Molecular Insights into the Coding Region Determinant-Binding Protein-RNA Interaction through Site-Directed Mutagenesis in the Heterogeneous Nuclear Ribonucleoprotein-K-homology Domains*

Mark Barnes[1,4], Gerrit van Rensburg[1,4], Wai-Ming Li[1,4], Kashif Mehmood[1], Sebastian Mackedenski[1], Ching-Man Chan[2], Dustin T. King[1], Andrew L. Miller[2,3], and Chow H. Lee[1]

[1]Chemistry Program, University of Northern British Columbia, 3333 University Way, Prince George, British Columbia V2N 4Z9, Canada.
[2]Division of Life Science and The Key State Laboratory for Molecular Neuroscience, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.
[3]Marine Biological Laboratory, Woods Hole, MA 02543, USA.
[4]These authors contributed equally to this work.

*Running title: Regulation of RNA-binding by KH domains of CRD-BP

To whom correspondence should be addressed: Chow H. Lee, Chemistry Program, University of Northern BC, 3333 University Way, Prince George, BC V2N 4Z9, Canada. Phone: (250) 960-5413; Fax: (250) 960-5845; E-mail: chow.lee@unbc.ca.

Keywords: CRD-BP; RNA binding proteins; mutagenesis; KH domain; zebrafish; granule formation; ribonucleoprotein, RNA-protein interaction; mRNA; molecular biology

Background: Coding Region Determinant – Binding Protein (CRD-BP) interacts physically with oncogenic mRNAs.
Results: Point mutation in the K-homology (KH) domains of CRD-BP abolishes its RNA-binding ability.
Conclusion: Two KH domains of CRD-BP are required for efficient binding to oncogenic mRNAs, and for granule formation in zebrafish embryos.
Significance: Learning how the KH domains interact with mRNAs is crucial for understanding oncogenic function of CRD-BP.

ABSTRACT

The ability of its four heterogeneous nuclear ribonucleoprotein-K-homology (KH) domains to physically associate with oncogenic mRNAs is a major criterion for the function of Coding Region Determinant-Binding Protein (CRD-BP). However, the particular RNA binding role of each of the KH domains remains largely unresolved. Here, we mutated the first glycine to an aspartate in the universally conserved Glycine-X-X-Glycine (GXXG) motif of the KH domain as an approach to investigate their role. Our results show that mutation of a single GXXG motif generally had no effect on binding but the mutation in any two KH domains, with the exception of the combination of KH3 and KH4 domains, completely abrogated RNA-binding *in vitro* and significantly retarded granule formation in zebrafish embryos, suggesting that any combination of at least two KH domains cooperate in tandem to bind RNA efficiently. Interestingly, we found that any single point mutation in one of the four KH domains significantly impacted CRD-BP binding to mRNAs in HeLa cells, suggesting that the dynamics of CRD-BP-mRNA interaction vary over time *in vivo*. Furthermore, our results suggest that different mRNAs bind preferentially to distinct CRD-BP KH domains. The novel insights revealed in this study have important implications on the understanding of the oncogenic mechanism of CRD-BP a well as in the future design of inhibitors against CRD-BP function.

Coding region determinant-binding protein (CRD-BP; mouse), also known as IMP1 (human), belongs to a highly conserved family of RNA-binding proteins called VICKZ (Vg1 RBP/Vera,

IMP-1,2,3, CRD-BP, KOC, ZBP-1) (1,2). Other orthologous or paralogous members include IMP2 (human), IMP3 (human), ZBP1 (chicken), and Vg1-RBP/Vera (*Xenopus*). All members of VICKZ have two N-terminal RNA recognition motifs (RRM) followed by four C-terminal heterogeneous nuclear Ribonucleoproteins-K-homology (KH) domains, and have been implicated in the post-transcriptional regulation of a number of different mRNAs (1,2).

CRD-BP was first discovered due to its ability to physically associate with a specific coding region of c-*myc* mRNA called the coding region determinant (3,4), and its ability to influence c-*myc* mRNA stability in a cell-free model system (5). Several subsequent studies have provided *in-vivo* evidence for the role of CRD-BP in controlling c-*myc* mRNA half-life (6-9). In addition to c-*myc*, CRD-BP also has high affinity for a number of mRNAs whose gene products have been implicated in cancer, which include mRNAs for *βTrCP1* (8), *CD44* (10), *IGF-II* (11), *GLI1* (12), *β-catenin* (13), *K-Ras* (14), *MAPK4* (15), *MITF* (16) and *MDR1* (17). The oncogenic role of CRD-BP is further exemplified by the following two observations: (i) over-expression of CRD-BP in many types of human cancers (2), and (ii) mice genetically engineered to over-produce CRD-BP in mammary glands developed mammary adenocarcinoma (18). Although the exact oncogenic function of CRD-BP is still unclear, many studies using cell lines have demonstrated that its ability to physically interact with target mRNAs is at least one major contributing factor. For instance, it was shown that CRD-BP binds with high affinity to several regions at the 3'untranslated region (UTR) of *CD44* mRNA to stabilize it, leading to cell adhesion, cytoplasmic spreading and invadopodia formation (10). CRD-BP binds to the coding region of *βTrCP1* mRNA and over-expression of CRD-BP led to the stabilization of *βTrCP1* mRNA and elevation of βTrCP1 expression levels in colorectal cancer cells (8). CRD-BP has recently been shown to bind to the coding region of *β–catenin* (13) and *GLI1* mRNAs (12) to stabilize the transcripts, providing further evidence for the role of CRD-BP in the Wnt/ β–catenin signaling pathway. CRD-BP also binds to the coding region and 3' UTR of *K-Ras* mRNA and its overexpression led to increases in K-Ras expression as well as colon cancer cell proliferation (14). Furthermore, the 3'UTR of *MITF* mRNA is also a binding site for CRD-BP and such an interaction is shown to be critical for protecting the *MITF* transcript from degradation by miR-340, a mechanism believed to be important in melanocytes and malignant melanoma (16).

To date, most if not all of our understanding on the structure and function of CRD-BP have been derived from studies on its paralogs. Deletion analysis studies of the human (IMP1), frog (Vg1RBP), and chicken (ZBP1) paralogs of CRD-BP have all concluded that the two RRMs are not important for RNA binding (19-21). However, there are some conflicting results regarding the significance of each KH domain with respect to their interaction with RNA. For instance, while the KH3&4 di-domain of ZBP1 appears to be sufficient to bind to its RNA substrate, the 54-nt zipcode β–actin RNA (21), this was not the case for IMP1 and Vg1RBP. Unlike the full-length and KH1-4 recombinant proteins, the KH3&4 didomain of IMP1 showed no affinity for the 173-nt H19 RNA (20). Similarly, the recombinant protein containing only the KH3&4 di-domain of Vg1RBP exhibited more than a 10-fold decrease in affinity for VLE RNA when compared to the wild-type protein (19). The crystal structure of the KH3&4 di-domain of IMP1 has recently been determined, and it has been proposed that RNA looping induced by IMP1 binding is the mechanism whereby IMP1 provides specific recognition for its RNA substrates (21). Using nuclear magnetic resonance spectroscopy, it was demonstrated that the C-terminal KH3 and KH4 domains can recognize a bipartite RNA sequence, with KH4 binding to a 5' CGGAC element and KH3 binding to a 5' C/A-CA-C/U element (22). In deciphering the role of the RNA-binding domains of CRD-BP paralogs, all the above mentioned studies have utilized isolated domains or employed deletion variants and hence it is still unclear to what extent each KH domain plays in the context of the entire CRD-BP protein in binding to RNA.

The minimal eukaryotic (type 1) KH domain is characterized by a $\beta_1\alpha_1\alpha_2\beta_2$ topology, with the maxi KH domain containing an extra β'α' at its C-terminus. Typically, the KH domain fold consists of three anti-parallel β-strands in the order β_1, β' and β_2 creating the core of the domain. The core β

sheet is framed on three sides by the α-helices, forming the outer face of the domain. The RNA binding platform is located on the outer, solvent exposed face of the $\alpha_1\alpha_2$ region, where the RNA substrate is positioned between two flexible loops (23). Essential to the KH domain RNA binding platform is a strictly conserved GXXG motif, located within a flexible loop between α_1 and α_2 (23). This GXXG motif, in concert with a variable loop located between β_2 and β', functions to clamp an RNA substrate in place upon binding. It is thought that the glycine residues in the GXXG motif are essential due to their conformational flexibility and small steric size (24). Correspondingly, the RNA binding function of the KH domains of NusA, GLD-1, Sam68, and hnRNP K were all severely impeded upon mutating the first glycine in the GXXG motif to an aspartate (24-27). In fact, a recent study showed the feasibility of double mutating the GXXG loop, namely from GXXG to GDDG, in the KH domain of the RNA-binding protein KSRP and the KH3 and KH4 domains of IMP1, as an effective tool for the investigation of the nucleic acid-binding function of individual KH domains (28).

In the present study, we use site-directed mutagenesis to mutate the first Gly residue in each of the GXXG motifs in the KH domains of CRD-BP as an approach to understand to what degree each of the four domains are involved in binding to RNA substrates *in vitro* and in cells. We show that, with the exception of KH3-4, mutations in any two GXXG motifs in the four KH domains leads to a complete abrogation of CRD-BP binding to c-*myc* and *CD44* RNAs *in vitro*. We also show that these *in vitro* RNA-binding profiles correlate with CRD-BP granule formation in intact zebrafish (*Danio rerio*) embryos. However, we find that even a single GXXG point mutation in each KH domain individually can lead to a significant decrease in the amount of c-*myc* and *CD44* mRNAs associated with FLAG-CRD-BP in HeLa cells. Taken together, our results reveal important insights into how CRD-BP physically associates with its RNA binding partners.

EXPERIMENTAL PROCEDURES

Plasmid construction - The plasmids pET28b(+)-CRD-BP and pcDNA-CRD-BP-FLAG containing mouse CRD-BP cDNA were generous gifts from Dr. Jeffrey Ross, University of Wisconsin. The pET28b(+)-CRD-BP was used to generate recombinant WT CRD-BP while the pcDNA-CRD-BP-FLAG was used to over-express FLAG-CRD-BP in HeLa cells. For the generation of various KH point mutation variants, we used the PCR-based site-directed mutagenesis method employing pET28b(+)-CRD-BP and pcDNA-CRD-BP-FLAG as templates. To generate the single KH variants, the following primer pairs were used. KH1: forward primer 5'-GGCGCT ATCATTG**A**CAAGGAGGGTGCC-3', reverse primer 5'-GGCACCCTCCTTG**T**CAATGAT AGCGCC-3'; KH2: forward primer 5'-GGGCGA CTCATTG**A**CAAGGAAGGGCGG-3', reverse primer 5'-CCGCCCTTCCTTG**T**CAATGAG TCGCCC-3'; KH3: forward primer 5'-GGCGCC ATCATTG**A**CAAGAAGGGCCAG-3', reverse primer 5'-CTGGCCCTTCTTG**T**CAATGAT GGC GCC-3'; KH4: forward primer 5'-CCGC GTC ATCG**A**CAAAGGCGGCAAAAC-3', reverse primer 5'-GTTTTGCCGCCTTTG**T**CGATGAC GCGG-3'. Bases altered from the CRD-BP cDNA are shown in bold and the targeted codons are underlined. KH variants containing a point mutation at the GXXG motif in two KH domains were sequentially generated. For instance, to generate the KH1-2 variant, the KH1 plasmid was used as a template and the primer set to point mutate the GXXG motif in the KH2 domain were used in PCR-based site-directed mutagenesis.

For the generation of the WT CRD-BP-EGFP-pSp64TNE plasmid, the open reading frame of CRD-BP was amplified by PCR using the plasmid pcDNA-CRD-BP-FLAG as a template and the following primer pairs: CRD-BP-Kpn1-F (5'-ACCAGGTACCATGAACAAGCTTTACATCG GCA-3') and CRD-BP-Age1-R (5'-ACCAACCGGTAACTTCCTCCGAGCCTGGGC CA-3'). The amplified fragment was then inserted into the *Kpn*I and *Age*I restriction sites of the EGFP-pSp64TNE plasmid. Similarly, for the generation of the EGFP-pSp64TNE KH point mutation variants, the open reading frame of the KH point mutation variants were amplified using the appropriate pcDNA-CRD-BP-FLAG KH variants as templates and CRD-BP-Kpn1-F and CRD-BP-Age1-R primers, and the PCR amplified fragments were inserted into the same sites of EGFP-pSp64TNE as described above. For the generation of truncated CRD-BP in EGFP-pSp64TNE plasmids, the following primer pairs

were used: RRM1&2 (forward primer CRD-BP-Kpn1-F, reverse primer 5'ACCAACCGGTAACCCTGCTGCCACGGGC GACCCTT-3'); KH1 to 4 (CRD-BP-KH1-Kpn1-F 5'ACCAGGTACCATGATCCCTCTCCGGCTCC TGGT-3', reverse primer CRD-BP-Age1-R); KH1 (CRD-BP-KH1-Kpn1-F, CRD-BP-KH1-Age1-R 5'ACCAACCGGTAACAAGATCATCTTGCAC GCGGA-3'); KH2 (CRD-BP-KH2-Kpn1-F 5'CCAGGTACCATGCTGAAGATCCTGGCTC ATAAC-3', CRD-BP-KH2-Age1-R 5' ACCAACCGGTAACTCTCGAACTTTCTTCAT GAT-3'); KH3 (CRD-BP-KH3-Kpn1-F 5'ACCAGGTACCATGGTACAAGTGTTCAT-3', CRD-BP-KH3-Age1-R 5' ACCAACCGGTGTAGCCTCTGGGGGTCCAG T-3'); KH4 (CRD-BP-KH4-Kpn1-F 5' ACCAGGTACCATGAAGCTAGAGACCCACA TACGG-3', CRD-BP-Age1-R); KH1&2 (CRD-BP-KH1-Kpn1-F, CRD-BP-Age1-R); KH2&3 (CRD-BP-KH2-Kpn1-F, CRD-BP-KH3-Age1-R); KH3&4 (forward primer 5'ACCAGGTACCATGGCTCCCTATAGCTCCT TCATGCA-3', CRD-BP-Age1-R). Two sub-cloning steps were used to generate the di-KH domains KH1&3, KH1&4, and KH2&4 EGFP constructs. For generating KH1&3-EGFP-pSp64TNE, KH1 was amplified by PCR and subcloned into the Kpn1 site at the 5' end of KH3-EGFP-pSp64TNE construct. The same fragment was subcloned into the 5'end of KH4-EGFP-pSp64TNE construct to generate KH1&4-EGFP-pSp64TNE. For generating KH2&4-EGFP-pSp64TNE, KH2 was amplified by PCR and subcloned into the 5' end of KH4-EGFP-pSp64TNE.

All of the above generated constructs were verified by DNA sequencing performed by Macrogen (Seoul, Korea).

Generation and purification of recombinant CRD-BP and its variants - Recombinant CRD-BP was purified from *Escherichia coli* BL21 (DE3) using a 1 mL bed volume of nickel-NTA (QIAGEN) column under denaturing conditions. Proteins eluted from the column at pH 5.4 were subjected to three steps of dialysis. The first step was for 24 hours in pH 7.4 buffer containing 200 mM NaCl, 20 mM Tris-HCl, 1 mM reduced glutathione, 0.1 mM oxidized glutathione, 10% (v/v) glycerol, 2 M urea, and 0.01% (v/v) Triton X-100. The protein was then dialyzed twice each for 2 hours in the same buffer as above but without urea and the glutathiones. Following dialysis, samples were centrifuged at 13,200 rpm for 30 min to remove any precipitated proteins. The purified protein solutions were quantified and analyzed for purity using Coomassie brilliant blue-stained 12% SDS-PAGE.

Radiolabelled in vitro transcription - Plasmid pUC19-CRD*myc*-1705-1886 was used to amplify DNA template for use to synthesize internally radiolabeled 182 nts c-*myc* CRD RNA. The PCR primers used to amplify c-*myc* DNA corresponding to nts 1705-1886 are: forward primer, GGATCCTAATACGACTCACTATAGGACCA GATCCCGGAGTTGG; reverse primer, TAGCTGTTCAAGTTTGTG. The T7 RNA promoter sequences are underlined. The plasmid pCYPAC2-CD44 which contains the last CD44 exon was a gift from Dr. Finn C. Nielsen (University of Copenhagen, Denmark), and was used as template for PCR amplification. The PCR primers that were used to amplify the CD44 DNA nts 2862-3055, corresponding to the 3´UTR of *CD44* mRNA are: forward primer, GGATCCTAATACGACTCACTATAGGAAATT AGGGCCCAATTAA; reverse primer, AAATTTCCTCCCAGGGAC. PCR amplified DNA templates were used directly for *in vitro* transcription by T7 RNA polymerase. One µg of DNA template was incubated for 1 hour at 37°C in a 20-µl reaction containing 1 x transcription buffer (Promega, Madison, Wisconsin), 10 mM dithiothreitol, 1 unit RNasin (Promega), 0.5 mM ATP, 0.5 mM CTP, 0.5 mM GTP, 12.5 µM UTP, 1.5 units T7 RNA polymerase (Promega, Madison, WI), and 40 µCi [α-^{32}P] UTP (3000 Ci/mmol). Following incubation, 3 units RNase-free DNase I (Promega) were added and the reaction was further incubated for 10 minutes at 37°C. Upon addition of 10 µl Stopping dye (9 M urea, 0.01% bromophenol blue, 0.01% xylene cyanole FF, 0.01% phenol), the entire sample was electrophoresed on a 8% polyacrylamide/7M urea gel and the band containing internally-radiolabeled RNA was excised and eluted with elution buffer (10 mM Tris-HCl pH 7.5, 0.1M NaCl, 1 mM EDTA, 0.01% SDS) at 45^{0}C for 6 hours. The purified radiolabeled RNA was then phenol/chloroform extracted followed by ethanol

precipitation. Specific activity of the RNA was then determined by scintillation counting.

Electrophoretic mobility shift assay - The electrophoretic mobility shift assay (EMSA) binding buffer (5 mM Tris-Cl pH 7.4, 2.5 mM EDTA pH 8.0, 2 mM DTT, 5% glycerol, 0.1 mg/ml bovine serum albumin, 0.5 mg/ml yeast tRNA, 5 units RNasin) (17) was prepared on ice prior to each experiment. In order to facilitate RNA denaturation and renaturation, the [^{32}P]-labeled RNA sample was heated to 55°C for 5 min and cooled to room temperature before adding the EMSA binding buffer and the appropriate amount of purified recombinant CRD-BP to a final volume of 20-µL. Reactions were incubated at 37°C for 10 minutes and transferred to ice for 5 min. The reaction was again incubated at 37°C for 10 minutes and transferred to ice for an additional 5 min. A total of 2 µL EMSA loading dye (250 mM Tris-Cl pH 7.4, 0.2% bromophenol blue, 0.2% xylene cyanol, 40% sucrose) was added to each reaction and 15 µL of the EMSA reaction was loaded onto a 4% native polyacrylamide gel and resolved at 25 mA for 60 minutes. Following electrophoresis, the gel was exposed overnight at -80°C and subjected to autoradiography using the Cyclone PhosphorImager and Optiquant Software.

EMSA saturation binding experiments were carried out as described above and the dissociation constant (K_d) for the CRD-BP-RNA interaction was determined using the Hill equation. The saturation binding data was analyzed by densitometry of the autoradiograph using the Cyclone Storage Phosphor-System and Optiquant software. For each reaction, the total activity in each lane was determined; this involved summing the total activity in bound complexes with the total activity present in the unbound fraction. The percentage of bound RNA and the protein concentration (nM) were inserted into the Hill equation and the results were expressed graphically.

Cell culture, transfection and immuno-precipitation - HeLa human cervical cancer cells purchased from ATCC were cultured in Minimum Essential Media supplemented with 10% fetal bovine serum at 37°C in 5% CO_2. The day before transfection, 10 x 10^4 cells/mL were plated onto 100 mm dishes. Transient transfection of 10 µg of pcDNA-CRD-BP-FLAG plasmids was carried out using Lipofectamine 2000 reagent (Invitrogen, Burlington, ON) as according to the manufacturer's instructions. As a negative control, cells were transfected with the plasmid vector pcDNA-FLAG. Forty eight hours after transfection, cells were lysed with 1 mL Total Cell Lysis (TCL) buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.1% Triton X-100) supplemented with 1 mM vanadyl ribonucleoside, 0.5 mM DTT, 0.05 units RNasin, and protease inhibitor tablet (Roche Diagnostics, Laval, Quebec). After 5 minutes on ice, lysed cells were aspirated using a 25.5 gauge needle five times to break the nuclei. After a 30 minute incubation on ice, the cell lysate was centrifuged at 14,000 rpm for 10 minutes. The collected supernatant was subjected to preclearing by incubating with 50 µL equilibrated-protein-G agarose beads (50% slurry) at 4°C for at least 1 hour. The resin was spun down at 3000xg for 1 minute and the lysate was collected. The pre-clearing step was repeated once after which the pre-cleared lysate was added to 5 µL anti-FLAG antibody (#F1804, Sigma-Aldrich, Oakville, Ontario) and mixed overnight at 4°C. Protein G-agarose beads were then added to the antibody-lysate mixture and mixed for four hours at 4°C to capture the FLAG antibody. The agarose beads were then washed four times with TCL buffer followed by five washes with TCL buffer containing 1 M urea. Following the fourth wash with TCL buffer containing urea, 50% of the agarose beads were collected and spun down at 3000xg for 1 minute followed by resuspension in 16 µL of water for Western blot analysis. The remaining 50% of the agarose beads was subjected to a final wash followed by resuspension in 100-µL TCL buffer containing 0.3 mg/mL proteinase K and 0.1% SDS. The sample was then incubated at 50°C for 30 minutes. Following the incubation, RNAs physically associated with the resin were extracted by the phenol-chloroform-isoamyl alcohol method and quantified using a Nanodrop Spectrometer (Wilmington, Delaware). One µg of RNA from each treatment group was treated with DNase (DNA-free™ kit, Ambion) before subjecting the sample for cDNA synthesis and quantitative real-time PCR as described below. We also used an equal volume of RNA samples for normalization purposes and found no differences in the quantitative real-time PCR results.

Western blot analysis - Protein samples were separated using a 12.5% polyacrylamide/SDS Lammeli gel system, transferred to a nitrocellulose membrane, and incubated against anti-FLAG antibody (#200472-21, Stratagene, La Jolla, California). PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific, Rockford, Illinois) was used to identify the molecular weight of FLAG-CRD-BP bands. The FLAG-CRD-BP bands were detected using the standard chemiluminescent technique, and visualized using Alpha Innotech FluorChem 5500. The AlphaEaseFC software program was used to assign an Integrated Density Value to each of the FLAG-CRD-BP bands.

Quantitative real-time PCR - The first strand cDNA synthesis was performed using iScript cDNA Synthesis kit (QBio-Rad) on 1 μg of total RNA or equi-volume of RNA samples, and the qPCR was performed using iQ SYBR Green Supermix (Bio-Rad) on an iQ5 Multicolor Real-Time PCR Detection System (Bio-Rad). The PCR primers synthesized by IDT Inc. were: CD44 forward primer, 5'-CAT CAG TCA CAG ACC TGC CCA ATG C-3', and CD44 reverse primer, 5'-ATG TAA CCT CCT GAA GTG CTG CTC C-3'; c-*myc* forward primer, 5'-ACG AAA CTT TGC CCA TAG CA-3', and c-*myc* reverse primer, 5' GCA AGG AGA GCC TTT CAG AG-3'; β-actin forward primer, 5'-TTG CCG ACA GGA TGC AGA AGG A-3', and β-actin reverse primer, 5'-AGG TGG ACA GCG AGG CCA GGA T-3'. The cycling protocol consisted of 95°C for 3 min and 40 cycles of denaturation at 95°C for 10 s, annealing at 52°C for 30 s. To confirm amplification specificity, we performed a melting curve analysis at the end of each cycling. Each sample was run in triplicate, and the data were analyzed using iQ5 optical system software. Serial dilutions were carried out for each total RNA sample and reverse-transcribed under the above-mentioned conditions for each primer set to ensure amplification with efficiencies near 100%. The C_T values for target genes (*CD44*, c-*myc*, and *β-actin*) were then used in the comparative C_T method or commonly known as the $2^{-\Delta\Delta CT}$ method (29) to determine the expression level of the target gene in pcDNA CRD-BP-FLAG-transfected cells (WT and variants) and pcDNA-FLAG-transfected cells. To ensure that there was no genomic DNA contamination in the cDNA samples used for the IP qPCR experiments, no-Reverse Transcriptase (no-RT) controls were used in each of the cDNA synthesis reactions. These samples were subjected to the same RT-qPCR reactions as the test samples and genomic DNA contamination was assessed.

Data collected from four biological replicates were pooled and one-way ANOVA statistical analysis was performed to compare mRNA levels in the variants to the wild-type CRD-BP.

Circular dichroism analysis - Circular dichroism (CD) spectra of all CRD-BP variants were attained using a Jasco J-815 CD spectrometer. Temperature during scanning was maintained at 25°C using a Peltier temperature-control device and proteins sampled in a 0.1 cm path-length quartz cell. CD measurements were recorded using wavelength scans of 195-240 nm, with all protein sample concentrations between 1 and 2 μM resuspended in 20 mM Tris buffer pH 7.4, 10% glycerol, 200 mM NaCl. Measurement parameters were standard resolution 100 mdeg, bandwidth 1 nm, D.I.T. 2 seconds, scan rate 100 nm/min, and accumulation of 8 scans. Data analysis: Secondary structure estimates were de-convoluted using the K2D3 neural networking algorithm (30).

Embryo collection - The AB strain of zebrafish (*Danio rerio*) (from the Zebrafish International Resource Centre; University of Oregon, Eugene, OR, USA) were maintained according to standard laboratory practices and all experiments were approved by the HKUST Institutional Animal Care and Use Committee. Fishes were maintained on a 14 hour light/10 hour dark cycle to stimulate spawning, and their fertilized eggs were collected as described previously (31). Embryos were maintained in 30% Danieau's solution (17.4 mM NaCl, 0.21 mM KCl, 0.18 mM $Ca(NO_3)_2$, 0.12 mM $MgSO_4.7H_2O$, 1.5 mM Hepes, pH 7.2) at ~28.5°C throughout development and during all experiments.

Transient expression of EGFP-CRD-BP, microinjection, microscopy and image acquisition and analysis of zebrafish embryos - For mRNA production, the WT CRD-BP-EGFP and various mutant CRD-BP-EGFP plasmids were linearized and transcribed using an mMESSAGE mMACHINE™ (Ambion) *in vitro* transcription kit. For the transient expression of EGFP-CRD-BP fusion proteins, zebrafish embryos were injected with mRNA at the one-cell stage, unless otherwise

stated. For acquisition of live images, embryos were immobilized in grooves made in a 1% agarose gel and bathed in 30% Danieau's solution (31). Most images were acquired using a Nikon C1 confocal system mounted on a Nikon 90i upright microscope using Nikon Fluor 40x/0.80 W or 60x/1.00W water immersion objectives and a pinhole size of 60 µm. Alternatively, some images were acquired using a Carl Zeiss LSM 780 system configured on an inverted Observer Z1 microscope and a 40x oil immersion objective. All images shown are single confocal sections of 512 x 512 pixels. Subsequent color and contrast enhancement was done using Corel PHOTOPAINT 11. Granules were quantified in all images using the "Analyze Particles" function in ImageJ software with a filter to detect granule size of 6-200 pixels^2 and circularity of 0.40-1.00. The number of granule is expressed per cell and the number of cells in each image was counted manually.

RESULTS

Point mutations in the GXXG motif of two KH domains generally abolish the ability of CRD-BP to bind c-myc and CD44 RNAs in vitro - To assess the importance of each KH domain in binding to RNA in the context of the full-length protein, we generated single point mutations within the GXXG motif of each individual KH domain. We mutated the first glycine of each GXXG motif to an aspartate as a way to remove the function of each individual KH domain (Fig. 1A). For simplicity, we have used the following abbreviation to name the point mutations as KH variants. For instance, KH1 variant has the point mutation in the GXXG motif within the KH1 domain whereas the KH2 variant has the point mutation in the GXXG motif within the KH2 domain. The KH1-2 variant has the point mutations at the GXXG motifs of both the KH1 and KH2 domains, whereas, the KH3-4 variant has the point mutations in the GXXG motifs of the KH3 and KH4 domains, and so on (Fig. 1A). As negative controls, we also generated four CRD-BP variants with point mutations that are not expected to affect the RNA-binding ability. CRD-BP variants with point mutations in the RRM1 (Y5A) or RRM2 domain (Q84A) were selected (Fig.1A). The E445D and D526E variants containing a point mutation in the variable loop located between β_2 and β_3 in KH3 and KH4 domains (21) respectively, were also generated (Fig. 1A). The recombinant mouse wild-type (WT) CRD-BP and mutant proteins were purified to ~95% homogeneity (Fig. 1B).

Using electrophoretic mobility gel-shift assays, we compared the binding profiles of the WT CRD-BP and its variants on two high affinity target RNAs: the 182 nts ^{32}P-labeled c-*myc* RNA (nts 1705-1886) (Fig. 2) and 194 nts ^{32}P-labeled CD44 RNA (nts 2862-3055) (10, 32) (Fig. 3). Fig. 2A shows representative results from gel-shift assays for binding the 182 nts c-*myc* RNA. It is clear that the KH1-2, KH1-3, KH1-4, KH2-3, and KH2-4 variants had completely abolished ability to bind the c-*myc* RNA. In contrast, KH3-4 and the single point mutation variants KH1, KH2, KH3, and KH4, all exhibited binding to c-*myc* RNA (Fig. 2A). Results from three biological replicates were pooled and the saturation-binding data were fitted to the Hill equation and expressed graphically to determine the dissociation constant (K_d) and Hill coefficient. A summary, showing K_d for the interaction of each of the recombinant protein variants with c-*myc* RNA is shown in Table 1, and the results are plotted for direct comparison in Fig. 2B. As shown in the quantitative data, KH3-4 and KH1 displayed slightly higher K_d values as compared to the WT CRD-BP, whereas KH2, KH3, KH4, Y5A, Q84A, D526E, and E445D, all displayed binding profiles that are comparable to the WT CRD-BP. On the other hand, KH1-2, KH1-3, KH1-4, KH2-3 and KH2-4 displayed significantly reduced or no binding to c-*myc* RNA based on data from three biological replicates (Fig. 2B).

Fig. 3A shows representative results from gel-shift assays for binding of the CRD-BP variants to the 194 nts *CD44* RNA. From the results, it is clear that the KH1-2, KH1-3, KH1-4, KH2-3, and KH2-4 mutants are unable to bind *CD44* RNA. In contrast, the Y5A, D526E, KH1, KH2, KH3, and KH3-4 mutants all displayed the ability to bind *CD44* RNA. The slightly different shift in the complex for KH3-4 as compared to the WT was because the two groups of samples were electrophoresed on different gels. KH4 appeared to bind *CD44* RNA very weakly (Fig. 3A), and hence the saturation-binding curve as well as the K_d could not be obtained. A summary of the K_d values for the WT and mutant CRD-BP proteins that were obtained from the saturation-binding

data is shown in Table 1 and plotted for direct comparison in Fig. 3B. Both the Y5A and D526E mutants displayed slightly lower K_d values indicating their higher affinity for *CD44* RNA while KH1, KH2 and KH3 had slightly higher K_d values indicating their lower affinity for *CD44* RNA. Results from three biological replicates confirmed that the KH4, KH1-2, KH1-3, KH1-4, KH2-3 and KH2-4 mutants have significantly reduced or no binding to CRD-BP (Fig. 3B).

In summary, we found that with the exception of KH3-4, simultaneous point mutation of the GXXG motif in two KH domains generally abolishes the ability of CRD-BP to bind c-*myc* and *CD44* RNAs. We also found that while the KH4 variant has high affinity for c-*myc* RNA, it was very inefficient in binding *CD44* RNA. Similarly, the KH2 and KH3 variants have higher K_d values than that of the WT CRD-BP in binding *CD44* RNA, but not for binding c-*myc* RNA.

No global secondary structure changes occur in CRD-BP upon point mutation of the KH domain GXXG motif - To determine if there were major global changes in the secondary structure upon point mutation at the GXXG motifs, we performed circular dichroism spectroscopy on the WT CRD-BP as well as the KH domain mutants in the wavelength range of 195-240 nm. The 'w-shaped' spectrum with a clear trough close to 208 nm and a less distinctive trough near 222 nm are indicative of the presence of α–helical structure. The CD spectrum of the entire CRD-BP protein is remarkably similar to that generated for the recombinant KH3&4 di-domain of ZBP1 (22). All of the point mutation variants exhibited CD spectra that are very similar, if not identical, to that of the WT CRD-BP (data not shown). We also estimated the percentages of α–helix and β–sheet present in the WT CRD-BP and the KH variants and found no significant differences (data not shown). Hence, we conclude that point mutations in the GXXG motifs of the KH domains do not alter the global secondary structure of CRD-BP.

CRD-BP granule formation in zebrafish embryos is dependent on newly transcribed zygotic RNAs - Next, we were interested in determining whether our findings that the RNA-binding ability of CRD-BP is abrogated upon introducing the GXXG motif point mutations in any two KH domains were translatable to an *in vivo* cellular system. Since the ability of IMP1 to bind RNA has been shown to be a pre-requisite for the formation of IMP granules in NIH 3T3 cells (20), we decided to use CRD-BP granule formation as our first cellular system. We turned to the zebrafish embryo as an *in vivo* model because the optical clarity of this system is microscopically superior and thus facilitates the rapid identification of CRD-BP granules. Furthermore, CRD-BP granule formation in zebrafish embryos is likely to be visible within a few hours upon translation, allowing visualization of the protein-RNA interaction at early time points which is not possible using a regular mammalian transfection system.

We first performed characterization experiments to explore and understand mouse CRD-BP granule formation in intact zebrafish embryos. We found that during the first 3 hours in the absence of zygotic bio-molecules, CRD-BP granules were absent (Fig. 4Ai). However, at 4 hours post-fertilization (hpf), CRD-BP granules were clearly visible (Fig. 4Aii), suggesting that CRD-BP granule formation strictly coincides with the mid-blastula transition that indicates a switch from maternal to zygotic transformation. We also assessed granule formation upon injection with the human IMP1-EGFP mRNA. As shown in Figs. 4Aii and iii, the granule formation by the mouse CRD-BP is remarkably similar to the human IMP1, indicating that granule formation by the two orthologs is very similar in zebrafish embryos. To determine if CRD-BP granule formation is dependent on newly transcribed RNA, we injected the transcriptional inhibitor Actinomycin D immediately after CRD-BP-EGFP mRNA injection. We found that Actinomycin D completely abolished the formation of CRD-BP granules (Fig. 4Bii), confirming that expression of zygotic genes is required for CRD-BP granule formation. To determine whether zygotic RNAs or proteins are important for granule formation, we injected the protein synthesis inhibitor cycloheximide either immediately after injection of CRD-BP-EGFP mRNA to inhibit total protein synthesis, or at the 16-cell stage to allow expression of CRD-BP-EGFP before inhibition of zygotic protein synthesis. We found that the expression of CRD-BP-EGFP was inhibited by cycloheximide treatment if given immediately after mRNA injection (Fig. 4Biii). However, if cycloheximide was given at a later stage to block

translation of the earliest zygotic transcripts, CRD-BP granules were still present (Fig. 4Biv), suggesting that zygotic RNAs rather than proteins are critical for the formation of CRD-BP granules.

The KH3&4 di-domain is sufficient for CRD-BP granule formation in zebrafish embryos - To determine if the requirement for CRD-BP granule formation in zebrafish embryos is similar to that previously observed in mouse NIH3T3 embryo fibroblast cells (20), we first assessed the deletion variants of CRD-BP-EGFP. A schematic diagram showing all deletion variants is illustrated in Fig. 5A and representative results of CRD-BP granule formation are shown in Fig. 5B. Injection of RRM1&2-EGFP or KH1&2-EGFP mRNAs failed to result in any granule formation (Fig. 5B and C) which is consistent with previous results in NIH3T3 cells (20) and in several studies which concluded that the RRM domains and KH1&2 domains alone are insufficient for interaction with target RNAs (19-21). Fig. 5B shows that KH1 to 4-EGFP granule formation is comparable to that of the WT CRD-BP-EGFP, and this is further confirmed by quantifying the number of granules per cell as shown in Fig. 5C. This is in good agreement with the previous finding that KH domains 1-4 is sufficient to assemble granules (20) and bind target RNAs (19, 20). Injection of Di-KH3&4-EGFP mRNA, but not KH4-EGFP mRNA, also resulted in granule formation (Fig. 5B). Surprisingly, we found a significantly higher number of granules per cell in Di-KH3&4-EGFP-expressed embryos as compared to embryos expressing the WT CRD-BP-EGFP (Fig. 5C). This result is consistent with findings that KH domains 3-4 can bind target RNA efficiently (21). However, a previous study showed no granule formation in NIH3T3 cells upon expressing GFP-KH3-4, and KH3-4 of IMP1 could not bind H19 RNA (20). The deviation in granule formation results could be due to one or a combination of the following: (i) cellular differences between zebrafish embryos and NIH3T3 cells, (ii) time differences and hence differences in the dynamics of CRD-BP-RNA interaction, or (iii) granules are less distinguishable in NIH3T3 cells. The deviation in RNA binding ability of KH3-4 is most likely due to differences in RNA targets used in the different studies, as we have observed the different contribution of each KH domain of CRD-BP in binding different RNA molecules (Figs. 2 and 3; unpublished observation). There was an accumulation of Di-KH3&4-EGFP in the nuclei of zebrafish embryos, which was not seen with KH1 to 4-EGFP (Fig. 5B), suggesting that there may be a nuclear export signal located within KH domains 1-2. In support of this hypothesis was the observation of a similar nuclear accumulation of KH4-EGFP (Fig. 5B). Indeed, Nielsen and colleagues have reported that there are IMP1 nuclear export signals located in KH2 and KH4 (33). We also examined the expression of other single KH domains, namely KH1-EGFP, KH2-EGFP and KH3-EGFP, but again no granule formation in zebrafish embryos was observed (Fig. 5B and C). We also noted that KH1-EGFP accumulated in the nuclei while KH3-EGFP was accumulated in the cytoplasm and KH2-EGFP appeared to be evenly distributed throughout the nuclei and cytoplasm (Fig. 5B). However, we cannot rule out possible differences in cell cycle when these images were acquired. Finally, to determine if the granule formation by a combination of two KH domains is unique to di-KH3&4-EGFP, we examined all other possible di-domain combinations. Fig. 5B and C show that di-KH1&3-EGFP, di-KH1&4-EGFP, di-KH2&3-EGFP, and di-KH2&4-EGFP did not show any distinct granule formation although some of these di-domains appeared to be more patchy (Fig. 5B). Hence, we conclude that KH3&4 di-domain is the minimal region of CRD-BP necessary for granule formation.

Point mutations in the GXXG motif in two KH domains generally reduces CRD-BP granule formation in zebrafish embryos - Having shown that CRD-BP granule formation in zebrafish embryos indeed correlated with the known RNA-binding ability amongst CRD-BP deletion variants, we next investigated the role of each KH domain in CRD-BP granule formation in zebrafish embryos in the context of the full length mutant proteins. To do this, we generated pSp64TNE-CRD-BP plasmid constructs containing the point mutation in the GXXG motif of each of the individual KH domain of CRD-BP. The KH domain point mutation variants generated were those used in the EMSA studies described earlier (Figs. 2 and 3). As shown in Fig. 6A, granule formation for the single point mutants was indistinguishable from that of the WT CRD-BP. Results corresponding to CRD-BP granule

formation for the KH domain double mutants are shown in Fig. 6A and the quantitative data are shown in Fig. 6B. Interestingly, KH3-4 was the only double mutant that exhibited comparable granule formation to the WT CRD-BP (Fig. 6A and B). Granule formation was significantly reduced when zebrafish embryos were injected with KH1-3-EGFP, KH1-4-EGFP, KH2-3-EGFP or KH2-4-EGFP mRNAs (Fig. 6A and B). Granule formation with KH1-2 was also reduced although did not reach statistically significant levels (Fig. 6A and B). In summary, we observed that the CRD-BP granule formation in zebrafish embryos by the KH point mutation variants correlated remarkably with their ability to bind c-*myc* and *CD44* RNA *in vitro* (Table 1).

A point mutation at the GXXG motif in KH domains generally reduces the ability of CRD-BP to associate with c-myc and CD44 mRNAs in HeLa cells - Next, we were interested in determining whether the point mutations in the GXXG motif of the KH domains had any effect on the ability of CRD-BP to physically interact with *CD44* and c-*myc* mRNAs in mammalian cells. We transfected HeLa cells with pcDNA-CRD-BP-FLAG plasmids containing the point mutations as described in Fig. 1A. The pcDNA-FLAG and pcDNA-CRD-BP-FLAG plasmids containing Y5A, D526E, Q84A or E445D mutations were used as negative controls. No genomic DNA contamination was observed in the immuno-precipitated RNA samples. This was indicated by a lack of PCR product amplified in the no-RT qPCR reactions (data not shown).

As shown in Fig. 7A, with the exception of E445D and Q84A variants, equal amounts of FLAG-CRD-BP were immuno-precipitated from cells transfected with different CRD-BP variants with single or double point mutations. All expressed FLAG-CRD-BP at approximately equal amounts based on similar volumes of cell lysate used for immunoprecipitation. The equal Integrated Density Values obtained from quantification using the AlphaEaseFC software program confirmed this observation. The reason for the lack of FLAG-CRD-BP detected in the Q84A and E445D variants is unknown, but could possibly be the result of a masked FLAG-tag as a consequence of structural changes. Using equal amounts of immuno-precipitated proteins, we measured *CD44*, c-*myc* and *β–actin* mRNAs that were physically associated with FLAG-CRD-BP (Fig. 7B). As shown, the level of all three transcripts associated with FLAG-CRD-BP in cells transfected with Y5A and D526E variants were similar to those in WT CRD-BP. In contrast, the level of *CD44* mRNA associated with FLAG-CRD-BP was significantly reduced in cells transfected with any of the KH domain point mutation variants. Similarly, with the exception of the KH4 variant, the amount of c-*myc* mRNA associated with FLAG-CRD-BP was also significantly reduced in cells transfected with KH variants. On the contrary, the level of β–actin mRNA physically associated with FLAG-CRD-BP in the WT CRD-BP- and KH mutant-transfected cells were very similar except for the KH3, KH1-4, and KH2-4 variants. In summary, we conclude that the GXXG mutants generally retarded the ability of CRD-BP to bind CD44 and c-*myc* mRNAs in HeLa cells. The results also suggested that as observed *in vitro*, there is a differential contribution of the CRD-BP KH domains to binding its various mRNA targets in HeLa cells.

DISCUSSION

A number of animal studies have now confirmed the critical role played by CRD-BP during embryogenesis and tumorigenesis (18,34-37), but unfortunately the exact molecular mechanism of this onco-fetal protein is still unclear. However, cumulative evidence from studies *in vitro* and in cell lines suggests that the ability to physically interact with a subset of mRNAs play a significant part in the function of CRD-BP and its orthologs. Hence, it is of the utmost importance to clearly understand how CRD-BP interacts with its RNA substrates. To date, *in vitro* and granule formation studies in mammalian cells using CRD-BP orthologs have shown that the KH domains, and not the RRM domains, are directly involved in binding RNA substrates (19,20). However, the extent to which each of the KH domains of CRD-BP contributes to the physical interaction with RNA is less clear. This is partly because most of the studies were based on deletion analysis. The RNA binding function of the KH domains for a number proteins have been shown to be severely impeded upon mutating the first glycine in the GXXG motif to an aspartate (24-27). Using a similar approach, we

investigated the role of each of the KH domains in the context of the full-length protein in binding c-*myc* and *CD44* RNAs *in vitro* and in cells, in addition to granule formation in intact zebrafish embryos.

Using electrophoretic mobility gel-shift assays, we found that a single point mutation in the GXXG motif at each of the four KH domains generally had no impact on the ability of CRD-BP to bind c-*myc* and *CD44* RNAs (Figs. 2 and 3). The only exception to this was an inefficient binding of the KH4 variant to *CD44* RNA. However, point mutations in the GXXG motif of any combination of two KH domains, with the exception of the simultaneous mutation at both KH3 and KH4 domains (KH3-4), resulted in complete abrogation of RNA binding (Figs. 2 and 3). CD spectral analysis showed no global secondary structural changes for the mutant proteins. This is also consistent with a recent finding that the GXXG-to-GDDG mutation in the four KH domains of KSRP as well as in the KH3 and KH4 domains of ZBP1 did not destabilize the KH domains (28). However, we do not rule out the possibility of local structural changes in the KH domains of the point mutants, which cannot be detected using CD spectroscopy. This may explain why the KH4 single mutant has significantly reduced binding while the KH3-4 double mutant binds *CD44* RNA efficiently. It is possible that for this specific double mutant, local structural changes at regions outside the GXXG motif within KH3 and KH4 facilitate binding to *CD44* RNA. Further structural experiments that are beyond the scope of this study are required to resolve this apparent anomaly.

The ability of IMP1 to bind RNA has been considered a pre-requisite for the formation of IMP1 granules in mammalian cells (20). The IMP1 granules are ribonucleoprotein complex that are believed to form leading to functional IMP1 involved in mRNA stability (9) and granule localization/transport (20). The granules collected from HEK293 human embryonic kidney cells 48 hours after transfection with plasmid carrying FLAG-IMP1 are distinct from stress granules and P bodies, and are enriched in mRNAs encoding proteins involved in endoplasmic reticulum and Golgi function (38). Therefore, we investigated whether our *in vitro* EMSA data with the KH point mutants correlates with CRD-BP granule formation by the variants. Here, we provide the first description of CRD-BP granule formation in intact zebrafish embryos and demonstrate that CRD-BP granule formation is dependent on zygotic mRNAs and not maternal mRNAs (Fig. 4). The zygotic mRNA dependence of CRD-BP granule formation is in-line with the onco-fetal role of this protein, which is overexpressed in neonatal tissues (39). Also, we find that the ability of the various mutant proteins to bind RNA *in vitro* is a good indication for CRD-BP granule formation as demonstrated using intact zebrafish embryos. Truncated CRD-BP-EGFP, such as RRM1&2 and KH1&2, which have been shown to have no ability to bind RNA (19,20), could not form granules in zebrafish embryos. However, KH1to4 CRD-BP-EGFP which binds RNA *in vitro* (19,20), resulted in granule formation that was comparable to the WT CRD-BP (Fig. 5). We found a significantly greater number of granules per cell in embryos expressing the truncated KH3&4 CRD-BP-EGFP as compared to the WT CRD-BP (Fig. 5C), which is consistent with its ability to bind RNA efficiently as previously reported (21). Such an effect is unique to the KH3&4 di-domains because all of the other possible KH di-domain combinations did not show granule formation (Fig. 5). On examining the KH point mutants for granule formation, we found a remarkable correlation between the ability to bind c-*myc* and *CD44* RNAs *in vitro* and CRD-BP granule formation in zebrafish embryos. For instance, the single point mutants generally did not display reduced granule formation or RNA-binding (Table 2). In contrast, with the exception of KH3-4, the KH domain double mutants displayed a significant reduction in both granule formation and RNA-binding. Based on these observations, we propose that at least two KH domains of CRD-BP bind to RNA in tandem. This is consistent with a recent x-ray crystallography study, whereby the individual domains in the KH3&4 didomain were shown to be arranged in an intramolecular anti-parallel pseudodimer conformation, which requires that the bound RNA must loop to contact both KH domains simultaneously (21). Also, in agreement with the previous study is the observation that the KH3&4 di-domain is important for binding RNA, and that sites other than the GXXG motif within the KH3 and KH4 domains play a critical role in binding

RNA substrates. This is because we found that a point mutation in the GXXG motif in both the KH3 and KH4 domains in combination, had no effect on *CD44* and c-*myc* RNA-binding as well as granule formation.

However, based on our point mutation studies, we further observe that KH1 and KH2 also play a critical role in binding RNA substrates. Our study revealed that KH1 and KH2 domains also participate with at least another KH domain in binding RNA. The role for KH1 and KH2 in RNA-binding is also supported by two earlier indirect observations: (i) unlike the full-length protein, KH3&4 do not associate with RNA below 100 nM (11), and (ii) the KH1&2 di-domain modulates binding of IMP1 to *β–actin* and c-*myc* RNA *in vitro* (2). Furthermore, using a filter-binding assay, significant reduction in RNA-binding was only observed upon GXXG to GEEG mutation in all four KH domains of ZBP1 (40). Although there is generally a similar trend in binding c-*myc* and *CD44* RNAs by the KH variants, there are a few notable deviations. While the KH4 variant binds c-*myc* RNA with a K_d similar to that of WT CRD-BP, it did not bind *CD44* RNA efficiently and hence no K_d can be calculated. Similarly, the KH2 and KH3 variants bind to *CD44* RNA less efficiently than to c-*myc* RNA. These results suggest that indeed CRD-BP utilizes distinct KH domains when binding to different RNA substrates.

The investigation into the physical interaction between CRD-BP mutants with c-*myc*, *CD44* and *β-actin* mRNAs using immuno-precipitation (IP) coupled with the qPCR method, yielded somewhat surprising results. Consistent with the *in vitro* EMSA and CRD-BP granule formation results, significant reduction in both c-*myc* and *CD44* mRNAs, but not the β–actin mRNA, were found associated with the KH domain double mutants. Unexpectedly, we found that the KH3-4 variant also had a significantly reduced amount of *CD44* and c-*myc* mRNAs physically associated with it. Similarly, a single point mutation in just one KH domain, significantly reduced the level of *CD44* and c-*myc* mRNAs binding to it, with the exception of the KH4 variant, which was still able to bind c-*myc* mRNA. We reasoned that the biochemical method employed in this study is valid because Y5A and D526E variants which bind RNA as effectively as the WT CRD-BP *in vitro* (Figs. 2 and 3) had levels of c-*myc* and *CD44* mRNAs comparable to that of WT CRD-BP in the immuno-precipitation experiments. Furthermore, except for KH1-4, KH2-4 and KH3 variants, there was no significant reduction in the amount of *β-actin* mRNA associated with all the other KH variants. The surprising results with KH3-4 and the single point mutants which were different from the *in vitro* EMSA and granule formation results may be potentially explained by the following: CRD-BP is known to associate with various RNA-binding proteins and mRNAs in ribonucleoprotein complexes (9,38). Hence, it is not unexpected to find differences in the dynamics of CRD-BP-RNA interactions 48 hours after transfection as performed in the HeLa cells experiments as opposed to the direct *in vitro* EMSA and in the granule formation experiments in zebrafish which occurred 4 hours after injection of mRNAs (Table 2). Such a proposal is supported by an earlier report of the lack of granule formation in NIHT3 cells 48 hours after transfection with the KH3&4-GFP didomain (20). The immuno-precipitation experiments also revealed the following novel findings: (i) different mRNAs bind to CRD-BP mutants variably in cells as exemplified by the interaction of c-*myc* mRNA with the KH4 variant and *β–actin* mRNA binding to the KH3, KH1-4 and KH2-4 variants, and (ii) the individual KH domains play an important role in the physical interaction between CRD-BP and specific mRNA in cells. Indeed, our data supports an earlier proposal that individual KH domains may have distinct functions. This was based on multiple sequence alignment analyses of CRD-BP and its orthologs, which indicates a higher degree of amino acid sequence similarity between the equivalent KH domains from different proteins than between KH domains within the same protein (19).

Understanding the RNA sequences and/or structures that are recognized by CRD-BP is important to uncovering its cellular function. Consensus primary RNA sequences have been identified using various methods, and were proposed to be targets of IMP1 or ZBP1. Using the SELEX method, the consensus sequence 5'-ACACCC-3', was found in RNAs bound to the full-length ZBP1 as well as to the truncated KH3-4 domains of ZBP1 (41). Using the immuno-precipitation method coupled with microarray

analysis, it was found that 307 transcripts associated with IMP1 granules were enriched in CCYHHCC (where Y = C or U and H = C, U or A) motif (38). More recently, the PAR-CLIP method was used to assess RNA-binding protein recognition elements for FLAG/HA-tagged IMP1, IMP2 and IMP3 (42). An RNA recognition element CAUH (H = A, U or C) was found in more than 75% of the top 1000 clusters from 100,000 sequence clusters recognized by the IMPs (43). Interestingly, our initial analysis of the secondary structure of the regions of c-*myc* (43) and *CD44* (44) RNA that bind with high affinity to CRD-BP showed that sequences located in single-stranded regions have some potential overlap with the previously reported consensus sequences recognized by IMP1/ZBP1. For instance, 5'-AAACA-3' at nts 20-24, 5'-ACAG-3' at nts 63-66, and 5'-ACA-3' at nts 70-72 are located in single-stranded regions of c-*myc* RNA and contain a partial match to the consensus sequence ACACCC. Similarly, 5'-CCCAAUU-3' is located at nts 11-17 in a single-stranded region of *CD44* RNA and is close to matching the consensus sequence CAUH. This shows that there is no strict 'consensus' sequence for binding RNA and suggests that secondary and tertiary structural elements are likely key for interaction. Therefore, further *in-silico* analysis of structural elements within target RNAs and crystallographic studies on the CRD-BP-RNA complex are required to fully elucidate the factors governing molecular recognition.

Using site-directed mutagenesis in the context of the entire protein, this study shows for the first time that at least two KH domains of CRD-BP act in tandem for efficient binding to two oncogenic mRNAs, c-*myc* and CD44, as well as for early RNP granule formation in zebrafish embryos. These results show that *in vitro* EMSA is indeed a valid method for assessing specific CRD-BP-RNA interactions that is translatable to an *in vivo* system, as shown for RNP granule formation in zebrafish embryos. This study supports the notion that mutating the GXXG motif within the KH domains (28) is a powerful tool for investigation into the function of individual KH domains in RNA-binding proteins. We show that mutating the first glycine to an aspartate in the GXXG motif is sufficient for such purposes. We also show for the first time that each KH domains is required for binding c-*myc* and *CD44* mRNAs at the later stage of RNP formation as observed in the HeLa cells. Based on our overall data, we propose that the dynamics of the CRD-BP-mRNA interaction vary over time *in-vivo*. This study has important implications for our understanding of the oncogenic mechanism of CRD-BP as well as in the future design of inhibitors targeted against CRD-BP function.

REFERENCES

1. Yisraeli, J. K. (2005) VICKZ proteins: a multi-talented family or regulatory RNA-binding proteins. *Biol. Cell* **97,** 87-96.
2. Bell, J. L., Wachter, K., Muhleck, B., Pazaitis, N., Kohn, M., Lederer, M., and Huttelmaier, S. (2013) Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell. Mol. Life Sci*. **70,** 2657-2675.
3. Prokipcak, R. D., Herrick, D. J., and 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,** 9261-9269.
4. Doyle, G.A.R., Betz, N. A., Leeds, P. L., Fleisig, A. J., Prokipcak, R. D., and 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.
5. Bernstein, P. L., Herrick, D. J., Prokipcak, R. D., and Ross, J. (1992) Control of c-*myc* mRNA half-life *in vitro* by a protein capable of binding to a coding region determinant. *Genes Dev.* **6,** 642-654.

6. Coulis, C. M., Lee, C., Nardone, V., and Prokipcak, R. D. (2000) Inhibition of c-*myc* expression in cells by targeting an RNA-protein interaction using antisense oligonucleotides. *Mol. Pharmacol.* **57,** 485-494.
7. Ioannidis, P., Mahaira, L. G., Perez, S. A., Gritzapis, A. D., Sotiropoulou, P. A., Kavalakis, G. J., Antsaklis, A. I., Baxevanis, C. N., and Papamichail, M. (2005) CRD-BP/IMP1 expression characterizes cord blood CD34+ stem cells and affects c-*myc* and IGF-II expression in MCF-7 cancer cells. *J. Biol. Chem.* **280,** 20086-20093.
8. Noubissi, F. K., Elcheva, I., Bhatia, N., Shakoori, A., Ougolkov, A., Liu, J., Minamoto, T., Ross, J., Fuchs, S. Y., and Spiegelman, V. S. (2006) CRD-BP-mediates stabilization of *βTrCP1* and c-*myc* mRNA in response to β–catenin signaling. *Nature* **441,** 898-901.
9. Weidensdorfer, D., Stohr, N., Baude, A., Lederer, M., Kohn, M., Schierhorn, A., Buchmeier, S., Wahle, E., and Huttelmaier, S. (2009) Control of c-*myc* mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA* **15,** 104-115.
10. Vikesaa, J., Hansen, T. V. O., Jonson, L., Borup, R., Wewer, U. M., Christiansen, J., and Nielsen, F. C. (2006) RNA-binding IMPs promote cell adhesion and invadopodia formation. *EMBO J* **25,** 1456-1468.
11. Nielsen, J., Kristensen, M. A., Willemoes, M., Nielsen, F. C., and Christiansen, J. (2004) Sequential dimerization of human zipcode-binding protein IMP1 on RNA: a cooperative mechanism providing RNP stability. *Nucleic Acids Res.* **32,** 4368-4376.11.
12. Noubissi, F. K., Goswami, S., Sanek, N. A., Kawakami, K., Minamoto, T., Moser, A., Grinblat, Y., and Spiegelman, V. S. (2009) Wnt signaling stimulates transcriptional outcome of the hedgehog pathway by stabilizing GLI1 mRNA. *Cancer Res.* **69,** 8572-8578.
13. Gu, W., Wells, A. L., Pan, F., and Singer, R. H. (2008) Feedback regulation between zipcode binding protein 1 and β–catenin mRNAs in breast cancer cells. *Mol. Cell. Biol.* **28,** 4963-4974.
14. Mongroo, P. S., Noubissi, F. K., Cuatrecasas, M., Kalabis, J., King, C. E., Johnstone, C. N., Bowser, M. J., Castells, A., Spiegelman, V. S., and Rustgi, A. K. (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.
15. Stohr, N., Kohn, M., Lederer, M., Glab, M., Reinke, C., Singer, R. H., and Huttelmaier, S. (2012) IGF2BP1 promotes cell migration by regulating MK5 and PTEN signaling. *Genes Dev.* **26,** 176-189.
16. Goswami, S., Tarapore, R. S., TeSlaa, J. J., Grinblat, Y., Setaluri, V., and Spiegelman, V. 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,** 20532-20540.
17. Sparanese, D., and Lee, C. H. (2007) CRD-BP shields c-*myc* and MDR-1 RNA from endonucleolytic attack by a mammalian endoribonuclease. *Nucleic Acids Res.* **35,** 1209-1221.
18. Tessier, C. R., Doyle, G. A., Clark, B.A., Pitot, H.C., and Ross, J. (2004) Mammary tumor induction in transgenic mice expressing an RNA-binding protein. *Cancer Res.* **64,** 209-214.15.
19. Git, A., and Standart, N. (2002) The KH domains of xenopus Vg1RBP mediate RNA binding and self-association. *RNA* **8,** 1319-1333.
20. Nielsen, F. C., Nielsen, J., Kristensen, M. A., Koch, G., and Christiansen, J. (2002) Cytoplasmic trafficking of IGF-II mRNA binding protein by conserved KH domains. *J. Cell Sci.* **115,** 2087-2097.
21. Chao, J. A., Patskovsky, Y., Patel, V., Levy, M., Almo, S. C., and Singer, R. H. (2010) ZBP1 recognition of β–actin zipcode induces RNA looping. *Genes Dev.* **24,** 148-158.
22. Patel, V.L., Mitra, S., Harris, R., Buxbaum, A. R., Lionnet, T., Brenowitz, M., Girvin, M., Levy, M., Almo, S. C., Singer, R. H., and Chao, J. A. (2012) Spatial arrangement of an RNA zipcode identifies mRNAs under post-transcriptional control. *Genes Dev.* **26,** 43-53.
23. Valverde, R., Edwards, L., and Regan, L. (2008) Structure and function of KH domains. *FEBS J* **275,** 2712-2726.

24. Lewis, H. A., Musunuru, K., Jensen, K. B., Edo, C., Chen, H., Darnell, R. B., and Burley, S. K. (2000) Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* **100,** 323-332.
25. Lin, Q., Taylor, S. J., and Shalloway, D. (1997) Specificity and determinants of Sam68 RNA binding. *J. Biol. Chem.* **272,** 27274-27280.
26. Zhou, Y., Mah, T., Greenblatt, J., and Friedman, D. I. (2002) Evidence that the KH RNA-binding domains influence the action of the E.coli NusA protein. *J. Mol. Biol.* **318,** 1175-1188.
27. Paziewska, A., Wyrwicz, .L. C., Bujnicki, J. M., Bomsztyk, K., and Ostrowski, J. (2004) Cooperative binding of the hnRNP K three KH domains to mRNA targets. *FEBS Letts.* **577,** 134-140.
28. Hollingworth, D., Candel, A. M., Nicastro, G., Martin, S. R., Briata, P., Gherzi, R., and Ramos, A. (2012) KH domains with impaired nucleic acid binding as a tool for functional analysis. *Nucleic Acids Res*. **40,** 6873-6886.
29. Schmittgen, T. D., and Livak, K. J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nat. Protocol* **3,** 1101-1108.
30. Louis-Jeune, C., Andrade-Navarro, M. A., and Perez-Iratxeta, C. (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins* **80,** 374-381.
31. Webb, S. E., Lee, K. W., Karplus, E., and Miller, A. L. (1997) Localized calcium transients accompany furrow positioning, propagation, and deepening during the early cleavage period of zebrafish embryos. *Dev. Biol.* **192,** 78-92.
32. King, D. T., Barnes, M., Thomsen, D., and Lee, C. H. (2014) Assessing specific oligonucleotides and small molecule antibiotics for the ability to inhibit the CRD-BP-CD44 RNA interaction. *PLoS ONE* **9,** e91585.33.
33. Nielsen, J., Adolph, S. K., Meyts, E. R. D., Lykke-Andersen, J., Koch, G., Christiansen, J., and Nielsen, F. C. (2003) Nuclear transit of human zipcode-binding protein IMP1. *Biochem. J.* **376,** 383-391.
34. Yaniv, K., Fainsod, A., Kalcheim, C., and Yisraeli, J. K. (2003) The RNA binding protein Vg1 RBP is required for cell migration during early neural development. *Development* **130,** 5649-5661.
35. Hansen, T. V. O., Hammer, N. A., Nielsen, J., Madsen, M., Dalbaeck, C., Wewer, U. M., Christiansen, J., and Nielsen, F. C. (2004) Dwarfism and impaired gut development in insulin-like growth factor II mRNA-binding protein 1-deficient mice. *Mol. Cell. Biol.* **24,** 4448-4464.
36. Hamilton, K. E., Noubissi, F. K., Katti, P. S., Hahn, C. M., Davey, S. R., Lundsmith, E. T., Klein-Szanto, A. J., Rhim, A. D., Spiegelman, V. S., Rustgi, A. K. (2013) IMP1 promotes tumor growth, dissemination, and a tumor-initiating cell phenotype in colorectal cancer cell xenografts. *Carcinogenesis* **34,** 2647-2654.
37. Gutschner, T., Hammerle, M., Pazaitis, N., Bley, N., Fiskin, E., Uckelmann, H., Heim, A., Grob, M., Hofman, N., Geffers, R., Skawran, B., Longerich, T., Breuhahn, K., Schirmacher, P., Muhleck, B., Huttelmaier, S., and Diedrichs, S. (2014) Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) is an important protumorigenic factor in hepatocellular carcinoma. *Hepatology* **59,** 1900-1911.
38. Jonson, L., Vikesaa, J., Krogh, A., Nielsen, L. K., Hansen, T. V. O., Borup, R., Johnsen, A. H., Christiansen, J., and Nielsen, F. C. (2007) Molecular composition of IMP1 ribonucleoprotein granules. *Mol. Cell. Proteomics* **6,** 798-811.
39. Leeds, P., Kren, B. T., Boylan, J. M., Betz, N. A., Steer, C. J., Gruppuso, P. A., and Ross, J. (1997) Developmental regulation of CRD-BP, an RNA-binding protein that stabilizes c-*myc* mRNA *in vitro. Oncogene* **14,** 1278-1286.
40. Wachter, K., Kohn, M., Stohr, N., and Huttelmaier, S. (2013) Subcellular localization and RNP formation of IGF2BPs (IGF2 mRNA binding proteins) is modulated by distinct RNA-binding domains. *Biol. Chem.* **394,** 1077-1090.

41. Farina, K.L., Huttelmaier, S., Musunuru, K., Darnell, R., and Singer, R.H. (2003) Two ZBP1 KH domains facilitate beta-actin mRNA localization, granule formation, and cytoskeletal attachment. *J. Cell Biol.* **160,** 77-87.
42. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M.Jr., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., and Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141,** 129-141.
43. Tafech, A., Bennett, W.R., Mills, F., and Lee, C.H. (2007) Identification of c-*myc* coding region determinant RNA sequences and structures cleaved by an RNase1-like endoribonuclease. *Biochim. Biophys. Acta* **1769,** 49-60.
44. Kim, W.C., King, D., and Lee, C.H. (2010) RNA-cleaving properties of human apurinic-apyrimidinic endonuclease 1 (APE1). *Int. J. Biochem. Mol. Biol.* **1,** 12-25.

Acknowledgements - We thank Dr. Finn Nielsen for the generous gifts of CD44 DNA template and corresponding primers. We are grateful to Dr. Jeffrey Ross for the mouse CRD-BP plasmids. We thank the Division of Life Science at the Hong Kong University of Science and Technology for providing facilities for our zebrafish embryos experiments.

FOOTNOTES

*This research was supported in part by a Discovery Grant (# 227158) from Natural Sciences & Engineering Research Council (NSERC) (to CHL), University of Northern British Columbia Research Project Awards (to MB, GVR, KM and SM), and the French National Research Agency/Hong Kong Research Grants Council Joint Research Scheme (# A-HKUST601/13) (to ALM). DTK was a recipient of NSERC Undergraduate Student Research Awards and a BC Cancer Agency Summer Studentship.
[1]To whom correspondence should be addressed: Chemistry Program, University of Northern BC, 3333 University Way, Prince George, BC V2N 4Z9, Canada. Phone: (250) 960-5413; Fax: (250) 960-5845; E-mail: chow.lee@unbc.ca.
[2,3]Division of Life Science and The Key State Laboratory for Molecular Neuroscience, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; Marine Biological Laboratory, Woods Hole, MA 02543, USA.
[4]The abbreviations used are: CRD-BP, coding region determinant-binding protein; IMP1, insulin-like growth factor-2 mRNA-binding protein1; EMSA, electrophoretic mobility shift assay; KH, K-homology; WT, wild-type; CD, circular dichroism.

FIGURE LEGENDS

FIGURE 1. Generation of wild type CRD-BP and CRD-BP point mutation variants. *A,*Schematic representation of the mouse CRD-BP and its variants used in the gel-shift experiments described in Figures 2 and 3. The two RRM and four KH domains are shown. As shown, the first Gly of the GXXG motif of the KH domain was mutated to an aspartate. Single (KH1, KH2, KH3, KH4) or double (KH1-2, KH1-3, KH1-4, KH2-3, KH2-4, KH3-4) point mutation was performed to generate the KH CRD-BP variants. Single point mutation was performed to generate the Y5A, Q84A, E445D and D526E variants. *B,* Purification of recombinant His-tagged wild-type CRD-BP and its point mutation variants. Two µg of purified recombinant wild-type (WT) CRD-BP and its variants were run on a 12% SDS-PAGE followed by staining with Coomassie brilliant blue.

FIGURE 2. The binding profile of CRD-BP and its point mutation variants on the coding region determinant of c-*myc* RNA. *A,* Electrophoretic mobility gel-shift assay on the binding of purified recombinant WT CRD-BP and its point mutation variants to [^{32}P] c-*myc* CRD RNA nts 1705-1886.

Various concentrations of proteins, as indicated, were incubated with 40 nM of the radiolabeled c-*myc* RNA. The positions of protein-RNA complexes (Bound) and unbound RNA (Unbound) are indicated. Samples within each panel indicate the same experiment. *B,* A summary of dissociation constants (K_d) of the WT CD-BP and its variants. The K_d values were taken from saturation binding curves (n = 4). The asterisk indicates that the p-value is less than 0.05 based on Student's t-test in comparing to the K_d of WT CRD-BP. For KH3-4 the P = 0.0321, and for KH1 the P = 0.0122.

FIGURE 3. The binding profile of CRD-BP and its point mutation variants on a specific 3' untranslated region of *CD44* RNA. *A,* Electrophoretic mobility gel-shift assay on the binding of purified recombinant WT CRD-BP and its point mutation variants to [^{32}P] *CD44* RNA nts 2862-3055. Various concentrations of proteins, as indicated, were incubated with 40 nM of the radiolabeled *CD44* RNA. The positions of protein-RNA complexes (Bound) and unbound RNA (Unbound) are indicated. Samples within each panel indicate the same experiment with the exception of the panel showing WT and KH3-4. *B,* A summary of dissociation constants (K_d) of the WT CD-BP and its variants. The K_d values were taken from saturation binding curves (n = 4). The asterisk indicates that the p-value is less than 0.05 based on Student's t-test in comparing to the K_d of WT CRD-BP. P = 0.0371 for Y5A, P = 0.005 for D526E, P = 0.0121 for KH1, P = 0.0091 for KH2, and P = 0.0143 for KH3.

FIGURE 4. CRD-BP granule formation in zebrafish embryos. *A,* Granule formation coincides with maternal zygotic transcription. Embryos were injected with mouse CRD-BP-EGFP mRNA at the one-cell stage and then imaged at 3 hpf (i) or 4 hpf (ii). For comparison, the human IMP1-EGFP mRNA was injected and imaged at 4 hpf (iii). *B,* Effect of transcriptional inhibitor or protein synthesis inhibitor on CRD-BP granule formation. After injection with CRD-BP-EGFP mRNA at the one-cell stage, the embryos were injected with actinomycin D (Act D) (20 µg/ml) immediately (ii). Some embryos were injected with cycloheximide (CHX) (20 µg/ml) immediately (iii) or at 16-cell stage (iv). Embryos were then imaged at 4 hours post-fertilization to examine the effect of drug treatment on CRD-BP granule formation. Data shown are representatives of three biological replicates. Scale = 10 µm.

FIGURE 5. Comparing truncated CRD-BP variants on granule formation in zebrafish embryos. *A,* Schematic representation of the mouse CRD-BP and its truncated variants used in this study. *B,* Embryos were injected with the WT CRD-BP-EGFP or various truncated variant mRNAs as shown at the one-cell stage. Embryos were then imaged at 4 hpf. Data shown are representatives of three biological replicates. Scale = 10 µm. *C,* The number of granules per cell were counted as described in Experimental Procedures and expressed as shown in the bar graph. At least three images per sample were counted from two separate mRNA injections. One-way ANOVA was performed as statistical analysis. The asterisk indicates $p < 0.01$ when compared with the WT CRD-BP.

FIGURE 6. Comparing CRD-BP point mutation variants on granule formation in zebrafish embryos. *A,* Embryos were injected with mRNA for the single KH point mutation CRD-BP, KH double point mutation CRD-BP, or WT CRD-BP-EGFP as shown at the one-cell stage. Embryos were then imaged at 4 hpf. Data shown are representatives of three biological replicates. Scale = 10 µm. *B,* The number of granules per cell were counted as described in the Experimental Procedures. At least three images per sample were counted from three separate mRNA injections. One-way ANOVA was performed as statistical analysis. The asterisk indicates $p < 0.05$ when compared with the WT CRD-BP.

FIGURE 7. Comparing mRNA levels that are physically associated with CRD-BP and its point mutation variants in HeLa cells. HeLa cells were transfected with plasmid vector pcDNA3-FLAG, pcDNA3-FLAG WT CRD-BP, or various pcDNA3-FLAG plasmids carrying CRD-BP variants as shown. Forty-eight hours after transfection, lysates from cells were isolated and subjected to immuno-precipitation using anti-FLAG antibody as described in the Experimental Procedures. *A,* Sixteen µl of immuno-precipitated FLAG-CRD-BP proteins were subjected to Western analysis using anti-FLAG antibody as shown. The

FLAG-CRD-BP bands from each transfection were quantified as Integrated Density Value (IDV) and expressed relative to the WT CRD-BP. *B,* RNAs that were physically associated with pulled down FLAG-CRD-BP were extracted and subjected to qRT-PCR for measurements of *CD44*, c-*myc* and *β–actin* mRNAs. mRNA levels were expressed relative to the pcDNA3-FLAG vector. The data collected from four biological replicates were then pooled and expressed relative to the WT CRD-BP which was taken as 1.0. One-way ANOVA statistical analysis was then performed to compare the single (middle panel) and double (right panel) KH variants to the WT CRD-BP: *CD44* mRNA/single KH variants, n = 5, F = 21.08, *P < 0.0001; *CD44* mRNA/double KH variants, n = 7, F = 22.43, *P < 0.0001; c-*myc* mRNA/single KH variants, n = 5, F = 10.36, **P = 0.0003; c-*myc* mRNA/double KH variants, n = 7, F = 13.01, *P < 0.0001; *β-actin* mRNA/single KH variants, n = 5, F = 0.909, P = 0.4837; *β-actin* mRNA/double KH variants, n = 7, F = 1.005, P = 0.4579.

TABLES

TABLE 1

A summary of the dissociation constants (K_d) of WT CRD-BP in comparison with its KH variants for binding to c-*myc* CRD and CD44 RNAs

CRD-BP variant	c-*myc* RNA nts 1705-1886		CD44 RNA nts 2862-3055	
	Dissociation constant (K_d)	Hill coefficient	Dissociation constant (K_d)	Hill coefficient
WT	398.0 ± 52.79	1.768 ± 0.011	149.32 ± 8.42	1.916 ± 0.0080
KH1	723.1 ± 75.13	1.539 ± 0.004	211.18 ± 34.94	1.824 ± 0.0173
KH2	360.3 ± 70.41	1.680 ± 0.013	250.27 ± 26.72	1.756 ± 0.0089
KH3	320.8 ± 42.75	1.790 ± 0.012	219.32 ± 19.35	1.826 ± 0.0090
KH4	320.4 ± 59.08	1.787 ± 0.017	NA	NA
KH1-2	NA	NA	NA	NA
KH1-3	NA	NA	NA	NA
KH1-4	NA	NA	NA	NA
KH2-3	NA	NA	NA	NA
KH2-4	NA	NA	NA	NA
KH3-4	586.9 ± 42.89	1.728 ± 0.004	133.33 ± 11.04	1.849 ± 0.0112
Y5A	467.6 ± 45.68	1.755 ± 0.007	100.11 ± 16.41	1.978 ± 0.0297
D526E	402.4 ± 37.25	1.767 ± 0.007	90.65 ± 1.64	1.986 ± 0.0034
Q84A	490.6±20.01	1.748 ± 0.003	ND	ND
E445D	459.5 ± 42.18	1.751 ± 0.006	ND	ND

NA indicates not applicable because significantly reduced or no binding occurred and hence binding curve cannot be plotted. ND indicates not determined. Data are means ± s.e.m.

TABLE 2

A summary of the electrophoretic mobility shift assay, zebrafish granule formation, and immuno-precipitation coupled qPCR experiments

RNA	Type of KH variants	*In-vitro* EMSA[a] (< 1 hour)	Granule formation in zebrafish embryos[b] (4 hours)	mRNA associated with FLAG-CRD-BP in HeLa cells[c] (48 hours)
c-*myc*	Single KH point mutation	All bind	+++	All reduced except KH4
	Double KH point mutation	All no binding except KH3-4	All defect except KH3-4	All reduced
CD44	Single KH point mutation	All bind except KH4	+++	All reduced
	Double KH point mutation	All no binding except KH3-4	All defect except KH3-4	All reduced

[a]Taken from Figs. 2 and 3.

[b]Taken from Fig. 6. The granule formation does not focus on specific mRNA species and the row for c-*myc* and CD44 is repeated. '+++' indicates normal granule formation compared to WT CRD-BP.

[c]Taken from Fig. 7B. Reduced means mRNA species were found to have reduced level associated with the corresponding FLAG-KH variants as compared to its association with the FLAG-WT CRD-BP.

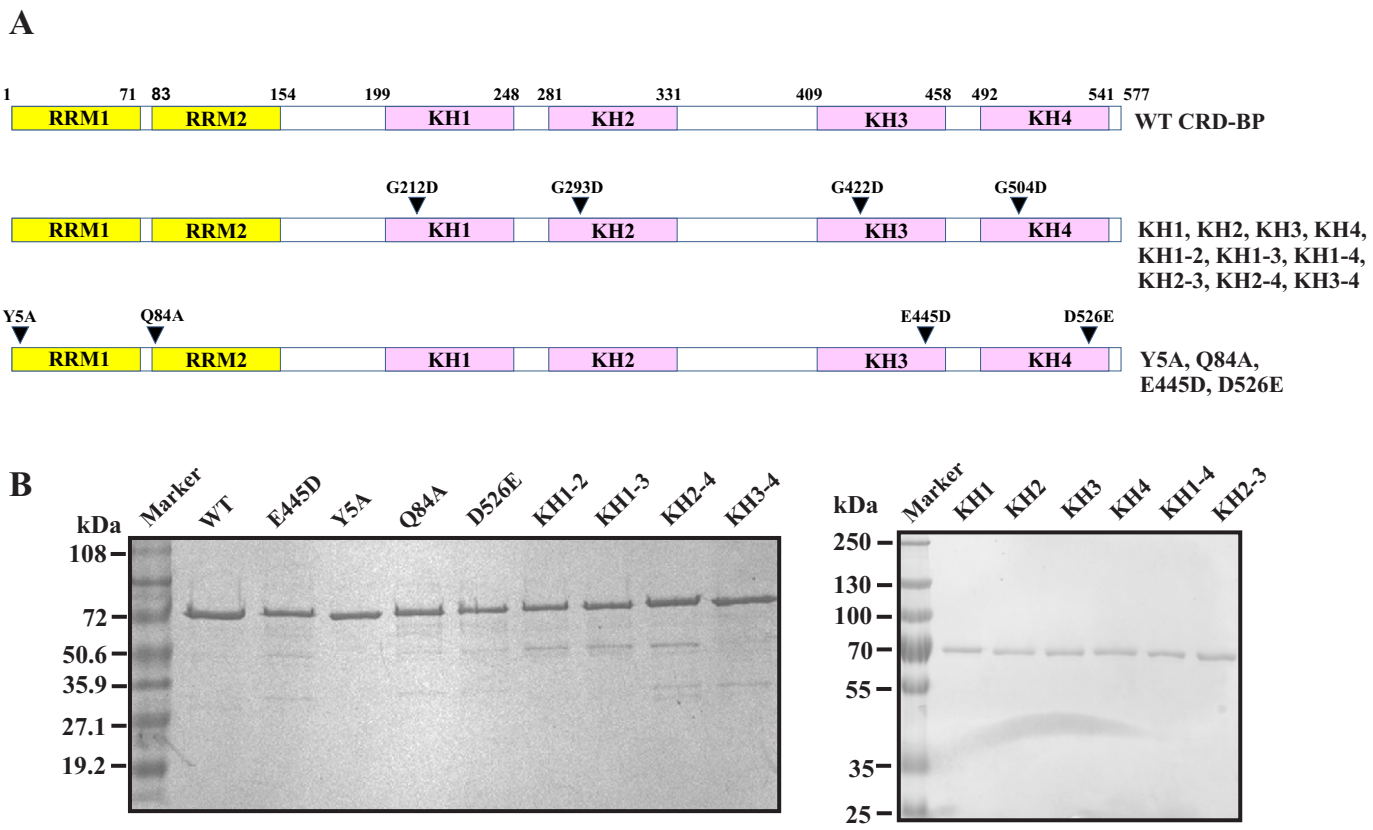

A
1
71
83
154
199
248
281
331
409
458
492
541
577
RRM1
RRM2
KH1
KH2
KH3
KH4
WT CRD-BP
G212D
G293D
G422D
G504D
KH1, KH2, KH3, KH4,
KH1-2, KH1-3, KH1-4,
KH2-3, KH2-4, KH3-4
Y5A
Q84A
E445D
D526E
Y5A, Q84A,
E445D, D526E
B
kDa
Marker
WT
E445D
Y5A
Q84A
D526E
KH1-2
KH1-3
KH2-4
KH3-4
108
72
50.6
35.9
27.1
19.2
kDa
Marker
KH1
KH2
KH3
KH4
KH1-4
KH2-3
250
130
100
70
55
35
25

Figure 1

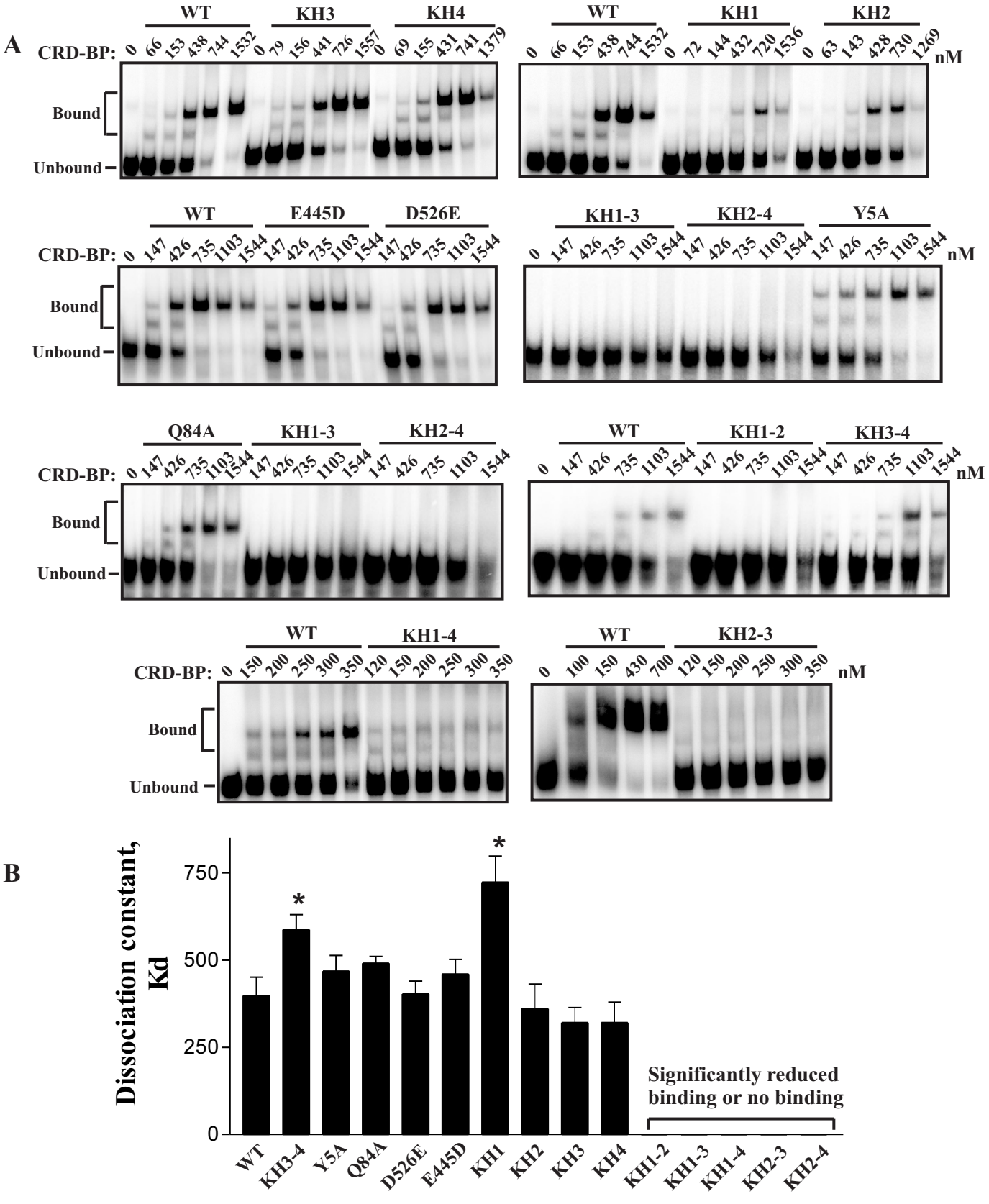
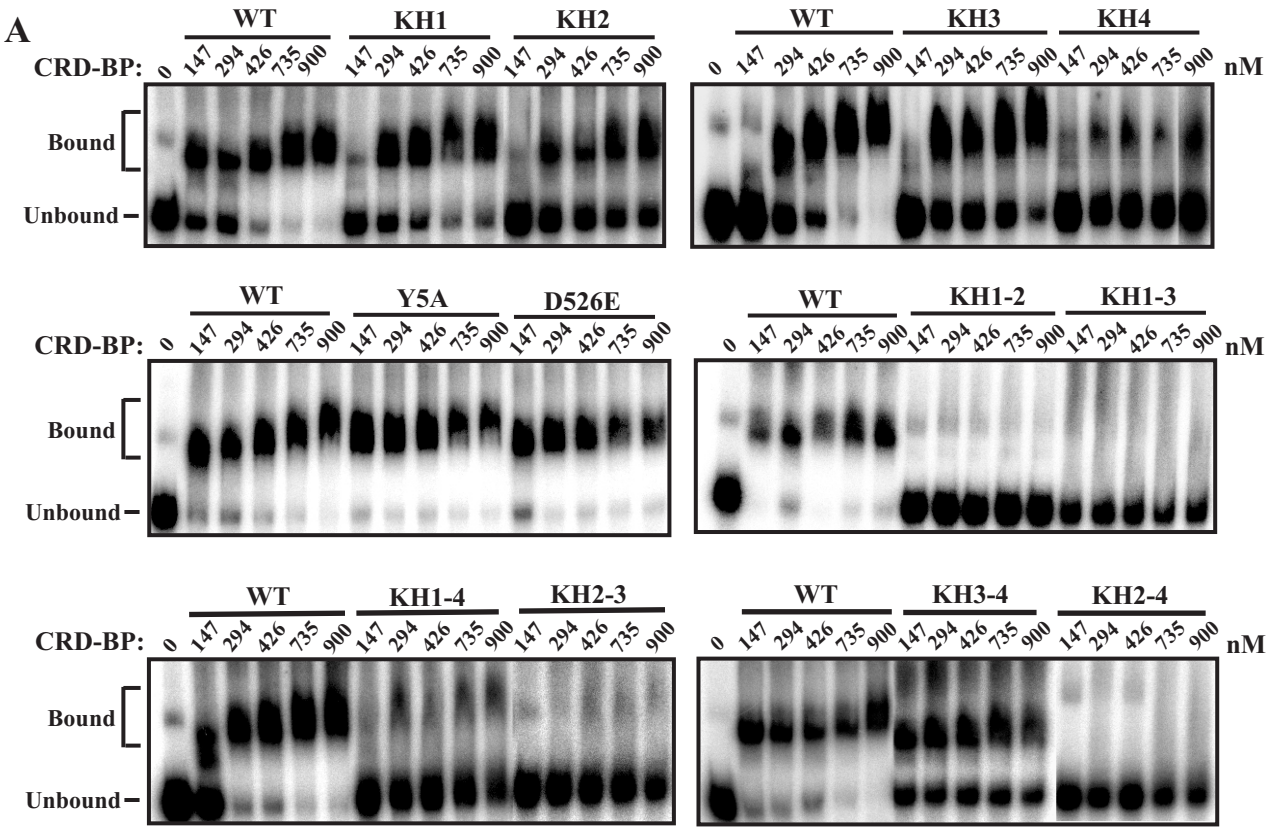

A
WT
KH3
KH4
CRD-BP: 0 66 153 438 744 1532 0 79 156 441 726 1557 0 69 155 431 741 1379
Bound
Unbound
WT
KH1
KH2
0 66 153 438 744 1532 0 72 144 432 720 1536 0 63 143 428 730 1269 nM
WT
E445D
D526E
CRD-BP: 0 147 426 735 1103 1544 147 426 735 1103 1544 147 426 735 1103 1544
KH1-3
KH2-4
Y5A
0 147 426 735 1103 1544 147 426 735 1103 1544 147 426 735 1103 1544 nM
Q84A
KH1-3
KH2-4
CRD-BP: 0 147 426 735 1103 1544 147 426 735 1103 1544 147 426 735 1103 1544
WT
KH1-2
KH3-4
0 147 426 735 1103 1544 147 426 735 1103 1544 147 426 735 1103 1544 nM
WT
KH1-4
CRD-BP: 0 150 200 250 300 350 120 150 200 250 300 350
WT
KH2-3
0 100 150 430 700 120 150 200 250 300 350 nM
B
Dissociation constant, Kd
750
500
250
0
*
*
Significantly reduced binding or no binding
WT
KH3-4
Y5A
Q84A
D526E
E445D
KH1
KH2
KH3
KH4
KH1-2
KH1-3
KH1-4
KH2-3
KH2-4

Figure 2


A
WT
KH1
KH2
WT
KH3
KH4
CRD-BP:
0 147 294 426 735 900 147 294 426 735 900 147 294 426 735 900
0 147 294 426 735 900 147 294 426 735 900 147 294 426 735 900
nM
Bound
Unbound
WT
Y5A
D526E
WT
KH1-2
KH1-3
CRD-BP:
0 147 294 426 735 900 147 294 426 735 900 147 294 426 735 900
0 147 294 426 735 900 147 294 426 735 900 147 294 426 735 900
nM
Bound
Unbound
WT
KH1-4
KH2-3
WT
KH3-4
KH2-4
CRD-BP:
0 147 294 426 735 900 147 294 426 735 900 147 294 426 735 900
0 147 294 426 735 900 147 294 426 735 900 147 294 426 735 900
nM
Bound
Unbound

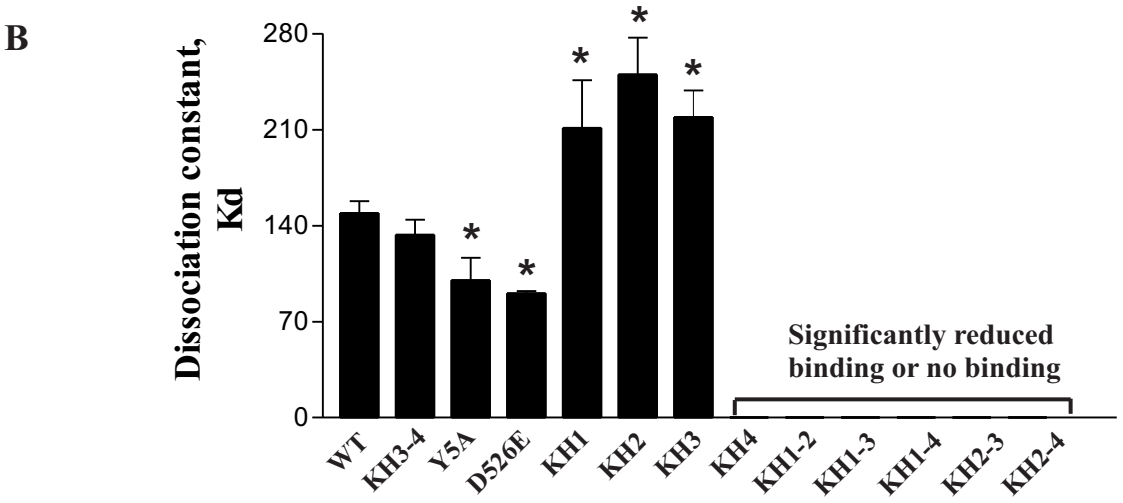

B
Dissociation constant, Kd
280
210
140
70
0
*
*
*
*
*
Significantly reduced binding or no binding
WT
KH3-4
Y5A
D526E
KH1
KH2
KH3
KH4
KH1-2
KH1-3
KH1-4
KH2-3
KH2-4

Figure 3

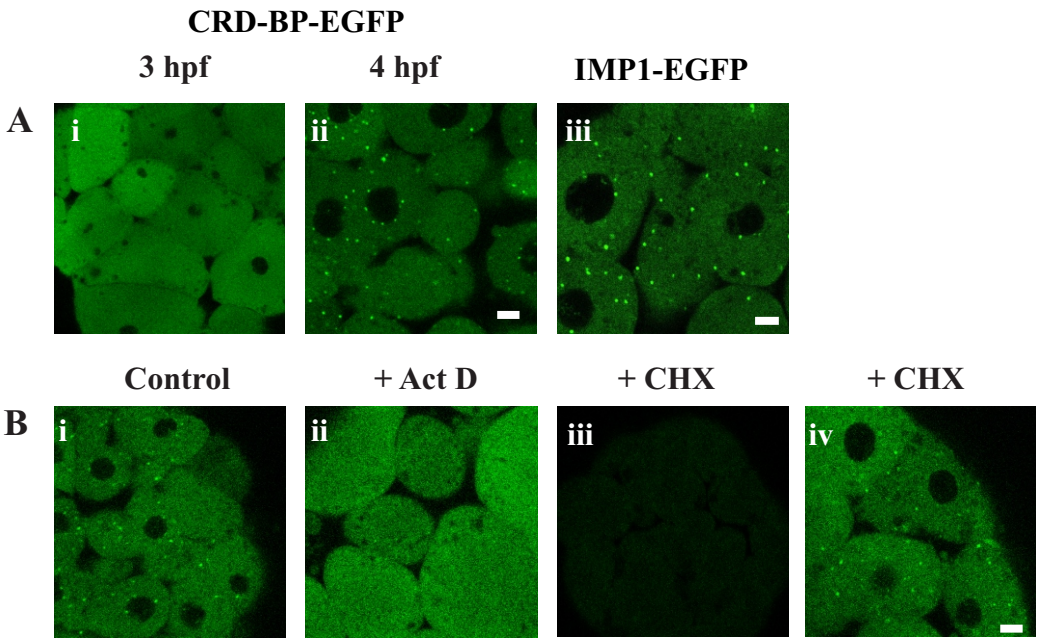

CRD-BP-EGFP
3 hpf
4 hpf
IMP1-EGFP
A
i
ii
iii
Control
+ Act D
+ CHX
+ CHX
B
i
ii
iii
iv

Figure 4

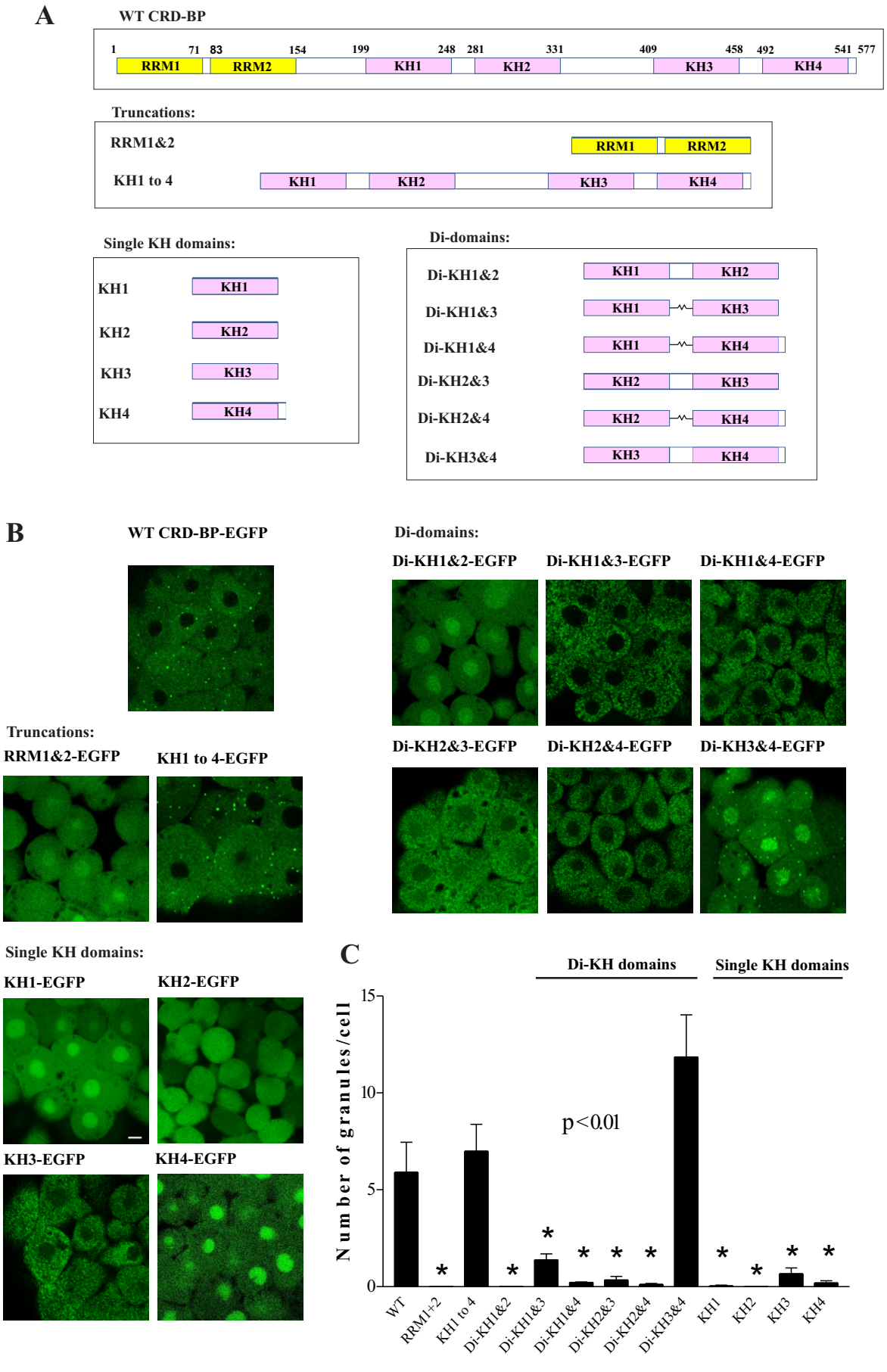

A
WT CRD-BP
1
71
83
154
199
248
281
331
409
458
492
541
577
RRM1
RRM2
KH1
KH2
KH3
KH4
Truncations:
RRM1&2
RRM1
RRM2
KH1 to 4
KH1
KH2
KH3
KH4
Single KH domains:
KH1
KH2
KH3
KH4
Di-domains:
Di-KH1&2
Di-KH1&3
Di-KH1&4
Di-KH2&3
Di-KH2&4
Di-KH3&4
B
WT CRD-BP-EGFP
Truncations:
RRM1&2-EGFP
KH1 to 4-EGFP
Single KH domains:
KH1-EGFP
KH2-EGFP
KH3-EGFP
KH4-EGFP
Di-domains:
Di-KH1&2-EGFP
Di-KH1&3-EGFP
Di-KH1&4-EGFP
Di-KH2&3-EGFP
Di-KH2&4-EGFP
Di-KH3&4-EGFP
C
Di-KH domains
Single KH domains
Number of granules/cell
15
10
5
0
p<0.01
WT
RRM1+2
KH1 to 4
Di-KH1&2
Di-KH1&3
Di-KH1&4
Di-KH2&3
Di-KH2&4
Di-KH3&4
KH1
KH2
KH3
KH4

Figure 5

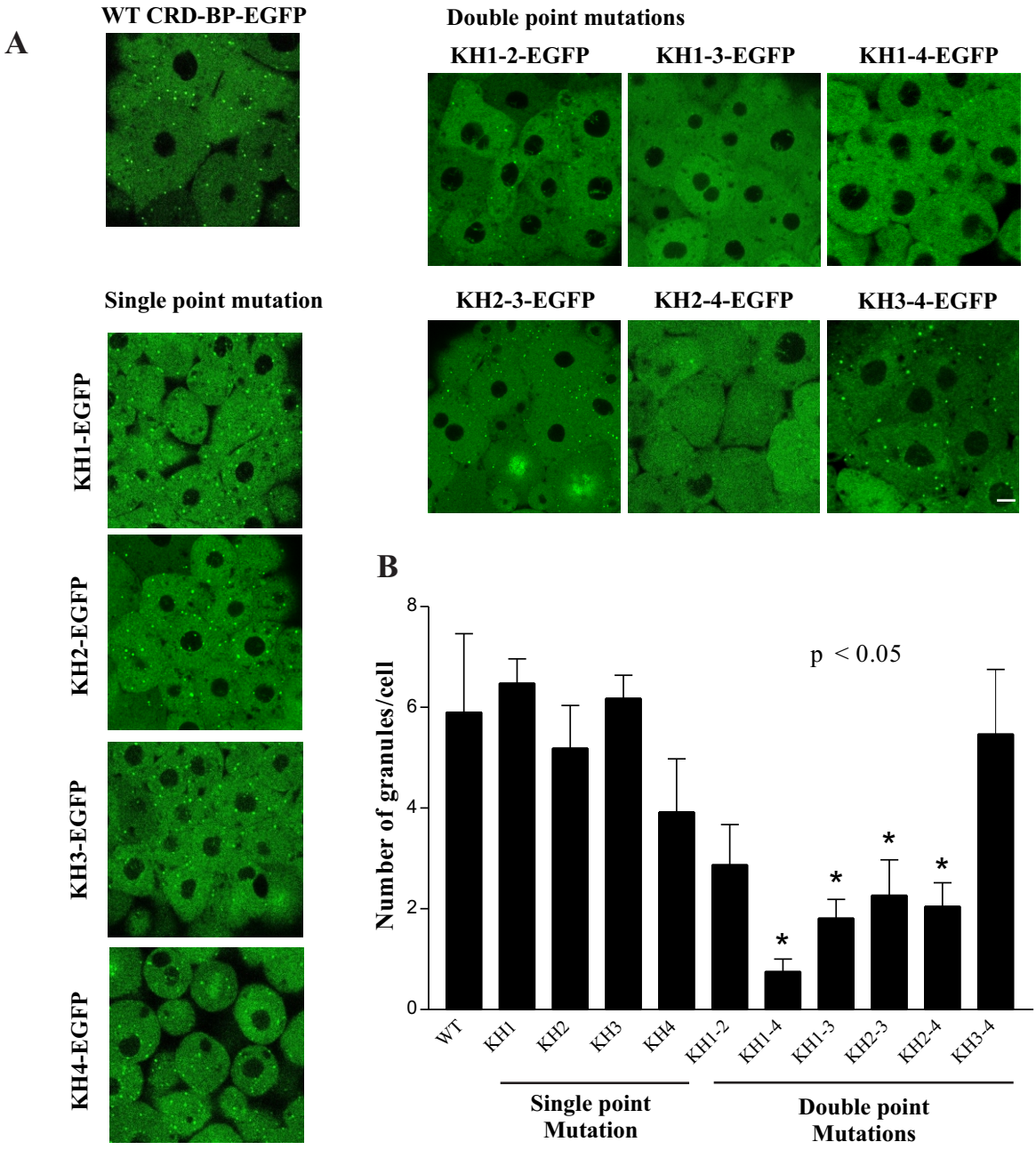

A
WT CRD-BP-EGFP
Double point mutations
KH1-2-EGFP
KH1-3-EGFP
KH1-4-EGFP
Single point mutation
KH2-3-EGFP
KH2-4-EGFP
KH3-4-EGFP
KH1-EGFP
KH2-EGFP
KH3-EGFP
KH4-EGFP
B
Number of granules/cell
p < 0.05
WT
KH1
KH2
KH3
KH4
KH1-2
KH1-4
KH1-3
KH2-3
KH2-4
KH3-4
Single point Mutation
Double point Mutations

Figure 6

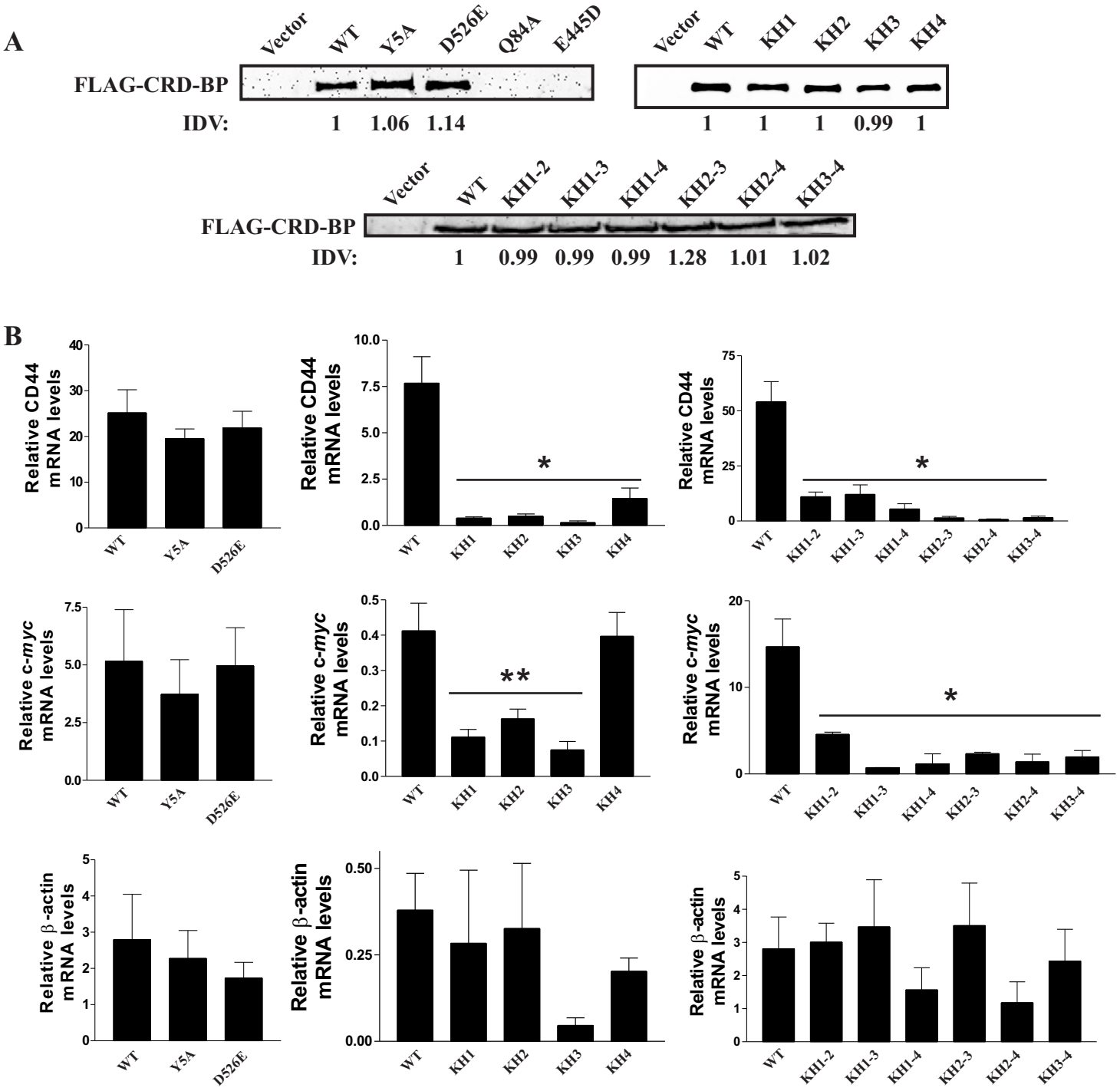

A
Vector
WT
Y5A
D526E
Q84A
E445D
FLAG-CRD-BP
IDV: 1 1.06 1.14
Vector
WT
KH1
KH2
KH3
KH4
1 1 1 0.99 1
Vector
WT
KH1-2
KH1-3
KH1-4
KH2-3
KH2-4
KH3-4
FLAG-CRD-BP
IDV: 1 0.99 0.99 0.99 1.28 1.01 1.02
B
Relative CD44 mRNA levels
Relative c-myc mRNA levels
Relative β-actin mRNA levels
WT
Y5A
D526E
KH1
KH2
KH3
KH4
KH1-2
KH1-3
KH1-4
KH2-3
KH2-4
KH3-4
*
**

Figure 7