result, with single KH mutations having no effect on binding (except KH4 did not bind CD44 mRNA). RNP granule formation in zebrafish was also analyzed to search for a correlation with in vitro binding data, with injection of the identical CRD-BP KH mutant's mRNAs directly into the embryo. These in vivo results better support the EMSA binding data, with the single KH mutant CRD-BP granules appearing in cells at the same rate as wild-type, and KH double mutants having much lower numbers of granules (except for KH1-2 and KH3-4; the CRD-BP didomains). KRAS mRNA, which did not bind strongly with any single KH-mutant, may therefore not contribute significantly to the CRD-BP granule core structure, as the granules continue to form under mutant conditions that do not allow KRAS mRNA to bind. While KRAS mRNA, at least in vitro and in HeLa cells does not have a pattern that correlates well with granule formation in zebrafish, its cellular relevance is

not lost. KRAS mRNA and many other mRNAs may not form the core structure of these granules, but may simply be present in lower copy numbers or inter-dispersed throughout the granule as terminal associations within the meshwork, not forming the core scaffolding structure. Interestingly as a side note, these zebrafish granule data allows for the possibility that an intact KH didomain possessing no mutations (either KH1-2 or KH3-4) is sufficient for granule formation, but that mutations in both (all other KH double mutants) didomains may most strongly impair function.

This conclusion is similarly supported by Nielsen *et al.* (2002) who showed with IMP-1 truncation studies (KH1to KH4, KH2 to KH4, and KH1 to KH3) that both intact KH didomains (KH1-2 and KH3-4) are required for IMP-1 to bind H19 RNA in similar EMSA studies. While CRD-BP truncations were not assessed for binding to KRAS mRNA, the results do complement the IMP-1-H19 binding results in that all KH domains are required. The group also found that only the wild-type and KH1to4 recombinant protein (no RRM domains) was able to form RNP granules and exhibit lamellipodia anchoring as determined by confocal microscopy in NIH 3T3 cells (mouse embryo), unlike the Barnes *et al.* (2015) paper where the KH3to4 truncation could still form granules. Given the evolutionary distance between zebrafish and mice, the compared KH3to4 data could be clues to a species-specific intracellular regulator of CRD-BP. These two groups also chose different amino-terminus truncation positions for the KH3to4 construct; Barnes *et al.* (2015): 409-577 and Nielsen *et al.* (2002): 344-577, respectively and perhaps the unique granule forming characteristics lie in the differential residues.

It has been demonstrated that in the *Xenopus* ortholog Vg1RBP, all four KH domains are required for high affinity binding to a 340 nucleotide region of Vg1 mRNA, and in fact the two didomains (KH1to2 and KH3to4) can act cooperatively, much like what appears to be the case

for KRAS mRNA (Git and Standart, 2002). On the other hand, only the KH3to4 didomain of IMP-1 is required to bind the beta-actin 54 nucleotide "zipcode" region with affinity similar to the full-length version (Chao *et al.* 2010). Similarly in chicken embryos, the ZBP1 ortholog utilizes only the two outermost C-terminal KH domains: KH3 and KH4 to bind beta-actin mRNA (Farina *et al.* 2003). While studies pertaining specifically to the KH-domain involvement in VICKZ proteins binding their various mRNA targets are limited, it can already be confidently concluded that different mRNAs require unique KH-domain combinations, with KRAS appearing to require all four domains to achieve a high affinity interaction.

Chapter 4

Assessing the structural impact on GXXG mutations in KH domains on CRD-BP, conformational change upon RNA binding, and KH1to4 crystallization

In this final experimental chapter, the CRD-BP molecule is subjected to multiple structural analyses. The effect mutation of the "GXXG" motif has on overall global secondary structure of CRD-BP is investigated by means of circular dichroism (CD) spectroscopy; the aforementioned single and double KH mutant CRD-BP variants analyzed and compared to a wild-type spectra. Using CD spectroscopy in a similar fashion, the possibility of CRD-BP undergoing significant conformational changes upon RNA binding is also addressed, as well as what effect said binding may have on overall CRD-BP thermal stability. Finally, attempts to produce a crystal structure of CRD-BP are also detailed and cover various aspects of the process including plasmid sub-cloning, recombinant protein expression, and the multiple purification techniques employed. While no diffraction data was ultimately achieved from a CRD-BP protein crystal, much was learned that may help pave the way to eventual success with this particular ongoing project; the findings appearing at the end of this chapter.

4.1 Methodology – Multiple structural analyses of the CRD-BP protein-RNA interaction

4.1.1 Production of CRD-BP KH variants

CRD-BP mutants harbouring the identical GXXG mutations previously described in section 3.3.1 and 3.3.2 (same clones), were used for structural analysis. De-natured samples stored in 8 M urea at -80 ^{o}C were dialyzed against refolding buffer: 200 mM NaCl, 20 mM Tris-Cl, 10% glycerol, 0.01% Triton-X100, pH 7.4 as described in section 2.1.1 however an additional dialysis buffer exchange step was included to optimize CD spectra data acquisition, as

unique buffer constraints exist for CD spectral scanning in the far-UV (185 – 260 nm). Final buffer for CD protein samples consisted of 20 mM Tris-Cl, 200 mM NaCl, 10% glycerol, pH 7.4 and lacked the addition of Triton-X100 which is known to contribute its own CD signal. Quantification of each protein sample was completed using a BCA Protein Assay Kit (Thermo Scientific) with CRD-BP protein samples typically ranging between 3 and 7 μM.

4.1.2 Circular dichroism spectra

Circular dichroism spectropolarimetry scanning was performed on wild-type CRD-BP as well as the KH single and double mutants using a Jasco J-815 CD spectropolarimeter (Jasco Inc.). All wavelength spectra were measured in a 0.1-cm path-length quartz cuvette; samples were scanned in the far UV range (260 – 195 nm) to assess α-helix and β-sheet content. CRD-BP protein samples were diluted into CD buffer (20 mM Tris-Cl, 200 mM NaCl, 10% glycerol, pH 7.4) to between 1 and 2 μM according to the BCA protein assay (Thermo Fischer Scientific Inc.) and scanned at a constant temperature of 25 °C maintained by a Peltier temperature controller. A blank reading consisting of CD buffer only was subtracted from each spectrum. Jasco J-815 CD spectra acquisition parameters were set according to the following (Table 4.1):

Table 4.1. Measurement parameters for CD spectra of CRD-BP KH variants.

Photometric mode:	CD, HT	**Scanning speed:**	20 nm/minute
Measure range:	240-195 nm	**Baseline correction:**	baseline
Data pitch:	0.5 nm	**Shutter control:**	manual
Sensitivity:	standard	**CD detector:**	PMT
D.I.T.:	2 seconds	**PMT voltage:**	auto
Bandwidth:	1.00 nm	**accumulations:**	8
Start mode:	immediately		

Final CD spectra data points were converted from raw CD signal (mdeg) to mean elipticity (deg•cm^2 dmol^{-1}) using the built-in SpectraManager software suite (Jasco Inc.) by

including the optical parameters of CRD-BP concentration and cuvette path-length; here small differences in CRD-BP quantified concentrations could be accounted for prior to a final CD spectra plot. Finally, molar elipticity values were converted to mean residue elipticity, the common unit to represent protein secondary structure analysis, using Microsoft Excel. Data plots were generated using KaleidaGraph™ software.

4.1.3 CRD-BP secondary structure estimation

In order to extract additional quantitative data from the CD analysis of the wild-type CRD-BP and the various single and double KH mutant variants, software algorithms were used to estimate the fraction of α-helix, β-sheet and un-coiled motifs present within each CRD-BP sample from the mean residue elipticity ($deg \bullet cm^2\ dmol^{-1}$) values plot. Numerical data points at 1 nm intervals were determined from a curve trace using built-in Jasco Spectra analysis software from 195-240 nm and converted to Δε, or molar circular dichroism (deciliter mol^-1 cm^-1), and used as the raw input values for secondary structure analysis with the K2D3 circular dichroism web server (Louis-Jeune *et al.* 2011) using generated protein CD reference data sets from DichroCalc (Bulheller and Hirst 2009).

4.1.4 CD monitoring of CRD-BP conformation upon RNA titration

CD spectropolarimetry was used to monitor secondary structure content of CRD-BP following titration with multiple target mRNA regions previously found to bind through EMSA. KRAS, C-myc and Gli all had binding and non-binding fragments that were used for CD comparison, the sequences of which follow below.

c-myc(+)	ACCAGAUCCCGGAGUUGGAAAACAAUGAAAAGGCCCCCAAGGUAGUUAUCCUUAAAAA
1705-1792	AGCCACAGCAUACAUCCUGUCCGUCCAAGC
c-myc(-)	GAAGAGGACUUGUUGCGGAAACGACGAGAACAGUUGAAACACAAACUUGAACAGCUAC
1815-1886	GGAACUCUUGUGCG
Gli (+)	CCAGCUCCCUCGUAGCUUUCAUCAACUCGCGAUGCAC
Gli(-)	CCAGCUCCCUAGCUUUUUUCAUCAACUCGCGAUGCAC

A truncated version of wild-type CRD-BP lacking the RRM domains, containing only KH1 to KH4, was also used in identical experiments with the rational being that if the RRM domains are inert and do not participate in RNA binding, removal of them will amplify any change in CD signal occurring within the KH domains observed upon RNA titration. For these experiments 0.1 cm quartz cuvettes were used for all CD sample analysis and the Jasco J-815 spectropolarimeter was used once again with similar parameters set; all fields as previously indicated (4.1.2) except scan speed was increased to 100 nm/minute and only 3 accumulations performed. This change in data acquisition parameters was made because the higher protein concentration used in these experiments compared to the mutant library analysis yielded lower noise and produced cleaner spectra. Concentrated stock CRD-BP KH1to4 samples were diluted into CD buffer (20 mM NaH_2PO_4, 200 mM NaF, pH 7.4) to a final concentration of 0.1 mg/ml (2.33 μM) in 110 μl samples that also included target RNAs at three different concentrations protein:RNA ratios: 16:1 (0.15 μM), 4:1 (0.58 μM), and 1:1 (2.33 μM). Protein-RNA mixtures were incubated at 35 °C for 10 minutes, followed by 5 minutes on ice, another 10 minutes at 35 °C and a final 5 minutes on ice immediately prior to CD scanning at 25 °C. For each RNA tested, three RNA-blank samples containing the target RNA molecule at identical final concentrations (three protein-RNA ratios each) were scanned, and the resulting RNA-only spectrum subtracted from the protein-RNA sample mixtures to judge the change specific to CRD-BP. For a more straight forward visual comparison of the various RNA's effect on CRD-BP conformation, the 222 nm CD value for each sample, proportionate to the α-helical content of the measured protein, was plotted as a bar graph to help distinguish any RNA that had a unique effect.

4.1.5 Generation of 18S RNA

18S RNA was generated from a MEGAscript T7 transcription kit (Ambion) using linearized pT7-18S plasmid template. Linear 18S template suitable for transcription was first digested in circular form with HindIII. A 40 µl reaction containing (2 µg) of pT7-18S plasmid DNA was incubated with 20 units of HindIII (New England Biolabs Inc.) for 4 hours at 37 °C followed by proteinase K digestion in 0.5% sodium docecyl sulfate for 1 hour at 50 °C. DNA was then purified by a phenol-chloroform extraction where 50 µl of phenol and 50 µl of chloroform was added and vortexed form 2 minutes at 13,000 rpm. To the upper phase (approximately 50 µl), 5 µl of 3 M sodium acetate and 100 µl of 100% ethanol was added, mixed, and incubated at -20 °C for one half hour to allow linearized DNA to precipitate. DNA was subsequently washed with 200 µl of 70% ethanol, and following centrifugation the DNA pellet was re-suspended in 20 µl of nuclease-free water. Purified linearized DNA was then used as a template in the MEGAscript kit to generate approximately 100 µg of pure RNA.

4.1.6 Thermal stability of CRD-BP-RNA complexes

Overall stability of CRD-BP was assessed using a thermal ramp (Peltier temperature controller) coupled with CD monitoring (Jasco J-815) at 222 nm. The wavelength 222 nm is proportional to the overall α-helical content of the protein and is a good measure of protein denaturation. Protein melting temperature was calculated by fitting the molar elipticity to equation 1, using the built-in Jasco Spectral Analysis software for determining Tm, as defined by the temperature where 50% of the total elipticity is lost.

Equation 1.

$$y = \frac{[\min + (\max - min)]}{1 + e^{[-k(x-Tm)]}}$$

CRD-BP protein was added to a 110 ul reaction to a final concentration of 0.4 mg/ml (9.3 μM) in a 20 mM Tris-Cl and 200 mM NaCl buffer. The switch back to NaCl was rationalized by the relaxed far-UV absorbance requirements for measuring only at 222 nm (chloride ions strongly absorb polarized light and interfere with spectra acquisition below 200 nm). Temperature was increased from 20 °C to 85 °C at a 2 °C/minute ramp-rate, 30 second equilibration delay, 0.5 °C pitch and otherwise identical data acquisition parameters as previously indicated (4.1.2). No buffer-only blank subtraction was carried out to establish a baseline, as only the change in CD signal is measured to determine melting temperature and its inclusion would have not changed results.

RNA samples were added to CRD-BP in 4 different ratios protein to RNA ratios (1:1, 4:1, 16:1 and 64:1). Thermal denaturation experiments included KRAS sub-region A34 as the positive binding RNA, and KRAS-A12, 18S and tRNA as negative binding test RNAs. Each experimental condition was repeated in triplicate.

4.1.7 Denaturing purification of CRD-BP KH1to4

BL21 (DE3) E.coli cells were transformed with pET-41c-KH1to4 plasmid and expressed in an identical fashion to section 2.1.3, as with the full-length CRD-BP. Frozen pellets (4) were re-suspended in a denaturing buffer containing 8 M urea, and purified using a Qiagen mini-column in a manner identical to the full-length CRD-BP protein production except "buffer E" (pH 4.9) was skipped, proceeding directly to "buffer F" (pH 4.5). This provided higher concentration in the elution step which was followed by dialysis into a 2 M urea refolding buffer and finally into a 0 M urea storage buffer.

4.1.8 Native purification of CRD-BP KH1to4

BL21 (DE3) E.coli cells were heat-shock transformed (2.1.1) with plasmid pET-41c (4.1.11) containing the truncated coding region of CRD-BP (KH1 to KH4) and plated on LB agar containing 25 μg/ml kanamycin anti-biotic and allowed to grow overnight in a 37 °C incubator. Approximately 20 colonies were picked from the successful transformants and were used to inoculate 1 L of LB broth containing 25 μg/ml kanamycin and incubated at 37 °C until OD_{600} = 0.5 at which point the culture was induced with 1 mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG, Bio Basic Inc.) followed by a 6 hour expression period at 37 °C. Cells were pelleted by centrifugation at 4 °C, 2,100 xg, for 15 minutes (sub-divided into 4 volumes of 250 ml) which produced 4 cell pellets which were stored at -80 °C.

Cell pellets were thawed on ice for 15 minutes and resuspended in 12 ml non-denaturing lysis buffer (table 4.2). Lysozyme (1 mg/ml) was added to the resuspended cell pellet and incubated for 30 minutes on ice followed by sonication lysis consisting of 6 cycles of 20 second bursts (power level 5) – 40 second cool-down time. Lysate was subsequently centrifuged for 30 minutes at 10,000 xg, 4 °C. Lysate supernatant was mixed with 2 ml of 50% Ni-NTA agarose resin slurry (Qiagen) and incubated on ice for 1 hour on a waver. The slurry was then loaded onto a Qiagen mini-column and the beads washed twice with 4 ml of wash buffer (table 4.2) followed by elution with 4 ml of elution buffer; collecting 0.5 ml fractions off the column. SDS-PAGE analysis was completed to assess the presence and purity of the fractions containing CRD-BP KH1to4 with a final 3-stage dialysis step against a storage buffer consisting of 50 mM NaH_2PO_4, 200 mM NaCl, pH 7.4 using 100 μl mini-dialysis units (Thermo Scientific).

Table 4.2. Buffers used in native purification of CRD-BP KH1to4

Non-denaturing lysis buffer	Wash buffer	Elution buffer
50 mM NaH_2PO_4	50 mM NaH_2PO_4	50 mM NaH_2PO_4
200 mM NaCl	200 mM NaCl	200 mM NaCl
10 mM imidazole	52.8 mM imidazole	250 mM imidazole
pH 8.0 using NaOH	pH 8.0 using NaOH	pH 7.4 using NaOH

4.1.9 Denaturing on-column refolding purification of CRD-BP KH1to4

High concentration CRD-BP KH1to4 protein required the use of an "on-column" protein refolding method whereby proteins are denatured and then re-folded while still bound to a Ni-NTA column to avoid inter-molecular aggregates forming. The process is described here and also a visual guide can be referred to (figure 4.1). BL21 (DE3) E.coli cells were heat-shock transformed (2.1.1) with plasmid pET-41c (4.1.11) containing the truncated coding region of CRD-BP (KH1 to KH4) and plated on LB agar containing 25 µg/ml kanamycin anti-biotic and allowed to grow overnight in a 37 °C incubator. Approximately 20 colonies were picked from the successful transformants and were used to inoculate 4x 1 L (4 L total culture) of LB broth containing 25 µg/ml kanamycin and incubated at 37 °C until OD_{600} = 0.5-0.7 (each 1 L culture varied slightly at time of reading) at which point the cultures were induced with 1 mM IPTG (Bio Basic Inc.) followed by a 6 hour expression period at 37 °C. Cells were pelleted by centrifugation at 4 °C, 2100 xg, for 15 minutes and cell pellets stored at -80 °C. Cell pellets were resuspended and lysed in 48 ml of 8 M urea denaturing lysis buffer (table 4.3) and placed on a waver for 1 hour on ice. Following cell lysis, the lysate was centrifuged for one half hour at 10,000 xg (4 °C) followed by filtration with a 0.45 µm syringe filter (EMD Millipore) yielding approximately 45 ml of cleared lysate.

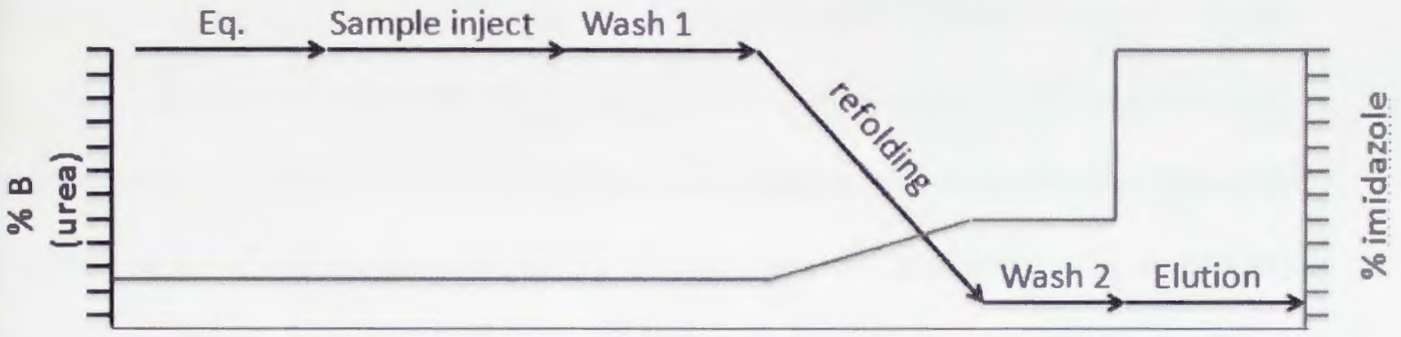


Figure 4.1. Diagram of on-column refolding KH1to4 protein purification scheme. Urea concentration ranges from a max of 8 M to minimum of 0 M (%B – denaturing lysis buffer) and imidazole ranges from a max of 250 mM to a minimum of 10 mM, with an intermediate wash concentration of 52.8 mM.

Cleared lysate was loaded onto an ÄKTAprime plus fast-protein liquid chromatography machine (FPLC, GE Healthcare Life Sciences) using a 50 ml ÄKTAprime Superloop™ (GE Healthcare Life Sciences) sample holder. Cleared lysate was then injected over a 5 ml Ni-NTA His-trap™ column (GE Healthcare Life Sciences) at a flow rate of 0.5 ml/min, with 50 ml of denaturing lysis buffer pumped through the Superloop™, behind the sample plunger to ensure no inadvertent mixing with non-denaturing buffer. CRD-BP bound to the His-trap™ column was washed with 5 column volumes (25 ml) of denaturing lysis buffer at 5 ml/min. On-column protein refolding was then initiated, where urea concentration was slowly reduced from 8 M urea to 0 M urea in the presence of oxidized/reduced glutathiones at a rate of 0.5 ml/min (100 ml program span) by progressively increasing the percentage of refolding/wash buffer (table 4.3) passed over the column. This was accomplished by utilizing the ÄKTAprime plus mixing valve (from 0 to 100% refolding/wash buffer). A second wash step passing 10 column volumes (50 ml) over the Ni-NTA column at 5 ml/min was programmed followed by the final elution step - passing 25 ml of elution buffer at 1 ml/min over the Ni-NTA His-trap™ column, collecting 1.0 ml fractions. Collected fractions were analyzed by SDS-PAGE and the six purest fractions pooled prior to a 3-stage dialysis against a final working buffer consisting of 100 mM Tris-Cl,

500 mM NaCl, pH 6.0 (using HCl) using two 3 ml 3,500 MWCO Slide-A-Lyzer™ dialysis cassettes (Thermo Scientific). The 43 kDa protein was doubly quantified with a Nanodrop spectrophotometer (Thermo Scientific) using a calculated extinction co-efficient (Protparam – EXPASY web tools; Gasteiger et al, 2005) of 8,940 M-1 cm-1 and additionally using a BCA protein quantification kit (Pierce Thermo Scientific). For crystallization trials, the protein was further concentrated to approximately 10 mg/ml using 10,000 NMWL (10 kDa) Amicon Ultra 2 ml centrifugal filters (EMD Millipore corp.).

Table 4.3. Buffers used in denaturing on-column refolding protein purification. GSH – reduced form glutathione, GSSG – oxidized form glutathione.

Denaturing lysis buffer	Refolding/wash buffer	Elution buffer
50 mM NaH_2PO_4	50 mM NaH_2PO_4	50 mM NaH_2PO_4
200 mM NaCl	200 mM NaCl	200 mM NaCl
10 mM imidazole	52.8 mM imidazole	250 mM imidazole
8 M urea	0 M urea	0 M urea
-	1 mM GSH/0.1 mM GSSG	-
pH 8.0	pH 8.0	pH 7.4

4.1.10 Fluorescence polarization analysis of CRD-BP binding KRAS RNA.

Fluorescence polarization (FP) spectroscopy was used to confirm the RNA binding activity of CRD-BP KH1to4 protein prepared by on-column refolding. A fluorescein-labelled KRAS probe was designed and synthesized by Integrated DNA Technologies Inc (IDT) based on the KRAS-A34(b) fragment found to bind using EMSA in section 2.3.2. The *mfold* RNA web server was used to fold the KRAS-A34(b) RNA fragment (57 nucleotides) in silico, and based on the secondary structure prediction, a 44-nucleotide fragment containing the nucleotides comprising the computed structured region was chosen. Binding reactions were assembled as 19 µl volumes containing CRD-BP KH1to4 protein (0-1000 nM), KRAS FP probe (10 nM) in a

buffer similar to EMSA binding assays (2.1.6); 20 mM Tris-Cl, 200 mM NaCl, pH 7.4. Immediately following reaction assembly, components were mixed by mild centrifugation at 3,000 rpm in a tabletop centrifuge for 10 seconds, mixed by flicking, and centrifuged again to pool the reaction. Binding reactions were then incubated at 35 °C for 10 minutes, on ice for 5 minutes, followed by another 35 °C for 10 minutes and 5 minutes on ice. Into a 384-well microplate, 17 µl of each sample reaction mix was added and subsequently the total fluorescence and parallel/perpendicular fluorescence values were measured using a Synergy 2 Multi-mode Microplate Reader and Gen5 Reader Control and Data Analysis Software (Biotek Instruments Inc.). Optical excitation filter and emission filter were set to 485/20 nm and 528/20 nm respectively and system gain set to auto-adjust with a read-height of 7.00 mm. Fluorescence anisotropy values were plotted for each CRD-BP KH1to4 concentration sample in KaleidaGraph™ graphing software to determine Kd. As a positive control, wild-type full-length CRD-BP was used for comparison with Bovine Serum Albumin (BSA, Sigma Aldrich) being used as a negative control in otherwise identical experimental conditions.

4.1.11 Construction of pET-41c(+)-CRD-BP KH1to4 plasmid

Crystallization of CRD-BP was also attempted with a truncated form containing only the KH1 to KH4 binding module (RRM1 and RRM2 removed) which was sub-cloned into a pET41c(+) that had been previously modified to contain no N-terminus GST-tag sequence. The plasmid vector was generously donated from Dr. Natalie Strynadka's X-ray crystallography laboratory at the University of British Columbia and is ideal for protein expression oriented towards structural studies due to a thrombin cleavable 8x his-tag and no GST-tag. NdeI and HindIII restriction enzymes (New England Biolabs Inc) were used to accomplish sub-cloning of KH1to4 CRD-BP into pET41c(+) plasmid.

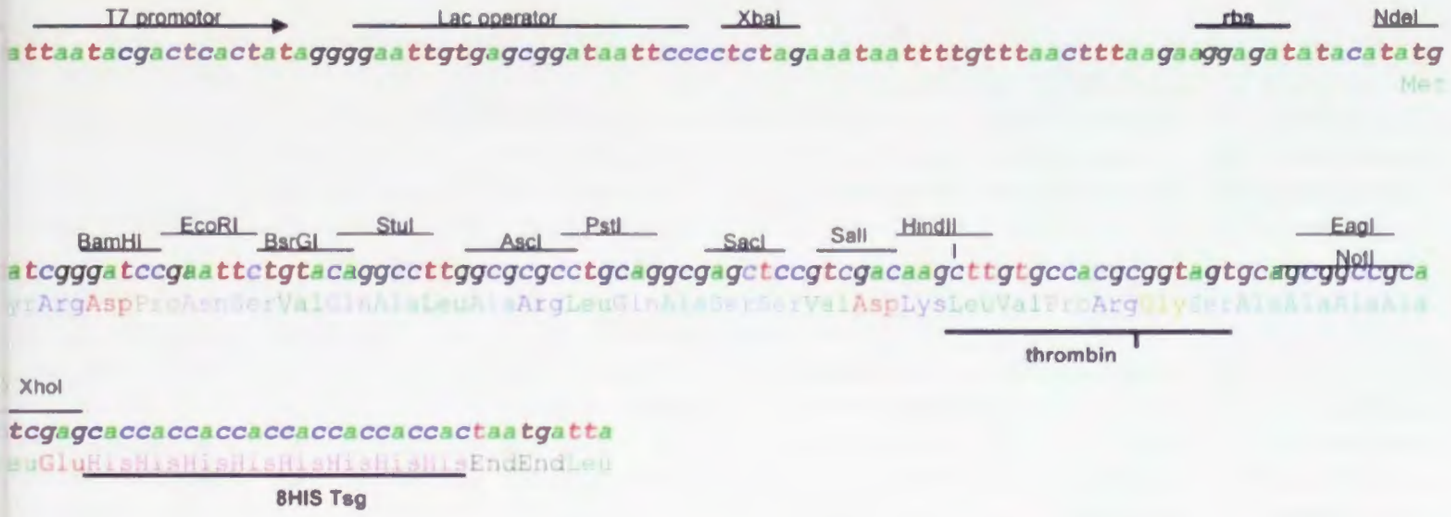


Figure 4.2. pET41c(+) –no GST plasmid multiple cloning site sequence. Note the thrombin cleavable 8x His-tag and lack of GST sequence. Nde1 and HindIII restriction sites were utilized for sub-cloning full-length and KH1to4 CRD-BP coding sequence.

CRD-BP KH1to4 insert DNA generation by PCR

PCR was utilized to generate the DNA insert (table 4.4); with wild-type CRD-BP plasmid pET28b backbone) serving as the reaction template. Primers were designed to amplify a 1,113 nucleotide region of the coding sequence encompassing KH1 to KH4 as outlined below.

CRD-BP KH1to4 forward primer:

Design scheme: 5' [handle]-[NdeI]-[complementary region] 3'
DNA sequence: 5' [ACCA]-[CATATG]-[ATCCCTCTCCGGCTCCTG] 3'
inal sequence: 5'-ACCACATATGATCCCTCTCCGGCTCCTG-3'

CRD-BP KH1to4 reverse primer:

Design scheme: 5'[complementary region]-[HindIII]-[handle] 3'
on-R.C. DNA sequence: 5'[AAGCAACAGCACCAGAAG]-[AAGCTT]-[ACCA] 3'
inal DNA sequence: 5'-ACCAAAGCTTCTTCTGGTGCTGTTGCTT-3'

*R.C. – reverse complement
*handle (ACCA) not reverse-complement

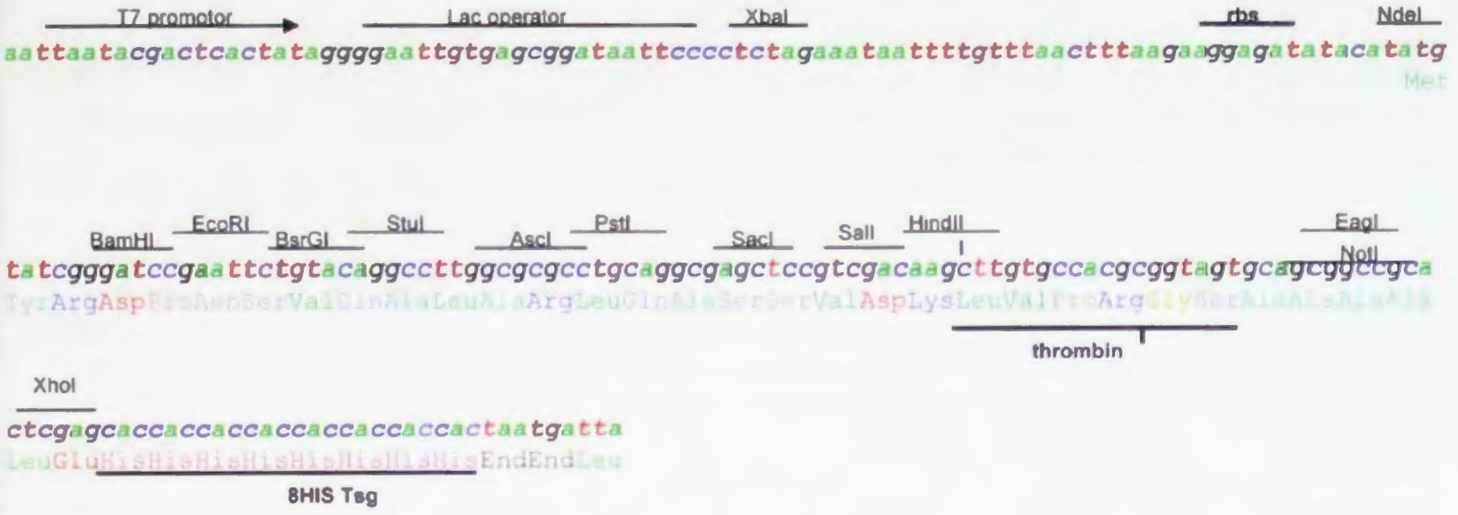


Figure 4.2. pET41c(+) –no GST plasmid multiple cloning site sequence. Note the thrombin cleavable 8x His-tag and lack of GST sequence. Nde1 and HindIII restriction sites were utilized for sub-cloning full-length and KH1to4 CRD-BP coding sequence.

CRD-BP KH1to4 insert DNA generation by PCR

PCR was utilized to generate the DNA insert (table 4.4); with wild-type CRD-BP plasmid (pET28b backbone) serving as the reaction template. Primers were designed to amplify a 1,113 nucleotide region of the coding sequence encompassing KH1 to KH4 as outlined below.

CRD-BP KH1to4 forward primer:

Design scheme: 5' [handle]-[NdeI]-[complementary region] 3'
DNA sequence: 5' [ACCA]-[CATATG]-[ATCCCTCTCCGGCTCCTG] 3'
Final sequence: 5'-ACCACATATGATCCCTCTCCGGCTCCTG-3'

CRD-BP KH1to4 reverse primer:

Design scheme: 5'[complementary region]-[HindIII]-[handle] 3'
non-R.C. DNA sequence: 5'[AAGCAACAGCACCAGAAG]-[AAGCTT]-[ACCA] 3'
Final DNA sequence: 5'-ACCAAAGCTTCTTCTGGTGCTGTTGCTT-3'

*R.C. – reverse complement
*handle (ACCA) not reverse-complement

CRD-BP KH1to4 amplicon DNA sequence:

```
                                                                    atcc
ctctccggct cctggtgcct acgcagtatg taggcgctat cattggcaag gagggtgcca
ccatccgaaa catcacaaaa cagacgcagt ccaaaataga cgtgcatagg aaggagaatg
cgggcgctgc ggagaaggcc atcagcgtgc attcaacccc tgaaggctgc tcctccgcgt
gcaagatgat cttggagatt atgcacaagg aggcaaagga caccaaaacg gcagatgaag
ttcccctgaa gatcctggct cataacaact tcgtcgggcg actcattggc aaggaagggc
ggaacctgaa gaaggtggag caggacacag agacgaagat caccatctca tcgctccagg
acctcacgct ctataaccct gagaggacca tcactgtgaa gggcgccatt gagaactgtt
gcagggccga gcaggagatc atgaagaaag ttcgagaggc ttacgagaac gacgtggccg
ccatgagctt gcagtcccac ctcatccctg ggcttaacct ggctgctgta ggtctcttcc
cagcttcatc cagcgctgtc cctcctcctc ccagcagtgt caccggggct gctccctata
gctccttcat gcaggctccg gagcaggaga tggtacaagt gttcatcccc gcccaggctg
tgggcgccat cattggcaag aagggccagc acatcaaaca actctcccgc ttcgccagcg
cctccatcaa gattgcacca ccagaaacac ctgactccaa agttcgaatg gtcgtcatca
ctggaccccc agaggctcag ttcaaggccc agggaagaat ctatggcaaa ctaaaagaag
agaatttctt tggtcccaag gaggaagtaa agctagagac ccacatacgg gttccggctt
cagcagccgg ccgtgtcatc ggcaaaggcg gcaaaacggt gaatgagctg cagaacttga
ccgcagctga ggtggtagtg ccaagagacc agaccccgga tgagaacgac caagtcattg
ttaagatcat cggacatttc tatgccagcc agatggctca gcggaagatc cgagacatcc
tggctcaagt taagcaacag caccagaag
```

PCR mixtures were assembled in triplicate and later combined to provide high working concentration. Each 25 µl PCR reaction was assembled according to table x and a thermocycler scheme programmed for optimal amplification. High-fidelity Phusion™ DNA polymerase (New England Biolabs Inc.) was used in order to reduce likelihood of mutation during insert amplification. PCR reactions upon completion were then run on a 1% agarose gel at 120 V stained with 0.5 µg/ml ethidium bromide and subsequently gel-purified using a QIAEX II Gel Extraction Kit (Qiagen) following the recommended protocol.

PCR thermocycle program:

Step	Condition		
1	98 °C	30 sec	
2	98 °C	15 sec	step 5 – repeat 25x
3	55 °C	30 sec	
4	72 °C	2 min	
6	72 °C	10 min	
7	4 °C	forever	

Table 4.4. PCR reagents used to generate KH1to4-pET-41c(-GST) plasmid vector.

Reagent	volume
5x High-fidelity (HF) Phusion buffer (NEB)	5 µl
dNTPs (2.5 mM)	2.5 µl
KH1to4 forward primer (10 µM)	1 µl
KH1to4 reverse primer (10 µM)	1 µl
CRD-BP template plasmid (41 ng/ µl)	2.4 µl
Phusion DNA polymerase (2,000 U/ml; NEB)	0.2 µl
H_2O	to 25 µl

Vector plasmid (pET41c) for sub-cloning, which was donated from UBC as a low concentration, small volume sample, was amplified in a 10 ml DH5α E.coli cell culture volume following heat-shock transformation (2.1.1) and plating onto an LB agarose plate with 25 µg/ml kanamycin antibiotic. Overnight-growth colonies were cultured in 10 ml LB broth volumes (25 µg/ml kanamycin) for 16 hours, followed by plasmid purification using a QIAprep Spin miniprep kit (Qiagen). Due to pET41c being a low-copy plasmid, two 5 ml culture volumes were prepared separately, each according to the manufacturer's recommendations and then combined by passing both plasmid preparations over a single QIAprep spin column and eluted together to yield approximately 5 µg of plasmid DNA (100 ng/µl, 50 µl).

Double-digest restriction endonuclease reactions were setup for the insert and pET41c empty vector DNA samples in preparation for sub-cloning. Digest reactions were setup in 20 µl volumes according to table 4.5 and allowed to continue for 2 hours at 37 °C. Digested DNA products were resolved on a 1% agarose gel and gel-purified from excised DNA bands using a QIAEX II Gel Extraction Kit (Qiagen) following the recommended protocol (note that KH1to4 insert DNA elution is at room temperature while the larger pET41c vector DNA is eluted at 50

Table 4.6. Reagents used for CRD-BP KH1to4 plasmid ligation. Plasmid pET41c(-GST) MW approximately 5 kb and KH1to4 insert DNA 1.2 kb; values used to calculated molar ratios.

Reagent	1:1	3:1	7:1	Vector only
10X T4 buffer (NEB)	2 µl	2 µl	2 µl	2 µl
KH1to4 DNA	10 ng	35 ng	85 ng	0 ng
pET41c DNA	50 ng	50 ng	50 ng	50 ng
T4 DNA ligase (NEB)	1 µl	1 µl	1 µl	1 µl
H_2O	to 20 µl	to 20 µl	to 20 µl	to 20 µl

Colony PCR was used to screen for successful ligations containing the CRD-BP KH1to4 insert. Colonies were sampled directly off the LB agarose plate using a sterile loop and used to inoculate small PCR reaction tubes. One reaction was assembled for each colony sampled according to table 4.6 and run on a 1% agarose gel for UV ethidium bromide detection of appropriate band. Successful transformants were then used to inoculate 10 ml LB broth cultures and plasmid purified using the QIAprep Spin miniprep kit (Qiagen) as described previously for the empty pET41c vector.

DNA sequencing was completed for three positive clone plasmids to further confirm the presence of correct DNA insert, and also to ensure no mutations were present in the sequence. Plasmid samples were diluted to 100 ng/µl and 10 µl volumes were shipped to Macrogen for next generation sequencing using three different sequencing primers for each plasmid: T7 promoter forward primer, T7 terminator reverse primer and a custom designed "KH1P" forward primer that anneals to the 3' end of KH1 DNA sequence.

Primers used for DNA sequencing of pET41c-KH1to4:

T7 promoter forward primer sequence: 5'- TAA TAC GAC TCA CTA TAG GG -3'

T7 terminator reverse primer sequence: 5'- GCT AGT TAT TGC TCA GCG G -3'

"KH1P" forward primer sequence: 5'- GGC GCT ATC ATT GGC AAG GAG GGT GCC -3'

Table 4.7. Reagents for colony PCR screening.

PCR reaction mix		Thermocycle program		
reagent	**volume**	**step**	**temperature**	**time**
10X PCR buffer	3.5 µl	1	95 °C	5 min
dNTPs (2.5 mM)	3.5 µl	2	95 °C	1 min
KH1to4_F primer	1 µl	3	54 °C	90 sec
KH1to4_R primer	1 µl	4	72 °C	2 min
Taq. Pol.	0.5 µl	5	to step 2 – 30X	
H_{2O}	25.5 µl	6	72 °C	5 min

4.2 Results and Discussion

4.2.1 Circular Dichroism analysis of CRD-BP KH mutant variants

CRD-BP KH single and double mutant variants harbouring the GXXG mutations used previously for EMSA binding experiments (3.3.1 and 3.3.2), were scanned using a circular dichroism (CD) spectropolarimeter and far-UV spectra acquired from 240-190 nm for each mutant protein. All protein samples were freshly dialyzed and refolded into CD buffer and were all between 1 and 3 µM concentration prior to loading into a 0.1 cm quartz cuvette for CD scanning. Acquired CD spectra for each KH mutant were remarkably similar to wild-type CRD-BP following conversion from raw CD milli-degrees to mean residue elipticity (figure 4.3). The mean residue elipticity unit accounts for differences in scanned protein concentration and returns the average CD signal per amino acid residue, making inter-protein sample comparisons more feasible. Far-UV wavelength scans of wild-type CRD-BP reveal a spectrum characteristic of a protein with mixed alpha-helix/beta-sheet structure with a strong broad peak at 222 nm, indicative of a large alpha-helical component. This finding matches known structures of KH

domain-containing proteins as well as other CD spectra derived from single and tandem KH domain proteins (Chao *et a.l* 2010; Valverde *et al.* 2007; Chmiel *et al.* 2006).

It is immediately clear upon inspection of the CRD-BP KH mutant CD spectra that the global structures are remarkably similar and do not show any large differences in the location of spectra peaks or the amplitude of CD signal. Subtle differences do exist between the spectra upon close inspection, most notably at wavelengths lower than 205 nm, and higher than 230 nm. These differences are small and strongly suggest that the overall fold of the various KH GXXG mutants appears to be maintained. However, as far-UV spectra analysis is not a measure of the tertiary structure of proteins, it remains possible that relative changes in alpha-helix/beta-sheet orientation have occurred as a result of introducing GXXG mutations. As well, shifts in the precise α-helix and β-sheet residue boundaries are possible, so long as the overall fraction of each motif remains constant. Wild-type CRD-BP was scanned at 90 °C as a control for denatured protein (figure 4.3D), and far-UV spectra showed weaker CD signal as expected, notably in the 208 nm and 222 nm regions where protein secondary structure contributes strongly.

Further structural analysis involving secondary structure calculations from CRD-BP CD spectra were also completed with the aid of the K2D3 secondary structure estimation web server, and the Dichrocalc online CD reference-spectra database. Inputting the Δε values, or molar circular dichroism (deciliter mol^-1 cm^-1) at 1 nm intervals for each mutant produced highly similar final secondary structure estimations, varying by less than a percent between each sample (table 4.8). Given the tight overlap in CD spectra between wild-type CRD-BP and the various mutant forms, especially at the critical wavelengths used for computing secondary structure estimations, this result was expected and brings numerical scope as to the likeness of proteins within the CRD-BP mutant library.

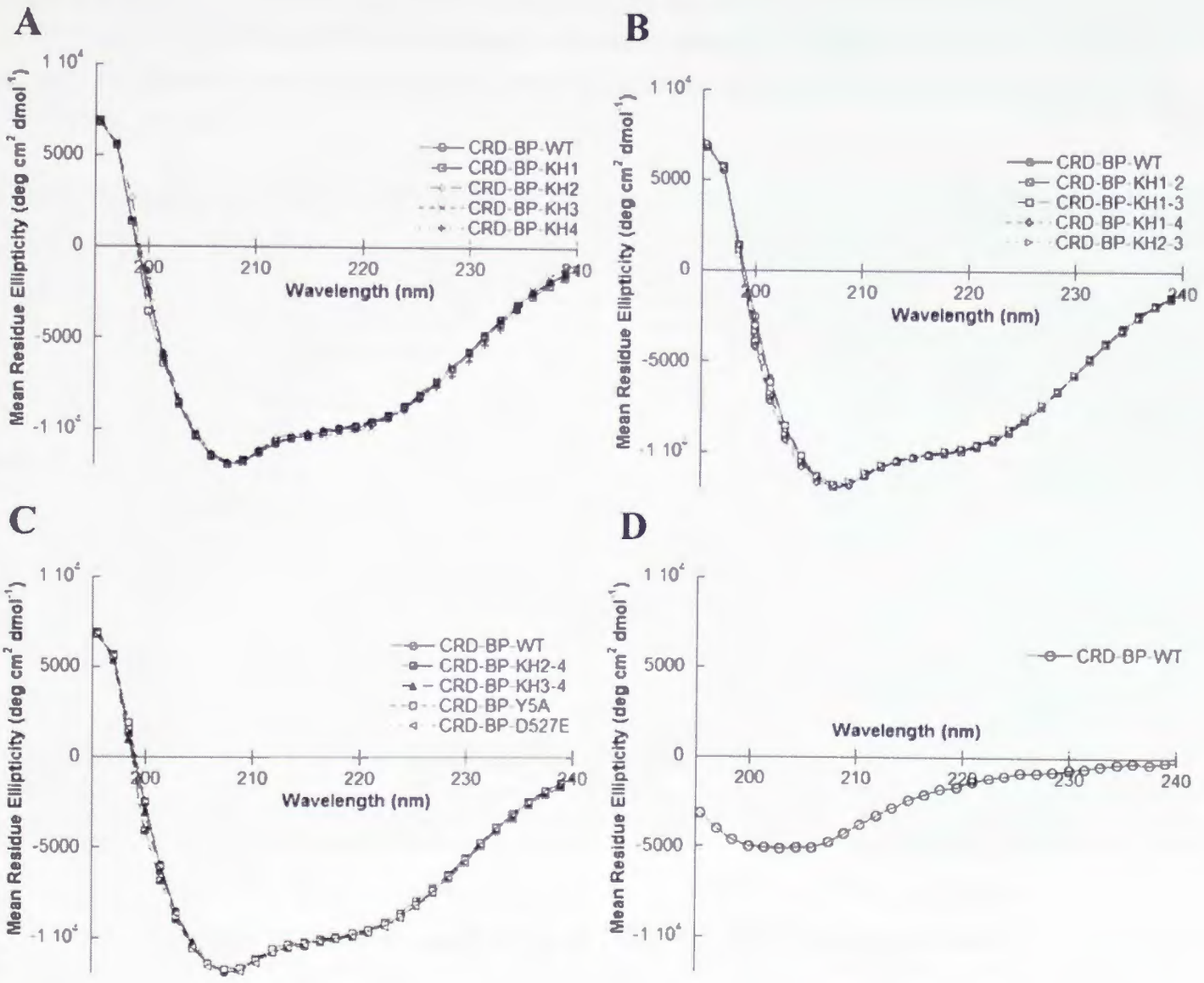


Figure 4.3. CD spectra of wild-type CRD-BP and KH mutant variants. Circular dichroism spectra of (A) wild-type, KH1, KH2, KH3 and KH4; (B) wild-type, KH1-2, KH1-3, KH1-4, KH2-3; (C) wild-type, KH2-4, KH3-4, Y5A, and D527E. (D) Wild-type CRD-BP spectral scan at 90 °C as positive control for denatured protein. All spectra generated from 8 accumulations using a Jasco J-815; proteins scanned in 20 mM Tris-Cl, 10% glycerol, 200 mM NaCl, pH 7.4 buffer.

Table 4.8. Protein secondary structure estimations for CRD-BP variants. K2D3 neural net circular dichorism structural determination web server was used in conjunction with Dichrocalc online CD reference spectra database to generate estimations from Δε (deciliter mol^-1 cm^-1) CD spectra values.

CRD-BP variant	% α-helix	% B-sheet	% Random coil
Wild-type	67.51	9.00	23.49
KH1	67.52	8.89	23.59
KH2	67.73	9.10	23.17
KH3	67.51	9.02	23.47
KH4	67.51	9.00	23.49
KH1-2	67.55	9.12	23.33
KH1-3	67.53	9.00	23.47
KH1-4	67.48	9.06	23.46
KH2-3	67.51	9.00	23.49
KH2-4	67.51	9.00	23.49
KH3-4	67.51	9.02	23.47

4.2.2 Circular dichroism monitoring of CRD-BP-RNA binding in vitro

Assessing whether or not CRD-BP undergoes a conformation change upon RNA interaction is important, as it may serve as a measure or indicator for RNA binding. Also, it sheds light on how the protein functions and can highlight the different relevancies of crystallographic structures published with and without RNA bound. To determine if CRD-BP has such dynamic character, CD spectroscopy was used as before in a series of RNA titration experiments with both high and low affinity RNA substrates. In addition to the full-length wild-type form, a truncated version of CRD-BP was created containing only the KH1 to KH4 binding module (referred to from now on simply as "KH1to4"), and used extensively in these experiments. The rationale behind this move to include KH1to4 was that the functionally inert

RRM domains have been determined not to contribute to RNA binding affinity, and thus are not likely to shift from one conformation to another. Hence, inclusion of the static RRM domains was anticipated as potentially dulling any observed changes in CD signal stemming from the KH domains and therefore removed. Full-length CRD-BP was examined first, measuring the complete far-UV spectrum from 190-260 nm and titrated with KRAS A34, KRAS A12, the c-myc CRD-region, a non-binding c-myc region, as well as a pair of GLI competitor oligonucleotides with relatively strong competition and weak competition, labelled GLI (+) and GLI(-) respectively. Furthermore, as a negative control for CD-observable structural changes, BSA was titrated with KRAS A34 RNA and spectra obtained in an otherwise identical fashion. Each sample was prepared in terms of protein:RNA molar ratio equivalents; 4:1 and 1:1 preparations were completed and scanned for each RNA substrate. It can be seen in figure 4.4, that the addition of any RNA substrate - high affinity or low affinity, appears to induce a change in the CD spectra in a similar fashion. At all wavelengths, the measured dichroism reduced upon titration with RNA, however while there is a lack of specificity for the general effect, it can be seen that in every experiment (figure 4.4 A-C), the stronger binding fragment was able to induce a larger change at the same concentration. As RNA alone possesses CD character, it is important to subtract the RNA spectra measured at the same concentration from the titration samples; an RNA-CD blank. In figure 4.4D, BSA was measured with KRAS A34 and A12, and the RNA spectra subtracted, producing essentially identical CD spectra reflecting both the unchanging nature of BSA in the presence of RNA, and also validating the spectral subtraction technique in principle. A reduction in absolute molar elipticity, as observed in all CRD-BP-RNA titrations, reflects an un-coiling effect of the various alpha-helices comprising the protein as well as beta-sheets. The effect is not absolute, as the CD signal does not base-line losing only about 25%

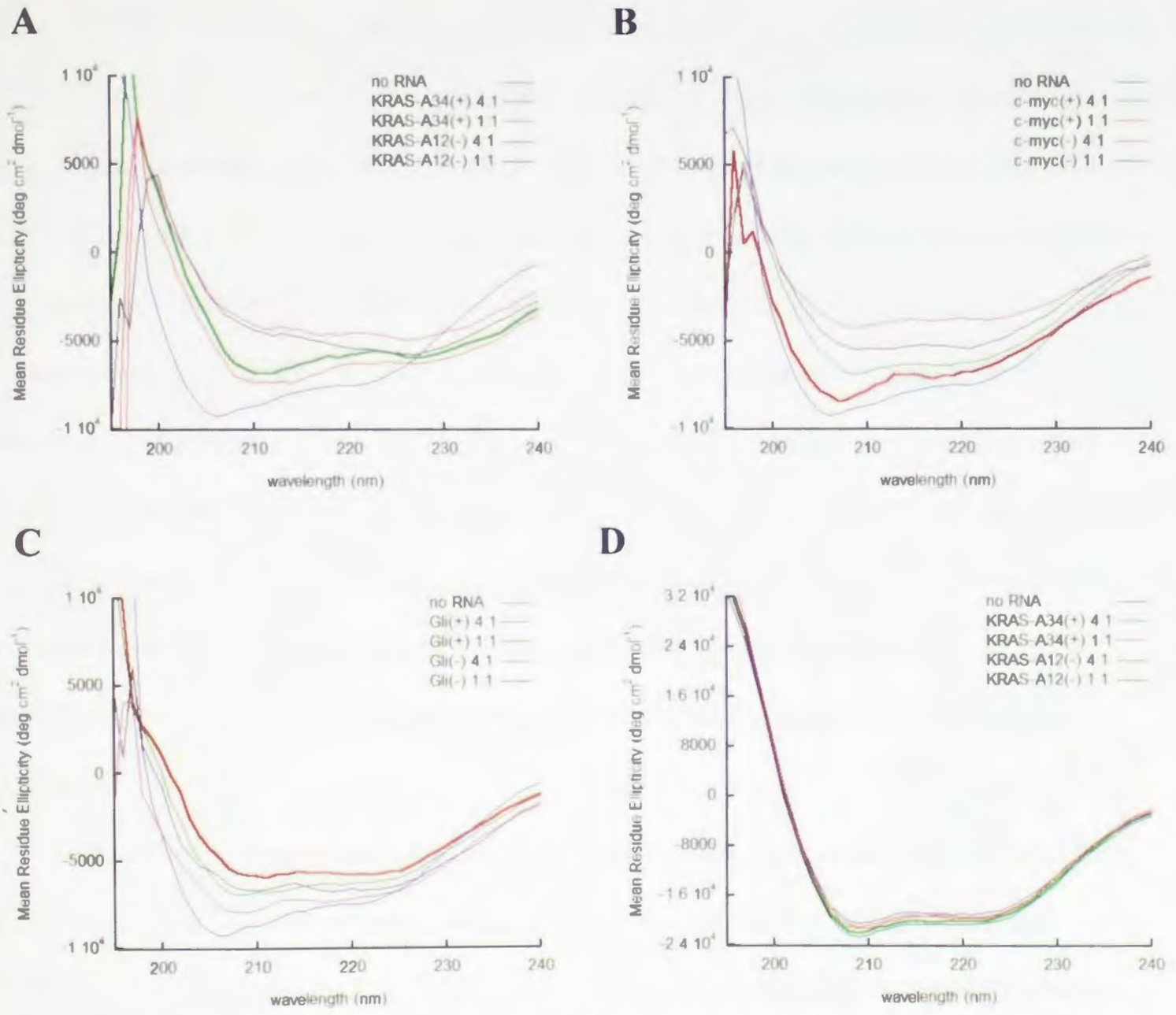


Figure 4.4. CD spectra of CRD-BP titrated with various substrates. All samples are presented as a molar ratio of protein to RNA, with RNA concentration increasing from 4:1 to 1:1. (A) KRAS A34 is a high affinity substrate while KRAS A12 binds poorly. RNA substrates also included: (B) c-myc (+) and c-myc (-) are strong and weak binding substrates respectively. (C) Gli (+) and Gli (-) are strong and weak oligonucleotide competitors respectively with (D) containing scans of BSA acting as a negative conformational shift control.

structure, however a loss in structure does reflect and less stable and higher free energy conformation. This is by itself an unlikely circumstance as RNA binding is spontaneous and does not require energy input, however this energy barrier may be overcome by the free-energy contribution of RNA binding, satisfying charges and possibly orienting hydrophobic residues into more energy favourable positions.

Further experimental data was extracted from these spectra, comparing most directly the alpha-helical content of each titration sample as 222 nm bar charts (figure 4.5). As each CD spectra was measured in triplicate and graphed as the average of all accumulations, the 222 nm value in figure 4.5 is an average with error bars representing standard deviation. This extracted 222 nm data allows for relatively easy comparison of the RNA substrates effect on CRD-BP protein structure. It can be easily seen in the bar charts that indeed the stronger binding substrates, most notably c-myc, labelled with "(+)" are reliably more able to reduce the CD signal strength, albeit on marginally more than their weaker binding "(-)" counterparts. As can be quickly deduced from the raw CD-spectra in figure 4.4D, BSA here shows no distinguishable pattern in protein secondary structure changes, with neither substrate affinity (KRAS-A34 vs KRAS-A12) nor RNA concentration correlating with the relatively small changes in 222 nm elipticity.

Another important accompanying piece of data in CD experiments is the recorded high tension voltage (HTV), essentially a measure of the required signal amplification or gain required to distinguish noise from CD signal. Certain buffers or proteins may absorb strongly in the far UV, resulting in higher voltages being required to acquire spectra. However, proteins precipitating from solution also typically show strong increases in HTV. Patel *et al.* (2012) proclaim that ZBP1 bound to RNA possesses reduced solubility compared to the protein alone, which has been a major stumbling block in the quest for a full-length crystal structure of ZBP1 or any of its homologs including CRD-BP. Reduced solubility of the CRD-BP-RNA complexes in the experiments presented here would indeed show loss of CD signal, as is observed.

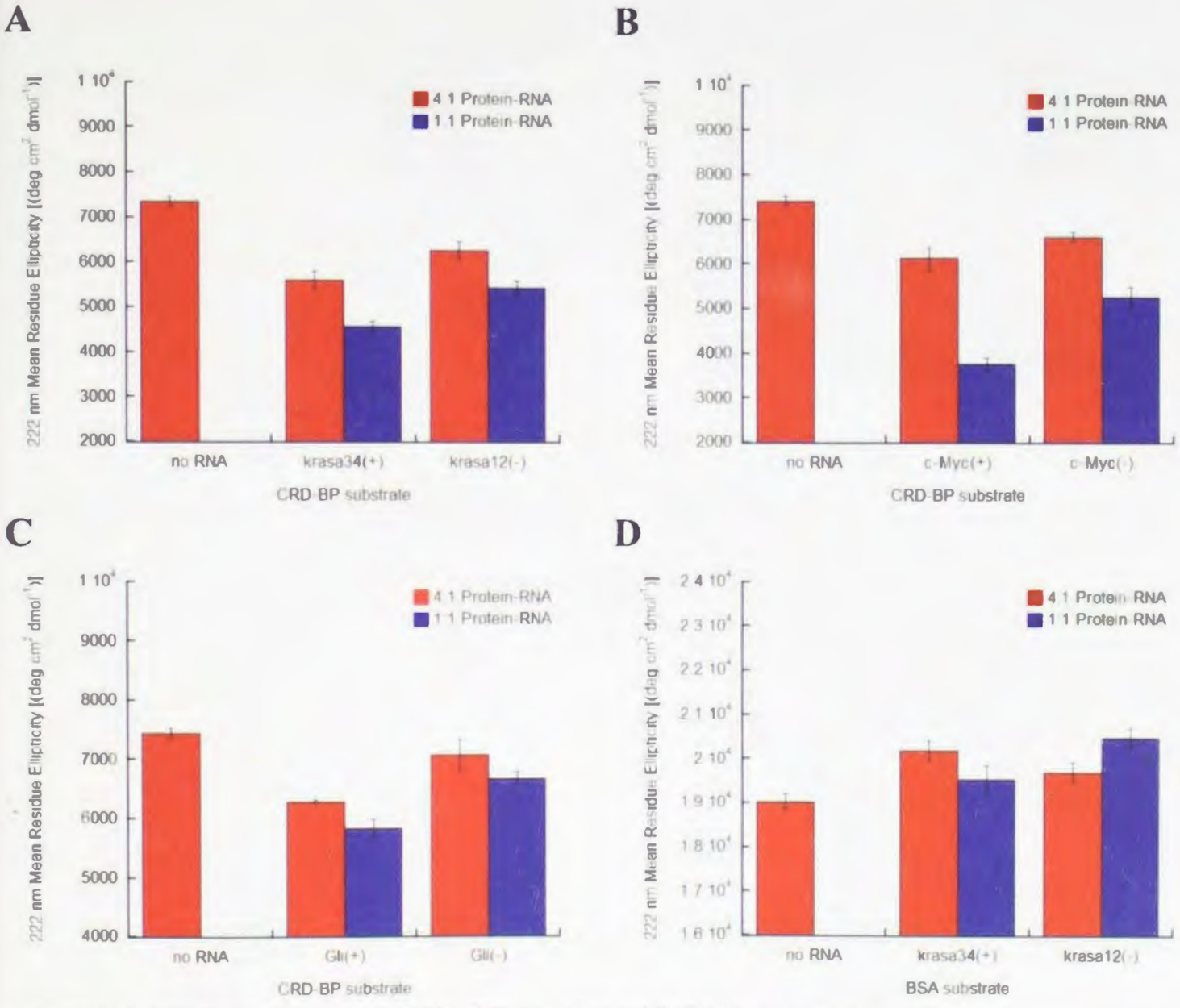


Figure 4.5. MRE at 222 nm of CRD-BP titrated with RNA substrates. All samples are presented as a molar ratio of protein to RNA with 4:1 having lower RNA concentration than 1:1 and mean residue elipticity (MRE) presented as absolute values. (A) KRAS A34 was determined to be a positive binding fragment while KRAS A12 binds poorly. (B) c-myc (+) and c-myc (-) are strong and weak binding substrates respectively. (C) GLI (+) and GLI (-) are strong and weak oligonucleotide competitors respectively with (D) containing 222 nm trace of BSA acting as a negative conformational shift control.

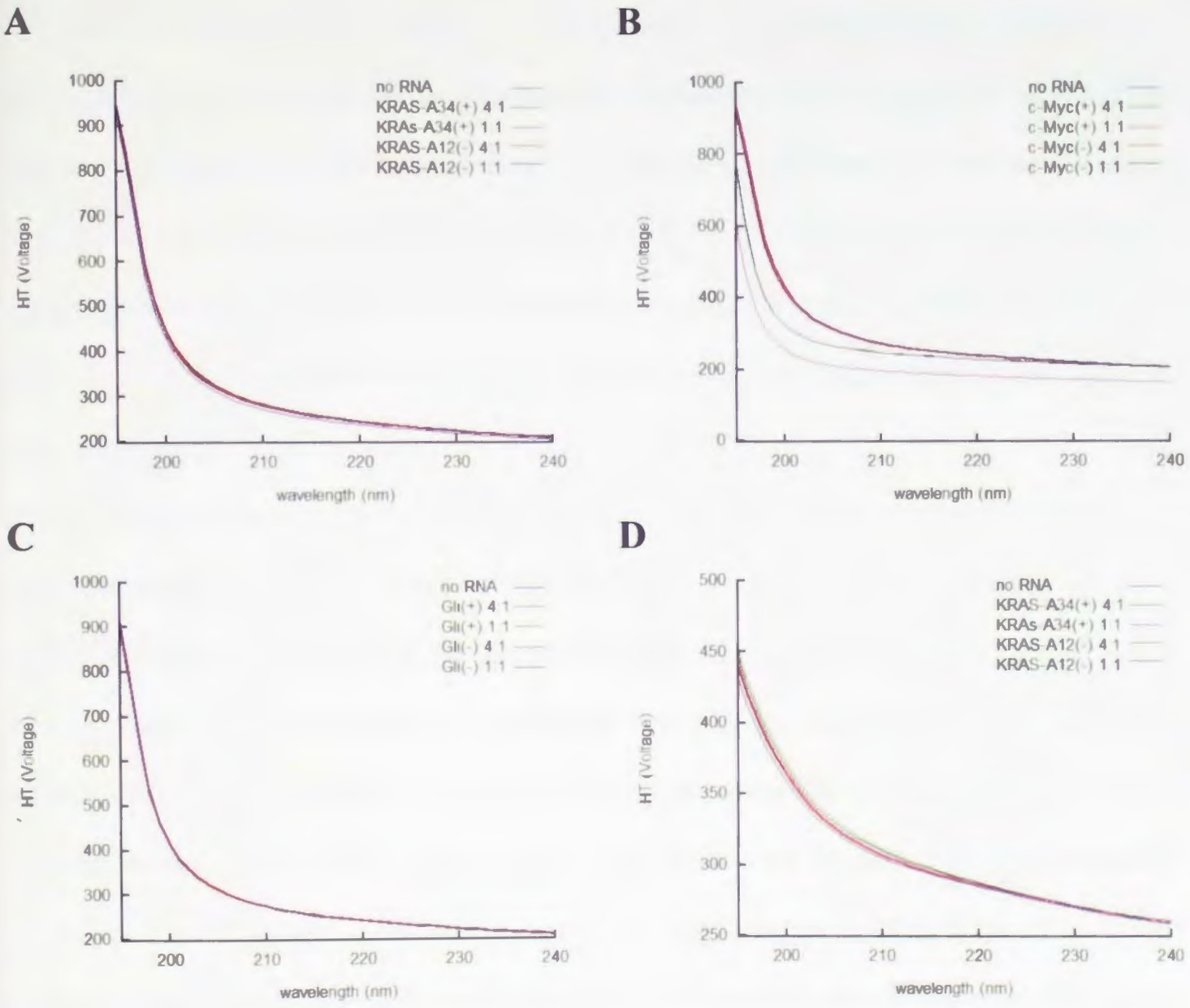


Figure 4.6. CD-HTV plots of CRD-BP samples titrated with RNA substrates. All samples are presented as a molar ratio of protein to RNA with 4:1 having lower RNA concentration than 1:1 and measurements plotted as high tensions voltage (HT-voltage). (A) KRAS A34 was determined to be a high affinity substrate while KRAS A12 binds poorly. (B) c-myc (+) and c-myc (-) are strong and weak binding substrates respectively. (C) Gli (+) and Gli (-) are strong and weak oligonucleotide competitors respectively with (D) BSA acting as a negative conformational shift control.

However, the HTV or gain would increase proportionally and appear different than that of wild type alone. As can be seen in figure 6, no such differences are evident suggesting that low complex solubility is not likely to be responsible for the CD reduction observed. However, considering CRD-BP forms large multimer complexes as visible granules in cells, it cannot be

discounted that the formation of these larger complexes could be responsible for loss of CD signal, therefore potentiating the data as a measure of granule formation and not actual structural change within the protein. A structural change in CRD-BP and its homologs however is likely, as its interaction with itself as a dimer, and with other proteins such as dynein and kinesin within micro-tubule associate granules is RNA dependent (Boylan *et al.* 2008; Havin *et al.* 1998; Ioannidis *et al.* 2005; Nielsen *et al.* 2004) and perhaps a protein-protein interaction motif is revealed upon RNA binding as suggested by Wang *et al.* (1999).

Further structural studies were similarly carried out with a KH1 to KH4 CRD-BP truncation comprising only the RNA-binding module. These experiments were done in an otherwise identical fashion except for an additional 16:1, low RNA concentration sample was included to test for a lower threshold concentration boundary for the CD effect. KRAS-A34(+) and KRAS-A12(-) were employed again as well as several alternative substrates that were not expected to bind: '18S', 'Orf1b', 'Spike' and 'tRNA' which were all used to broaden the search for potential RNA substrates that do not cause the observed changes in CD spectra, which would offer much stronger support for the notion that these conformation changes are induced by RNA binding. CD monitoring of KH1to4 secondary structure titrated with any of the RNA substrates however did not uncover any particular RNA that was ineffective at producing the previously observed reduction in CD signal (figure 4.7 and 4.8).

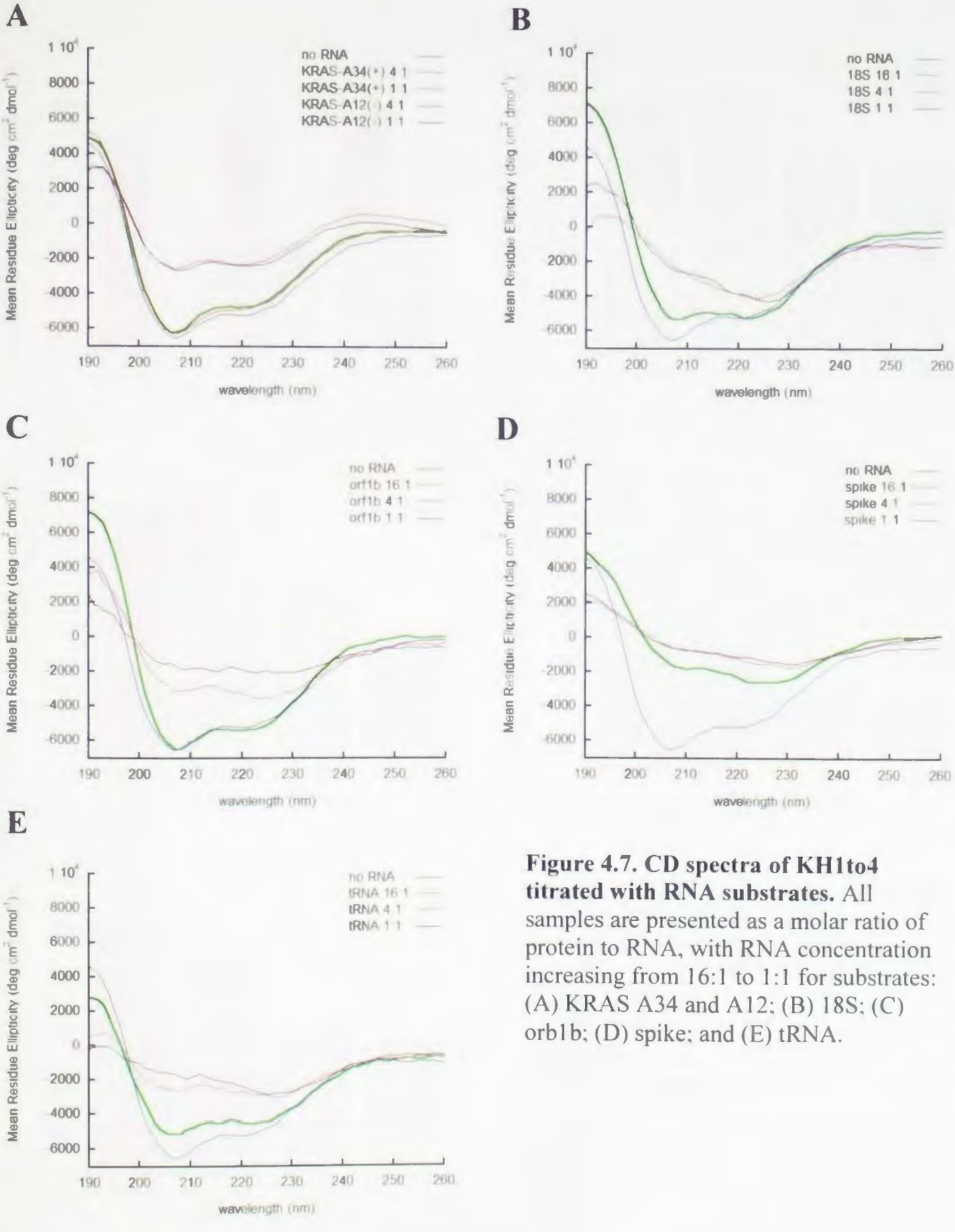


Figure 4.7. CD spectra of KH1to4 titrated with RNA substrates. All samples are presented as a molar ratio of protein to RNA, with RNA concentration increasing from 16:1 to 1:1 for substrates: (A) KRAS A34 and A12; (B) 18S; (C) orb1b; (D) spike; and (E) tRNA.

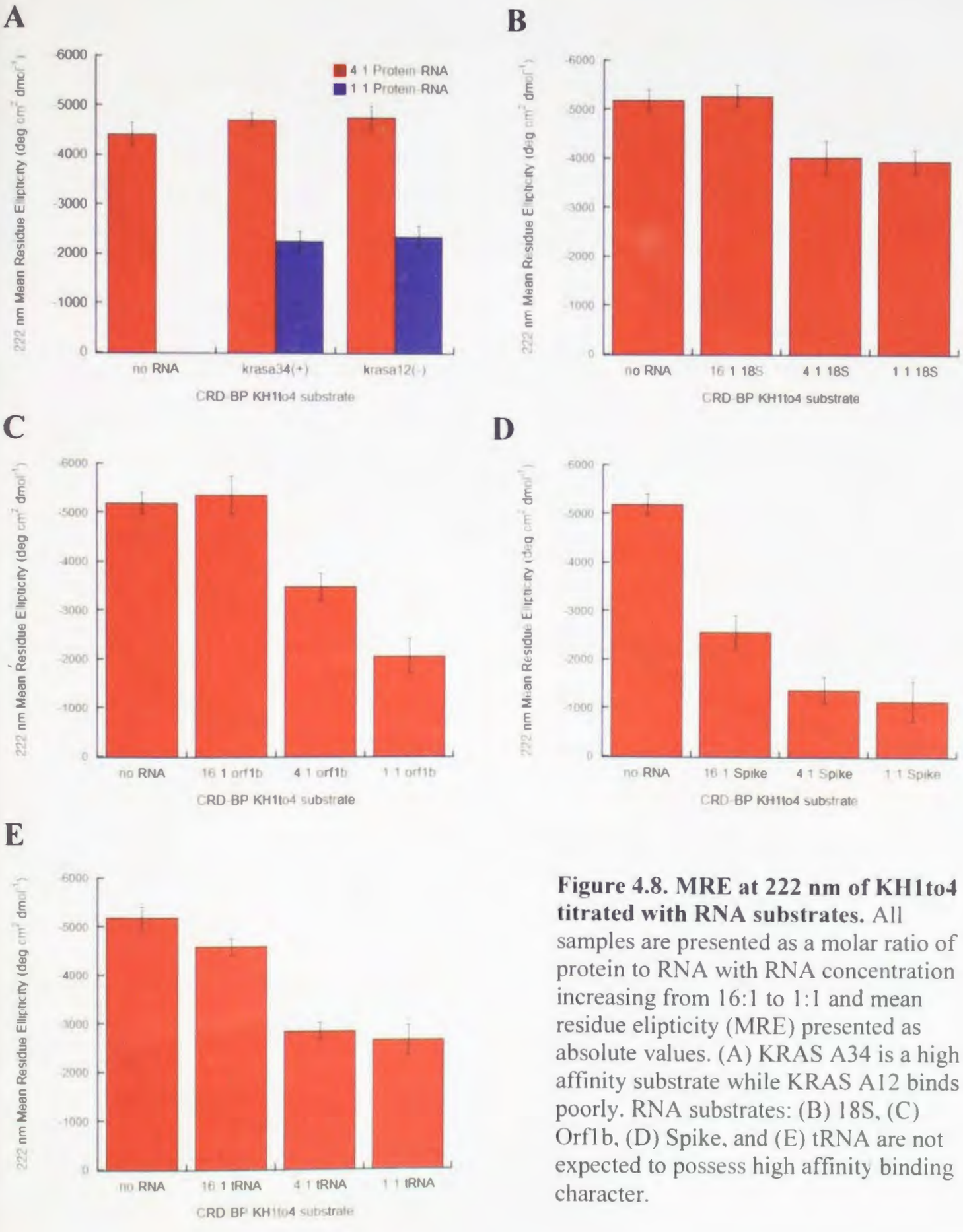


Figure 4.8. MRE at 222 nm of KH1to4 titrated with RNA substrates. All samples are presented as a molar ratio of protein to RNA with RNA concentration increasing from 16:1 to 1:1 and mean residue elipticity (MRE) presented as absolute values. (A) KRAS A34 is a high affinity substrate while KRAS A12 binds poorly. RNA substrates: (B) 18S, (C) Orf1b, (D) Spike, and (E) tRNA are not expected to possess high affinity binding character.

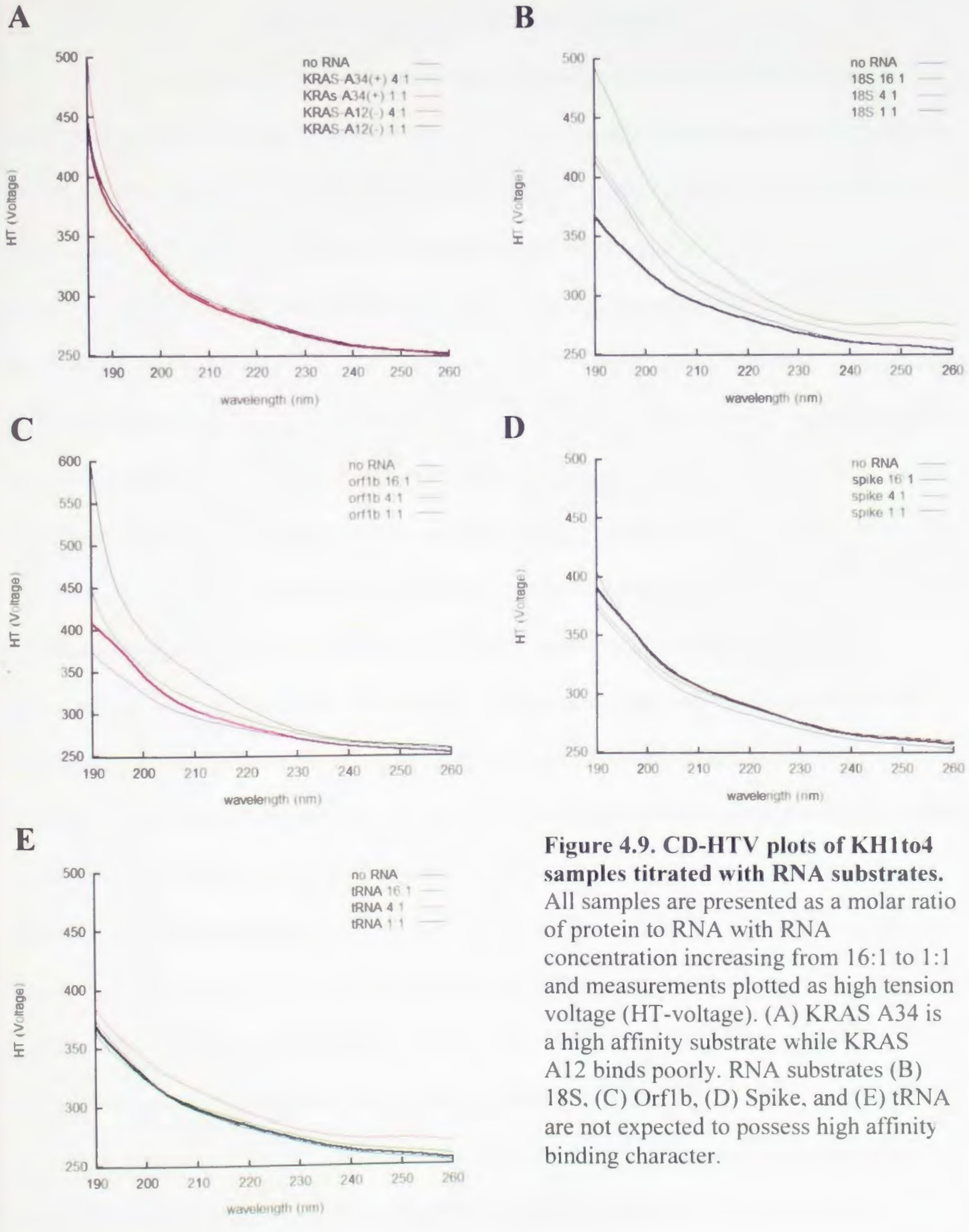


Figure 4.9. CD-HTV plots of KH1to4 samples titrated with RNA substrates. All samples are presented as a molar ratio of protein to RNA with RNA concentration increasing from 16:1 to 1:1 and measurements plotted as high tension voltage (HT-voltage). (A) KRAS A34 is a high affinity substrate while KRAS A12 binds poorly. RNA substrates (B) 18S, (C) Orf1b, (D) Spike, and (E) tRNA are not expected to possess high affinity binding character.

4.2.3 Thermal stability of CRD-BP-RNA RBP complexes

To further analyze the potential for circular dichroism to be used diagnostically in characterizing CRD-BP binding with various RNA targets, melting temperature of CRD-BP, as defined as 50% unfolded protein, was determined in the presence of various RNA targets. In theory, high affinity RNA targets of CRD-BP may have the potential to increase the overall stability of the complex by reducing free-energy as a result of spontaneous binding and increase the melting temperature of CRD-BP. CD spectroscopy was again utilized to monitor the protein secondary structure at 222 nm in the presence of different RNA species, and the temperature was increased sufficiently to melt the protein. KRAS-A12 and KRAS-A34 were analyzed in complex with KH1to4 truncated CRD-BP, as well as negative control substrates 18S and tRNA RNA molecules. As with the previous CD monitoring experiments, multiple RNA concentrations were tested including 1:1, 4:1, 16:1, and 64:1 molar ratios. Importantly, protein concentration was increased to 0.4 mg/ml (~ 10 μM) for scanning. This decision was made due to the observed dramatic reduction in CD signal at 222 nm (figure 4.7) with increasing RNA concentration. Because the baseline CD signal prior to melting would be quite low and prone to noise as a result of RNA presence already, the overall quantity of protein-RNA complex was increased sufficiently to achieve approximately 30 mdeg of raw CD at 222 nm. Results can be seen in figure 10, as determined melting temperatures (Tm) which were computed by Jasco Spectral Manager Thermal-denaturation software. Only the 64:1 and 16:1 samples (figure 4.10 A and B, respectively) were scanned due to visible precipitation of the CRD-BP-RNA complexes in the 4:1 and 1:1 sample sets. This precipitation was not present in the CD spectral analysis samples and is attributable mainly to the increased concentration of both protein and RNA used and the apparent reduced solubility of the complexes compared to protein alone.

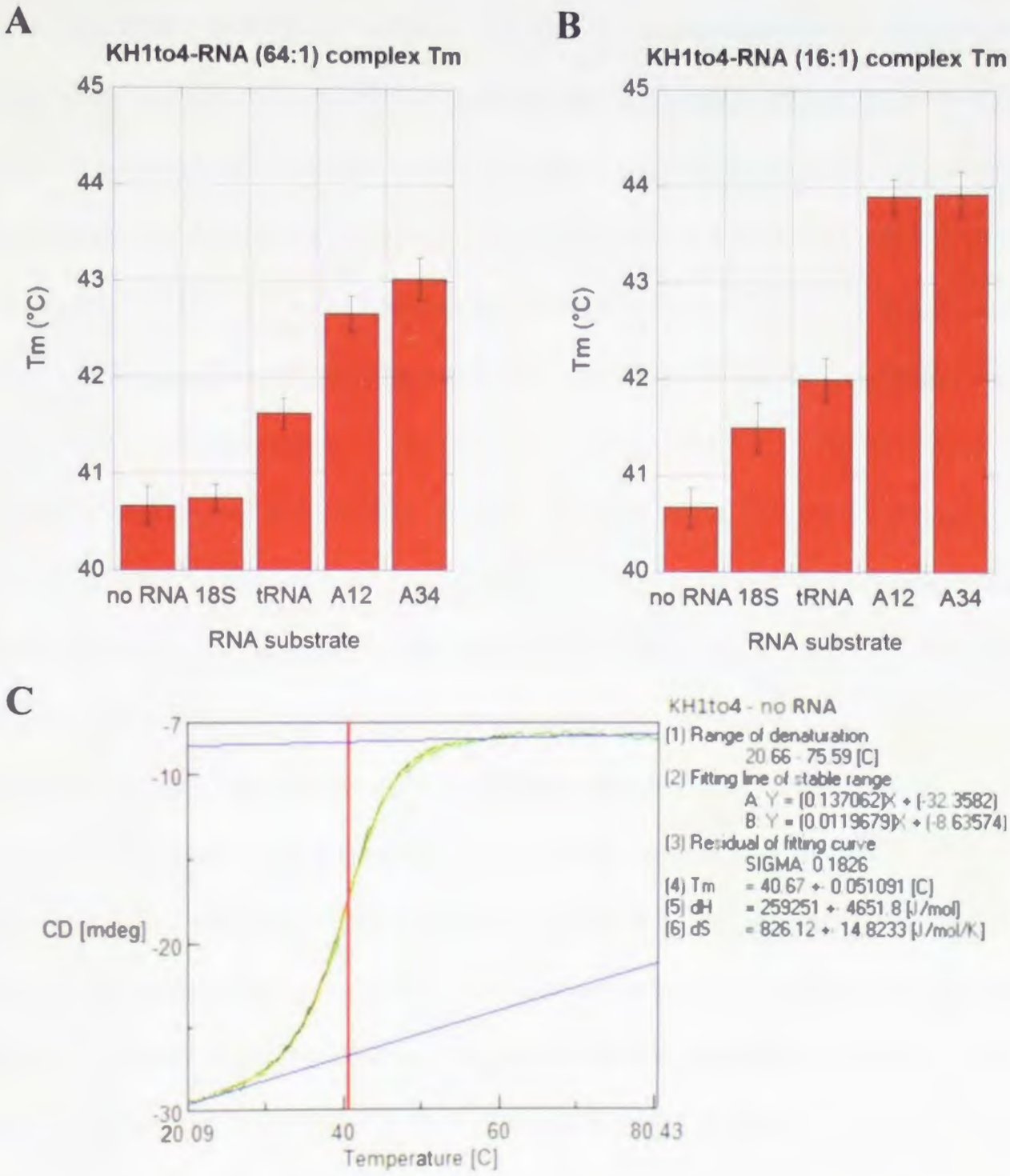


Figure 4.10. TMs of KH1to4 CRD-BP protein in complex with RNA substrates. Plotted values were determined from 222 nm CD sample traces over temperature-ramping from 20 to 80 ^{0}C and Tm calculated based on loss of half CD signal. (A) RNA substrates present in a 64:1 protein:RNA molar ratio and (B) 16:1 in otherwise identical conditions. (C) Example Tm calculation from thermal melting curve of CRD-BP KH1to4 protein sample with associated computed values.

Overall, there was very little difference in melting temperature between the protein alone and protein-RNA samples (less than 4 °C). As expected, the greatest difference between protein alone and samples with RNA was produced at the higher RNA concentration (16:1 protein to RNA). However, the differences in melting temperature between KRAS-A12 and KRAS-A34 were most pronounced at the lower RNA concentration (64:1 protein to RNA). While some selectivity can be seen in both concentration ranges, with both 18S and tRNA addition producing negligible differences in measured Tm, both KRAS-A12 and KRAS-A34 samples generally appeared to be within error of each other with possibly a slight bias towards the stronger binding substrate KRAS-A34 in the lower RNA concentration samples. To explain the more prevalent differences in KRAS-A12 and KRAS-A34 at lower concentrations compared to higher, it could be that under the CD experimental conditions, there is a degree of non-specific binding, although it cannot be completely non-specific as 18S and tRNA complexes have significantly lower Tm. In the context of using melting temperature to assess CRD-BP complex stability and RNA substrate binding, the reduced solubility of the complexes at the high concentrations required for Tm determination render the method limited. Improvements may follow by improving complex solubility by experimenting with alternative buffer contents, or perhaps engineering a more soluble CRD-BP variant by genetically fusing a tag such as Fh8 or GST.

4.2.4 CRD-BP KH1to4 protein crystal

Crystallization of CRD-BP was first attempted using our lab's existing wild-type, full-length CRD-BP plasmid construct and standard protein production protocol, whereby a denaturing protein purification scheme is employed and refolded by removing urea with a 3-stage buffer exchange process. However, refolding in the dialysis units appeared to have an effective concentration limit of 0.7 mg/ml with additional protein precipitating out of solution.

Combined with increased difficulty concentrating (clogged filter-concentrators) when this limit was approached, this standard preparation method was determined to be unsuitable for crystal protein preparation. Additionally for good measure, a 15 ml denatured sample at 0.7 mg/ml was sent to Dr. Strynadka's laboratory at the University of British Columbia X-ray crystallography facility where attempts to refold the protein via dialysis also failed for similar reasons.

I decided to construct an alternative construct whereby the RRM domains were removed (figure 4.11C), as there are large flexible regions connecting the RRM domains to the KH1to4 RNA-binding module that I reasoned may affect the ability to refold by our standard method. A PCR amplicon covering the KH1 to KH4 domain region was generated for use in sub-cloning and proved to be the expected size (1,122 nt) and of high purity as determined by agarose gel analysis (figure 4.11A). Both the DNA insert and pET41c(+) plasmid vector were successfully digested with NdeI and HindIII producing sticky ends and ligated with DH5α E.coli transformants able to grow on LB-agar kanamycin plates. Presence of the KH1to4 insert was confirmed using a digest analysis again on mini-prepped plasmid samples from successful transformants (figure 4.11B).

Expression in BL21 E.coli cells successfully produced protein with similar purity and expression levels as with the full-length construct. Direct size comparison of the full-length CRD-BP (67 kDa) and KH1to4 truncation confirms the absence of RRM1 and RRM2 domains, appearing at the expected 43 kDa (figure 4.12A). However, dialysis-based refolding using our 3-stage buffer exchange method again proved limited to approximately 0.7 mg/ml, implying that the limiting aggregation factor is likely present within the KH1 to KH4 region.

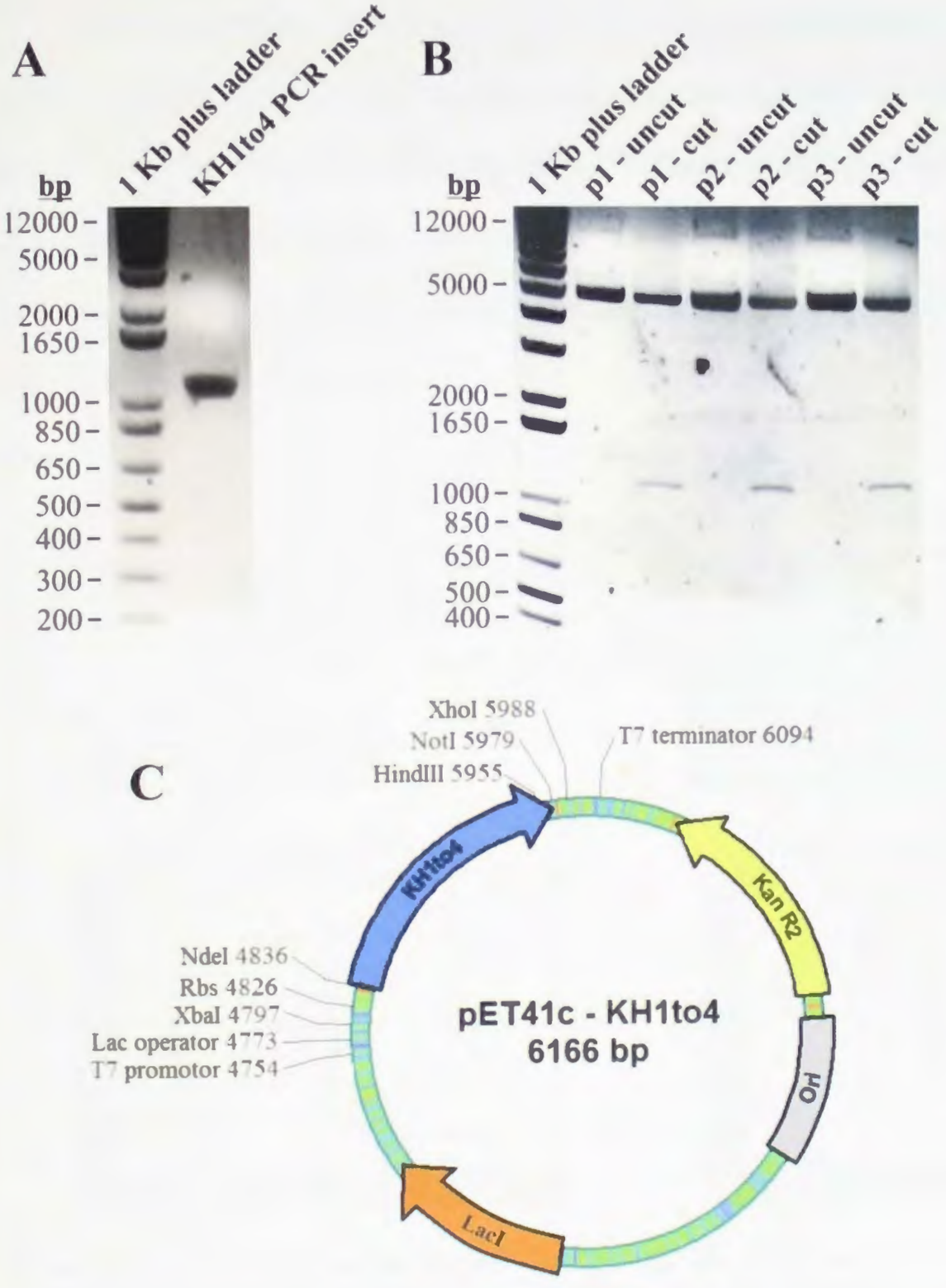


Figure 4.11. Plasmid pET41c(+)-CRD-BP KH1to4. (A) DNA agarose gel analysis of PCR product spanning domains KH1 to KH4 of CRD-BP gene. (B) Endonuclease digest analysis of pET41c(+)-CRD-BP KH1to4 using NdeI and HindIII resolved on 1.5% agarose gel. Uncut and cut vector bands can be seen at ~5 kb and insert DNA bands appears at expected 1122 bp (C) Plasmid map of pET41c(+)-CRD-BP KH1to4.

Native protein purification was subsequently pursued as described in 4.1.8. Fractions purified by this method contained significantly less protein, with most of the protein in the insoluble fraction found in the pellet (figure 4.12B). Similar restrictive results for native purification of CRD-BP in recombinant systems have been demonstrated previously with Prokipcak *et al.* (1993) and Du *et al.* (2008) that specifically indicate that most of the protein ends up in the insoluble fraction.

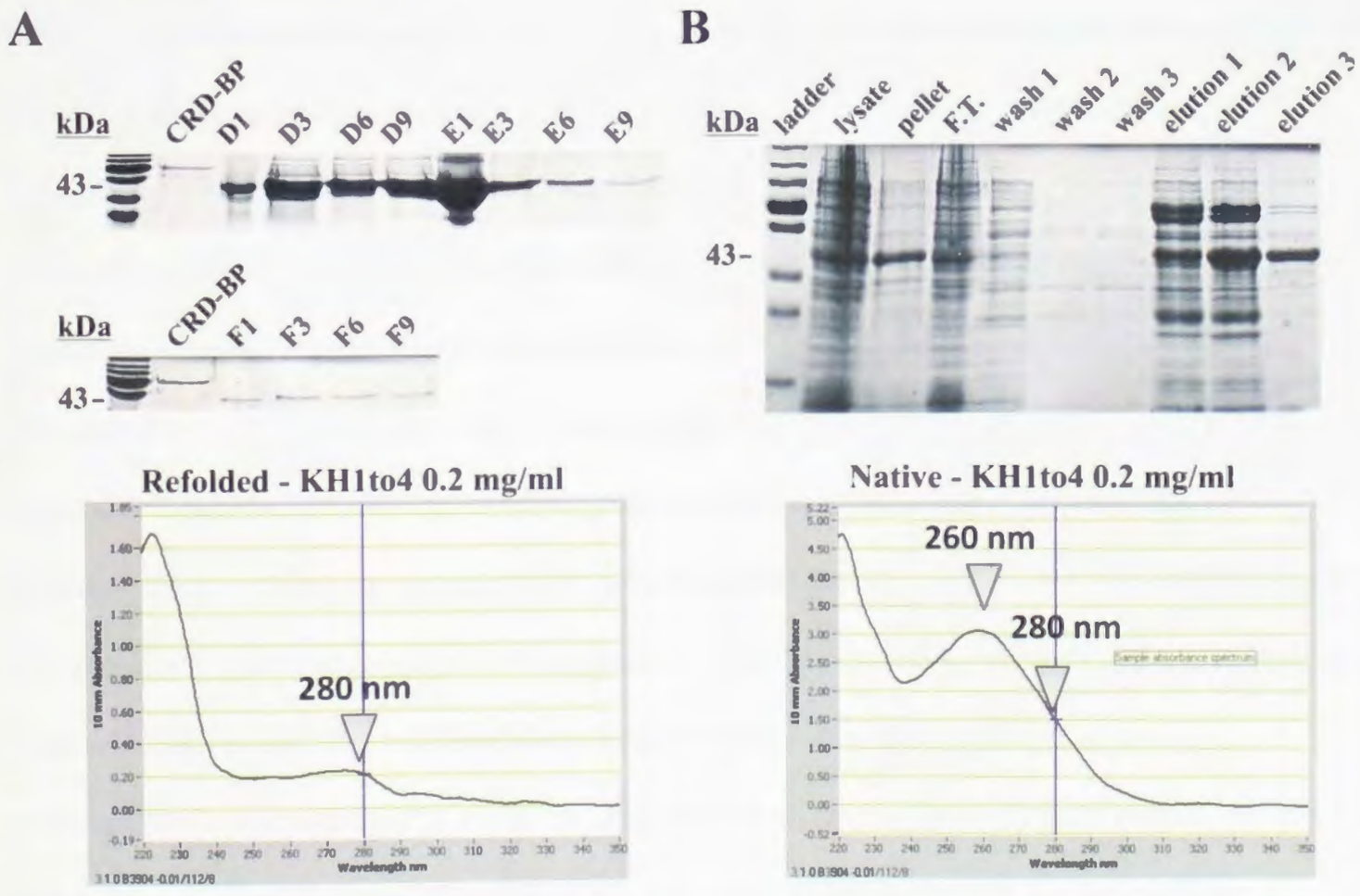


Figure 4.12. Comparison of native and denaturing KH1to4 protein purifications. (A) Denaturing protein purification – refolding facilitated by gradual removal of urea via dialysis in 3-stage buffer exchange scheme. Below: UV-spectrum of protein sample showing dominant 280 nm peak; 260/280: 0.61 (B) Native protein purification with imidazole elution. Below: UV-spectrum of purified protein showing dominant 260 nm peak; 260/280: 2.08.

While the reduced yield was not entirely limiting, endogenous bacteria RNA were found to co-purify with the protein in large amounts as determined by UV-spectroscopy (figure 4.12B)

which would likely interfere with crystallization and reduce solubility. Approaches including non-specific RNases such as benzonase treatment to rid the protein of RNA were considered, however combined with the lower yield of this method, experiments continued to find another approach with superior results. A third approach called on-column refolding was tried. Here the denaturing purification steps ensure solubilisation of the large otherwise insoluble fraction of the protein. His-tag binding to the Ni-NTA column occurs in 8 M urea, which is then slowly reduced to 0 M urea over a smooth gradient prior to elution using concentrated imidazole as described in section 4.1.9. Results from this method were excellent. Off the column, protein concentration approached 2 mg/ml with no signs of precipitation. As well, optimization of the process eventually yielded >95% purity samples with more stringent wash-steps, and further concentration up to 10 mg/ml proved simple with no clogged filters (figure 4.13C). To ensure this alternative method was providing active protein still capable of binding RNA, fluorescence polarization (FP) technique was employed. This technique exploits the time-scale whereby a fluorophore re-emits plane-polarised light, and can detect the increasingly random distribution of un-bound, free tumbling labelled-RNA compared to protein-bound RNA measured as anisotropy. A fluorescein-labelled KRAS RNA probe based on previous EMSA work (2.3.2) was synthesized by IDT and used to assess its interaction with the CRD-BP KH1to4 protein sample. It is evident according to the FP analysis in figure 13A that unlike the BSA negative binding control, the degree of anisotropy with full-length CRD-BP and the KH1to4 construct increases considerably with concentrations similar to those seen in EMSA experiments. Indeed upon close inspection, it can be seen that the KH1to4 variant appears to bind the RNA substrate with higher affinity than the natural form. Calculation of the dissociation constant from an RNA bound fraction plot (figure 4.13B) reveals a 3-fold increase in substrate affinity for the KH1to4 variant

(311.8 ± 67.1 nM compared to 93.6 ± 34.3 nM), possibly implicating the RRM domains as a restricting element in binding.

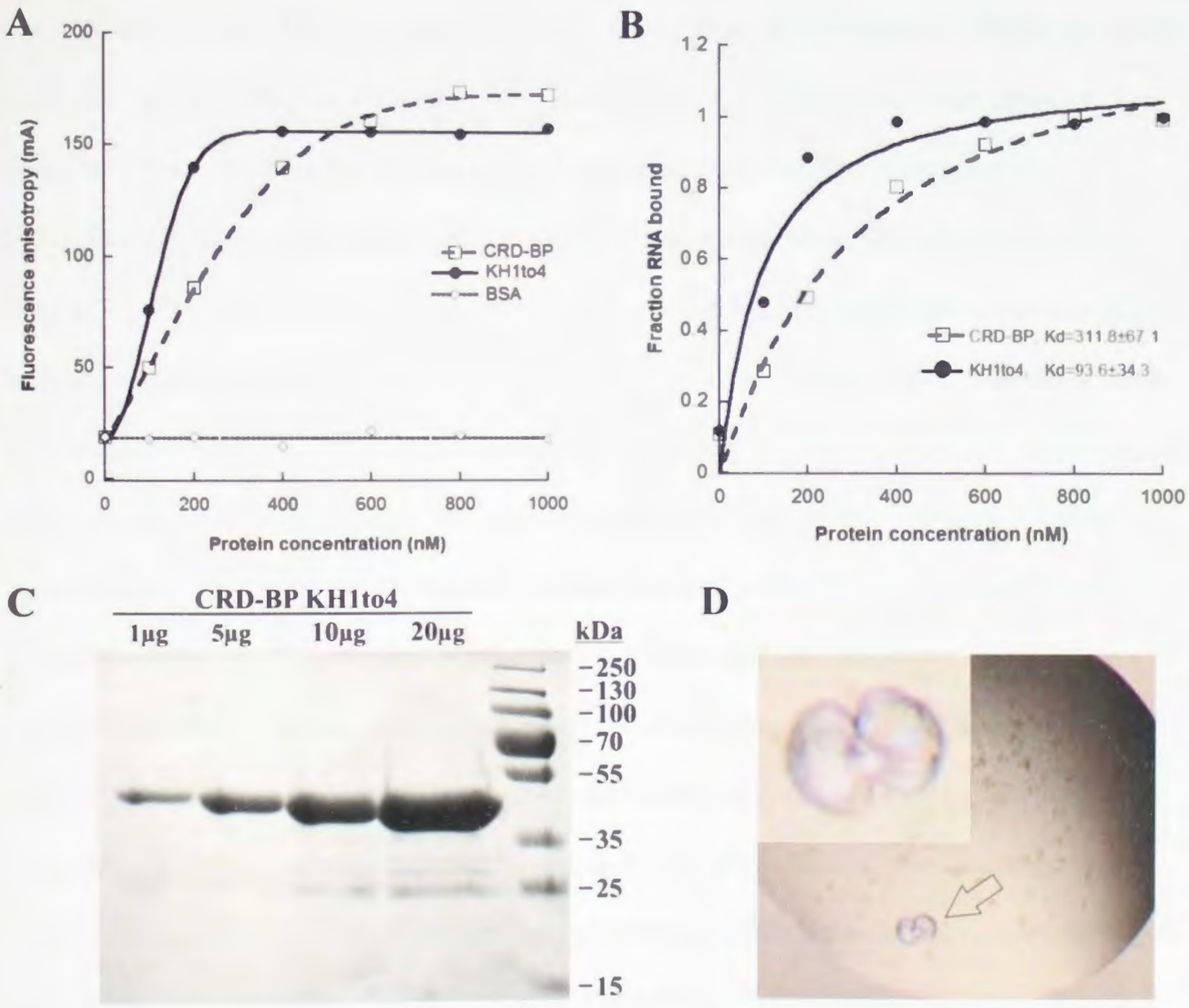
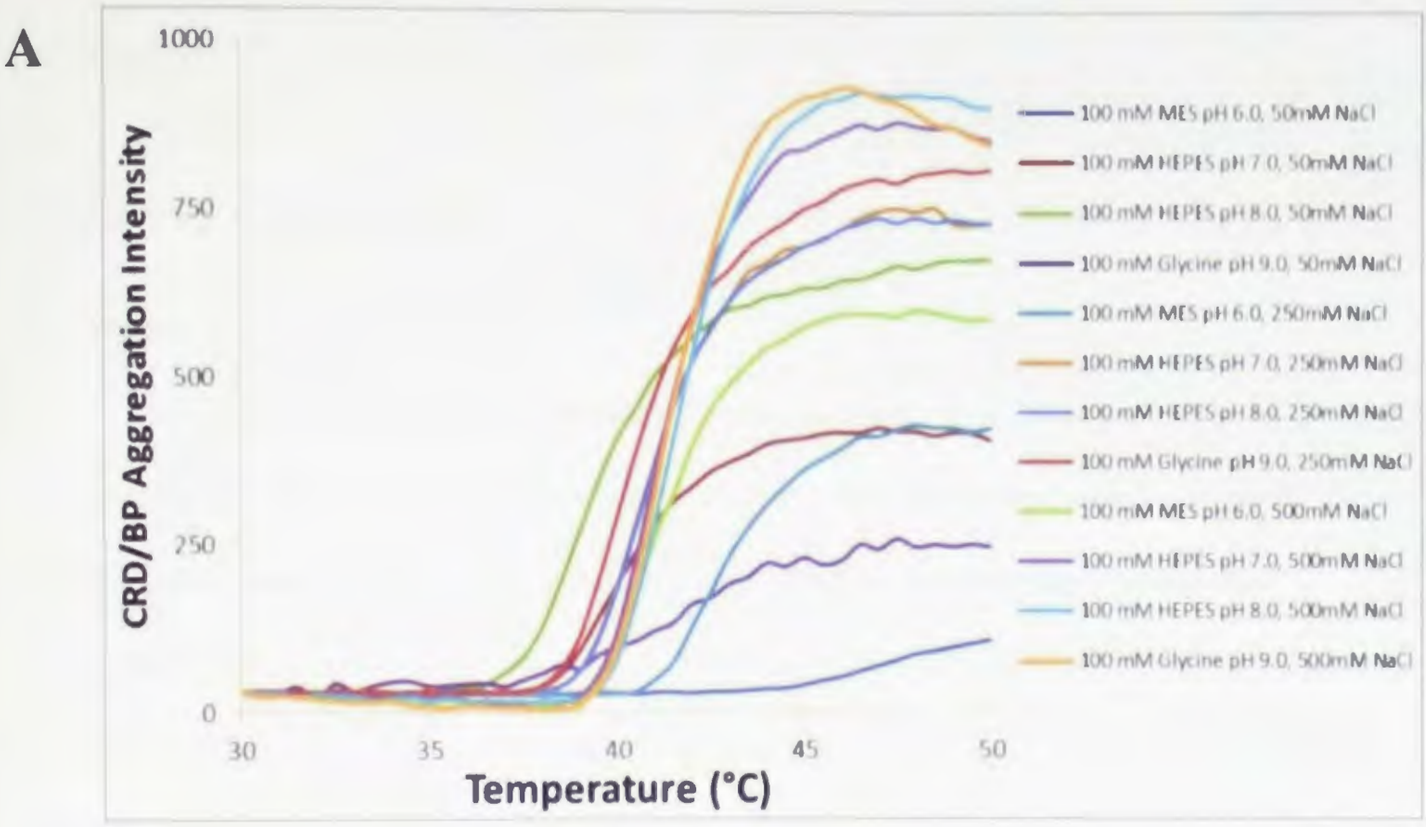


Figure 4.13. CRD-BP KH1to4 construct binding activity and protein crystal. All figures pertain to the CRD-BP construct containing only the KH1 to KH4 domain binding module, excluding the N-terminal RRM1 and RRM2 domains. (A) Fluorescence polarization/anisotropy RNA-binding activity assay. CRD-BP KH1to4 confirmed to bind fluorescein-labelled KRAS RNA probe using full-length CRD-BP as a positive control and BSA as a negative binding control. (B) KH1to4 and full-length CRD-BP Fluorescence anisotropy data using KRAS RNA probe converted to fraction RNA bound plot for dissociation constant determination. (C) SDS-PAGE analysis of column-refolded CRD-BP KH1to4 protein. (D) Positive hit protein crystal of KH1to4 in 0.4 µl mother liquor.

Perhaps a function of the RRM domains is to confer enhanced selectivity for a given RNA substrate, acting as negative regulators of binding in some circumstances. It has been demonstrated that the RRM domains do not contribute to a lower Kd, however their conservation across the various orthologs implies a useful function, one of which could be the negative regulation postulated here. Numerous KH1to4 protein preparations were delivered to Dr. Strynadka's X-ray crystallography laboratory for crystallization trials. High concentration samples of 10-11 mg/ml protein were shipped from our lab for direct entry into crystallization, as well as a 1 mg/ml preparation where a subsequent thrombin digest allowed for removal of the 8 X his-tag prior to concentration and crystallization. Two trial concentrations were setup (4 mg/ml and 8 mg/ml) for crystal screening with three separate crystal screening kits: Qiagen Classic Suite, Qiagen PACT suite, and Molecular Dimensions JCSG-Plus™ MD1-37 suite. No protein crystals were produced from samples in our original buffer (20 mM NaH_2PO_4, 200 mM NaCl, pH 7.4), however a single crystal was achieved that was dialyzed into 0.2M lithium sulphate, 0.1M Tris-Cl pH 8.5, 40% v/v PEG400 buffer and confirmed to fluoresce under UV light, indicating that it was comprised of protein (figure 4.13D). Unfortunately the sphereolyte type crystal bore unsuitable characteristics for further study, including a twinned structure and thus no diffraction pattern was obtained despite the lengthy attempt. Further attempts to reproduce the crystal did not prove successful, however data obtained from subsequent pre-crystallization optimization tests did reveal some useful information pertinent to future attempts at crystallizing this protein (figure 4.14). Buffer factors including salt, pH, and the buffer chemical itself were varied in a thermal aggregation experiment, whereby stabelizing elements were systematically tested for ability to prevent non-specific protein interaction. From these experiments conducted at the Strynadka laboratory, we discovered that KH1to4 CRD-BP is stabilized significantly by

lower pH; pH 6 conferred a profound stabilizing effect. High salt up to 500 mM NaCl was also determined to also increase the temperature of aggregation, albeit to a lesser extent than the pH effect, and the effect of the salt was also dependent on the buffer system used. The specific buffer system choice had a strong impact on protein stability as well, with MES providing the greatest degree of stabilization according to the temperature of aggregation experiments. Future endeavors into crystallizing CRD-BP will certainly benefit from this knowledge, as it is likely an ideal storage and transport buffer, as well as being optimal for crystallization trials.



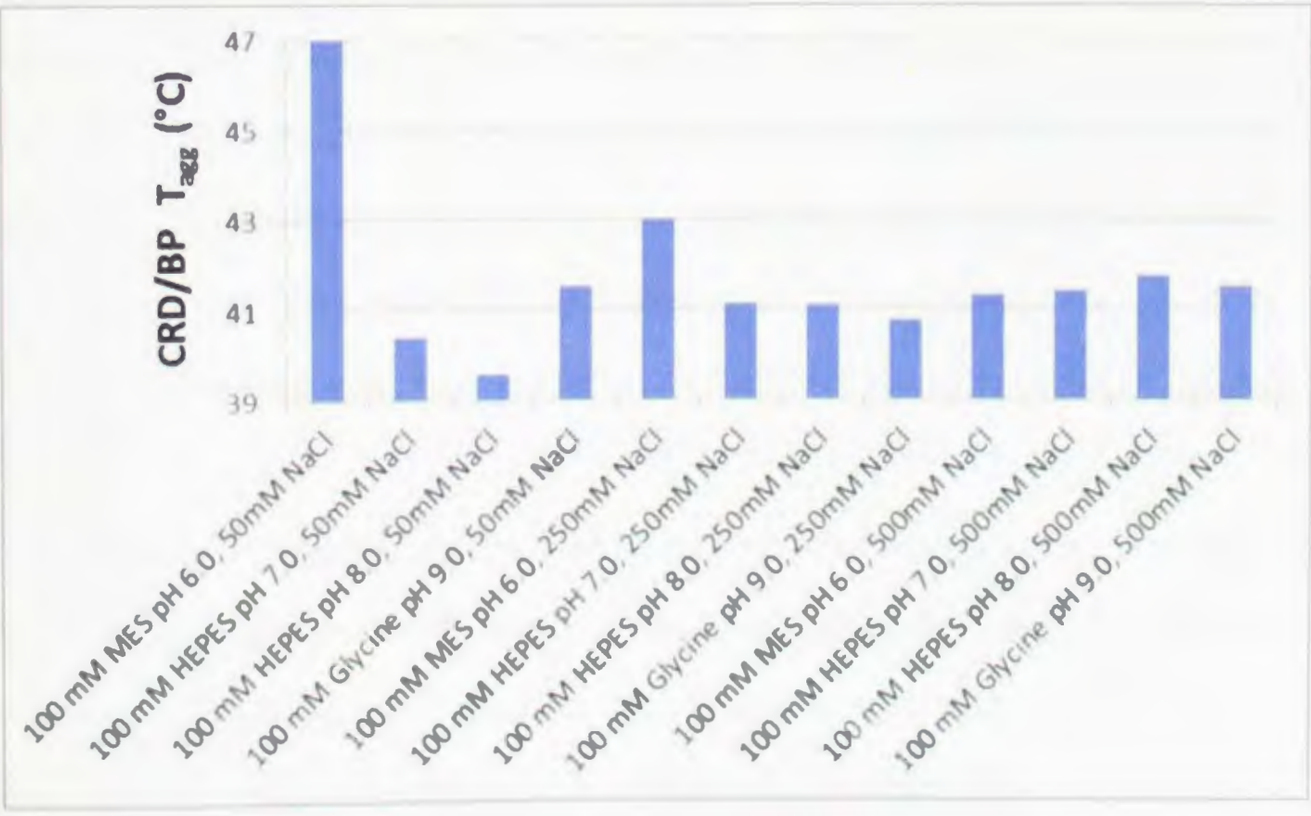


Figure 4.14. Thermal aggregation buffer optimization for CRD-BP KH1to4. (A) Protein aggregation as a function of temperature in various buffer solutions. Data was produced from spectroscopy experiments conducted at the Strynadka X-ray crystallography laboratory. (B) T_{agg} (temperature of aggregation, defined as temperature at which 50% of protein is aggregated) calculated from aggregation intensity trace for CRD-BP in different buffer systems.

Chapter 5

General Discussion

5.1 Project general overview

The VICKZ family of RNA-binding proteins includes CRD-BP as well as its orthologous proteins (IMPs, KOC, ZBP1, and Vg1 RBP/Vera) have multifaceted roles including intracellular mRNA localization, nuclear export of mRNA, transcript stabilization and translational regulation. Their function is achieved through the interaction of 2 N-terminal RRM domains and 4 C-terminal hnRNP (heterogenous nuclear ribonucleoprotein)-K-homology domains (KH domain for short) which function in combination to bind a variety of mRNA targets including c-myc, GLI, MDR1, IGF-II amongst others (Prokipcak et al. 1994; Noubissi et al. 2009; Sparanese and Lee 2007; Nielsen et al. 1999). CRD-BP is spatiotemporally regulated, normally expressed during the first 12 weeks of embryogenesis and disappearing into complete absence or scarcity in adulthood (Leeds et al. 1997). Human CRD-BP (IMP-1) is often re-expressed in many cancers and has been demonstrated to be important for tumour cell survival as summarized by Bell et al. (2013).

KRAS, a well established genetic driver of many cancers including colon, breast and lung tumours has recently been suggested to be an important oncogenic mRNA target of human CRD-BP, increasing expression of KRAS protein in colon cancer cell lines (Mongroo et al. 2011). Furthermore, while many studies focus on mutated gain-of-function KRAS variants such as $KRAS^{G12V}$ which constitutively activates downstream targets such as the PI3K, EGFR, and MAPK pathways, gene amplification and overexpression of wild-type KRAS also associate with aggressiveness and poor response to anti-EGFR therapy in colorectral, endometrial and non-

small cell lung cancers (Valtorta et al. 2013; Birkeland et al. 2012; Sasaki et al. 2011). Despite numerous and diverse attempts to target the KRAS protein using small molecules, antioncogene ribozymes, and post-translational modification inhibitors, there remains no effective means of treatment for KRAS driven cancers (reviewed in Jancik et al. 2011; Collins et al. 2014). Hence, there is need for novel approaches to anti-KRAS therapies. The goal of this investigation was to confirm and further characterize the molecular interaction between CRD-BP and KRAS in vitro as in cells, as well as spearhead an attempt to interrupt the CRD-BP-KRAS mRNA interaction specific antisense oligonucleotide molecules.

5.2 Mapping of the CRD-BP binding sites on the coding region and 3' UTR of KRAS mRNA using the electrophoretic mobility shift assay

Determining the binding potential of KRAS mRNA and determining a smaller segment of the mRNA molecule that retained affinity for CRD-BP was an important endeavor. It offered the opportunity to confirm previously reported results regarding the interaction between KRAS mRNA and CRD-BP protein using an alternative experimental method, and also to narrow down a more precise sub-region containing the necessary sequence information required for binding. Knowledge of a minimum binding sequence not only contributes to the goal of determining an mRNA target consensus sequence, but opens up options for other methodologies to study CRD-BP-KRAS mRNA interaction. For instance, to use the fluorescence polarization spectroscopy to study nucleic acid-protein interaction, nucleic acids must be larger than 90-mer (Hafner *et al.* 2008).

Through the use of PCR and in vitro transcription, [^{32}P]-labelled fragments of KRAS mRNA were synthesized and tested for CRD-BP affinity using EMSA. Of the six course

fragments, each approximately 200 nucleotides in length, four were discovered to bind with moderate affinity in the nanomolar range (section 2.3.2). This finding confirmed previously published results that indicate IMP-1 (99% identity to CRD-BP) binds KRAS mRNA in the coding region and also the 3' UTR (Mongroo *et al.* 2011), but also elucidated multiple attachment sites within both of these mRNA regions. Numerous CRD-BP-KRAS mRNA attachment sites is not surprising and this notion is supported by Nielsen et al. (1999) and Vikessa *et al.* (2006) where several IMP-1 binding sites were found on the 5' UTR of IGF-II mRNA and 3' UTR of CD44 mRNA respectively. Similar conclusions were drawn regarding FMR1 mRNA, in which multiple IMP1 proteins were suggested to bind in the coding region and 3' UTR (Rackham and Brown 2004).

KRAS mRNA sub-region 'A' possessed the highest affinity with a measured kD of 177.9±18.4 nM, with the other binding fragments: C, D and F having at least 2-fold lower affinity (figure 2.6). This fragment size compares well to literature defined targets (table 5.1). Further truncation analysis of the KRAS-A sub-region identified 'KRAS-A34', a 93 nucleotide fragment which binds CRD-BP with similar affinity. KRAS-A34(b), a smaller 57 nucleotide region located near the 3' end of the parent KRAS-A fragment, also retained ability to form a larger molecular weight complex with CRD-BP, albeit with reduced affinity as determined by EMSA (figure 2.18). It also became apparent that with the progressive whittling down of KRAS-A to KRAS-A34(b) sequence length has a general impact on the overall affinity of this RNA target and specificity only becomes apparent when comparing RNA fragments of similar size as in KRAS-A12/KRAS-A34 (92 and 93 nt, respectively – figure 2.13) and KRAS-A34(e)/KRAS-A34(b) (57 and 57 nt – figure 2.17 and 2.18). Indications that other targets may follow this pattern stem from a similar truncation study using PSI, a KH domain containing protein in

Drosophila related to CRD-BP, where additional sequence flanking a core element contributed to a lower Kd (Chmiel et al. 2006).

Table 5.1. VICKZ mRNA target sizes

RNA-binding protein	mRNA target	Target size (nt)	Reference
IMP1	IGF-II	162	Nielsen (1999)
CRD-BP	c-myc	182	Bernstein (1992)
IMP1	Tau	600	Atlas (2004)
ZBP1	β-actin	54	Farina (2003)
IMP1	FMR1	37	Rackham (2004)
IMP-1	PABP	22	Patel (2005)
Vg1RBP	Vg1 (VLE)	340	Mowry (1992)
IMP1	H19	173	Runge (2000)
PSI	Δ11	69	Siebel (1990)

KRAS-A34(b) RNA sequence was analyzed using *mfold* RNA-folding software to assess possible structure recognition in the form of stem loop motifs. A structured core of the RNA molecule combined with inclusion of putative binding elements identified from literature: 5'-ACAC -3', 5'- UUCUCGACAC -3' and 5'-ACACCC-3' (2.3.4) subsequently allowed the identification of a 44-mer fragment using fluorescence polarization spectroscopy (Patel *et al.* 2012, Git and Standart 2002, Ross *et al.* 1997). Development of the KRAS-A34(b)-based fluorescence polarization probe was instrumental in the binding capacity assessment for the truncated CRD-BP KH1to4 protein, as well as other molecular endeavors in our laboratory including small molecule inhibitor screening. Notably none of the mentioned sequence elements appear to be able to bind the protein by itself, however it has been proposed that they do contribute in tandem with other elements, to a high affinity interaction likely each mediated by a separate KH domain opening the possibility of many combination binding sequences (Chao *et al.* 2010).

While a small subset of sequence elements proposed do help explain KRAS binding, many recognition elements proclaimed for the VICKZ protein family do not explain KRAS-A34(b) binding with CRD-BP. Notably, a fairly consistent notion is that RNA targets contain high C/U nucleoside content and Git and Standart (2002) showed that indeed poly-C/poly-U homopolymers are more capable of competing away Vg1RBP-VLE mRNA interaction in vitro. As KRAS-A34(b) attributes only 44% of its composition to uridine or cytosine, an alternative mechanism than simply high C/U content must be responsible for the binding of this region. It is likely that multiple mechanisms exist for both binding and granule formation in vivo. KRAS as well as c-myc both appear to form a homodimer with an intermediate band between free RNA and the larger complexes (figure 3.2) while ZBP1 shows no such intermediate band and corresponding size exclusion chromatography data also suggest a monomeric species (Chao *et al.* 2010). The same group also found no cooperativity with a Hill coefficient of approximately 1, in contrast to KRAS RNA binding (Hill coefficient approximately 1.9 – see figure 2.8) as well binding to c-myc and CD44 RNAs (Barnes *et al.* 2014). Krissinel and Henrick (2007) computed a free-energy of dissociation calculation and also suggest that homodimerization of VICKZ proteins is unfavourable and thus the phenomenon may be an artifact of the experimental conditions used in EMSA as suggested previously by Nielsen *et al.* (2004).

5.3 Testing the specific anti-sense oligonucleotide inhibitors against the CRD-BP-KRAS mRNA interaction *in vitro* using EMSA

The benefit of eliminating CRD-BP interactions with its various target mRNA are largely apparent (Iannidis *et al.* 2005; Elcheva *et al.* 2008; Mongroo *et al.* 2011), however IMP1 knockdown in T47D and MDA231 human breast cancer cells caused increased invasiveness due

to reduced localization of E-cadherin mRNA (Gu *et al.* 2011). E-cadherin, along with another CRD-BP target β-actin, has a cell-cell adhesion role, forming the core structure of focal adhesions that reduce migration and mitigate invasiveness (Berx *et al.* 1995; Birchmeier and Behrens 1994) and because IMP1 is able to bind and translationally repress E-cadherin mRNA, VICKZ proteins may have anti-proliferative function as well. Furthermore, direct knockdown of CRD-BP in human K562 leukemia cells resulted in increased cell proliferation (Liao *et al.* 2004). While complicating a therapeutic strategy considerably, the mixed nature of CRD-BP's pro-growth and anti-growth factors justifies the target-specific approach our lab pursues, as opposed to simply trying to down-regulate or otherwise inhibit CRD-BP function entirely, although many cancers would still respond positively from a general inhibition of CRD-BP due to increased chemo-sensitivity (Craig *et al.* 2012). A target specific approach would allow a personalized treatment of CRD-BP-positive cancers, preferentially driving down specific oncogene expression.

Work on identifying effective, specific molecular inhibitors of the CRD-BP KRAS mRNA interaction was an ultimate goal of this research and the attempt was spearheaded by designing multiple 23 to 24-mer DNA-based anti-sense oligonucleotide molecules that were complementary to sub-regions spanning the entire length of the KRAS-A region in a contiguous fashion (Table 2.8). EMSA assays were used to test the effectiveness of each oligonucleotide at breaking the interaction in vitro, as evidenced by the loss of super shifted [^{32}P]-labelled bound fraction RNA in complex with CRD-BP; an approach adopted from King *et al.* (2014). By titrating increasing amounts of anti-sense oligonucleotide molecules into a CRD-BP-KRAS mRNA binding reaction, from 1 to 1000 fold molar excess, two candidates were revealed as good molecular inhibitors. Labelled as SM-6 and SM-7, each proved considerably more effective

than any of the other candidate molecules, spanning KRAS nucleotide positions 116-138 and 139-161 of the coding region respectively (figure 2.19). SM-2 displayed virtually no ability to compete away the binding interaction as determined by EMSA and is considered evidence of high sequence specificity for interrupting the KRAS-CRD-BP interaction under these experimental conditions (see figures 2.21 and 2.22, lanes 13-17). Complementary annealing of the oligonucleotides to the SM-6 and SM-7 positions effectively masked the RNA nucleotides in the KRAS-A transcript and thus prevented proper CRD-BP access to the binding sequence. Alternatively, a significant RNA structure change induced as a result of the oligonucleotides binding could also explain the effect, however this model requires that CRD-BP recognize specific RNA structural elements as a mechanism of binding which is as of yet unconfirmed. Regardless of the mechanism, the re-appearance of this region as a potential CRD-BP interaction site further supports earlier KRAS mapping findings, which suggested that the KRAS-A34(b) region covering nucleotide positions 129-185 falls within the SM-6 and SM-7 coverage, and likely contains important binding information (refer to figure 2.24 for graphic representation). The SM-6 and SM-7 DNA molecules were subsequently used as positive controls for inhibitors of the KRAS-CRD-BP interaction in small molecule screens, as well as serving as potential proto-drugs in cell-based KRAS inhibition assays with SM-2 continuing to serve as a negative control (Wang *et al.* 2015).

5.4 Examining CRD-BP KH domain requirements for binding KRAS mRNA *in vitro*

VICKZ family of proteins contains 2 N-terminal RRM domains (labelled RRM 1 and 2) and 4 C-terminal KH domains (labelled KH 1 through 4). The importance of determining the sub-set of KH domains required for high affinity binding to KRAS mRNA stems from a basic

science perspective in understanding the oncogenic function of CRD-BP. Such knowledge is also important in applied research as it provides critical information for the design of specific inhibitors against the oncogenic function of CRD-BP. As different RNA transcripts appear to have different KH domain requirements for binding, molecules effective at blocking a specific KH domain may predictably only affect a subset of RNA targets. Knowledge of these unique KH domain requirements for each RNA molecule would allow classification of the various RNAs and potentiate prediction of the multiple targets a given inhibitor or combination thereof would have.

To approach the specific KH domain requirements for CRD-BP to bind KRAS mRNA, multiple plasmid constructs containing CRD-BP variants each harbouring either a single GXXG mutation in one of the four KH domains (KH1, KH2, KH3 and KH4), or combinations of 2 GXXG mutations – each simultaneously present in 2 separate KH domains (KH1-2, KH1-3, KH1-4, KH2-3, KH2-4 and KH3-4) as described in greater detail (section 3.1.1). Each GXXG motif exists as a flexible loop between the 1st and 2nd α-helix of type I KH domains and directly interacts with the RNA targets of KH domain containing proteins (Grishin *et al.* 2001). Mutation of the first glycine to an aspartate (table 3.1) has previously been demonstrated to abolish the ability to bind RNA, presumably through steric hindrance from the larger side chain, or from charge repulsion - both aspartate and the RNA backbone possess negative charge (Lewis *et al.* 2000; Lin *et al.* 1997; Paziewska *et al.* 2004; Fenn *et al.* 2007; Jones and Schedl 1995). Furthermore, limiting the mutation to 2 GXXG motifs is rationalized by the fact that no RNA target observed thus far has been demonstrated to bind with anything less than 2 KH domains. Each mutant protein was then subjected to EMSA binding assay using the KRAS-A probe and

the effect each GXXG mutation compared to wild-type then assessed. The c-myc CRD (coding region determinant – where CRD-BP binds) RNA was utilized as a positive control.

Experiments using any of the single CRD-BP KH mutants alone abrogated the ability of CRD-BP to bind the [^{32}P]-labelled KRAS RNA almost entirely (figure 3.1 and 3.2). Mutation in KH1 was repeatedly seen to impact the binding to a slightly lesser extent, but was still required for the high affinity binding displayed by the wild-type control. Hollingworth *et al.* (2012) similarly concluded that the KH1 domain played a mitigated role for KSRP RNA binding (an RNA binding protein similar to ZBP1 in the study, containing 4 type I KH domains). The KH4 mutation on the other hand appeared to have the largest effect, independently reducing the bound fraction of RNA almost entirely. Introduction of a second mutation in any combination of KH domains proved sufficient to entirely inhibit any trace of binding under conditions used in the electrophoretic assays. This result was expected given the impact of a single KH mutation on KRAS binding.

5.5 IP-qRT PCR analysis of CRD-BP KH domain requirements for KRAS mRNA binding in HeLa cells

Complementary experiments using CRD-BP-FLAG immune-precipitation and RT-qPCR also strongly agreed with the KRAS EMSA data. Mammalian expression vectors (pcDNA) harbouring each of the KH mutant variant genes were transfected into HeLa cells and subjected to anti-flag CRD-BP pull-down combined with RT-qPCR using KRAS specific primers (section 3.2.3). Results revealed a similar pattern as with EMSA experiments, with single and double KH mutants displaying relatively less association with KRAS mRNA in cells than the wild-type form ($P=0.01$). Notably while the KH1 mutant precipitated approximately 10-fold less KRAS mRNA compared to the wild-type, the other single KH mutants were nearly 1000-fold less, mirroring the

previously discussed EMSA result (figure 3.5). Interestingly the conclusion drawn from both these experiments supports a notion that all KH domains are required for RNA binding, at least to the high affinity region of KRAS mRNA examined in this study.

This conclusion is echoed with Vg1RBP reportedly binding the VLE sequence in Vg1 mRNA using all four KH domains (Git and Standart 2002). Similarly, IMP-1 truncation experiments by Nielsen *et al.* (2002) demonstrate clearly that neither the RRM1-2, KH1-2 or KH3-4 didomains were capable of shifting H19 RNA using EMSA, with the minimum requirements being all four KH domains. In the same study, removal of any KH domain also had inhibitory effects on granule formation in NIH 3T3 cells. Bridging the evolutionary gap to the Drosophila ortholog PSI, once again only the presence of all four KH domains was conducive to RNA binding, and an extended interaction surface along the tandem KH domains was proposed that highlighted the inter-KH domain geometry as in important element in binding (Chmiel *et al.* 2006). In contrast other groups have focussed on the KH3-4 didomain and show that KH3 and KH4 are necessary and sufficient for binding of some mRNA targets (Farina *et al.* 2003; Nielsen *et al.* 2004; Chao *et al.* 2010). Interestingly, Barnes *et al.* (2014) was able to show that simultaneous mutation in KH3 and KH4 did not abrogate binding of CRD-BP with CD44 mRNA, yet taken together, the culmination of CRD-BP/VICKZ experiments focussing on KH domain involvement in RNA binding strongly suggest a substrate specific model of binding with perhaps two classes: i) KH3-4 didomain dependency and ii) KH1, 2, 3 and 4 dependency.

5.6 Structural analysis of CRD-BP KH variants containing point mutation in the GXXG motif at the KH domain

As the introduction of mutations in a protein can have drastic folding consequences for α/β topology, an important feature of the CRD-BP KH mutation experiments was structural

analysis using circular dichroism (CD) spectropolarimetry of each mutant variant. In section 4.2.1 the CD spectra for the wild-type CRD-BP, single KH mutants and double KH mutants can be seen to vary minimally between the proteins with only small variation predominantly at wavelengths below 205 nm and higher than 230 nm (figure 4.3). Each CRD-BP mutant was scanned 8 times (8 accumulations) and experiments completed in triplicate; the resulting far-UV spectral curves are representative of 24 scans and while the averaged spectra did show very small deviations between samples, the overall fold is likely identical and within error of pipetting. Hence the CD evidence for proper KH domain folding strongly suggests that the observed changes in KRAS affinity with the various GXXG mutations (section 3.3.1) were not due to perturbed global structure or destabilization of neighboring CRD-BP domains and further validate the relative roles for each KH domain in binding KRAS mRNA.

GXXG to GDDG mutation in the KH3 and KH4 domains within ZBP1 (GKKG and GKGG to GDDG respectively, to be more specific) has been demonstrated to not have a significant impact on domain structure as determined by superimposition of 15N–1H HSQC spectra and CD thermal denaturation experiments. In the same paper, all four KH domains in KSRP, a related RNA binding protein with high sequence similarity to CRD-BP, were similarly shown to maintain structural integrity with the introduction of the GDDG double mutation (Hollingworth *et al.* 2012). In PSI protein KH domains, a highly conserved isoleucine positioned immediately adjacent to the canonical GXXG motif was mutated to asparagine, and CD analysis of both secondary structure content and thermal stability also revealed that no major structural changes resulted. Structural stability as evidenced from the CRD-BP mutation experiments presented here and from other groups, strongly suggests that while the ubiquitous GXXG loop is

absolutely critical for stable RNA interaction, it does not appear to contribute to the structural integrity of KH domains.

5.7 Monitoring ligand-induced conformational change in full-length CRD-BP and the KH1to4 truncation variant

Protein conformation changes are an inherent property of many nucleic acid (NA) binding proteins such as transcription factors, which can undergo large scale shifts in α-helical content upon binding DNA, such as CREB - a b/Zip type transcription factor that binds a 32 bp DNA sequence (Moll *et al.* 2002). CD spectroscopy is a choice technique for analyzing conformation changes in biomolecules, especially proteins where concentration and stability restraints can be limiting factors. CD spectroscopy also allows the examination of secondary structure shifts in many proteins with regard to binding of ligands, including small molecules, bio-polymers and nucleic acids and is used extensively in this area (Gray 2012). As CRD-BP is an RNA binding protein, we investigated the possibility of it undergoing a conformational shift in secondary structure upon ligand interaction with KRAS RNA. This avenue, if fruitful would open up additional means of assessing sequence specificity for CRD-BP by CD monitoring secondary structural changes associated with high affinity RNA targets, as determined by EMSA or other methods. Furthermore, characterizing CRD-BP structural changes upon binding would complement crystallographic and NMR data with and without RNA present for the entire VICKZ family of proteins.

To approach the assessment of a possible CRD-BP structural change upon RNA binding, far-UV (190-250 nm) CD spectra were gathered for both the wild-type protein and a KH1to4 truncated mutant containing only the four KH domain RNA binding module. Truncated CRD-BP was generated for this purpose, as it has been shown that the RRM domains do not by themselves

bind RNA to any appreciable affinity (mM range) and thus it was rationalized that presence of the RRM domains would likely only mitigate any changes in CD signal (Git and Standart 2002; Nielsen *et al.* 2002). For both CRD-BP variants, a library of potential RNA targets were titrated into solution at protein:RNA molar ratios varying from 16:1 to 1:1 and subsequently scanned using a Jasco J-815 CD spectropolarimeter to search for evidence of secondary structure changes. Due to the inherent CD activity of nucleic acids, solutions of the RNA substrates alone were also scanned and these spectra subtracted from the protein-RNA combination spectra using an in-house software (Jasco SpectraManager) so as to isolate changes associated with CRD-BP.

Examination of the various full-length and KH1to4 CRD-BP spectra revealed interesting findings. Firstly, both versions of CRD-BP displayed a rapid reduction of CD signal at the 208 nm and 222 nm peaks upon addition of RNA (figure 4.4 and 4.7). The change in the spectrum equates to a reduction in overall α/β-topology, achieving a more unstructured form. Such change, while less common among NA-binding proteins (as opposed to undergoing an increase in secondary structure) has been reported previously for the human translin protein. Translin binds both RNA and DNA targets with a particularly high affinity for microsatellite repeats and telomeric repeats and reduces in total α-helicity from 67% to 50% upon binding as determined using CD spectroscopy (Jacob et al. 2004; Kaluzhny *et al.* 2005). However unlike translin, the reduction in total α-helicity in CRD-BP was mostly non-specific, as non-binding negative control RNAs were also able to generate the effect.

Upon close examination of the protein-RNA CD spectra however, there was a trend for the higher affinity RNAs to more effectively induce the loss of α-helicity at equivalent concentrations (figures 4.5 and 4.8). This was the case for all tested RNAs, and these results suggest that an equilibrium exists between bound and un-bound CRD-BP conformational forms,

with higher affinity targets shifting CRD-BP to a bound state more readily. However because the overall effect was generally non-specific, it is also likely that CRD-BP is capable of structurally preorganizing itself in the presence of any RNA prior to settling down on a high affinity sequence, perhaps for 1-dimensional scanning or other purposes. Indeed other nucleic acid-binding proteins have been shown to have the proposed non-specific preorganizing capacity, such as the eukaryotic yeast transcription factor GCN4 in which no relation was evident between the DNA substrate affinity and induction of increased α-helicity. The reported increase in GCN4 α-helicity did however vary from 22-53% depending on substrate binding affinity, much like the case appears to be for CRD-BP (Chan *et al.* 2007).

Despite the large changes in CD signal, it still cannot be said with certainty that CRD-BP actually changes structurally upon RNA binding because similar effects are observed when complexes begin to precipitate or aggregate. As Patel *et al.* (2012) reported a reduction in ZBP1 protein solubility when bound to RNA, it is possible that the observed reduction in structure presented here could result from micro-precipitation, aggregation, or granule formation, although no precipitate was detected after centrifugation of the samples (Correa and Ramos 2008). Furthermore, Chmiel *et al.* (2006) performed similar experiments with the Drosophila PSI protein (73% identity, 80% similarity in the KH1 to KH4 region) and found no substantial changes in secondary structure upon binding target RNA, although the CD data was not shown in the published work for further scrutiny. Interestingly, regardless of how the observed CD change is occurring, the KH1to4 truncation did not appear to have the same substrate specificity for induction of the α-helical loss, as both KRAS-A34 and KRAS-A12 at 1:1 molar ratios were able to achieve the effect with equal competitiveness (figure 4.7A), perhaps implicating the RRM domains as binding regulators.

5.8 RRM1-RRM2 didomain as a negative regulator of RNA-binding

I propose that the RRM1-RRM2 didomain may increase target selectivity by functioning as a negative regulator of RNA binding. This was first evident after fluorescence polarization binding experiments were conducted on the KH1to4 truncation protein. These experiments repeatedly confirmed (this binding assay was completed for every KH1to4 protein sample delivered to the UBC X-ray Crystallography facility) that the KH1to4 variant was able to bind KRAS RNA with enhanced affinity, with a Kd nearly $1/3^{rd}$ that of full-length CRD-BP; KH1to4 Kd = 93.6 ± 34.3 nM and full-length Kd = 311.8 ± 67.1 nM respectively (figure 4.13B). While RRM1 and RRM2 domains in CRD-BP are not capable of binding RNA independently (Git and Standart 2002; Nielsen *et al.* 2002), they may recognize certain sequences and affect the KH1to4 binding module through inter-domain cooperativity. Alternatively they may play a more static role, providing additional steric restrictions to a lock and key fit for CRD-BP RNA targets. However, further experiments with additional RNA targets will be needed to test this hypothesis. If substantiated, a regulatory role for the RRM domains may explain the aggressive clonogenic growth of several breast cancer cell lines that have only recently been discovered to express a KH1 to KH4 CRD-BP isoform (Fakhraldeen *et al.* 2015).

5.9 Assessing CRD-BP-RNA complex stability using thermal shift analysis

Stability of a protein is another factor that can change with ligand binding and is often used as a tool to aid protein crystal formation (Vedadi *et al.* 2006). An increase in the melting temperature of a protein, or a thermal shift, is a common measure of protein stability and changes in Tm are typically measured using either the indirect thermofluor method, which uses non-

specific fluorescent dyes that bind hydrophobic protein surfaces, or the more direct CD monitoring of secondary structure (reviewed in Lavinder *et al.* 2009). Measuring the CD signal at 222 nm (wavelength most representative of α-helical content) over a temperature range allows the protein unfolding to be monitored accurately and a Tm (midpoint of thermal denaturation curve) can be subsequently calculated from the melting curve to compare mutant proteins, homologs, and also protein stability changes as a result of ligand binding (Layton *et al.* 2011). As the previously explored CD structural analysis revealed no clear way of distinguishing whether or not CRD-BP was binding a potential RNA target, and although a minor correlation between RNA affinity and α-helicity reduction was apparent, and no ligand-induced thermal shift experiments have been reported for VICKZ proteins, the CD thermal shift assay was pursued in another attempt to develop a novel method of assessing RNA affinity for CRD-BP. In theory a high affinity interaction may stabilize the protein significantly and be detectable with this method as a shift in melting temperature.

Initial work in developing this method proved difficult, as the solubility of the protein-RNA complex appeared to reduce with increasing amounts of RNA approaching a 1:1 protein:RNA ratio and large amounts of precipitate could be recovered from centrifugation. An increased concentration of the complex compared to the CD structural scanning was necessary however for determining melting temperature, due to the fact that the necessary CD signal at 222 nm was so strongly reduced upon addition of the RNA (refer to section 4.2.2). Thus the ratio of protein to RNA was kept quite high at 16:1 and 4:1 with protein concentration at approximately 10 μM. Determination of melting temperature was accomplished as per section 4.1.6 and the resulting Tm's compared in figure 4.10 where it can be seen that while there was a small increase in melting temperature of 3.5 ^{o}C upon addition of KRAS-A34 RNA, the effect was not specific

as KRAS-A12 also induced a similar change at both concentrations (Tm = 43.9 °C for both). Compared to the 18S and tRNA (Tm = 41.5 °C and 42.0 °C, respectively) negative control RNAs however, there was a larger increase in Tm indicating that there is some degree of specificity. It is worth noting that while EMSA was able to show that KRAS-A34 had higher affinity for CRD-BP than KRAS-A12 (figure 2.13), it was still lower than that of the initial parent fragment KRAS-A, indicating that KRAS-A12 may have contained a weak binding site not evident in the EMSA assays and could be responsible for the similar Tm changes observed between the two RNAs. While increasing the RNA closer to a 1:1 ratio may better distinguish the effect of RNAs with different CRD-BP affinities, the difference in Tm between the high and low affinity RNAs can also be seen to decrease with increasing concentration of the RNA, suggesting that the method becomes even less sensitive as more RNA is added. This ran somewhat counter-intuitive to our initial expectations and implies a non-specific effect that may be explainable by non-specific RNA binding or ubiquitous losses in solubility, although precipitation in these samples was not observed. In general, thermal shift analysis of CRD-BP as a method for assessing RNA target affinity proved limited, as the differences in Tm were very small and did not correlate well with RNA binding.

5.10 CRD-BP KH1to4 protein crystal

No complete crystal structure for CRD-BP or any of the VICKZ proteins currently exists, with the closest being a structure solved for the KH3 and KH4 didomain of IMP1 bound to a small segment of RNA (Chao et al. 2010). While this structure has elucidated an RNA-looping binding mechanism, with anti-parallel binding surfaces on opposite sides of the di-domain, it has not explained the diversity of mRNA targets or been sufficient to generate a consensus binding

sequence, necessitating the need for a more complete model. For this purpose, a KH1to4 CRD-BP truncation variant was synthesized containing amino acids 196 to 566 (NCBI Reference Sequence: AAC72743.1), covering the four KH domains. The precise positions of truncation were chosen based on previous studies; the C-terminus position based on the successful Chao et al. (2010) IMP1 crystal construct, and the N-terminal based on KH1 to KH4 functional studies (Nielsen et al. 2002). Physiological relevance of a KH1 to KH4 crystal structure stems from a recent work revealing the presence of a CRD-BP splice variant in human breast cancer cells that only contains the KH1 to KH4 region and is necessary for clonogenic growth (Fakhraldeen et al. 2015). Strikingly this discovery suggests that CRD-BP is underrepresented in many cancer studies, as antibodies specific to the N-terminal deletion will not detect this CRD-BP isoform. Also, the KH1 to KH4 variant retains full RNA binding capacity as well as the full-length's propensity for protein-protein dimerization (Nielsen et al. 2004).

Conventional native preparation of the full-length CRD-BP and KH1to4 truncation yielded low quantities of soluble protein, a complication previously reported by Du et al. (2008), Git and Standart (2002) and Prokipcak et al. (1993) as well as high levels of RNA contamination (figure 4.12B), presumably from endogenously bound E.coli RNA. Furthermore, our lab's existing denaturing protein purification protocol also proved limited to refolding concentrations of less than 0.7 mg/ml as well as issues concentrating to high levels. Eventually a denaturing, on-column refolding method was developed that allowed for both high concentration of CRD-BP with no RNA contamination (4.1.9) and this method was used to begin crystallization trials in collaboration with the Dr. Strynadka X-ray crystallography laboratory at UBC. Though several attempts produced only false positives (mostly in phosphate buffer), a single hit with conditions: 0.2M lithium sulphate, 0.1M Tris pH 8.5, 40% v/v PEG400 produced a UV-fluorescent crystal

but was not of high enough quality for further study (figure 4.13D). Optimization experiments calculated a temperature of aggregation for CRD-BP KH1to4 of near 41 °C in most buffers tested (figure 4.14), which matches perfectly with the melting temperature as determined by CD thermal shift experiments (4.1 °C), confirming the high temperature sensitivity of this protein which may explain the difficulties encountered in these crystallization attempts. In addition, the KH1-KH2 didomain has been studied using solution NMR, and was determined to possess inherent flexibility not present in the KH3-KH-4 didomain, possibly impacting the ability to form a stable crystal nucleation site (Du et al. 2008). While the crystal hit was ultimately unsuitable for X-ray diffraction data collection, thermal aggregation experiments found that pH 6, high salt and MES buffer (4.2.4) were all stabilizing factors and are highly relevant for future CRD-BP crystal pursuits.

5.11 Summary & concluding remarks

CRD-BP and the orthologous VICKZ protein family members, including human IMP-1, have been implicated in several forms of cancer, as well as playing critical roles in the early stages of embryogenesis (Elcheva *et al.* 2008; Fakhraldeen *et al.* 2015; Ioannidis *et al.* 2001;Nielsen *et al.* 1999). CRD-BP achieves its effects (conferring enhanced transcript stability, translational regulation, and mRNA localization) through binding mRNA in the cytoplasm and nucleus (Doyle *et al.* 2000). Despite demonstrated importance in gene regulation and expansive international research efforts, there remains no model that fully characterizes VICKZ proteins. In this work, the molecular interaction between CRD-BP and target mRNA was studied, focusing on KRAS – a proto-oncogene that when mutated (G12V) or overexpressed leads to misregulation of cell growth through induction of the downstream effector pathways: mTOR, PI3K, MAPK, and EGFR (Radzioch *et al.* 2010; Birkeland *et al.* 2012; Valtorta *et al.* 2013;

Sasaki *et al.* 2011). It has been demonstrated by Mongroo *et al.* (2011) that CRD-BP positively regulates KRAS expression in colorectal cancer cell lines. Knockdown with CRD-BP-siRNA displayed a 53% reduction in KRAS in SW-480 and induction of CRD-BP with a mammalian expression vector nearly tripled KRAS levels in LIM2405 cells. Uncoupling CRD-BP expression with KRAS signalling using inhibitor molecules may therefore prove to be a valuable strategy in battling KRAS-driven cancers.

Molecular details pertaining to the CRD-BP-KRAS mRNA interaction show direct binding with the coding region and 3'UTR, as determined through UV-crosslinking (Mongroo *et al.* 2011). No further specificity of a binding sequence or indication of actual binding affinity is known, which prompted the KRAS mapping experiments for CRD-BP binding (section 2.3.2). KRAS mRNA was found to bind CRD-BP in both the coding region and 3'UTR using EMSA, confirming previous findings of Mongroo *et al.* (2011). Indeed multiple binding sites were discovered within both regions, with at least two present in the coding region (1-185 and 388-610) and two present in the 3'UTR (568-793 and 971-1155). The highest affinity RNA fragment "KRAS-A" covered nucleotide positions 1-185 in the coding region and the dissociation constant calculated to be 177.9 ± 18.4 nM. Further EMSA analysis of KRAS-A revealed a 57 nucleotide minimum CRD-BP binding sequence (129-185) KRAS-A34(b), as well as fragments possessing virtually no affinity for CRD-BP, confirming some specificity of binding. The predicted *mfold* secondary structure of the KRAS-A34(b) sub-region, covering nucleotides 129-185, as well as the raw primary RNA sequence, will add to a growing library of searchable RNA structures and sequences for determination of a consensus binding motif for CRD-BP. Furthermore, the small size of the KRAS binding region allows for future fluorescence polarization based high-throughput screens for potential small molecule inhibitors.

CRD-BP like many RBPs utilizes multiple RNA-attachment sites to achieve high affinity, high specificity interactions. Four K-homology (KH) domains constitute the primary RNA binding module, with two N-terminal RRM domains also present; the precise purpose of these two domains remains unclear. Specific KH domain requirements for binding appear unique to each mRNA transcript, implying that various short binding elements on the RNA act in combination with the protein. The high affinity KRAS-A sub-region (nts 1-185) was examined in conjunction with KH mutant variants of CRD-BP to determine which if any domain was predominantly responsible for binding. Both EMSA and immune-precipitation coupled with RT-qPCR analysis in HeLa cells revealed a similar pattern of KRAS association, strongly supporting the notion that KRAS requires all four KH domains to stabilize the interaction, with a somewhat mitigated role for KH1 domain. This result optimistically suggests that drugs found capable of blocking any of the four KH domains will strongly inhibit the CRD-BP-KRAS mRNA interaction, in cells and in vitro, making KRAS a lucrative opportunity among CRD-BP targets.

Traditional drug discovery often proceeds with random testing of generated small molecule libraries, searching for a desired effect among thousands or millions of candidates (reviewed in Hughes *et al.* 2010). The use of specific anti-sense oligonucleotide inhibitors that bind mRNA and reduce translation, or interaction with nascent proteins is relatively new, and requires specific knowledge of the target mRNA but can indirectly inhibit a protein target (Henning and Best, 2002). To date there is no published work on inhibitors designed to abrogate the interaction between CRD-BP and KRAS mRNA, in vitro or in vivo. This thesis presents two DNA-based anti-sense oligonucleotide inhibitors (section 2.3.3, SM6 and SM7) against the CRD-BP-KRAS mRNA interaction that were developed with the knowledge gained from the KRAS minimum binding sequence (section 2.3.1), the most effective with an IC_{50} of 11.8 nM.

The most probable mechanism of their inhibition is complementary Watson-Crick base pairing, masking the KRAS RNA binding sequence/structure from CRD-BP and thereby inhibiting interaction. Overlap of the minimum binding sequence discovered in section 2.3.1 with the anti-sense oligonucleotide inhibitor target regions doubly implicate the KRAS-A34(b) region (nts: 129-185) as important for CRD-BP interaction with KRAS RNA. The research into DNA-based oligonucleotide inhibitors presented here has laid the groundwork for future cell-based assays using SM6 and SM7 to target KRAS, and may join the other hopeful anti-KRAS therapeutic strategies with further development. Additionally, SM7 and SM2 have found further use as a positive control in fluorescence polarization based screening of proto-drugs for anti-CRD-BP activity in our laboratory.

Secondary structure analysis of the CRD-BP protein revealed a highly structured protein with primarily α-helical (69%) composition, in-line with other KH domain-containing proteins (Chao *et al.* 2010; Valverde *et al.* 2007; Chmiel *et al.* 2006). Further analysis of the CRD-BP KH mutant library concluded that introduction of the GXXG mutations used to study KH domain requirements in RNA binding did not compromise the global structure of the protein. Remarkably little variation was apparent between the CRD-BP KH mutant proteins CD spectra and provides compelling evidence that the GXXG loop region between the first and second α-helix does not contribute any structural integrity to the protein. Further CD analysis showed quantifiable changes in secondary structure upon binding RNA implying that CRD-BP may have multiple forms, one conformation stabilized by RNA binding. The effect was generally non-specific with all RNA able to induce the shift, but the degree of change did appear to correlate with RNA target affinity suggesting some specificity (4.2.2). Interestingly with the removal of the RRM domains, the specificity was mitigated between KRAS-A12 and KRAS-A34 RNA

targets (high and low affinity respectively) suggesting some role in specificity for the RRM domains. Supporting this notion is the fluorescence polarization binding assays completed, comparing the full-length CRD-BP to a generated KH1to4 truncation variant (4.2.4). Kd comparison reveals a striking 3-fold increase in affinity with the removal of the RRM domains. As no function has yet been attributed to the RRM domains in VICKZ proteins, confirmation of this effect with other RNA targets may ascribe a novel negative binding regulator function to the RRM domains.

Interdomain geometry within the "KH1to4" binding module is key to understanding and predicting CRD-BP mRNA targets, as suggested by Chao *et al.* (2010) and evidenced by the appearance of bipartite sequence requirements for β-actin "zipcode" RNA. Currently there is no crystal structure of full-length CRD-BP or any of the VICKZ proteins, with only a structure of the KH3-KH4 didomain available to scrutinize. However, as KRAS requires all four KH domains, a more complete structure is necessary to understand how binding occurs. Despite a thorough attempt to obtain a diffraction pattern and crystal structure of CRD-BP, no crystal suitable for structure determination were obtained. A single large protein crystal possessing twinned character did grow in 0.2M lithium sulphate, 0.1M Tris-Cl pH 8.5, 40% v/v PEG400 buffer, however it was deemed not suitable for X-ray diffraction. Formation of a crystal does however suggest that the specific KH1to4 construct described in this thesis can potentially crystallize and future efforts in complex with an RNA target may be more successful.

References

Atlas, R., Behar, L., Elliott, E. Ginzburg, I. (2004). The insulin-like growth factor mRNA binding-protein IMP-1 and the Ras-regulatory protein G3BP associate with tau mRNA and HuD protein in differentiated P19 neuronal cells. *J. Neurochem.* **89**: 613-626.

Appels, N. M., Beijnen, J. H., and Schellens, J. H. (2005). Development of farnesyl transferase inhibitors: a review. *Oncologist.* **10**: 565–578.

Baines, A., Xu, D., Der, C. (2011). Inhibition of Ras for cancer treatment: the search continues. *Future Med. Chem.* **3**: 1787-1808.

Barnes, M., Van Rensburg, G., Li, W., Mehmood, K., Mackedenski, S., Chan, C., King, T., Miller, A., Lee, H. (2014). Molecular insights into the coding region determinant-binding protein-RNA interaction through site-directed mutagenesis in the heterogeneous nuclear ribonucleoprotein-K-homology domains. *J. Biol. Chem.* **290**: 625-639.

Bell, J., Wächter, K., Mühleck, B., Pazaitis, N., Köhn, M., Lederer, M., Hüttelmaier, S. (2013). Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression?. *Cell. Mol. Life Sci.* **70**: 2657-2675.

Bernstein, P., Herrick, D., Prokipcak, R., Ross, J. (1992). Control of c-myc mRNA half-life in vitro by a protein capable of binding to a coding region stability determinant. *Genes Dev.* **6**(4): 642-654.

Berx, G., Cleton-Jansen, A. M., Nollet, F., de Leeuw, W. J., van de Vijver, M., Cornelisse, C. and van Roy, F. (1995). E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. *EMBO J.* **14**: 6107-6115.

Birchmeier, W., Behrens, J. (1994). Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. *Biochim. Biophys. Acta.* **1198**: 11-26.

Birkeland, E., Wik, E., Mjos, S., Hoivik, E., Trovik, J., Werner, H,. Kusonmano, K., Peterson, K., Raeder, M., Holst, F., Oyan, A., Halland, K., Akslen, L., Simon, R., Krakstad, C., Salvesen, H. (2012). KRAS gene amplification and overexpression but not mutation associates with aggressive and metastatic endometrial cancer. *Br. J. Cancer.* **107**: 1997-2004.

Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlaan–de Vries, M., van Boom, J. H., van der Eb, A. J., & Vogelstein, B. (1987). Prevalence of ras gene mutations in human colorectal cancers. *Nature.* **327**(6120): 293-297.

Boyerinas, B., Park, S., Shomron, N., Hedegaard, M., Vinther, J., Anderson, J., Feig, C., Xu, J., Burge, C., Peter, M. (2008). Identification of Let-7-regulated oncofetal genes. *Cancer Res.* **68**(8): 2587-2591.

Boylan, K., Mische, S., Li, M., Marques, G., Morin, X., Chia, W., Hays, T. (2008). Motility screen identifies Drosophila IGF-II mRNA-binding protein-zipcode-binding protein acting in oogenesis and synaptogenesis. *PLoS Genet.* **4**(2): e36.

Brewer, G. (2003). Messenger RNA decay during ageing and development. *Ageing Res.* **1**: 607-625.

Bulheller, B., Hirst, J. (2009). DichroCalc-circular and linear dichroism online. *Bioinformatics.* **25**: 539-540.

Burmer, G., Loeb, L. (1989). Mutations in the KRAS2 oncogene during progressive stages of human colon carcinoma. Proc. *Natl. Acad. Sci.* **86** (7): 2403–7.

Chan, I., Shahravan, S., Fedorova, A., Shin, J. (2008). The bZIP targets overlapping DNA subsites within a half-site, resulting in increased binding affinities. *Biochemistry.* **46**: 1663-1671.

Chandra, A., Grecco, H. E., Pisupati, V., Perera, D., Cassidy, L., Skoulidis, F., et al. (2012). The GDI-like solubilizing factor PDEdelta sustains the spatial organi- zation and signalling of Ras family proteins. *Nat. Cell Biol.* **14**: 148–158.

Chao, J., Patskovsky, Y., Patel, V. (2010). ZBP1 recognition of β-actin zipcode induces RNA looping. *Genes Dev.* **24**: 148-158.

Chen, C. Y., Del Gatto–Konczak, F., Wu, Z., & Karin, M. (1998). Stabilization of interleukin-2 mRNA by the c-Jun NH2-terminal kinase pathway. *Science.* **280**: 1945-1949.

Chmiel, N., Rio, D., Doudna, J. (2006). Distinct contributions of KH domains to substrate binding affinity of drosophila P-element somatic inhibitor protein. *RNA.* **12**: 283-291.

Cogoi, S., Zorzet, S., Rapozzi, V., Géci, I., Pedersen, E., Xodo, L. (2013). MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res.* **41**(7): 4049-4064.

Collins, M., Magliano, M (2014). Kras as a key oncogene and therapeutic target in pancreatic cancer. *Front. Physiol.* **4**: 407.

Condeelis, J., Singer, R. H., & Segall, J. E. (2005). The great escape: when cancer cells hijack the genes for chemotaxis and motility. Annu. Rev. *Cell Dev. Biol.* **21**: 695-718.

Costa, S. (2013). Development of a Novel Fusion System for Recombinant Protein Production and Purification in Escherichia coli. Ph.D. Thesis, University of Minho, Braga, Portugal.

Coulis, C., Lee, C., Nardone, V., Prokipcak, R. (2000). Inhibition of c-myc expression in cells by targeting an RNA-protein interaction using antisense oligonucleotides. Mol. Pharmacol. **57**(3): 485-494.

Craig, E. A., & Spiegelman, V. S. (2012). Inhibition of coding region determinant binding protein sensitizes melanoma cells to chemotherapeutic agents. Pigment Cell Melanoma Res. **25**(1): 83-87.

Dai, N., Christiansen, J., Nielsen, C., Avruch, J. (2013). mTOR complex 2 phosphorylates IMP1 cotranslationally to promote IGF2 production and the proliferation of mouse embryonic fibroblasts. *Genes Dev.* **27**: 301-312.

Davies, M.A., and Samuels, Y. (2010). Analysis of the genome to personalize therapy for melanoma. *Oncogene.* **29**: 5545–5555.

Dean, J. L., Sully, G., Clark, A. R., & Saklatvala, J. (2004). The involvement of AU-rich element-binding proteins in p38 mitogen-activated protein kinase pathway-mediated mRNA stabilisation. *Cell. signal.* **16**(10): 1113-1121.

Dergham, S. T., Dugan, M. C., Kucway, R., Du, W., Kamarauskiene, D. S., Vaitkevicius, V. K., ... & Sarkar, F. H. (1997). Prevalence and clinical significance of combined K-ras mutation and p53 aberration in pancreatic adenocarcinoma. Int. J. Pancreatol. **21**(2): 127-143.

Derrigo, M., Cestelli, A., Savettieri, G., Diliegro, I. (2000). RNA-protein interactions in the control of stability and localization of mRNA. *Int. J. Mol. Med.* **5**(2): 111-123.

Deshler, J., Highett, M., Abramson, T., Schnapp, B. (1998). A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates. *Curr. Biol.* **8**: 489-496.

Dimitriadis, E., Trangas, T., Milatos, S., Foukas, P. G., Gioulbasanis, I., Courtis, N., ... & Ioannidis, P. (2007). Expression of oncofetal RNA-binding protein CRD-BP/IMP1 predicts clinical outcome in colon cancer. J. Int. Cancer. **121**(3): 486-494.

Doyle, G., Betz, N., Leeds, P., Fleisig, A., Prokipcak, R., Ross, J. (1998). The c-myc coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins. *Nucleic Acids Res.* **26**: 5036-5044.

Doyle, G., Bourdeau-Heller, J., Coulthard, S., Meisner, L., Ross, J. (2000). Amplification in human breast cancer of a gene encoding a c-myc mRNA binding protein. *Cancer Res.* **60**: 2756-2759.

Du, Z., Fenn, S., Tijhen, R., James, T. (2008). Structure of a construct of a human poly(C)-binding protein containing the first and second KH domains reveals insights into its regulatory mechanisms. *J. Biol. Chem.* **283**: 28757-28766.

Elcheva, I., Tarapore, R., Bhatia, N., Spiegelman, V. (2008). Overexpression of mRNA-binding protein CRD-BP in malignant melanomas. *Oncogene.* **27**: 5069-5074.

Elcheva, I., Goswarmi, S., Noubissi, F., Spiegelman, V. (2009). CRD-BP protects the coding region of βTrCP1 mRNA from miR-183-mediated degradation. *Mol. Cell.* **35**(2): 240-246.

Fakhraldeen, S., Clark, R., Roopra, A., Chin, E., Huang, W., Castorino, J., Wisinski, K., Kim, T., Spiegelman, V., Alexander, C. (2015). Two isoforms of the RNA binding protein, Coding Region Determinant-binding Protein (CRD-BP/IGF2BP1), are expressed in breast epithelium and support clonogenic growth of breast tumor cells. *J. Biol. Chem.* **290**(21): 13386-13400.

Farina, K., Hüttelmaier, S., Musunuru, K., Darnell, R., Singer, R. (2003). Two ZBP1 KH domains facilitate β-actin mRNA localization, granule formation, and cytoskeletal attachment. *J. Cell Biol.* **160**(1): 77-87.

Fenn, S., Du, Z., Lee, J., Tjhen, R., Stroud, M., James, T. (2007). Crystal structure of the third KH domain of human poly(C)-binding protein-2 in complex with a C-rich strand of human telomeric DNA at 1.6 Å resolution. *Nucleic Acids Res.* **35**(8): 2651-2660.

Ferrara, N. (2010). Pathways mediating VEGF-independent tumor angiogenesis. *Cytokine Growth Factor Rev.* **21**: 21–26.

Filipowicz, W., Bhattacharyya, S., Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?. *Nat. Rev. Genet.* **9**(2): 102-14.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., Bairoch, A. ; *Protein Identification and Analysis Tools on the ExPASy Server*; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607.

Georges, R. N., Mukhopadhyay, T., Zhang, Y., Yen, N., & Roth, J. A. (1993). Prevention of orthotopic human lung cancer growth by intratracheal instillation of a retroviral antisense K-ras construct. Cancer Res. **53**(8): 1743-1746.

Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., and Chen, C.Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol. Cell.* **14**: 571–583.

Git, A., Standart, N. (2012) The KH domains of Xenopus Vg1RBP mediate RNA binding and self-association. *RNA*. **8**: 1319-1333.

Goolsby, K., Shapiro, D. (2003). RNAi-mediated depletion of the 15 KH domain protein, vigilin, induces death of dividing and non-dividing human cells but does not initially inhibit protein synthesis. *Nucleic Acids Res.* **31**(19): 5644-5653.

Goswarmi, S., Tarapore, R., TeSlaa, J., Grinblat, Y., Setaluri, V., Spiegelman, S. (2010). MicroRNA-340-mediated degradation of microphthalmia-associated transcription factor mRNA is inhibited by the coding region determinant-binding protein. *J. Biol. Chem.* **285**(27): 20532-20540.

Gray, D. Circular dichroism of protein-nucleic acid interactions. In Comprehensive chiroptical spectroscopy, Volume 2: Application in stereochemical analysis of synthetic compounds, natural products, and biomolecules Berova, N., Polavarapu, P., Nakanishi, K., Woody, R., Ed.; Wiley: New York, 2012; vol. 2; p 615-633.
Grishin, N. (2001). KH domain: one motif, two folds. *Nucleic Acids Res.* **29**: 638–643.

Gu, W., Katz, Z., Wu, Bin., Park, H., Li, D., Lin, Stanley., Wells, A., Singer, R. (2011). Regulation of local expression of cell adhesion and motility-related mRNAs in breast cancer cells by IMP1/ZBP1. *J. Cell Sci.* **125**: 81-91.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A., Munschauer, M., Ulrich, A., Wardle, G., Dewell, S., Zavolan, M., Tuschl, T. (2010). Transciptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell.* **141**: 129-141.

Hall, M., Huang, S., and Black, D. (2004). Differentiation-induced colocalization of the KH-type splicing regulatory protein with polypyrimidine tract binding protein and the c-src pre-mRNA. *Mol. Biol. Cell.* **15**: 774–786.

Hamilton, K., Noubissi, F., Katti, P., Hahn, C., Davey, S., Lundsmith, E., Klein-Szanto, A., Rhim, A., Spiegelman, V., Rustgi, A. (2013). IMP1 promotes tumor growth, dissemination and a tumor-initiating cell phenotype in colorectal cancer cells xenografts. *Carcinogenesis*. **34**(11): 2647-2654.

Hananan, D., Weinberg, R. (2011). Hallmarks of cancer: the next generation. *Cell.* **144**(5): 646-674.

Hansen, T., Hammer, N., Nielsen, J., Madsen, M., Dalbaeck, C., Wewer, U., Christiansen, J., Nielsen, F. (2004). Dwarfism and impaired gut development in insulin-like growth factor II mRNA-binding protein 1-deficient mice. Mol Cell Biol **24**(10): 4448–4464

Havin, L., Git, A., Elisha, Z., Oberman, F., Yaniv, K., Schwartz, S., Standart, N., Yiseraeli, J. (1998). RNA-binding protein conserved in both microtubule- and microfilament-based RNA localization. *Genes Dev.* **12**: 1593-1598.

Haynes, S. (1999). RNA-Protein interactions. Totowa, New Jersey, Humana Press.
Hennessy, B. T., Smith, D. L., Ram, P. T., Lu, Y., & Mills, G. B. (2005). Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat. Rev. Drug Discovery*. **4**(12): 988-1004.

Henning SW, Beste G (2002). Loss-of-function strategies in drug target validation. Curr. Drug Discov. May: 17–21.

Hollingworth, D., Candel, A., Nicastro, G., Martin, S., Briata, P., Gherzi, R., Ramos, A. (2012). KH domains with impaired nucleic acid binding as a tool for functional analysis. *Nucleic Acids Res*. **40**(14): 6873-6886.

Hüttelmaier, S., Zenklusen, D., Lederer, M., Dictenberg, J., Lorenz, M., Meng, X., Bassell, G., Condeelis, J., Singer, R. (2005). Spatial regulation of β-actin translation by Src-dependent phosphorylation of ZBP1. *Nature*. **438**: 512-515.

Ioannidis, P., Trangas, T., Dimitriadis, E., Samiotaki, M., Kriazoglou, I., Tsiapalis, C., Kittas, C., Agnantis, N., Nielsen, F., Nielsen, J., Christiansen, J., Pandis, N. (2001). C-myc and IGF-II mRNA-binding protein (CRD-BP/IMP-1) in benign and malignant mesenchymal tumors. *Int. J. Cancer*. **94**: 480-484.

Ioannidis, P., Mahaira, L., Papadopoulou, A., Teixeira, M. R., Heim, S., Andersen, J. A., ... & Trangas, T. (2003). 8q24 Copy number gains and expression of the c-myc mRNA stabilizing protein CRD-BP in primary breast carcinomas. *Int. J. Cancer*. **104**(1): 54-59.

Ioannidis, P., Mahaira, L., Perez, S., Gritzapis, A., Sotiropoulou, P., Kavalakis, G., Antsaklis, A., Baxevanist, C., Papamichail, M. (2005). CRD-BP/IMP1 expression characterizes cord blood $CD34^+$ stem cells and effects c-myc and IGF-II expression in MCF-7 cancer cells. *J. Biol. Chem.* **280**(20): 20086-20093.

Jacob, E., Pucshanski, L., Zeruya, E., Baran N., Manor, H. (2004). The human protein translin specifically binds single-stranded microsatellite repeats, d(GT)n, and G-strand telomeric repeats, d(TTAGGG)n: a study of the binding parameters. *J. Mol. Biol.* **344**: 939-950.

Jančík, S., Drábek, J., Radzioch, D., and Hajdúch, M. (2010). Clinical Relevance of KRAS in Human Cancers. *J Biomed Biotechnol.* vol. 2010, Article ID 150960, 13 pages, 2010. doi:10.1155/2010/150960

Johnson, S., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K., Brown, D., Slack, F. (2005). RAS is regulated by the let-7 microRNA family. *Cell.* **120**: 635-647.

Jones, A., Schedl, T. (1995). Mutations in gld-1, a female germ cell-specific tumor suppressor gene in *Caenorhabditis elegans*, affect a conserved domain also found in Src-associated protein Sam68. *Genes Dev.* **9**: 1491-1504.

Jonson, L., Vikesaa, J., Krogh, A., Nielsen, L., Hansen, T., Borup, R., Johnsen, A., Christiansen, J., Nielsen, F. (2007). Molecular composition of IMP1 ribonucleoprotein granules. *Mol. Cell. Proteomics*. **6**: 798-811.

Kaluzhny, D., Laufman, O., Timofeev, E., Borisova, O., Manor, H., Shchyolkina, A. (2005). Conformational changes induced in the human protein translin and in the single-stranded oligodeoxynucleotides d(GT)(12) and d(TTAGGG)(5) upon binding of these oligodeoxynucleotides by translin. *J. Biomol. Struct. Dyn.* **23**: 257-265.

Kato, T., Hayama, S., Yamabuki, T., Ishikawa, N., Miyamoto, M., Ito, T., Tsuchiya, E., Kondo, S., Nakamura, Y., Daigo, Y. (2007). Increased expression of insulin-like growth factor-II

messenger RNA-binding protein 1 is associated with tumor progression in patients with lung cancer. *Clin. Cancer Res.* **13**(2): 434-442.

Keene, J. D. (2007). RNA regulons: coordination of post-transcriptional events. *Nature Reviews Genetics*, **8**(7): 533-543.
Keith, B., Johnson, R. S., & Simon, M. C. (2011). HIF1α and HIF2α: sibling rivalry in hypoxic tumour growth and progression. Nat. Rev. Cancer. **12**(1): 9-22.

Kim, M., Slack, F. (2014) MicroRNA-mediated regulation of KRAS in cancer. *J. Hematol. Oncol.* **7**: 84.

King, D., Barnes, M., Thomsen, D., Lee, C. (2014). Assessing specific oligonucleotides and small molecule antibiotics for the ability to inhibit the CRD-BP-CD44 RNA interaction. *PLoS ONE*. **9**(3): e91585.

Kjersem, J., Ikdahl, T., Guren, T., Skovlund, E., Sorbye, H., Hamfjord, J., Pfeiffer, P., Glimelius, B., Kersten, C., Solvang, H., Tveit, K., Kure, E. (2012). Let-7 miRNA-binding site polymorphism in the KRAS 3'UTR; colorectal cancer screening population prevalence and influence on clinical outcome in patients with metastatic colorectal cancer treated with 5-flourouracil and oxaliplatin +/- cetuximab. *BMC Cancer*. **12**: 534.

Köbel, M., Weidensdorfer, D., Reinke, C., Lederer, M., Schmitt, W. D., Zeng, K., ... & Hüttelmaier, S. (2007). Expression of the RNA-binding protein IMP1 correlates with poor prognosis in ovarian carcinoma. Oncogene. **26**(54), 7584-7589.

Laheru, D., Shah, P., Rajeshkumar, N. V., Mcallister, F., Taylor, G., Goldsweig, H., et al. (2012). Integrated preclinical and clinical development of S-trans, trans- Farnesylthiosalicylic Acid (FTS, Salirasib) in pancreatic cancer. *Invest. New Drugs* **30**: 2391–2399.

Lavinder, J., Hari, S., Sullivan, B., Magliery, T. (2008). High-throughput thermal screening: a general, rapid dye-binding thermal shift screen for protein engineering. *J. Am. Chem. Soc.* **131**: 3794-3795.

Layton, C., Hellinga, H. (2011). Quantitation of protein-protein interactions by thermal stability shift analysis. *Protein Sci.* **201**: 1439-1450.

Le, Hang., Sorrell, A., Siddle, K. (2012). Two isoforms of the mRNA binding protein IGFBP2 are generated by alternative translational initiation. *PLoS ONE*. **7**(3): e33140.

Lebedeva, S., Jens, M., Theil, K., Schwanhäusser, B., Selbach, M., Landthaler, M., & Rajewsky, N. (2011). Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol. cell*. **43**(3): 340-352.

Leeds, P., Kren, B., Boylan, J., Betz, N., Steer, C., Gruppuso, P., Ross, J. (1997). Developmental regulation of CRD-BP, an RNA-binding protein that stabilizes c-myc mRNA in vitro. *Oncogene*. **14**: 1279-1286.

Lewis, H., Musunuru, K., Jensen, K., Edo, C., Chen, H., Darnell, R., Burley, S. (2000). Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell.* **100**: 323-332.

Liao, B., Patel, M., Hu, Y., Charles, S., Herrick, D., Brewer, G. (2004). Targeted knockdown of the RNA-binding protein CRD-BP promotes cell proliferation via an insulin-like growth factor II-dependent pathway in human K562 cells. *J. Biol. Chem.* **279**(47): 48716-48724.

Lin, Q., Taylor, S., Shalloway, D. (1997). Specificity and determinants of Sam68 RNA binding. *J. Biol. Chem.* **272**(43): 27274-27280.

Lin, F., Chen, Y., Lin, Y., Tsai, J., Chen, J., Wang, H., Lin, S. (2006). The Role of Human Antigen R, an RNA-binding Protein, in Mediating the Stabilization of Toll-Like Receptor 4 mRNA Induced by Endotoxin A Novel Mechanism Involved in Vascular Inflammation. Arterioscler. Thromb. Vasc., **26**(12): 2622-2629.

Linker, K., Pautz, A., Fechir, M., Hubrich, T., Greeve, J., and Kleinert, H. (2005). Involvement of KSRP in the post-transcriptional regula- tion of human iNOS expression-complex interplay of KSRP with TTP and HuR. *Nucleic Acids Res.* **33**: 4813–4827.

Louis-Jeune, C., Andrade-Navarro, M., Perez-Iratxeta, C. (2011). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins*: Structure, Function, and Bioinformatics*. **80**(2): 374-381.

Mascaux, C., Iannino, N., Martin, B., Paesmans, M., Berghmans, T., Dusart, M., ... & Sculier, J. P. (2004). The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br. J. Cancer.* **92**(1): 131-139.

Maurer, T., Garrenton, L., Oh, A., Pitts, K., Anderson, D., Skelton, N., Fauber, B., Malek, S., Stokeoe, D., Ludlam, M., Bowman, K., Gianetti, A., Starovasnik, M., Mellman, I., Jackson, P., Rudolph, J., Wang, W., Fang, G. (2012). Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc Natl. Acad. Sci. USA.* **109**(14): 5299-5304.

Miura, Y., Ohnami, S., Yoshida, K., Ohashi, M., Nakano, M., Ohnami, S., Aoki, K. (2005). Intraperitoneal injection of adenovirus expressing antisense K-ras RNA suppresses peritoneal dissemination of hamster syngeneic pancreatic cancer without systemic toxicity. Cancer Lett. **218**(1): 53-62.

Moll, J., Acharya, A., Gal, J., Mir, A., Vinson, C. (2002). Magnesium is required for specific DNA binding of the CREB B-ZIP domain. *Nucleic Acids Res.* **30**: 1240-1246.

Mongroo, P., Noubissi, F., Cuatrecasas, M. (2011) IMP-1 displays cross-talk with K-Ras and modulates colon cancer cell survival through the novel proapoptotic protein CYFIP2. *Cancer Res.* **71**: 2172-2182.

Morin, P., Sparks, A., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., Kinzler K. (1997). Activation of β-catenin-Tcf signaling in colon cancer by mutations in β-catenin or APC. *Science.* **275**: 1787–1790.

Mowry, K., Melton, D (1992). Vegetal messenger RNA localization directed by a 340-nt RNA sequence element in Xenopus oocytes. *Science.* **255**: 991–994.

Müeller-Pillasch, F., Lacher, U., Wallrapp, C., Micha, A., Zimmerhackl, F., Hameister, H., Varga, G., Friess, H., Büchler, M., Beger, H., Vila, M., Adler, G., Gress, T. (1997). Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein. *Oncogene.* **14**: 2729-2733.

Neugerbauer, K. (2002). On the importance of being co-transcriptional. *J. Cell. Sci.* **115**: 3865-3871.

Nielsen, F., Nielsen, J., Kristensen, M., Koch, G., Christiansen, J. (2002). Cytoplasmic trafficking of IGF-II mRNA-binding protein by conserved KH domains. *J. Cell Sci.* **115**: 2087-2097.

Nielsen, J., Kristensen, A., Willemoës, M., Nielsen, F., Christiansen, J. (2004). Sequential dimerization of human zipcode-binding protein IMP1 on RNA: a cooperative mechanism providing RNP stability. *Nucleic Acids Res.* **32**(14): 4368-4376.

Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A., Wewer, U., Nielsen, F. (1999). A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. *Mol. Cell. Biol.* **19**(2): 1262-1270.

Nielsen, J., Adolph, S., Rajpert de Meyts, E., Lykke-Andersen, J., Koch, G., Christiansen, J., Nielsen, F. (2003). Nuclear transit of human zipcode-binding protein IMP1. *Biochem J.* **376**: 383-391.

Noubissi, F., Elcheva, I., Bhatia, N., Shakoori, A., Ougolkov, A., Liu, J., Minamoto, T., Ross, J., Fuchs, S., Spiegelman, V. (2006). CRD-BP mediates stabilization of *β-TrCP1* and *c-myc* mRNA in response to β-catenin signalling. *Nature.* **44**: 898-901.

Noubissi, F., Goswami, S., Sanek, N., Kawakami, K., Minamoto, T., Moser, A., Grinblat, Y., Spiegelman, S. (2009). Wnt signaling stimulates transcriptional outcome of the hedgehog pathway by stabilizing GLI1 mRNA. *Cancer Res.* **69**(22): 8572-8578.

Oleynikov, Y., Singer, H. (2003). Real-time visualization of ZBP1 association with β-actin mRNA during transcription and localization. *Curr Biol.* **13**: 199-207.

Ostrem, J., Peters, U., Sos, M., Wells, J., Shokat, K. (2013). K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature.* **503**(7477): 548-551.

Pan, F., Hüttelmaier, S., Singer, R., Gu, W. (2007). ZBP2 facilitates binding of ZBP1 to β-actin mRNA during transcription. *Mol. Cell. Biol.* **27**(23): 8340-8351.

Patel, G., Bag, J. (2006). IMP1 interacts with poly(A)-binding protein (PABP) and the autoregulatory translational control element of PABP-mRNA through the KH III-IV domain. *FEBS J.* **273**: 5678-5690.

Patel, V., Mitra, S., Harris, R., Buxbaum, A., Lionnet, T., Brenowitz, M., Girvin, M., Levy, M., Almo, S., Singer, R., Chao, J. (2012). Spatial arrangement of an RNA zipcode identifies mRNAs under post-transcriptional control. *Genes Dev.* **26**: 43-53.

Patgiri, A., Yadav, K., Arora, P., Bar-Sagi, D. (2011). An orthos- teric inhibitor of the Ras-Sos interaction. *Nat. Chem. Biol.* **7**: 585–587.

Paziewska, A., Wyrwicz, L., Bujnicki, J., Bomsztyk, K., Ostrowski, J. (2004). Cooperative binding of the hnRNP K three KH domains to mRNA targets. *FEBS Lett.* **577**: 134-140.

Poincloux, R., Collin, O., Lizárraga, F., Romao, M., Debray, M., Piel, M., & Chavrier, P. (2011). Contractility of the cell rear drives invasion of breast tumor cells in 3D Matrigel. *Proc. Natl. Acad. Sci. U.S.A.* **108**(5): 1943-1948.

Prokipcak, R., Herrick, D., Ross, J. (1994). Purification and properties of a protein that binds to the C-terminal coding region of human c-myc mRNA. *J. Biol. Chem.* **269**(12): 9261-9269.

Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D (2011). RAS oncogenes: weaving a tumorigenic web. Nat Rev Cancer **11**(11): 761–774.

Rackham, O., Brown, C. (2004). Visualizaton of RNA-protein interactions in living cells: FMRP and IMP1 interact on mRNAs. *EMBO J.* **23**: 3346-3355.

Raineri, I., Wegmueller, D., Gross, B., Certa, U., & Moroni, C. (2004). Roles of AUF1 isoforms, HuR and BRF1 in ARE-dependent mRNA turnover studied by RNA interference. *Nucleic acids res.* **32**(4): 1279-1288.

Rodriguez-Viciana, P., Warne, P. H., Vanhaesebroeck, B., Waterfield, M. D., & Downward, J. (1996). Activation of phosphoinositide 3-kinase by interaction with Ras and by point mutation. EMBO J. **15**(10): 2442-2451.

Ross, A., Oleynikov, Y., Kislauskis, E., Taneja, K., Singer, R. (1997). Characterization of a β-actin mRNA zipcode-binding protein. *Mol. Cell. Biol.* **17**(4) 2158-2165.

Ross, J., Lemm, I., & Berberet, B. (2001). Overexpression of an mRNA-binding protein in human colorectal cancer. *Oncogene.* **20**(45): 6544.

Sasaki, H., Hikosaka, Y., Kawano, O., Moriyama, S., Yano, M., Fuji, Y. (2011). Evaluation of Kras gene mutation and copy number gain in non-small cell lung cancer. *J. Thorac. Oncol.* **6**(1): 15-20.

Shaw, G., Kamen, R. (1986). A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell.* **46**(5): 659-667.

Shaw, A. T., Winslow, M. M., Magendantz, M., Ouyang, C., Dowdle, J., Subramanian, A., ... & Jacks, T. (2011). Selective killing of K-ras mutant cancer cells by small molecule inducers of oxidative stress. Proc. *Natl. Acad. Sci. U.S.A.* **108**(21): 8773-8778.

Shima, F., Yoshikawa, Y., Ye, M., Araki, M., Matsumoto, S., Liao, J., Hu, L., Sugimoto, T., Ijiri, Y., Takeda, A., Nishiyama, Y., Sato, C., Murako, S., Tamura, A., Osoda, T., Tsuda, K., Miyakawa, T., Fukunishi, H., Shimada, J., Kumasaka, T., Yamamoto, M., Katoaka, T. (2013). In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras-effector interaction. *Proc. Natl Acad. Sci. USA.* **110**: 8182-8187.

Shin, K., Bae, S., Hong, H., Kim, R., Kang, M., Park, N. (2011). miR-181a shows tumor suppressive effect against oral squamous cell carcinoma cells by downregulating K-ras. *Biochem. Biophys. Res. Commun.* **404**: 896-902.

Siebel, C., Rio, D. (1990). Regulated splicing of the Drosophila P transposable element third intron in vitro: Somatic repression. *Science.* **248**: 1200–1208.

Sparanese, D., Lee, C. (2007). CRD-BP shields c-myc and MDR-1 RNA from endonucleolytic attack by a mammalian endoribonuclease. *Nucleic Acids Res.* **35**(4): 1209-1221.

Sun, Q., Burke, J., Phan, J., Burns, M., Olejniczak, E., Waterson, A., Lee, T., Rossanese, O., Fesik, S. (2012). Discovery of small molecules that bind to K-Ras and inhibit sos-mediated activation. *Angew. Chem. Int. Edn.* **51**: 6140-6143.

Suzuki, M., Imanishi, S., Dohmae, N., Asanuma, M., Matsumoto, S. (2010). Identification of a male-specific RNA binding protein that regulates sex-specific splicing of Bmdsx by increasing RNA binding activity of BmPSI. *Mol. Cell. Biol.* **30**(24): 5776-5786.

Takahashi, Y., Kitadai, Y., Bucana, C. D., Cleary, K. R., and Ellis, L. M. (1995). Expression of vascular endothelial growth factor and its receptor KDF, correlates with vascularity, metastasis, and proliferation of human colon cancer. *Cancer Res.* **55**: 3964-3968.

Tanic, M., Yanowsky, K., Rodriguez-Antona, C., Andres, R., Marquez-Rodas, I., Osorio, A., Benitez, J., Martines-Delgado, B. (2012). Deregulated miRNAs in hereditary breast cancer revealed a role for miR-30c in regulating KRAS oncogene. *PLos ONE.* **7**(6): e38847.

Tessier, C., Doyle, G., Clark, B., Pitot, H., Ross, J. (2004). Mammary tumor induction in transgenic mice expressing an RNA-binding protein. *Cancer Res.* **64**(1): 209-214.

Tuveson, D., Shaw, A., Willis, N., Silver, D., Jackson, E., Chang, S., Mercer, K., Grochow, R., Hock, H., Crowley, D. et al. (2004). Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. *Cancer Cell.* **5**: 375–387

Valtorta, E., Misale, S., Sartore-Bianchi, A., Nagtegaal., I., Paraf, F., Lauricella, C., Dimartino, V., Hobor, S., Jacobs, B., Ercolani, C., Lamba, S., Scala, E., Veronese, S., Laurent-Puig, P., Siena, S., Tejpar, S., Mottelese, M., Punt, C., Gambacorta, M., Bardelli, A., Nicolantonio, F. (2013). KRAS gene amplification in colorectal cancer and impact on response to EGFR-targeted therapy. *Int. J. Cancer.* **133**: 1259-1266.

Valverde, R., Edwards, L., Regan, L. (2008). Structure and function of KH domains. *FEBS J.* **275**: 2712-2726.

Valverde, R., Pozdnyakova, I., Kajander, T., Venkatraman, J., Regan, L. (2007). Fragile X mental retardation syndrome: structure of the KH1-KH2 domains of fragile X mental retardation protein. *Structure.* **15**: 1090-1098.

Vedadi, M., Niesen, F., Allali-Hassani, A., Fedorov, O., Finerty, P., Wasney, G., Yeung, R., Arrowsmith, C., Ball, L., Berglund, H., Hui, R., Marsden, B., Nordlund, P., Sundstrom, M., Weigelt, J., Edwards, M. (2006). Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *PNAS.* **103**: 15835-15840.

Vikessa, J., Hansen, T., Jonson, L. (2006). RNA-binding IMPs promote cell adhesion and invadopodia formation. *EMBO J.* **25**(7):1456-1468.

Wang, Z., Day, N., Trifillis, P., Kiledjian, M. (1999). An mRNA stability complex functions with Poly(A)-Binding Protein to stabilize mRNA in vitro. *Mol. Cell. Biol.* **19**(7): 4552-4560.

Wang, Q., Zhang, Z., Blackwell, K., Carmichael, G. (2005). Vigilins bind to promiscuously A-to-I-edited RNAs and are involved in the formation of heterochromatin. *Curr. Biol.* **15**: 384–391.

Wang, C., Mackedenski. S., Lee, C. Inhibitors of CRD-BP-KRAS mRNA interaction. RiboWest conference, Edmonton AB (2015)

Weisz, B., Giehl, K., Gana-Weisz, M., Egozi, Y., Ben-Baruch, G., Marciano, D., et al. (1999). A new functional Ras antagonist inhibits human pancreatic tumor growth in nude mice. *Oncogene.* **18**: 2579–2588.

Wilkie, G., Davis, I. (2001). Drosophila wingless and pair-rule transcripts localize apically by dynein-mediated transport of RNA particles. *Cell.* **105**: 209-219.

Yanagihara, M., Ishikawa, S., Naito, M., Nakajima, J., Aburatani, H, Hatakeyama, M. (2005). Paired-like homeoprotein ESXR1 acts as a sequence-specific transcriptional repressor of the human K-ras gene. *Oncogene.* **24**: 5878-5887.

Yang, J., and Weinberg, R.A. (2008). Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. *Dev. Cell.* **14**: 818–829.

Yaniv, K., & Yisraeli, J. K. (2002). The involvement of a conserved family of RNA binding proteins in embryonic development and carcinogenesis. *Gene.* **287**(1): 49-54.
Yeh, C. H., Hung, L. Y., Hsu, C., Le, S. Y., Lee, P. T., Liao, W. L., ... & Tseng, J. T. (2008). RNA-binding Protein HuR Interacts with Thrombomodulin 5′ Untranslated Region and Represses Internal Ribosome Entry Site–mediated Translation under IL-1β Treatment. *Mol. Biol. Cell.* **19**(9): 3812-3822.

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

Yu, S., Lu, Z., Liu, C., Meng, Y., Ma, Y., Zhao, W., Liu, J., Yi, J., Chen, J. (2010). miRNA-96 suppresses KRAS and functions as a tumor suppressor gene in pancreatic cancer. *Cancer Res.* **70**(14): 6015-6025.

Zhang,Y., Nemunaitis,J., Samuel,S.K., Chen,P. and Shen,Y. (2006). Antitumor activity of an oncolytic adenovirus-delivered oncogene small interference RNA. *Cancer Res.* **66**:9736–9743.
Zimmermann, G., Papke, B., Ismail, S., Vartak, N., Chandra, A., Hoffmann, M., et al. (2013). Small molecule inhibition of the KRAS-PDEdelta inter- action impairs oncogenic KRAS signalling. *Nature.* **497**: 638–642.