Defining the RNA binding functions of the translational repressor Bicaudal-C

By

Megan E. Dowdle

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

At the

UNIVERSITY OF WISCONSIN-MADISON

2020

Date of Final Oral Examination: 8/12/2020

This dissertation is approved by the following members of the Final Oral Committee:
Michael Sheets, Professor Biomolecular Chemistry
Catherine Fox, Professor, Biomolecular Chemistry
Melissa Harrison, Associate Professor, Biomolecular Chemistry
Joshua Coon, Professor, Biomolecular Chemistry
Francisco Pelegri, Professor, Genetics and Medical Genetics

ABSTRACT

Bicaudal-C (Bicc1) is a conserved RNA binding protein and developmental regulator that functions in metazoans to selectively repress the translation of specific mRNAs. This repression is essential for controlling the synthesis of proteins that regulate developmental decisions. Despite Bicc1's biological relevance, there is a lack of experimental structure-function data that addresses how this protein binds to specific mRNA targets. Studies in *Xenopus* indicate that the N-terminal half of Bicc1, which contains three canonical KH-domains and two KH-like (KHL) domains, is both necessary and sufficient for mRNA binding. Furthermore, analysis of different mutant Bicc1 proteins indicates that removal of any KH or KHL domains eliminates RNA binding, indicating that the KH domains are critical for Bicc1's ability to interact with RNA. Additional analysis of variants in which individual KH domains were modified to abolish their RNA binding activity revealed that the second KH domain (KH2) is the critical feature for Bicc1's mRNA binding activity, while the KH1 and KH3 domains make minimal contributions to mRNA binding. Bicc1 binds to a 32-nucleotide element in the 3'UTR of the CR1 mRNA. However, it is unknown the exact sequence motif(s) that Bicc1 recognizes. Previous studies have shown that mutations in the leader and loop sequences of the element adversely affect Bicc1's ability to stably bind the CR1 mRNA. In order to determine the exact sequence motif that Bicc1 binds to in this 32-nucleotide structure, we have systemically changed the nucleotides in both the leader and the loop and monitored the binding via electrophoretic mobility shift assays (EMSAs) and fluorescent polarization. Our results have indicated that Bicc1 binding to the CR1 target is complex. Sequences in both the leader and the loop are crucial for mRNA binding

indicating a possible bipartite motif. Additionally, while several RNA mutants impair Bicc1 binding, no single mutation abolishes Bicc1 binding to the CR1 mRNA target. Results from these experiments are critical for understanding the mechanisms by which Bicc1 influences embryonic development as well as understanding how RNA-binding domains function together to identify specific target RNA sequences.

ACKNOWLEDGEMENTS

When I set my thesis defense date I imagined a nice spring day in late April, surrounded by family, friends, and colleagues. I could have never imagined that I would be writing and defending my thesis during a global pandemic. And while that has continued to bum me out for so many different reasons, the unique circumstances of my defense have made me really appreciate all the people in my life and all those who made my defense possible even more. It is too easy to let these bad circumstances get to you, and I cannot tell you how many times I have cried during this quarantine. But there are so many people who have uplifted me that it made it possible to push through everything and I hope that I can mention them all here.

First I'd like to thank my advisor Mike Sheets. There are so many little things that Mike has done to show that he cares about my life as a graduate student and my life as a person. I mean how many PIs send you an e-mail during the Women's World Cup to tell you that The Sett is playing all the games? Mike pushes me to be a better writer, scientist, and thinker and that's just the tip of the iceberg. I have been so fortunate and so grateful to be a student in Mike's lab and I want to thank him so much for giving me the opportunity to be his student.

I'd also like to thank every member of the Sheets lab. We have never been very big, but I have been lucky to know each one of you. First I'd have to thank Sookhee Park. Sookhee was my only lab mate for the first five years of my graduate school career and she helped me with everything. I hoped she would never leave but I know she's doing great things out in Boston. I also have to thank Susanne Blaser Imboden.

There's nobody who is more a whiz at making RNA and just keeping everything going than Susanne. Next, I'd like to thank Cole and Sam. New members of the lab and just as valuable. You two really made me realize how much I'd grown just by how many questions I have to answer. Trust me that's not a bad thing at all. And now all my many many undergraduates: Andy, Nithin, Ian, Thomas, Tom, Sonam, Tommaso, and Krystal. In a small lab, undergraduates really bring a lot of life and energy. You all kept me on my feet and so much of your work is represented in this thesis. You've kept me up on all the trends when I've felt old. And kept me updated on what undergrad life is life and made me realize I don't miss it for a second.

My committee has been outstanding: Catherine Fox, Melissa Harrison, Joshua Coon, and Francisco Pelegri. You have always pushed me to be a more critical thinker and have had really great suggestions every meeting. Catherine Fox especially has been a great person to rant with about whatever situation is going on in the world. And this is where I want to shout out the entire Fox lab for just accepting all my endless interruptions to their day when I want to talk about Pop Culture and gossip and there's nobody in my lab to listen.

To IPIB. The faculty and staff and everybody. I especially want to thank Kate Ryan who keeps everything going smoothly. I want to thank the media lab, especially Laura Vanderploeg without whom the Sheets Lab would all have to learn Illustrator. Our papers and posters are fantastic from her input. To Robin for taking great photos of me. I'd like to thank all the BMC staff for answering so many logistical questions. And I have to thank Joe for fixing every piece of equipment I've ever had a problem with.

have so many friends and I'd like to thank them all here. I only regret that for brevity's sake I cannot give them a paragraph each. To Kati, Erynn, Madeline, and Chris for their awful nicknames and endless conversations about weird stuff. I'd also like to thank Erynn's parent Val and Gary who allowed me to stay at their home and loaned me a car when mine broke down. To Wendy who I think is the only person who likes to gossip more than I do. To Emily Sellers who has always kept me on my toes. To my IPiB cohort. We are so close to all being finished. To Ana who is always down for a fancy dinner and drinks and a political discussion. To Haley, my roommate for my first two years and who shouldered a lot of emotional talks and teasing. To Stephanie who was in this 7th year with me and I can't say enough about how important that was for my mental health and also thanks for all the Bachelor/ette/in Paradise nights. To Alex and Greg even though you left me to move to Boston. To Tyler who is both one of the grumpiest people I know and one of the most caring. To Karl who has dared me to do to so many things. To Markus and Derek who are probably as eccentric as I am. I'd have to thank my trivia teams and my Dungeons and Dragons group, who almost entirely overlap. Dylan, Brian, Harriet, Dana, Sebastien, Sophie, I didn't mention you above but I'm glad I was able to make friends with people younger than me so that I'd have enough people in Madison to hang out with and play games with when my entire cohort graduated and left before me. And to all my other IPiB colleagues and friends who I can't mention by name or this list would be too long.

To all my friends from outside of IPiB who I have made in Madison: Alexa, Sydney, Kim, Katrina, Angie, Beth, Hayley, Jenny, and Chet (I think that's everybody), and the staff of Planned Parenthood both current and past. Thank you. Without

volunteering I don't think I would have stayed sane and you were the coolest people to volunteer and travel with. And all my internet friends, particularly Dee, Elliott, and Andrew who've also been essential in keeping me sane.

To my Promega group: Danette, Jacqui, Kristin, Sarah, and Nancy. You all made my internship wonderful and I learned so much from all of you.

To my cat Kara. I have always thought if you lived alone you should have a pet and quarantine really showed me why. I spent probably hours a day just watching you knock everything off my shelves.

And last, I'd like to thank my family. I'd like to thank my parents Zack and Vera, my sister Emily, and my brother Ben (sorry Will). You have all been so supportive in my endeavors. I'd like to thank my grandparents: Grandma Kane, Papaw Bill, and Nana Jennie for being my biggest fans. I'd like to thank the rest of my family: Aunt Lisa, Aunt Denise, Uncle Gary, Aunt Kathy, Uncle Richard, Uncle Steve, Alyssa, Bailey, Mark, Laura, Allyson, and Andrew. You all are some of the wackiest people I know, and I wouldn't trade any of you for the world.

Table of Contents

Chapter 1:

Post-transcriptional control by the translational repressor Bicaudal-C

Introduction

The regulation of gene expression is fundamental for all biological processes as the genes expressed by a cell dictate its proteome, its identity and functions. When proper regulation is disturbed, this can lead to a variety of abnormalities caused by either the reduced or over expression of the gene at the proper time. Regulation can occur at a variety of levels that include the selective transcription of genes to produce messenger RNAs (mRNAs) to post-transcriptional events that determine which mRNAs are eventually translated into proteins. Post-transcriptional regulation is especially crucial in early animal development when the embryo is not transcribing its own genes and all regulation must occur at steps after RNA synthesis (Sheets et al., 2017). This type of control is also critical in cells such as neurons that have long cytoplasmic projections referred to as dendrites. mRNAs that encode proteins that need to be expressed at the end of dendrites as opposed to close to the nucleus must be translationally suppressed until they reach their destination (Holt et al., 2019). In these and similar contexts, post-transcriptional regulatory processes are mediated by RNA-binding proteins (RBPs) (Mitchell and Parker, 2014).

This thesis focuses on the Bicaudal-C RNA binding protein that regulates post-transcriptional gene expression by binding to mRNAs and repressing their translation. In order to give context and background to my studies I will describe the importance of RNA binding proteins and the role they play in common mechanisms of post-transcriptional regulation. I will then discuss translational repression and mechanisms of repression and the role this plays in *Xenopus* embryonic development. Finally, I will review what is currently known about the Bicaudal-C protein including its function as a

translational repressor, importance in embryonic development, and its impact on human disease.

RNA binding proteins

RNA binding proteins have a central role in post-transcriptional regulatory events (Mitchell and Parker, 2014). Therefore, in order to understand post-transcriptional regulation, it is important to understand the function of RNA-binding proteins. There are approximately 1,500 human proteins known to bind to RNA and these RNA-binding proteins serve a variety of functions in the cell (Gerstberger et al., 2014). There are many types of RNAs present in eukaryotic cells; messenger RNAs which encode genes to be translated into proteins, non-coding RNAs which include RNAs that perform critical roles in ribonucleoprotein (RNP) complexes such as the spliceosome and ribosome, transfer RNAs (tRNAs) which are responsible for codon recognition during translation and regulatory RNAs such as micro RNAs which regulate gene expression at the transcriptional and post-transcriptional levels. While all these RNAs are critical for cell function, the focus of my thesis is on the regulation of mRNAs by the Bicaudal-C RNA binding protein.

mRNA structure and interactions with RNA-binding proteins

To understand how RNA-binding proteins interact with mRNAs, it is important to understand the basic structure of eukaryotic mRNAs **(Fig. 1)**. Most mature mRNAs have a 5'cap followed by a 5' untranslated region (UTR) upstream of the coding sequence or open reading frame (ORF) which encodes the protein. Following the ORF, the mRNA has a 3'UTR and a poly(A) tail. The 5' and 3' UTRs of specific mRNAs often contain

regulatory sequences that mediate mRNA localization, regulation of translation, and stability (Mayr, 2019; Hinnebusch et al., 2016). These regulatory elements are often sites for interaction with specific RNA binding proteins and it's the interaction of the proteins with these sites that specify an mRNA's behavior.

RNA binding domains

RNA-binding proteins have evolved many diverse domains responsible for forming specific protein-RNA interfaces. Often these domains recognize a specific nucleotide motif that consists of a few nucleotides. Such small nucleotide motifs recognized by individual RNA binding domains are often insufficient to confer specificity for recognition of certain targets in large RNAs and in complex environments, such as the cytoplasm of eukaryotic cells (Dominguez et al., 2018). Hence many RNA-binding proteins contain and utilize multiple domains to confer binding specificity and target a subset of mRNAs for regulation. Some of the most common RNA-binding domains include RNA recognition motifs (RRMs), Zinc Fingers (ZF), hnRNP K homology domains (KH domains), RGG/RG domains, double-stranded RNA binding domains (dsRBD) (Lunde et al., 2007; Dominguez et al., 2018).

An example of a well-studied RNA-binding protein is the Zipcode binding protein (ZBP1) expressed in neurons. ZBP1 binds to a sequence element present in the β-actin mRNA and represses its translation until the mRNA is transported to the dendrites of neurons (Farina et al., 2003). This protein has six total potential RNA-binding domains; two RRMs and four KH domains. However, only KH domains 3 and 4 are required for specific binding to the β-actin mRNA (Nicastro et al., 2017). These KH domains make contact to the phosphate backbone of the mRNA via GXXG motifs (where X can be any

amino acid).(Nicastro et al., 2017) Then surface exposed amino acids near the GXXG motif contact specific bases and confer nucleotide specificity (Nicastro et al., 2017).

The Polypyrimidine Tract Binding Protein (PTB) is another example of a protein with multiple potential RNA-binding domains. PTB is involved in a variety of processes including facilitating cap-independent translation through binding to internal ribosomal entry sites (IRES) and promoting alternative splicing by repressing the inclusion of exons (Romanelli et al., 2013). PTB contains four RRMs. RRM1 and RRM2 are connected by a long flexible linker which physically separate them from the rest of the protein. It is thought that this unique structure of the protein allows for its binding to different targets based on the RNA binding preferences of the difference RRMs (Romanelli et al., 2013). Experiments indicate that RRM1 and RRM2 bind to shorter polypyrimidine tracts of splicing substrates while RRM3 and RRM4 prefer to bind to longer polypyrimidine tracts (Clerte and Hall, 2009). In addition, PTB directly interacts with the RNA component of the splicing snRNP U1 through only its first two RRM domains (Sharma et al., 2011).

In addition to the recognition of short sequence element some RNA-binding proteins recognize specific structures of RNA substrates. For example, the RRM of the RBMY protein specifically recognizes and binds to a stem-loop structure present in its RNA substrates (Skrisovska et al., 2007). In addition, the sterile-alpha motif (SAM) domains in the Smaug protein in *Drosophila* and the Vts1 protein in yeast both recognize and bind to a distinct stem-loop structure (Aviv et al., 2003; Aviv et al., 2006). These examples demonstrate the both complexity of RNA-binding proteins and

highlight the challenges to characterizing the RNA binding mechanisms of specific proteins.

Cytoplasmic mRNA regulation

mRNA regulation in the cytoplasm is a critical component of cell metabolism as it dictates the composition of the proteome and in turn the identity and functions of a cell. mRNAs are regulated at several points throughout their lifespan in the cell. Transcriptional controls function such that only certain genes will be transcribed into pre-mRNAs. Then once a pre-mRNA is transcribed from a gene, it is subject to many levels of regulation. This includes pre-mRNA processing in the nucleus to form the mature mRNA and export of the mRNA to the cytoplasm. Once in the cytoplasm, mRNAs are subject to additional regulatory processes such as mRNA localization, stability, and translational control **(Fig. 2)** (Gebauer and Hentze, 2004). My work has focused on RNA regulation in the cytoplasm. In the following section I will give a brief and general overview of these cytoplasmic processes.

mRNA localization

In certain specialized cell types, once particular mRNAs enter the cytoplasm they are actively localized to particular sub-regions of the cell. This localization allows for spatial regulation of mRNA translation and as a consequence the deposition (or localization) of the encoded proteins in one region of the cell (Martin and Ephrussi, 2009). mRNA localization is particularly vital in embryonic development where mRNAs are localized to help establish embryo polarity (Du et al., 2007).

Two classic examples of mRNA regulation are the ASH1 mRNA in budding *Saccharomyces cerevisiae* and the VegT mRNA in *Xenopus laevis* embryos. ASH1 mRNA is localized to the bud tip of the budding yeast and is deposited in the daughter cell to ensure that the mother and daughter cells have different mating types (Paquin and Chartrand, 2008). The VegT mRNA is localized in the vegetal pole of *Xenopus* embryos (Zhang and King, 1996). After localization the mRNA is translated and the localized protein participates in cell-fate processes, guiding the vegetal cells to become the endodermal and mesodermal cells of the developing embryo (Zhang and King, 1996). These and other examples demonstrate how mRNA localization is critical for cell-fate decisions.

mRNA stability

A critical component of cytoplasmic mRNA regulation is mRNA stability. mRNAs must be present and translated at certain times in certain quantities and then either siphoned away or quickly degraded to prevent overexpression of the protein coded by the message. mRNA stability is often conferred by regulating the presence of its 5'cap and/or its 3' poly(A) tail. Fittingly, mRNA degradation typically starts with deadenylation by enzymes like PAN1-PAN2 or CCR4-NOT in yeast or CCR4-NOT and PARN in animals (Garneau et al., 2007). Following deadenylation, the mRNA can either be decapped by the Dcp1 and Dcp2 enzymes to undergo 5'3' degradation by the enzyme XRN1 or it can undergo 3'to 5' degradation by the exosome and then decapped (Garneau et al., 2007; Nickless et al., 2017).

The stability of specific mRNAs is highly regulated by RNA binding proteins. For example, many mRNAs contain AU-rich elements (AREs) that can be recognized by

specific RNA binding proteins (RBPs), such as KSRP (Gherzi et al., 2004) that in turn recruit deadenylation and decapping enzymes to promote degradation. Conversely, proteins such as poly(A) binding protein (PABP) or cytoplasmic polyadenylation element binding protein (CPEB) can bind to the 3'UTR of mRNAs to promote stability by counteracting the degradation processes (Garneau et al., 2007; Towler and Newbury, 2018; Nickless et al., 2017).

Regulation of mRNAs via microRNAs

miRNAs are small non-coding RNAs, approximately 21-nucleotides long, that bind to specific sequences in mRNA targets and recruit RNA-binding proteins such as the Argonaute (AGO) proteins to carry out their role in post-transcriptional regulation, principally the repression of translation and mRNA destabilization (Filipowicz et al., 2008; Jonas and Izaurralde, 2015). Classically, miRNAs are thought to bind to mRNAs and prime them for degradation by recruiting deadenylation factors, such as the CCR4-NOT complex (Filipowicz et al., 2008). An example of this is found in the zebrafish embryo where the miRNA miR-430 is responsible for the mass deadenylation of maternal mRNAs at the point in early development when zygotic transcription is initiated (Filipowicz et al., 2008; Jonas and Izaurralde, 2015).

Regulated mRNA translation

The regulation of mRNA translation is a key in post-transcriptional process as it significantly influences the proteome of the cell and as such has a significant influence on a cell's identity and functions (Gebauer and Hentze, 2004; Hershey et al., 2019; Tahmasebi et al., 2019). There are three main phases of translation which are all under tight regulation: initiation, elongation, and termination. Most often translation is regulated

at the point of initiation which requires the 5’ cap and 3’ poly(A) tail (Gebauer and Hentze, 2004). In the first steps of cap-dependent translation initiation, the cap-binding protein eIF4E binds along with other factors to recruit the ribosome and initiate translation. Many mRNAs possess 3’ poly(A) tails and these are bound by poly A binding protein (PABP) (Mangus et al., 2003). Once bound to a poly(A) tail PABP interacts with eIF4G which binds to eIF4E to potentially circularize the mRNA and facilitates the recruitment of ribosomal subunits (Wells et al., 1998; Tahmasebi et al., 2019; Tahmasebi et al., 2018; Hershey et al., 2019).

RNA-binding proteins have key roles in regulating the translation of specific mRNAs **(Fig. 3)**. Some of these proteins bind to mRNAs and promote translation either by recruiting initiation factors. Conversely, other RNA-binding proteins bind specific mRNAs and then interfere with translation and thus repress the process (Gebauer and Hentze, 2004; Hentze et al., 2018; Castello et al., 2013). Bicaudal-C functions as a translational repressor; therefore, I will provide a general overview of repression mechanisms in the follow section

Translational repression by RNA binding proteins

As stated above, in eukaryotes most translational regulation occurs at the initiation step. At initiation, eukaryotic initiation factor 4E (eIF4E) binds to the 5’cap of the mRNA and PABP binds to the poly(A) tail (Mangus et al., 2003; Gebauer and Hentze, 2004). Simultaneously these two can bind to eIF4G to effectively circularize the mRNA and enhance the recruitment of the large and small ribosomal subunits. Translational repressors can interfere with the process of initiation at each of these

steps and repress the mRNA from synthesizing protein. For example, specific repressor proteins can bind and prevent eIF4E from binding to the cap (Hinnebusch, 2017; Hinnebusch, 2014). Other proteins called 4E binding proteins (4EBPs), can bind to eIF4E and prevent it from binding to mRNA (Hinnebusch, 2014; Hinnebusch, 2017). Some mRNAs are deadenylated without being degraded. When the poly (A) tail is removed PABP cannot bind to the mRNAs, this prevents the recruitment of eIF4G and results in translational repression (Kawahara et al., 2008; Harvey et al., 2018).

In a variety of different biological contexts, specific sets of mRNAs are translationally repressed while the translation of most mRNAs is unaffected (Sonenberg and Hinnebusch, 2009; Groppo and Richter, 2009; Chekulaeva and Landthaler, 2016). These mRNA-specific repression processes are mediated by particular RNA binding proteins that confer specificity for the repression. There are many different types of these repressor proteins and in the following section I will provide some examples to illustrate some of the principals through which these proteins function.

A classic example of translational repressor is the *Drosophila* protein Bruno. In early *Drosophila* embryos Bruno recognizes specific binding sites in the 3'UTR of the *oskar* mRNA to repress translation (Chekulaeva et al., 2006). Mechanistically, Bruno interacts with the Cup protein and Cup directly binds to eIF4E to prevent it from binding to the 5'cap of the mRNA and initiating translation (Nakamura et al., 2004; Chekulaeva et al., 2006).

While translational regulation often happens at initiation, other steps of translation can be targets of repression as well. The Fragile X protein FMRP is an RNA-binding which binds to RNA targets in the neuron and represses their translation until

they can be properly localized (Siomi et al., 1993; Khandjian, 1999). FMRP contains two KH domains which are involved in RNA binding and mutations in these domains can cause Fragile X syndrome (Khandjian, 1999). While the exact mechanism of translational repression remains unknown, it has been shown that FMRP associates with polysomes and binds directly to the ribosome itself (Feng et al., 1997; Darnell et al., 2011). Thus, FMRP may repress translation by stalling polysome complexes and blocking the elongation phase of translation.

These and other examples also serve to illustrate two biological contexts where translational repression is a key regulator of gene expression: embryogenesis and neurons. In embryos before the maternal-to-zygotic transition, all regulation of gene expression must happen post-transcriptionally (Tadros and Lipshitz, 2009; Vastenhouw et al., 2019). This is because the embryo is not able to synthesize its own mRNAs until several hours after fertilization and thus all proteins expressed are translated from maternally deposited mRNAs. These mRNAs are thus under tight regulation both spatially and temporally: certain mRNAs must be expressed at specific times at specific locations in the embryo (Tadros and Lipshitz, 2009; Vastenhouw et al., 2019).

Bicaudal-C: an mRNA-specific translational repressor

Bicaudal-C is a conserved RNA-binding protein that was first discovered in *Drosophila* in analyses of mutations that caused defects in embryonic anterior-posterior patterning (Mohler and Wieschaus, 1986). Since its identification, homologs of the protein have been found throughout the animal kingdom. In embryos of the frog *Xenopus laevis*, the maternally deposited Bicaudal-C is localized to the vegetal

hemisphere of the oocyte and early embryo (Wessely and De Robertis, 2000). In adult organisms, defects in Bicaudal-C are associated with several human disease states such as major depressive disorder (Lewis et al., 2010; Bermingham et al., 2012; Ota et al., 2015) and polycystic kidney disease (Cogswell et al., 2003; Tran et al., 2010). Functional studies demonstrated that Bicaudal-C is a translational repressor; it binds to specific mRNA targets and disrupts their translation (Zhang et al., 2013; Zhang et al., 2014, Park et al., 2016).

In the section below, I will provide an overview of Bicaudal-C to give context to why this protein is the subject of my thesis. First, I will discuss the early studies of the protein and the first characterization of the protein as an RNA-binding protein. Then I will discuss the identification of specific mRNA targets identified in *Drosophila*, the frog *Xenopus laevis*, and mammals. I will then describe what is known about the structure and function of the protein. Finally, I will discuss the biological functions of Bicaudal-C in embryogenesis, organogenesis and highlight the biological consequences that result from defects in the protein and how these defects in many cases resemble specific human disease states.

Discovery of Bicaudal-C

One of the most critical topics in the field of oogenesis and embryogenesis is defining how the egg and embryo establish polarity to initiate body patterning. An early study published in 1966 on this topic focused on the genetic causes of the double-abdomen or “bicaudal” phenotype in *Drosophila* embryos (Bull, 1966). Through genetics, it was observed that this bicaudal phenotype could arise from maternal mutations that could be assigned to specific chromosomal regions, but not at high

resolution (Bull, 1966). While the exact mutations and genes responsible were not defined in these initial experiments, subsequent studies genetically identified the gene responsible, called it Bicaudal-C and cloned the cDNA for the encoded mRNA (Mohler and Wieschaus, 1986). Analysis of the protein sequence revealed that Bicaudal-C was a large protein composed of several putative RNA binding domains (Mahone et al., 1995). Consistent with this information, crude biochemical experiments demonstrated that the protein could efficiently bind to Poly (U) sequences *in vitro* (Mahone et al., 1995). The first evidence that Bicaudal-C functions as a translational repressor came from observations that *Drosophila* embryos lacking Bicaudal-C protein prematurely translated the *Oskar* mRNA (Saffman et al., 1998). Subsequent analysis revealed that Bicaudal-C was conserved throughout the animal kingdom and the first vertebrate homolog was identified and characterized from *Xenopus* (Wessely and De Robertis, 2000) and then shortly after a mammalian homologue was identified (Wessely et al., 2001).

Identification of Bicaudal-C target mRNAs

Though Bicaudal-C's importance in embryonic patterning was clear, there was little information from any organism about the mRNAs that it could potentially regulate. To address this deficit, an RNA immunoprecipitation (RIP) experiment was performed with Bicaudal-C with *Drosophila* embryos. Analysis of the Bicaudal-C associated mRNAs revealed an enrichment for mRNAs that encode proteins with varying functions in oogenesis and cytoskeletal regulation as well as Bicaudal-C itself (Chicoine et al., 2007). However, the Oskar mRNA was not identified as a

Bicaudal-C target mRNA and minimal characterization of these target mRNAs was performed (Chicoine et al., 2007).

The most extensive identification and characterization of Bicaudal-C target mRNAs has come from the Sheets lab using *Xenopus* embryos (Zhang et al., 2013). Specifically, 62 mRNAs were identified and validated as targets for Bicaudal-C mediated translational repression (Zhang et al., 2013). Importantly, the Cripto-1 (CR1) mRNA was identified among these targets as was predicted by previous studies from the Sheets lab (Zhang et al., 2013). Other studies from the Constam group of *Bicc1*$^{-/-}$ mice found that Bicaudal-C binds and regulates the adenylate cyclase-6 (AC6) mRNA in the kidneys to control cAMP levels (Piazzon et al., 2012).

The most well characterized target of Bicaudal-C is the Cripto-1 (CR1) mRNA in *Xenopus laevis* (Zhang et al., 2013; Zhang et al., 2014). In early *Xenopus* embryos, the CR1 mRNA is present throughout the embryo, however the mRNA is only translated in the animal pole leading to an animal accumulation of the CR1 protein (Dorey and Hill, 2006). Results from the Sheets lab indicated that this differential accumulation of protein was because translation of the CR1 mRNA was repressed in the vegetal pole of the embryo due to binding by Bicaudal-C (Zhang et al., 2013). Further studies revealed that a 32-nucleotide stem-loop located in the 3'UTR of *cripto-1* mRNA was both necessary and sufficient for Bicaudal-C binding and repression (Zhang et al., 2014). While it is known that the stem-loop structure is crucial for Bicaudal-C recognition, it is still unclear what specific nucleotide motifs within the 32nt binding site Bicaudal-C recognizes for specific and high affinity interaction.

Structural and functional analysis of Bicaudal-C protein

There are several conserved domains that can be identified in the Bicaudal-C protein Fig.4). The N-terminal region contains three KH and two KH-like domains and the C-terminal region contains an intrinsically disordered region (IDR) followed by a Sterile Alpha Motif (SAM) **(Fig. 4)**. The Sheets lab through a combination of MS2-tethered function assays and *in vivo* repression assays in *Xenopus* embryos, observed that the C-terminal region is both necessary and sufficient for translational repression, independent of the SAM domain. However, the exact mechanism of translational repression remains unclear. While the SAM domain has no direct role in repression it is involved in forming protein-protein interactions (Rothé et al., 2015); both in forming Bicaudal-C homopolymers and forming polymers with other proteins (Rothé et al., 2015; Rothé et al., 2018).

Results also from the Sheets lab demonstrated that the N-terminal is responsible for recognizing and binding specific mRNA targets (Zhang et al., 2014). In results that I will discuss in Chapter 3 of my thesis, we determined that of the three canonical KH domains, KH2 is the most critical for RNA binding, while KH1 and KH3 have limited importance for binding (Dowdle et al., 2019). However, it was also determined that the KH2 domain alone is not sufficient for specific binding to mRNA targets and thus specificity is likely encoded in other parts of the protein or potentially conferred through the binding of all three KH domains (Dowdle et al., 2019). Thus, there is still much to learn about how Bicaudal-C recognizes and binds to specific mRNA targets.

Additional studies in *Drosophila*, suggest that Bicaudal-C mediated repression may occur through regulation of mRNA polyadenylation. Specifically, it was observed that Bicaudal-C regulated translation of its own mRNA through binding to the 5’UTR of

the Bicaudal-C mRNA (Chicoine et al., 2007). Mechanistically focused experiments suggested that translational repression mediated by the NOT3/5 subunits of the CCR4-NOT regulatory complex to promote mRNA deadenylation (Chicoine et al., 2007). Interestingly, in contrast to *Drosophila* where Bicaudal-C binds to the 5'UTR of its own mRNA, in *Xenopus* Bicaudal-C has been shown to bind to the 3'UTR of its targets (Zhang et al., 2013; Zhang et al., 2014). This suggests that the mechanism of binding and repression may differ between invertebrate and vertebrate homologs of Bicaudal-C.

Biological role of Bicaudal-C in embryogenesis and adult disease

The maternally deposited Bicaudal-C functions in anterior-posterior patterning of the *Drosophila* embryo. In a heterozygous mutant of the Bicaudal-C gene of *Drosophila*, the embryos develop a bicaudal double abdomen phenotype (Mohler and Wieschaus, 1986). This phenotype can be at least partially attributed to premature translation of the *Oskar* mRNA in oocytes (Castagnetti and Ephrussi, 2003). Bicaudal-C mediated translational repression is critical to embryonic patterning in vertebrates as shown by experiments with *Xenopus* embryos (Park et al., 2016). In contrast to *Drosophila, Xenopus* embryos depleted of the maternal Bicaudal-C display expanded anterior structures as opposed to posterior structures. The formation of anterior structures and cell types in the absence of Bicaudal-C can be attributed to Bicaudal-C's regulation of specific maternal *Xenopus* mRNAs. Several of these Bicaudal-C regulated mRNAs encode cell fate determinant proteins; the previously discussed Cripto-1, a member of the Nodal signaling pathway, and Wnt11b, a member of the Wnt signaling pathway (Zhang et al., 2013; Park et al., 2016). Overexpression of these proteins results in the formation of excess anterior structures similar to phenotypes observed upon Bicaudal-C

depletion (Park et al., 2016) and consistent with the model that Bicaudal-C normally represses the mRNAs encoding these proteins **(Fig. 5)**.

Bicaudal-C protein is expressed during the later stages of development after the activation of zygotic transcription and has critical functions in organ homeostasis and left-right patterning. In *Bicc1*$^{-/-}$ mice, the homozygous knockout is often lethal (Maisonneuve et al., 2009). Analysis of the embryos that do survive reveals that they possess left-right patterning defects either showing complete situs inversions or situs ambiguus, that is the organs showing a reversal of their normal position in the body cavity (Maisonneuve et al., 2009). Left-right patterning in mice is controlled by the Nodal pathway (Blum et al., 2014; Schweikert et al., 2018) and further experiments in mice provided evidence that Bicaudal-C regulates Wnt signaling pathway indicating that Bicaudal-C regulatory activity is conserved through vertebrate species (Maisonneuve et al., 2009).

In addition, defects in zygotic Bicaudal-C has been linked to organ aberrations associated with human disease states, most notably polycystic kidney disease. In both mouse and *Xenopus*, Bicaudal-C is highly expressed in the kidneys and disrupting Bicaudal-C expression in embryos gives rise to defective kidneys (Cogswell et al., 2003; Tran et al., 2010). The specific defects include formation of fluid filled cysts and aberrant tubule formation, phenotypes similar to human patients suffering from polycystic kidney disease (Cogswell et al., 2003; Rothé et al., 2019). In *Xenopus* it was found that by reducing expression of the miRNA miR-17 that the phenotype was reduced indicating that Bicaudal-C may be antagonizing miR-17 in the adult *Xenopus* (Tran et al., 2010). More recently, Bicaudal-C has been identified in genome-wide association

studies (GWAS) when screening for genes that may impact diseases such as myopia (Yoshikawa et al., 2014) and major depression disorder (Lewis et al., 2010; Bermingham et al., 2012; Ota et al., 2015). While Bicaudal-C fusions with the FGFR2 receptor in the liver have been shown to be a causative agent for cholangiocarcinoma, a cancer of the bile duct (Wu et al., 2013; Arai et al., 2014; Li et al., 2019).

Overview and Significance

My thesis is focused on the mechanism of Bicaudal-C's RNA binding activities. Specifically, how does Bicaudal-C bind to its target mRNAs and what sequences in the 32-nucleotide *cripto-1* stem-loop does Bicaudal-C specifically recognize. In the first chapter a broad overview of RNA-binding proteins, post-transcriptional mechanisms of gene regulation, translational repression, and how all three of these topics come together in the Bicaudal-C protein. I then described what is currently known about Bicaudal-C function and its biological importance. In Chapter 2, I discuss a method used to characterize RNA binding proteins and its application to studying Bicaudal-C binding to specific RNAs. This was published as Dowdle *et al* 2017. Chapter 3 describes experiments to show that the KH2 domain in Bicaudal-C is a critical feature of the protein for mRNA binding. This was published as Dowdle *et al* 2019. Chapter 4 presents experiments that Bicaudal-C binds to mRNA substrates via a bipartite binding site. The results from this chapter will be combined with other results and submitted for publication. Finally, Chapter 5 summarizes the conclusions from my research and outlines future directions and experiments.

Figure 1. Diagram of the mature eukaryotic mRNA.
A mature eukaryotic mRNA with a hypothetical RNA-binding protein (RBP) bound to Binding Elements (BE) in the 3'UTR. The mature mRNA has a 5' cap and a 3' Poly(A) tail. Binding elements for RBP's are typically found in both the 5'UTR and 3'UTR of the RNA.

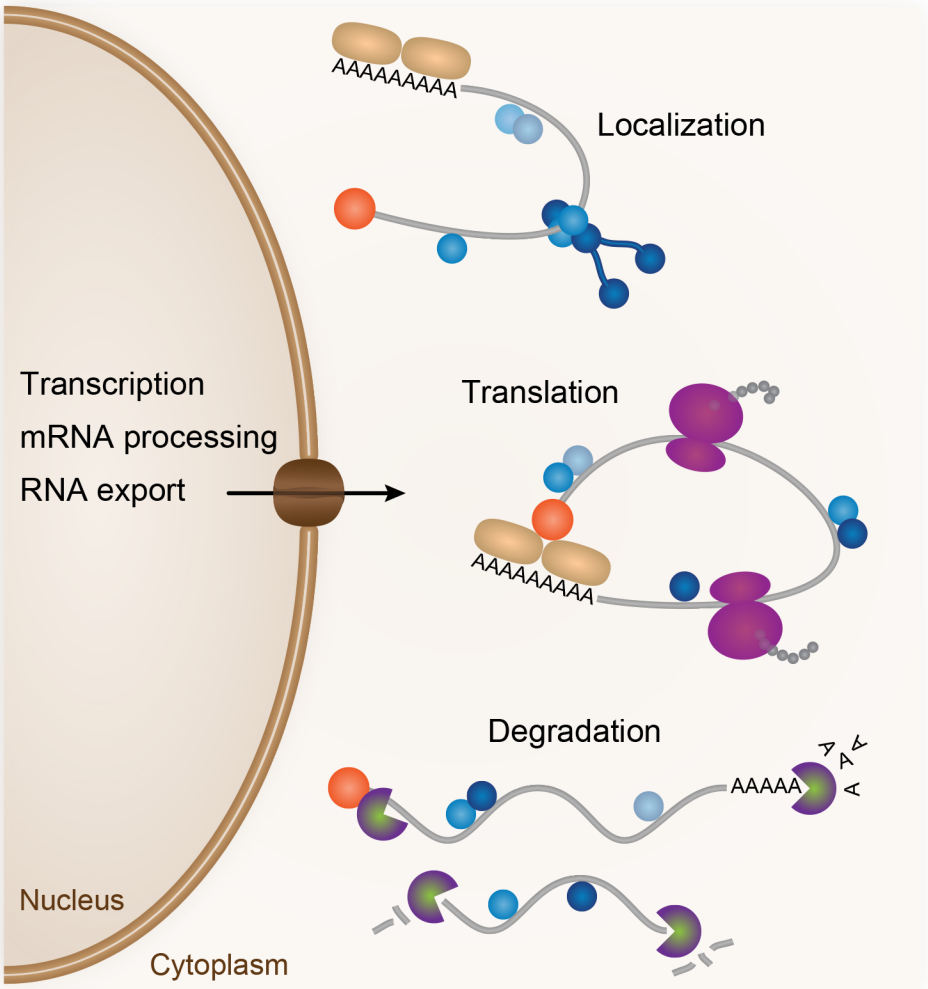

Figure 2. Examples of cytoplasmic RNA regulation in eukaryotic cells.
RNA binding proteins (blue) can bind to specific sequences in the RNA and control processes such as localization, translation, and degradation in the cytoplasm.

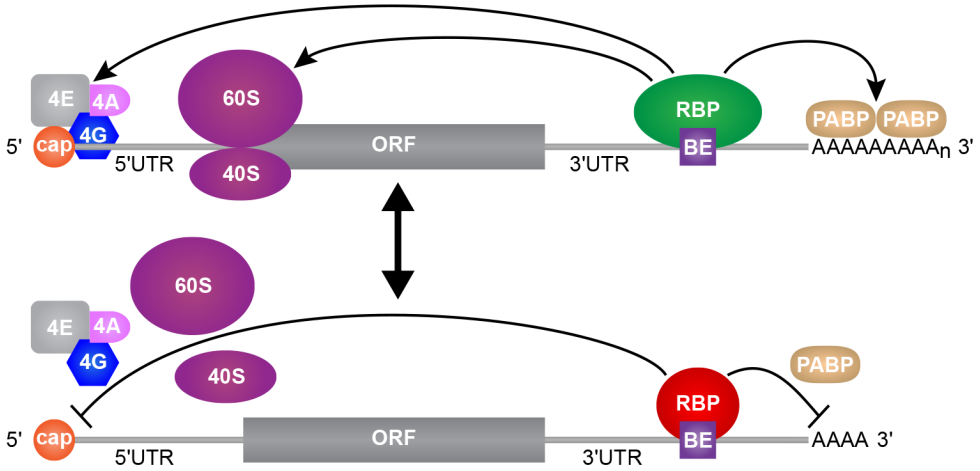

Figure 3. RNA binding proteins can activate or inhibit translation.
RNA binding proteins can act as activators (Green) and recruit factors needed for translation such as the 4E initiation complex, the 60S and 40S ribosomal subunits, or Poly(A) binding protein (PABP) (Wells et al., 1998; Tahmasebi et al., 2019; Tahmasebi et al., 2018; Hershey et al., 2019). RNA binding proteins can function as repressors (red) often prevent recruitment of these factors (Gebauer and Hentze, 2004; Hentze et al., 2018; Castello et al., 2013).

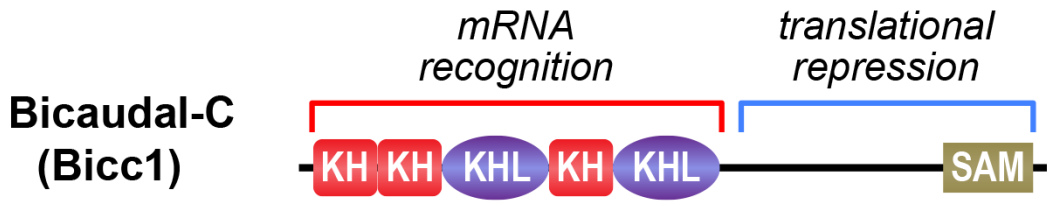

Figure 4. Diagram of Bicaudal-C (Bicc1).
A diagram of Bicc1 highlighting the mRNA recognition region and the translational repression region (Zhang et al., 2014). The canonical KH domains are in red while the KH-like domains are in purple. The SAM domain is in bronze.

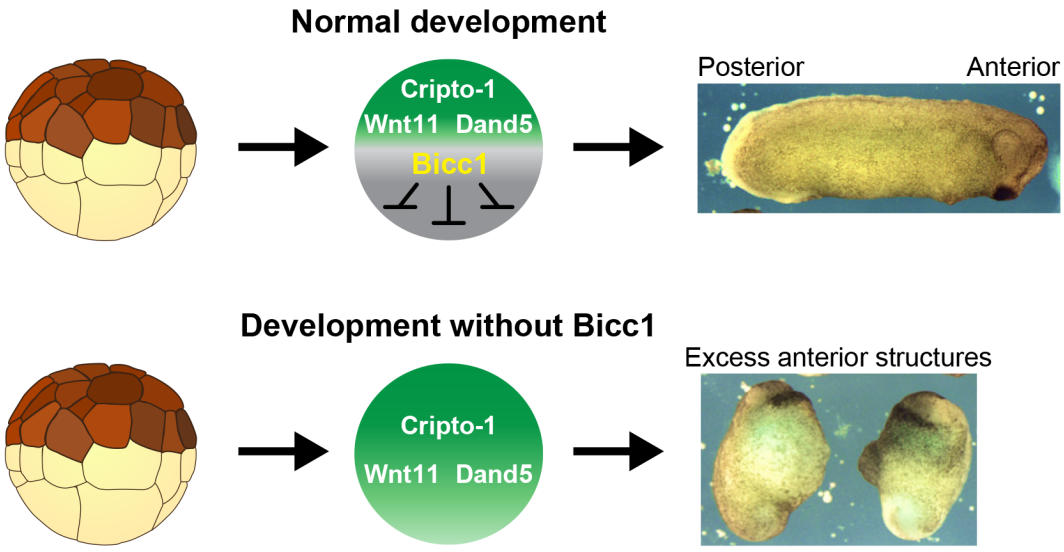

Figure 5. Translational repression mediated by Bicc1 has a critical impact on proper development of the *Xenopus* embryo.
Bicc1 is expressed in the bottom half of the *Xenopus* embryo and represses specific target mRNAs, such as Cripto-1 Park et al., 2016). When this repression occurs normally (wild type animal –at top) proper development of posterior and anterior structures occurs in the embryo (top). When this repression is disrupted by depleting Bicc1 from the embryo, the proteins from Bicc1's target mRNAs are expressed throughout the embryo resulting in an excess of anterior structures (bottom) (Park et al., 2016).

REFERENCES

Arai, Y., Totoki, Y., Hosoda, F., Shirota, T., Hama, N., Nakamura, H., Ojima, H., Furuta, K., Shimada, K., Okusaka, T., et al. (2014). Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology* 59, 1427–34.

Aviv, T., Lin, Z., Ben-Ari, G., Smibert, C. A. and Sicheri, F. (2006). Sequence-specific recognition of RNA hairpins by the SAM domain of Vts1p. *Nat Struct Mol Biol* 13, 168–76.

Aviv, T., Lin, Z., Lau, S., Rendl, L. M., Sicheri, F. and Smibert, C. A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nat Struct Biol* 10, 614–21.

Bermingham, R., Carballedo, A., Lisiecka, D., Fagan, A., Morris, D., Fahey, C., Donohoe, G., Meaney, J., Gill, M. and Frodl, T. (2012). Effect of genetic variant in BICC1 on functional and structural brain changes in depression. *Neuropsychopharmacology* 37, 2855–62.

Blum, M., Feistel, K., Thumberger, T., Schweikert, A. (2014) The evolution and conservation of left-right patterning mechanisms. *Development* 141, 1603-13.

Bull, A. L. (1966). Bicaudal, a genetic factor which affects the polarity of the embryo in Drosophila melanogaster. *Journal of Experimental Zoology* 161, 221–241.

Castagnetti, S. and Ephrussi, A. (2003). Orb and a long poly(A) tail are required for efficient oskar translation at the posterior pole of the Drosophila oocyte. *Development* 130, 835–43.

Castello, A., Fischer, B., Hentze, M. W. and Preiss, T. (2013). RNA-binding proteins in Mendelian disease. *Trends Genet* 29, 318–27.

Chekulaeva, M., Hentze, M. W. and Ephrussi, A. (2006). Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles. *Cell* 124, 521–33.

Chekulaeva, M. and Landthaler, M. (2016). Eyes on Translation. *Mol Cell* 63, 918–25.

Chicoine, J., Benoit, P., Gamberi, C., Paliouras, M., Simonelig, M. and Lasko, P. (2007). Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev Cell* 13, 691–704.

Clerte, C. and Hall, K. B. (2009). The domains of polypyrimidine tract binding protein have distinct RNA structural preferences. *Biochemistry* 48, 2063–74.

Cogswell, C., Price, S. J., Hou, X., Guay-Woodford, L. M., Flaherty, L. and Bryda, E. C. (2003). Positional cloning of jcpk/bpk locus of the mouse. *Mamm Genome* 14, 242–9.

Cotton, C. U., Hobert, M. E., Ryan, S. and Carlin, C. R. (2013). Basolateral EGF receptor sorting regulated by functionally distinct mechanisms in renal epithelial cells. *Traffic* 14, 337–54.

Darnell, J. C., Van, D. S. J., Zhang, C., Hung, K. Y., Mele, A., Fraser, C. E., Stone, E. F., Chen, C., Fak, J. J., Chi, S. W., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146, 247–61.

Dominguez, D., Freese, P., Alexis, M. S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N. J., Van, N. E. L., Pratt, G. A., et al. (2018). Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol Cell* 70, 854–867.e9.

Dorey, K. and Hill, C. S. (2006). A novel Cripto-related protein reveals an essential role for EGF-CFCs in Nodal signaling in Xenopus embryos. *Dev Biol* 292, 303–16.

Dowdle, M. E., Imboden, S. B., Park, S., Ryder, S. P. and Sheets, M. D. (2017). Horizontal Gel Electrophoresis for Enhanced Detection of Protein-RNA Complexes. *J Vis Exp*.

Dowdle, M. E., Park, S., Blaser, I. S., Fox, C. A., Houston, D. W. and Sheets, M. D. (2019). A single KH domain in Bicaudal-C links mRNA binding and translational repression functions to maternal development. *Development* 146 (10).

Du, T. G., Schmid, M. and Jansen, R. P. (2007). Why cells move messages: the biological functions of mRNA localization. *Semin Cell Dev Biol* 18, 171–7.

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.

Feng, Y., Absher, D., Eberhart, D. E., Brown, V., Malter, H. E. and Warren, S. T. (1997). FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association. *Mol Cell* 1, 109–18.

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

Garneau, N. L., Wilusz, J. and Wilusz, C. J. (2007). The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* 8, 113–26.

Gebauer, F. and Hentze, M. W. (2004). Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* 5, 827–35.

Gerstberger, S., Hafner, M. and Tuschl, T. (2014). A census of human RNA-binding proteins. *Nat Rev Genet* 15, 829–45.

Gherzi, R., Lee, K. Y., Briata, P., Wegmüller, 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, 571–83.

Groppo, R. and Richter, J. D. (2009). Translational control from head to tail. *Curr Opin Cell Biol* 21, 444–51.

Harvey, R. F., Smith, T. S., Mulroney, T., Queiroz, R., Pizzinga, M., Dezi, V., Villenueva, E., Ramakrishna, M., Lilley, K. S., & Willis, A. E. (2018). Trans-acting translational regulatory RNA binding proteins. *Wiley interdisciplinary reviews. RNA*, 9, e1465.

Hentze, M. W., Castello, A., Schwarzl, T. and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* 19, 327–341.

Hershey, J. W. B., Sonenberg, N. and Mathews, M. B. (2019). Principles of Translational Control. *Cold Spring Harb Perspect Biol* 11,

Hinnebusch, A. G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem Sci* 42, 589–611.

Hinnebusch, A. G. (2014). The scanning mechanism of eukaryotic translation initiation. *Annu Rev Biochem* 83, 779–812.

Hinnebusch, A. G., Ivanov, I. P. and Sonenberg, N. (2016). Translational control by 5’-untranslated regions of eukaryotic mRNAs. *Science* 352, 1413–6.

Holt CE, Martin KC, Schuman EM. (2019). Local translation in neurons: visualization and function. Nat Struct Mol Biol. 2019 Jul;26(7):557-566.

Jonas, S. and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* 16, 421–33.

Kawahara, H., Imai, T., Imataka, H., Tsujimoto, M., Matsumoto, K., & Okano, H. (2008). Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. *The Journal of cell biology*, 181, 639–653.

Khandjian, E. W. (1999). Biology of the fragile X mental retardation protein, an RNA-binding protein. *Biochem Cell Biol* 77, 331–42.

Lewis, C. M., Ng, M. Y., Butler, A. W., Cohen-Woods, S., Uher, R., Pirlo, K., Weale, M. E., Schosser, A., Paredes, U. M., Rivera, M., et al. (2010). Genome-wide association

study of major recurrent depression in the U.K. population. *Am J Psychiatry* 167, 949–57.

Li, F., Peiris, M. N. and Donoghue, D. J. (2019). Functions of FGFR2 corrupted by translocations in intrahepatic cholangiocarcinoma. *Cytokine Growth Factor Rev*.

Lunde, B. M., Moore, C. and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* 8, 479–90.

Mahone, M., Saffman, E. E. and Lasko, P. F. (1995). Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. *EMBO J* 14, 2043–55.

Maisonneuve, C., Guilleret, I., Vick, P., Weber, T., Andre, P., Beyer, T., Blum, M. and Constam, D. B. (2009). Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. *Development* 136, 3019–30.

Mangus, D. A., Evans, M. C. and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 4, 223.

Martin, K. C. and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136, 719–30.

Mayr, C. (2019). What Are 3’ UTRs Doing?. *Cold Spring Harb Perspect Biol* 11,

Mitchell, S. F. and Parker, R. (2014). Principles and properties of eukaryotic mRNPs. *Mol Cell* 54, 547–58.

Mohler, J. and Wieschaus, E. F. (1986). Dominant maternal-effect mutations of Drosophila melanogaster causing the production of double-abdomen embryos. *Genetics* 112, 803–22.

Nakamura, A., Sato, K. and Hanyu-Nakamura, K. (2004). Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. *Dev Cell* 6, 69–78.

Nicastro, G., Candel, A. M., Uhl, M., Oregioni, A., Hollingworth, D., Backofen, R., Martin, S. R. and Ramos, A. (2017). Mechanism of β-actin mRNA Recognition by ZBP1. *Cell Rep* 18, 1187–1199.

Nickless, A., Bailis, J. M. and You, Z. (2017). Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell Biosci* 7, 26.

Ota, K. T., Andres, W., Lewis, D. A., Stockmeier, C. A. and Duman, R. S. (2015). BICC1 expression is elevated in depressed subjects and contributes to depressive behavior in rodents. *Neuropsychopharmacology* 40, 711–8.

Paquin, N. and Chartrand, P. (2008). Local regulation of mRNA translation: new insights from the bud. *Trends Cell Biol* 18, 105–11.

Park, S., Blaser, S., Marchal, M. A., Houston, D. W. and Sheets, M. D. (2016). A gradient of maternal Bicaudal-C controls vertebrate embryogenesis via translational repression of mRNAs encoding cell fate regulators. *Development* 143, 864–71.

Piazzon, N., Maisonneuve, C., Guilleret, I., Rotman, S. and Constam, D. B. (2012). Bicc1 links the regulation of cAMP signaling in polycystic kidneys to microRNA-induced gene silencing. *J Mol Cell Biol* 4, 398–408.

Romanelli, M. G., Diani, E. and Lievens, P. M. (2013). New insights into functional roles of the polypyrimidine tract-binding protein. *Int J Mol Sci* 14, 22906–32.

Rothé, B., Leal-Esteban, L., Bernet, F., Urfer, S., Doerr, N., Weimbs, T., Iwaszkiewicz, J. and Constam, D. B. (2015). Bicc1 Polymerization Regulates the Localization and Silencing of Bound mRNA. *Mol Cell Biol* 35, 3339–53.

Rothé, B., Gagnieux, C., Leal-Esteban, L. C. and Constam, D. B. (2019). Role of the RNA-binding protein Bicaudal-C1 and interacting factors in cystic kidney diseases.. *Cell Signal* 109499.

Rothé, B., Leettola, C. N., Leal-Esteban, L., Cascio, D., Fortier, S., Isenschmid, M., Bowie, J. U. and Constam, D. B. (2018). Crystal Structure of Bicc1 SAM Polymer and Mapping of Interactions between the Ciliopathy-Associated Proteins Bicc1, ANKS3, and ANKS6. *Structure* 26, 209–224.e6.

Saffman, E. E., Styhler, S., Rother, K., Li, W., Richard, S. and Lasko, P. (1998). Premature translation of oskar in oocytes lacking the RNA-binding protein bicaudal-C. *Mol Cell Biol* 18, 4855–62.

Schweickert, A., Ott, T., Kurz, S., Tingler, M., Maerker, M., Fuhl, F., & Blum, M. (2017). Vertebrate Left-Right Asymmetry: What Can Nodal Cascade Gene Expression Patterns Tell Us?. *Journal of cardiovascular development and disease* 5, 1.

Sharma, S., Maris, C., Allain, F. H. and Black, D. L. (2011). U1 snRNA directly interacts with polypyrimidine tract-binding protein during splicing repression. *Mol Cell* 41, 579–88.

Sheets MD, Fox CA, Dowdle ME, Blaser SI, Chung A, Park S. (2017). Controlling the Messenger: Regulated Translation of Maternal mRNAs in *Xenopus laevis* Development. Adv Exp Med Biol. 2017;953:49-82. Review.

Siomi, H., Siomi, M. C., Nussbaum, R. L. and Dreyfuss, G. (1993). The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 74, 291–8.

Skrisovska, L., Bourgeois, C. F., Stefl, R., Grellscheid, S. N., Kister, L., Wenter, P., Elliott, D. J., Stevenin, J. and Allain, F. H. (2007). The testis-specific human protein RBMY recognizes RNA through a novel mode of interaction. *EMBO Rep* 8, 372–9.

Sonenberg, N. and Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136, 731–45.

Tadros, W. and Lipshitz, H. D. (2009). The maternal-to-zygotic transition: a play in two acts. *Development* 136, 3033–42.

Tahmasebi, S., Khoutorsky, A., Mathews, M. B. and Sonenberg, N. (2018). Translation deregulation in human disease. *Nat Rev Mol Cell Biol* 19, 791–807.

Tahmasebi, S., Sonenberg, N., Hershey, J. W. B. and Mathews, M. B. (2019). Protein Synthesis and Translational Control: A Historical Perspective. *Cold Spring Harb Perspect Biol* 11,

Towler, B. P. and Newbury, S. F. (2018). Regulation of cytoplasmic RNA stability: Lessons from Drosophila. *Wiley Interdiscip Rev RNA* 9, e1499.

Tran, U., Zakin, L., Schweickert, A., Agrawal, R., Döger, R., Blum, M., De Robertis, E. M. and Wessely, O. (2010). The RNA-binding protein bicaudal C regulates polycystin 2 in the kidney by antagonizing miR-17 activity. *Development* 137, 1107–16.

Vastenhouw, N. L., Cao, W. X. and Lipshitz, H. D. (2019). The maternal-to-zygotic transition revisited. *Development* 146.

Wells, S. E., Hillner, P. E., Vale, R. D. and Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* 2, 135–40.

Wessely, O. and De Robertis, E. M. (2000). The Xenopus homologue of Bicaudal-C is a localized maternal mRNA that can induce endoderm formation. *Development* 127, 2053–62.

Wessely, O., Tran, U., Zakin, L. and De Robertis, E. M. (2001). Identification and expression of the mammalian homologue of Bicaudal-C. *Mech Dev* 101, 267–70.

Wu, Y. M., Su, F., Kalyana-Sundaram, S., Khazanov, N., Ateeq, B., Cao, X., Lonigro, R. J., Vats, P., Wang, R., Lin, S. F., et al. (2013). Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov* 3, 636–47.

Yoshikawa, M., Yamashiro, K., Miyake, M., Oishi, M., Akagi-Kurashige, Y., Kumagai, K., Nakata, I., Nakanishi, H., Oishi, A., Gotoh, N., et al. (2014). Comprehensive replication of the relationship between myopia-related genes and refractive errors in a large Japanese cohort. *Invest Ophthalmol Vis Sci* 55, 7343–54.

Zhang, J. and King, M. L. (1996). Xenopus VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. *Development* 122, 4119–29.

Zhang, Y., Cooke, A., Park, S., Dewey, C. N., Wickens, M. and Sheets, M. D. (2013). Bicaudal-C spatially controls translation of vertebrate maternal mRNAs. *RNA* 19, 1575–82.

Zhang, Y., Park, S., Blaser, S. and Sheets, M. D. (2014). Determinants of RNA binding and translational repression by the Bicaudal-C regulatory protein. *J Biol Chem* 289, 7497–504.

Chapter 2:

Horizontal gel electrophoresis for enhanced detection of protein-RNA complexes

Megan E. Dowdle[1], Susanne Imboden Blaser[1], Sookhee Park[1], Sean P. Ryder[2] and Michael D. Sheets[1]

This chapter was published:

Dowdle, M. E., Imboden, S. B., Park, S., Ryder, S. P. and Sheets, M. D. (2017). Horizontal Gel Electrophoresis for Enhanced Detection of Protein-RNA Complexes. *J Vis Exp*.

I performed all the experiments and contributed all data. I wrote the paper with Sean Ryder and Michael Sheets.

1Department of Biomolecular Chemistry, University of Wisconsin – Madison
[2]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School

SHORT ABSTRACT:

Native polyacrylamide gel electrophoresis is a fundamental tool for analyzing RNA-protein interactions. Traditionally most experiments have used vertical gels. However, horizontal gels provide several advantages, such as the opportunity to monitor complexes during electrophoresis. We provide a detailed protocol for generating and using horizontal native gel electrophoresis.

LONG ABSTRACT:

Native polyacrylamide gel electrophoresis is a fundamental tool of molecular biology that has been used extensively for the biochemical analysis of RNA-protein interactions. These interactions have been traditionally analyzed with polyacrylamide gels generated between two glass plates and samples electrophoresed vertically. However, polyacrylamide gels cast in trays and electrophoresed horizontally offers several advantages. For example, horizontal gels used to analyze complexes between fluorescent RNA substrates and specific proteins can be imaged multiple times as electrophoresis progresses. This provides the unique opportunity to monitor RNA-protein complexes at several points during the experiment. In addition, horizontal gel electrophoresis makes it possible to analyze many samples in parallel and this can greatly facilitate time course experiments as well as analyzing multiple reactions simultaneously to compare different reaction components and conditions. Here we provide a detailed protocol for generating and using horizontal native gel electrophoresis for analyzing RNA-Protein interactions.

INTRODUCTION:

Electrophoretic mobility shift assays (EMSAs) have proven to be an invaluable biochemical tool to analyze specific protein-nucleic acid interactions (Ryder et al., 2008; Dahlberg et al., 1969; Hellman and Fried, 2007). These assays can provide important information regarding the binding affinity of proteins to RNA or DNA (Hellman and Fried, 2007), the component stoichiometry of nucleic acid-protein complexes (Ryder et al., 2008) and provide important new insights about the binding specificity of RNA binding proteins via substrate competition experiments (Ryder et al., 2008).

The traditional experimental setup for these assays consists of mixing purified protein with a radiolabeled RNA substrate. The resulting complexes are then analyzed with non-denaturing (native) polyacrylamide gels poured between two glass plates followed by sample electrophoresis in a vertical apparatus (Hellman and Fried, 2007). While this approach has been used exhaustively to provide important insights in the biochemical mechanisms that underlie the binding of proteins to nucleic acids, it also has several limitations. For example, this basic strategy has relatively limited throughput and it is not readily adaptable for applications that require analyzing many binding reactions in parallel. In addition, with the traditional vertical apparatus it is challenging to monitor complexes at multiple times during electrophoresis (Hellman and Fried, 2007; Pagano et al., 2009).

Here we present an adaptation of the EMSA assay that uses native polyacrylamide gels cast in a flatbed apparatus, horizontal electrophoresis and fluorescently labeled RNA substrates (Pagano et al., 2009; Royer and Scarlata, 2008). The incorporation of these relatively simple modifications to the basic strategy provides

some powerful advantages. In particular, the horizontal flatbed format easily lends itself to analyzing dozens of samples simultaneously (Pagano et al., 2009). Also, for some RNA-protein complexes, such as those formed between the Bicaudal-C protein and its RNA substrate electrophoresis in a horizontal gel provides an increased ability to resolve distinct RNA-protein complexes and discriminate these from unbound RNA substrate.

PROTOCOL:

Horizontal Electromobility Shift Assays

1.0 Preparation of the horizontal native polyacrylamide gel (Pagano et al., 2009).

1.1. Materials needed.

a. Horizontal gel apparatus. The apparatus that we use is an Owl Gel Box 37cmx24cm with a 27cm x 21cm tray (similar models are sold by Thermo Scientific). It has the capacity for two 24-well combs. This setup provides a total of 48 samples that can be analyzed simultaneously. Another example is the BioRad Sub-Cell GT system (170-4483) with a 15X20cm tray, gel caster that can hold at least three 20-well combs.

b. 40% Acrylamide-Bis 19:1

c. 5X TBE (Tris-Borate-EDTA pH 8.0)

d. TEMED

e. 10% APS (ammonium persulfate)

1.2 Generating a 10% native acrylamide gel:

a. In a 500mL flask add 80mL of 5xTBE, 100mL of 40% Acrylamide, and water to a final volume of 400ml.
b. Add 4mL of 10%APS and mix well.
c. Add 400uL of TEMED and mix well.
d. Pour into the horizontal casting tray and allow gel to polymerize for 30 minutes. Gels can be cast in a fume hood to speed up polymerization.
e. After polymerization a thin layer of liquid will remain on top of the gel. Pour the liquid off and rinse with running buffer.
f. Remove the comb(s) carefully and rinse the wells with running buffer using a syringe.

1.3 Preparation of running buffer.

a. The Owl Gel Box needs 2L of 1X TBE running buffer.
b. Dilute 400mL of 5X TBE with 1600mL and mix.
c. Place 1X TBE buffer in cold room to cool.

1.4 Pre-Run the gel:

a. Add 2L of cold running buffer to the gel box and insert the gel.
b. Pre-run the gel at 120V for 1 hour at 4º C in a cold room. During this time, prepare the binding reaction

2.0 Binding reaction (Zhang et al., 2014)

2.1 Reaction components

a. 20mM DTT
b. 10mM BSA
c. 10mM yeast tRNAs

d. 5X binding buffer (50mM HEPES pH 8.0, 5mM EDTA, 250mM KCl, 0.2% Tween)
e. 100nM fluorescent RNA
f. Concentrated Bicaudal-C protein
g. 50% glycerol in DEPC-treated H2O
h. DEPC treated H2O

2.2 Binding reaction assembly

a. Assemble components in an RNAse free microcentrifuge tube
b. 5uL 10mM BSA
c. 5uL 20mM DTT
d. 5uL 10mM yeast tRNAs
e. 10uL of 5x Binding Buffer
f. Add protein to a final concentration of 250nM in 50uL
g. Add water to a final volume of 45uL
h. Add 5uL of fluorescent RNA substrate
i. Mix and incubate in the dark for 30 minutes at room temperature.

3.0 Analyzing samples on the native horizontal polyacrylamide gel

a. To each reaction add 15uL of 50% glycerol and mix gently.
b. Carefully add 30uL of each reaction to the gel.
c. After connecting the power supply to the gel apparatus switch on the power and adjust the voltage.
d. For the Owl Gel apparatus, we use 120V.
e. To prevent damaging or bleaching of the fluorescently labeled RNA, electrophoresis is conducted in the dark.

f. Electrophoresis times will vary depending upon the specifics of the RNA-protein complex being analyzed.

g. A major advantage of the horizontal EMSA analysis is that electrophoresis can be stopped, the gel imaged and then the gel can be placed back into the apparatus and electrophoresis can be continued for longer times.

4.0. Imaging the Gel

a. Analysis can be done with any instrument designed for detection and imaging fluorescent substrates. We use a Typhoon FLA 9000 equipped with version 1.2 software.

b. Place the gel on the Fluor Stage.

c. Adjust imager settings:

1. FAM (wavelength 473)
2. 750 Photomultiplier voltage (PMT)
3. 100uM pixel size

d. Scan the gel to capture image.

REPRESENTATIVE RESULTS:

To demonstrate the power and versatility of the horizontal native gel electrophoresis we analyzed the binding of the *Xenopus* Bicaudal-C (Bicc1) protein to a fluorescently labeled RNA containing a Bicc1 binding site. Bicc1 proteins function as mRNA-specific translational repressors to control cell fate decisions during the maternal stages of animal development (Gamberi and Lasko, 2012; Saffman et al., 1998; Chicoine et al., 2007; Zhang et al., 2013) as well as the functions of specific organs in

adults (Maisonneuve et al., 2009; Tran et al., 2010). Our work in *Xenopus* identified the first binding site for a Bicc1 protein (Zhang et al., 2014), a 32-nucleotide region from the 3'UTR of the Cripto-1 mRNA (Zhang et al., 2013; Zhang et al., 2014). This binding site was defined using a combination of techniques; RNAse protection and footprinting assays, *in vivo* binding experiments and *in vitro* EMSA assays (Zhang et al., 2014).

Purified *Xenopus* Bicc1 protein was mixed with a fluorescently labeled 32 nucleotide RNA; the Bicc1 binding site from the Cripto-1 mRNA described above (Zhang et al., 2014). In a separate reaction Bicc1 protein was mixed with a fluorescently labeled 32 nucleotide RNA derived from the 3'UTR of the cyclin B1 mRNA. Bicc1 does not bind to the cyclin B1 mRNA (Zhang et al., 2013; Zhang et al., 2014) and this RNA serves as a negative control. Half of each reaction was analyzed on a vertical native polyacrylamide gel (**Fig.1A**) while the other portion was analyzed in parallel using horizontal native gel (**Fig.1B**). It is readily apparent using either method that Bicc1 binds to the RNA Cripto-1 RNA and forms a specific complex as it does not bind the cyclin B1 negative control RNA. However, the Bicc1-Cripto-1 complexes are significantly more distinct when analyzed with the horizontal gel, both facilitating their detection and allowing the discrimination of protein-RNA complexes from unbound RNA. After imaging, the gel was placed back into the apparatus and electrophoresed for an additional three and a half hours. After re-imaging the gel, the complexes had migrated further and were still readily detectable, indicating their stability (**Fig.1C**). This ability to continue electrophoresis and analysis after initial imaging is not possible or would be challenging with the vertical gel format.

DISCUSSION:

Native polyacrylamide gels are an invaluable tool for investigating protein-RNA interactions and traditionally these gels are electrophoresed vertically (Dahlberg et al., 1969; Hellman and Fried, 2007). We have used a modification of the protocol that substitutes native polyacrylamide gels created and electrophoresed horizontally (Ryder et al., 2008; Pagano et al., 2009; Pagano et al., 2011). These changes used in combination with fluorescently labeled RNA substrates provide numerous advantages for the biochemical analysis of protein-RNA binding interactions. These advantages include the following: mixing of buffers is simplified compared to a vertical apparatus which requires a pump to transfer buffer from the lower chamber to the upper, horizontal gels can be generated with larger wells that provide an increased capacity for analyzing larger samples, electrophoresis on horizontal gels often provide improved resolution of RNA-protein complexes compared to vertical gels, and the horizontal gel electrophoresis provides the ability to image experiments at multiple time points during analysis.

Another major advantage is that the horizontal format lends itself to creating gels with large number of wells and therefore provides the capacity to analyze multiple samples in parallel (Ryder et al., 2008) . This feature significantly expands their use as a biochemical tool for analyzing RNA-protein interactions. For example, the ability to analyze multiple samples facilitates investigating the quantitative parameters of protein-RNA complex formation through the use of competition experiments (Pagano et al., 2011), k_{off} analysis (Pagano et al., 2011) as well as experiments designed to define the stoichiometry of the binding complex components (Pagano et al., 2011). In addition, the

use of fluorescent RNA substrates makes it possible to perform multiple analyses, such as fluorescent polarization binding assays from a single reaction and obtain both quantitative and qualitative binding data (Royer and Scarlata 2008; Pagano et al., 2011).

To obtain the best results from the horizontal gel setup, there are a variety of parameters that can be adjusted to improve the quality of the complexes for imaging. These include changing the acrylamide percentage of the gel in order to make it easier to handle and enhance the separation of protein-RNA complexes from free RNA. However, it can be difficult to handle low percentage acrylamide gels due to the size. We find it easiest to use gels that have an acrylamide concentration of 7.5% or greater. Also, if necessary the electrophoresis conditions can be changed to improve stability of protein-RNA complexes; adjusting the electrophoresis voltage, running the gel at room temperate instead of at 4º and adjusting the pre-run conditions.

Horizontal native gel electrophoresis has the potential to be used as a tool for molecular screens to identify compounds that target clinically relevant RNA-protein interactions (Foley et al., 2016). These screens typically use solution biochemical assays to identify relevant compounds for further study. The high throughput potential of the horizontal format provides an opportunity to apply native gel electrophoresis either as part of the initial screening process or as a tool for secondary validation of potential candidates.

ACKNOWLEDGMENTS:

We thank Laura Vanderploeg for preparing figures. Work in the Sheets lab is supported by NSF grant 1050395 and NIH grant (R21HD076828). Work in the Ryder lab is supported by NIH grants R01GM117237 and R01GM117008. Megan Dowdle is supported by a SciMed GRS Advanced Opportunity Fellowship through University of Wisconsin-Madison Graduate School and Biotechnology Training Program through the National Institute of General Medical Sciences of the National Institutes of Health (T32GM008349).

DISCLOSURES:

The authors have nothing to disclose.

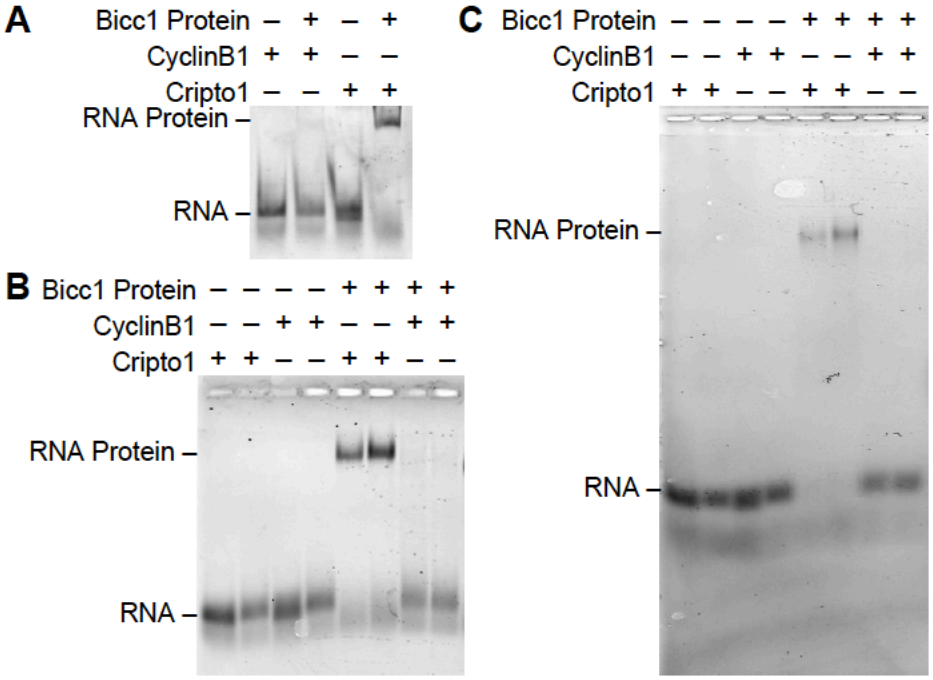

Figure 1. A comparison of the vertical and the horizontal native gels for analyzing the binding of Bicc1 to Cripto1 RNA.
Bicc1 binding experiments were conducted with a 32-nucleotide fluorescent CyclinB1 RNA negative control or a fluorescent 32-nucleotide Cripto1 RNA positive control. Each reaction was mixed with either 0nM or 250nM of SUMO-Bicc1 N-terminal and then analyzed on either a **(A)** vertical polyacrylamide native gel run for 30 minutes at 120V or as duplicate samples in the horizontal polyacrylamide native gel for **(B)** 2 hours or **(C)** 5.5 hours.

REFERENCES

Chicoine, J., Benoit, P., Gamberi, C., Paliouras, M., Simonelig, M., & Lasko, P. (2007). Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Developmental Cell*, **13**(5), 691–704.

Dahlberg, A. E., Dingman, C. W., & Peacock, A. C. (1969). Electrophoretic characterization of bacterial polyribosomes in agarose-acrylamide composite gels. *Journal of Molecular Biology*, **41**(1), 139–47.

Foley, T. L., Dorjsuren, D., Dexheimer, T. S., Burkart, M. D., Wight, W. C., & Simeonov, A. A (2016). Platform to Enable the Pharmacological Profiling of Small Molecules in Gel-Based Electrophoretic Mobility Shift Assays. *Journal of Biomolecular Screening*, **21**(10), 1125–1131.

Gamberi, C., & Lasko, P. (2012). The Bic-C family of developmental translational regulators. *Comparative and Functional Genomics*, **2012,** 141386.

Hellman, L. M., & Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols*, **2**(8), 1849–61.

Maisonneuve, C., Guilleret, I., Vick, P., Weber, T., Andre, P., Beyer, T., Constam, D. B. (2009). Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. *Development*, **136**(17), 3019–30.

Pagano, J. M., Clingman, C. C., & Ryder, S. P. (2011). Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA*, **17**(1), 14–20.

Pagano, J. M., Farley, B. M., Essien, K. I., & Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(48), 20252–7.

Royer, C. A., & Scarlata, S. F. (2008). Fluorescence approaches to quantifying biomolecular interactions. *Methods in Enzymology*, **450**, 79–106.

Ryder, S. P., Recht, M. I., & Williamson, J. R. (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods in Molecular Biology*, **488**, 99–115.

Saffman, E. E., Styhler, S., Rother, K., Li, W., Richard, S., & Lasko, P. (1998). Premature translation of oskar in oocytes lacking the RNA-binding protein bicaudal-C. *Molecular and Cellular Biology*, **18**(8), 4855–62.

Tran, U., Zakin, L., Schweickert, A., Agrawal, R., Döger, R., Blum, M., Wessely, O. (2010). The RNA-binding protein bicaudal C regulates polycystin 2 in the kidney by antagonizing miR-17 activity. *Development*, **137**(7), 1107–16.

Zhang, Y., Cooke, A., Park, S., Dewey, C. N., Wickens, M., & Sheets, M. D. (2013). Bicaudal-C spatially controls translation of vertebrate maternal mRNAs. *RNA.* **19**(11), 1575–82.

Zhang, Y., Park, S., Blaser, S., & Sheets, M. D. (2014). Determinants of RNA binding and translational repression by the Bicaudal-C regulatory protein. *The Journal of Biological Chemistry*, **289**(11), 7497–504.

Chapter 3:

A single KH domain in Bicaudal-C links mRNA binding and translational repression functions to maternal development

Megan E Dowdle[1*], Sookhee Park[1*], Susanne Blaser[1], Catherine A. Fox[1], Douglas W. Houston[2], and Michael D. Sheets[1]

This chapter was published:

Dowdle, M. E., Park, S., Blaser, I. S., Fox, C. A., Houston, D. W. and Sheets, M. D. (2019). A single KH domain in Bicaudal-C links mRNA binding and translational repression functions to maternal development. *Development* 146,

I conceived of all the experiments with Sookhee Park and Michael Sheets. I performed all the *in vitro* binding experiments, helped analyze the *in vivo* data, and wrote the manuscript with Sookhee Park, Catherine Fox, Doug Houston, and Michael Sheets

1Department of Biomolecular Chemistry, University of Wisconsin – Madison

[2]Department of Biology, University of Iowa

*Authors contributed equally to this work

Abstract

Bicaudal-C (Bicc1) is a conserved RNA binding protein that represses the translation of selected mRNAs to control development. In *Xenopus* embryos Bicc1 binds and represses specific maternal mRNAs to control anterior-posterior cell fates. However, it is not known how Bicc1 binds its RNA targets or how binding affects Bicc1-dependent embryogenesis. Focusing on the KH domains, we analyzed Bicc1 mutants for their ability to bind RNA substrates *in vivo* and *in vitro*. Analyses of these Bicc1 mutants demonstrated that a single KH domain, KH2 was critical for RNA binding *in vivo* and *in vitro*, while the KH1 and KH3 domains contributed minimally. The Bicc1 mutants were also assayed for their ability to repress translation, and results mirrored the RNA binding data, with KH2 being the only domain essential for repression. Finally, maternal knock-down and rescue experiments indicated that the KH domains were essential for Bicc1's regulation of embryogenesis. These data advance our understanding of how Bicc1 selects target mRNAs and provide the first direct evidence that Bicc1's RNA binding functions are essential for both Bicc1-dependent translational repression and maternal vertebrate development.

Introduction

The conserved Bicaudal-C (Bicc1) protein functions as a cell-fate regulator in metazoans in a variety of biological contexts. Bicc1 was first identified in *Drosophila* as a maternal gene required for the anterior-posterior patterning of the embryo (Bull 1966). Subsequently, studies in vertebrates revealed that Bicc1 contributed to the normal formation and function of several organs. For example, the heart, kidneys, and pancreas of *Bicc1-/-* homozygous mutant mice exhibit several abnormalities in structure and function (Maisonneuve et al. 2009; Piazzon et al. 2012; Lemaire et al. 2015). Consistent with these experimental genetic studies, mutations in human *Bicc1* are linked to cystic renal dysplasia, a kidney disease (Kraus et al. 2012). While these observations in vertebrate organ systems define roles for zygotic Bicc1, recent loss-of-function analyses of Bicc1 in *Xenopus* embryos establish that maternal Bicc1 is essential for anterior-posterior patterning of vertebrate embryos, indicating that Bicc1's importance throughout development is conserved (Park et al. 2016). Based on substantial experimental evidence, particularly and most recently from studies of maternal Bicc1 in *Xenopus*, as well as sequence comparisons to known RNA binding proteins, Bicc1's critical biological roles depend upon its ability to bind directly to target mRNAs and regulate their translation. Indeed, several maternal mRNA targets of *Xenopus* Bicc1 have been defined that encode known regulators of anterior-posterior patterning in vertebrate embryogenesis (Zhang et al. 2013). Thus, the mechanisms by which Bicc1 binds RNA are central to understanding its biological roles, but few experiments have addressed this issue.

The N-terminal portion of Bicc1 proteins contain regions with strong sequence conservation to conserved KH (hnRNP-**K** **h**omology) and KHL (KH-like) domains (Gamberi and Lasko 2012). KH domains are a prevalent RNA binding module used by eukaryotic regulatory proteins for post-transcriptional regulatory processes (Dominguez et al. 2018; Nicastro, Taylor, and Ramos 2015). In depth studies revealed that the N-terminal region of *Xenopus* Bicc1 is sufficient for binding specific mRNA target substrates both *in vivo*, based on RNA immunoprecipitation experiments (RIP), and *in vitro*, based on electrophoretic mobility shift assays (EMSAs; gel shifts) (Dowdle et al. 2017; Zhang et al. 2013). While these observations provide compelling evidence that the key RNA binding function of Bicc1 is encoded by this N-terminal portion of the protein, this region contains multiple KH and KHL domains and it is not possible to predict which ones are important for specific Bicc1 RNA binding. For example, the KH-domain containing protein, ZBP1 (***z***ip code ***b***inding ***p***rotein 1), which functions to localize and translationally repress target mRNAs in the dendrites of neurons, contains four KH domains but only two are critical for ZBP1-mRNA binding (Nicastro et al. 2017; Farina et al. 2003; Chao et al. 2010). From this and other examples it is clear that different biologically important multi-KH domain RNA binding proteins can use KH domains singly or in combination to bind relevant mRNA targets. Thus, direct experimental examination is required to define the mechanism by which a multi KH-domain containing protein such as Bicc1 binds RNA.

In this study, we analyzed how the *Xenopus* Bicc1 protein binds to specific mRNA target substrates, by examining several Bicc1 protein mutants for RNA binding *in vivo* and *in vitro*. Bicc1-RNA interactions could only be observed with the intact multi-KH domain N-terminal region, suggesting that a stable, functional RNA binding activity

required a folded architecture that could only be achieved by the individual KH and KHL domains working together. Using amino acid substitutions to abolish the RNA binding function of individual KH domains revealed that the KH2 domain was the major determinant of efficient RNA binding. Translation-reporter assays indicated that KH2 was also the only domain essential for translational repression, suggesting that target mRNA binding was the primary and possibly the only role for this region of Bicc1 in translational regulation. Evolutionary comparison revealed that the KH2 domain and its associated GKGG motif is one of the most highly conserved features of Bicc1 proteins. Finally, a maternal knock-down and rescue assay that we established in a previous study to define the essential maternal role for Bicc1 in vertebrate embryogenesis revealed that canonical KH domain function in RNA binding was essential to Bicc1's maternal role in development.

Results

Functional KH domains were required for both mRNA binding and translational repression by Bicc1 *in vivo*. In a previous study, we demonstrated that the N-terminal portion of Bicc1 that contained multiple KH and KH-like domains, was sufficient to direct selective target mRNA binding by Bicc1 *in vivo* (Zhang et al. 2013). Because this region of Bicc1 contained multiple domains predicted to be capable of RNA binding, we sought to define whether these KH domains were important for selective target mRNA binding in embryos. We exploited embryo expression of HA-tagged Bicc1 protein variants as an *in vivo* assay that we have previously established recapitulates biologically relevant Bicc1-mRNA target interactions (**Fig.1A**) (Zhang et al. 2013; Zhang et al. 2014; Park et al.

2016). A defining functional feature of KH domains is the presence of a conserved GXXG motif that is essential for KH-domains to bind RNA substrates because it makes critical contacts with the RNA backbone (Hollingworth et al. 2012; Nicastro, Taylor, and Ramos 2015; Valverde, Edwards, and Regan 2008). To begin to analyze the role of its KH domains we created a Bicc1 variant in which each of the conserved GXXG motifs was changed to GDDG, which has been shown to abolish RNA binding without altering the structure of the domains (Hollingworth et al. 2012)(**Fig.1B**; **KH1-2-3 GDDG**). mRNAs encoding HA-tagged wild-type and the KH1-2-3 GDDD variant were injected into the animal cap cells of 4-8 cell embryos. After the embryos developed to the blastula stage, extracts were prepared and subjected to anti-HA directed RNA immunoprecipitation (IP-Bicc1) followed by quantitative RT-PCR to detect specific associated mRNAs (**Fig.1A**). Two mRNAs were assessed: Cripto1 and Cyclin B1 (**Fig.1C**). The Cripto1 mRNA is a Bicc1 target that was described and validated in previous work and the Cyclin B1 mRNA is a negative control (Zhang et al. 2013). While the wild type Bicc1 protein selectively bound endogenous Cripto1 mRNA in embryos, the KH1-2-3 GDDG mutant did not (**Figs.1B and C**). The difference in binding was due to activity and not difference in expression as both proteins were efficiently expressed in embryos (**Fig.1D**). Therefore, the KH domains were essential for RNA selective binding by Bicc1 in embryos.

Our previous studies established that Bicc1 binding to specific mRNAs, such as Cripto1, targets them for translational repression (Zhang et al. 2013; Park et al. 2016). Therefore, any Bicc1 mutant defective for mRNA binding should also be defective for translational repression. To address this issue, we used a previously established translational reporter assay in which the luciferase coding region is fused to a region of

the Cripto1 3'UTR required for both Bicc1 binding and Bicc1-dependent translational repression (**Fig.2A and B**) (Zhang et al. 2009; Zhang et al. 2013). In this assay, the reporter mRNA was co-injected with an mRNA encoding Bicc1 into animal cells of 4-8 cell *Xenopus* embryos and the injected embryos were allowed to develop to the blastula stage when embryo extracts were prepared and assayed for luciferase. Bicc1-dependent luciferase expression was defined as the ratio of luciferase activity measured in embryos co-injected with both Bicc1 and the luciferase reporter mRNA to the luciferase activity measured in control embryos that were injected with only the reporter mRNA (Zhang et al. 2009; Zhang et al. 2013). These experiments revealed that while a Bicc1 protein containing the wild type RNA binding domain repressed the reporter mRNA 5–fold, as expected, the mutant Bicc1 protein in which each KH domain contained a GXXG→GDDG substitution was defective for repression (**Fig.2C**). In fact, the KH1-2-3 GDDG mutant was as ineffective at repression as the minus-Bicc1 control (**Fig.2C,** None). Therefore, the KH domains were essential for Bicc1-dependent translation repression.

The KH2 domain was the most critical determinant of mRNA binding by Bicc1. The experiments described above set the stage for determining whether all or just a subset of the KH domains were more important for Bicc1 mRNA selective binding and translational repression. To address this issue, individual and pairwise combinations of KH GXXG→GDDG mutant proteins (**Fig.3A**) were examined for RNA binding *in vivo* as described above (**Figs.1A and 3A**). The GXXG→GDDG substitution in the KH2 domain (KH2-GDDG) abolished Bicc1-RNA binding to levels equivalent to that observed in the control embryos (**Fig.3B,** None). In contrast, the KH1 and KH3 domains minimally contributed to RNA binding. Specifically, a GXXG→GDDG substitution in either KH1

(KH1-GDDG) or KH3 (KH3-GDDG) reduced Bicc1 association by about 2-fold. Combining both substitution mutants into one variant protein (KH1-3-GDDG) caused no further reduction in Bicc1-RNA association (**Fig.3B**). In contrast, the GXXG→GDDG substitution in the KH2 domain (KH2-GDDG) abolished Bicc1-RNA binding to levels equivalent to that observed in the control embryos (**Fig.3B, None**). Each Bicc1 variant was expressed similarly in these experiments (**Fig.3C**). Thus, the KH2 domain was the most critical for robust RNA binding by Bicc1.

The same mutants were examined for their ability to direct Bicc1-dependent translational repression as described above (**Fig.2**). If substrate RNA binding was the only role for the N-terminal region of Bicc1 in executing translational repression, then the results from the translation reporter assays should mirror the RNA binding experiments and this result was observed. Consistent with data presented in Figure 1 and the RNA binding analyses in Figure 3B, the GXXG→GDDG KH2 substitution was the only single substitution to reduce Bicc1-dependent translational repression (**Fig.3D**). In addition, a mutant containing GXXG→GDDG substitutions in both KH1 and KH3 domains remained capable of substantial Bicc1-dependent translational repression, consistent with the observation that this same mutant remained capable of selective RNA binding (**Fig.3B**). Because all variant proteins examined were expressed to similar levels (**Fig.S1**) and the reporter mRNA was equally stable in all cases (**Fig.S2**), the observed effects were caused by defects in Bicc1-RNA interactions. We conclude that the KH2 domain's ability to interact with RNA is a critical determinant of Bicc1's selective mRNA binding function, and that target mRNA binding is the primary role of the multi-KH domain containing region of Bicc1 in Bicc1-dependent translational repression.

The KH2 domain was the most critical determinant of Bicc1-RNA interactions *in vitro*. The embryo experiments demonstrated a pivotal role for the KH2 domain in selective Bicc1-RNA interactions and translational repression *in vivo*. However, they could not rule out the possibility that these effects were facilitated by Bicc1-protein interactions that might occur *in vivo*. Therefore, to test whether the *in vivo* effects could in fact be ascribed to direct Bicc1-RNA binding defects, Bicc1 mutants, expressed as recombinant proteins (Bicc1 RNA binding domain amino acids 1-506) were purified and used to assess RNA binding *in vitro* using protein-RNA gel shift assays (**Fig.4**). For these experiments, a 32-nucletide region from the Cripto-1 mRNA 3' UTR that was previously shown to be sufficient for selective Bicc1-RNA binding was used as a substrate (Zhang et al. 2014). The wild type Bicc1 protein and the two single GXXG→GDDG KH1 and KH3 mutants efficiently bound the Cripto1 RNA substrate in these experiments to efficiently form distinct RNA-protein complexes (**Fig.4A, lanes 1, 2, 4 and 6**). In contrast, the triple GXXG→GDDG KH1-2-3 substitution mutant abolished formation of a stable Bicc1-RNA complex, while the single GXXG→GDDG KH2 mutant reduced formation of this complex (**Fig.4A, lane 3**). Note that even the triple GXXG→GDDG KH1-2-3 substitution mutant exhibited some binding to the Cripto1 RNA, as indicated by the diffuse signals in the gel due to RNA-protein complexes dissociating during gel electrophoresis. Thus, while the triple mutant caused a significant RNA binding defect, it did not abolish RNA binding activity, indicating that other features of the Bicc1 RNA binding domain were contributing to Bicc1's RNA binding function. In addition, all the RNA binding activity that was detected showed specificity for Cripto1 RNA substrate: the same proteins exhibited no detectable binding to the negative control CyclinB1 RNA substrate (**Fig.4B**). The data from these *in*

vitro experiments together with the data from the *in vivo* experiments revealed that Bicc1 KH2 was the key domain responsible for efficient RNA binding by Bicc1.

The KH2 domain was not sufficient for RNA binding. For some multi-KH domain proteins, a select subset of their KH domains are both necessary and sufficient for RNA binding (Nicastro et al. 2017; Nicastro, Taylor, and Ramos 2015; Valverde, Edwards, and Regan 2008; Chao et al. 2010; Farina et al. 2003). Our results described above indicate that the KH2 domain is necessary for Bicc1 binding and we wanted to test whether it was also sufficient. We used the *in vivo* RNA binding assay to test sufficiency. mRNAs encoding HA-tagged variants of the N-terminal region of Bicc1 were injected into 4 to 8-cell embryos and binding to specific endogenous mRNAs was assayed as described above (**Fig.1**). The KH1-2-3 variant (KH1-2-3 1-506) that contained the entire Bicc1 N-terminus was the only protein that showed significant enrichment of the Cripto1 and GRG5 target mRNAs **(Fig.5B**), although all variants were expressed at similar levels (**Fig.5C**).

To complement these results, we sought to perform *in vitro* RNA binding studies. We observed that, while the subdomains of the Bicc1 N-terminus could be expressed in *E. coli*, they were generally insoluble and difficult to purify. However, two constructs, KH1-KH2 and KH2, could produce sufficient recombinant protein and these were tested for RNA binding using gel shift assays. The KH1-KH2 protein bound to the Cripto1 RNA, whereas the KH2 protein did not exhibit RNA binding **(Fig.6A)**. However, the binding activity observed with the KH1-KH2 protein was not specific as it also bound to the Cyclin B1 RNAs **(Fig.6B)**. Because the RNA-protein complexes formed with these variants were similar in size to the unbound RNA, we sought to confirm these results with a second

assay for RNA binding. We used solution-based fluorescence polarization assays in which the binding of a protein to a fluorescently labeled RNA results in an increase in polarization (Pagano, Clingman, and Ryder 2011). With this sensitive assay we also observed that while the KH1-KH2 protein could bind RNA, the binding lacked the specificity of the intact N-terminal region as it bound to both the Cripto1 and Cyclin B1 RNAs (**Fig.S3**).

Taken together with the results from RNA IP experiments, these data from *in vivo* and *in vitro* approaches provide evidence that sub-regions of the Bicc1 multi-KH domain do not bind RNA selectively in isolation but only in the context of the intact N-terminal region.

The KH2 domain is a conserved domain feature of vertebrate and invertebrate Bicc1 proteins. Bicc1 proteins share a common architecture: an N-terminal region that contains three predicted KH domains (Gamberi and Lasko 2012). Comparison of Bicc1 proteins from different species revealed that the N-terminal region is highly conserved (**Fig.7, Fig.S4**). In particular, the KH2 GXXG motif (GKGG) and surrounding amino acid residues were the most conserved amongst animal species conserved in all species while the GXXG motifs for KH1 and KH3 domains were divergent (**Fig.7, Fig.S4**). This evolutionary conservation supports our functional analysis and suggests that the KH2 domain is an important feature for RNA binding in all Bicc1 proteins.

The RNA binding functions of Bicc1 were essential for embryonic patterning. In a recent study, we established the essential embryonic role for maternal Bicc1 using the host-transfer method to generate a maternal knock-down of Bicc1 that depleted maternal sources of Bicc1 protein from eggs prior to fertilization (Park et al. 2016).

Embryos significantly depleted of maternal Bicc1 develop abnormally, with an excess of anterior structures, and exhibit an increase in organizer-specific gene expression accompanied by a reduction in ventral-posterior gene expression. Importantly, re-introduction of wild type Bicc1 into maternal Bicc1 knock-down embryos rescues the anterior-posterior developmental defects. We used this rescue as a functional assay to determine whether the RNA binding defects caused by the triple GXXG→GDDG KH1-2-3 substitution affected Bicc1's maternal role in embryogenesis (**Fig.8**). Specifically, as previously described, embryos were substantially depleted of detectable maternal Bicc1 mRNA (**Fig. 8A**) and protein (**Fig.S5)** (Park et al. 2016). The Bicc1-depleted embryos (*bicc1*-) developed with severe abnormalities compared to untreated control embryos (Untreated) or embryos that were co-injected with wild type *bicc1* mRNA during the host-transfer experiment (*bicc1*-; +*HA-bicc1*) **(Figs.8A-D**). If the RNA binding function of Bicc1 were critical for Bicc1's maternal role, as predicted based on the model for how Bicc1 functions in development by acting as an mRNA selected translational repressor, then an mRNA encoding a mutant Bicc1 defective in mRNA binding (*bicc1*-; +*HA-bicc1 KH1-2-3 GDDG*) should fail to rescue embryonic development in a host-transfer/rescue experiment. This result was observed (**Fig.8E, G**). Importantly, both the wild type and mutant HA-bicc1 mRNAs expressed protein to similar levels (**Fig.8F**), indicating that the inability of the *HA-bicc1 KH1-2-3 GDDG* mutant to rescue the maternal knock-down of Bicc1 was not simply the result of reduced stability of the mutant protein in embryos. We conclude that the RNA binding functions of maternal Bicc1 are essential for its control of vertebrate embryogenesis.

Discussion

It is well established that Bicc1 plays critical roles in metazoan development and it is generally thought that its RNA binding and translational repressor functions are relevant to these roles (Park et al. 2016; Bull 1966; Mahone, Saffman, and Lasko 1995; Saffman et al. 1998; Maisonneuve et al. 2009; Yaguchi, Yaguchi, and Inaba 2014). However, there is a paucity of mechanistic experiments addressing how Bicc1 directly binds to its relevant mRNA targets and how this binding affects either Bicc1-dependent translational repression or its roles as a cell-fate regulator. Here we defined a single KH domain, KH2 of Bicc1 that was critical for Bicc1's efficient binding to a key maternal mRNA target of Bicc1, the Cripto1 mRNA. The RNA binding function provided by this KH2 domain was also the only KH domain critical for Bicc1-dependent translational repression of a luciferase reporter that contained the relevant Cripto mRNA 3' UTR Bicc1-control region, providing evidence that the only critical mechanistic role of the N-terminal region of Bicc1 in translational repression *per se* is to localize Bicc1 to its mRNA target. Using host transfer knock-down and rescue experiments, we show that a Bicc1 mutant containing KH domains defective for contacting RNA failed to provide for Bicc1's maternal role in embryonic patterning, providing direct evidence that Bicc1's ability to bind RNA through canonical KH domain-RNA interactions is essential for its role in vertebrate development.

Multi-KH-domain containing regions are common features of biologically important RNA binding proteins, but the precise mechanisms by which they contact their RNA substrates varies between proteins (Gerstberger, Hafner, and Tuschl 2014; Hentze et al. 2018). For some, such as the KSR protein all four of its KH domains contribute to RNA binding (Hollingworth et al. 2012). For many others, such as the vigilin and ZBP1 proteins

RNA binding is conferred by a specific subset of KH domains (Nicastro et al. 2017; Nicastro, Taylor, and Ramos 2015; Cheng and Jansen 2017). Our results provide evidence that the Bicc1 N-terminal region functions similarly to the latter category of multi KH containing RNA binding domains. Specifically, while analyses of precise amino acid substitution variants indicated that only the KH2 domain was critical for Bicc1-RNA contacts. However, analyses of subdomains indicated that the KH2 domain was not sufficient to bind RNA. Thus, we suggest that the entire Bicc1-N-terminal region must fold into a structure that allows KH2 to contact Cripto1 RNA to drive stable Bicc1-RNA binding both *in vivo* and *in vitro*. In this view, direct RNA contacts by KH1 or KH3, or the KHL domains, are not critical but rather their interactions with each other in three-dimensional space are required to support KH2's ability to contact RNA. Structural work will be required to address this issue.

While the GXXG→GDDG substitution mutants in Bicc1 revealed that the KH2 domain was the critical domain required for stable Bicc1-Cripto1 RNA interactions, what drives specific Bicc1-target mRNA interactions remains an open issue. Studies of other KH domains establish that the GXXG motif is critical for contacting the RNA backbone and contributes minimally to sequence-specific RNA binding (Nicastro, Taylor, and Ramos 2015; Nicastro et al. 2017; Teplova et al. 2011). Consistent with these observations, our analyses of Bicc1-Cripto1 RNA interactions *in vitro*, which were sensitive enough to detect some residual RNA binding by the KH2 GDDG mutant, showed that this residual binding remained specific for Cripto1 mRNA. It is possible that another region of KH2, or other regions of the N-terminal region of Bicc1 provide specificity. In addition, because our previous work provides evidence that Bicc1 has many relevant target mRNAs in addition

to Cripto1 mRNA (Zhang et al. 2013), it is possible that the other KH domains are important for binding some of these targets. In other words, the Bicc1 multi KH and KHL domain may define a flexible RNA binding surface that can bind different target mRNAs using distinct RNA-protein interfaces. Additional target RNA substrate characterization and Bicc1-RNA interaction assays to address these interesting possibilities are underway. Finally, while the GXXD→GDDG KH1 and KH3 substitution variants provide evidence that the canonical RNA binding activity of these domains is not critical for Cripto1 mRNA binding or Cripto1-3'UTR directed translational repression, it remains possible that these domains are important for mediating protein-protein interactions critical for this translational repression, as KH domains in other proteins have been shown to direct protein-protein interactions (Zheng et al. 2014; Nakel et al. 2015; Du et al. 2007; Teplova et al. 2011). However, while such interactions may be necessary they are clearly not sufficient because the single GXXD→GDDG KH2 Bicc1 variant was unable to repress translation.

This work advances our understanding of how Bicc1 contacts its mRNA targets and establishes that such RNA contacts were indispensable for Bicc1-directed translation repression and linking these functions to Bicc1's role as a regulator of anterior-posterior patterning in vertebrate embryos. In addition, the host-transfer-rescue experiments used initially to define Bicc1's maternal role in embryogenesis, and here to test the importance of Bicc1's ability to make direct KH-RNA contacts to this role, provide a powerful context to further examine and refine our understanding of how Bicc1's biochemical functions relate to its biological role in cell fate decisions.

Methods

Xenopus laevis oocyte and embryo manipulations

X. laevis oocyte and embryos were obtained and injected as described (Sive et al., 2000). Host transfer experiments were performed using antisense oligodeoxynucleotides (oligos) against *bicc1* (Park et al. 2016). These were synthesized as HPLC-purified phosphorothioate-phosphodiester chimeric oligos (Integrated DNA Technologies) with the following sequences (*indicates phosphorothioate linkages): *bicc1 9460*, 5′-G*C*GTGTTTGTCTCTTC*C*A-3′ (nts 180-162 where the A of the AUG start codon is nucleotide 1); *bicc1 9463*, 5′- T*G*TAACATTGTCTCGAG*C*T -3′ (nts 374-357). Oocytes were injected in the vegetal pole and cultured for 24 hours at 18°C before being matured by treatment with 2.0 μM progesterone. Matured oocytes were colored with vital dyes, transferred to egg-laying host females, recovered and fertilized essentially as described (Heasman, Holwill, and Wylie 1991). For rescue experiments, *HA-bicc1* mRNA that encodes Bicc1 with an N-terminal HA epitope tag was injected into vegetal cells of *bicc1* depleted embryos shortly after fertilization. The injected *HA-bicc1* mRNA used for rescue was not affected by the antisense oligos because the oligos are degraded a few hours after injection.

Bicc1 variants

Xenopus Bicc1 variants were created via PCR and cloned into a pCS2+ plasmid as fusions with a 3xHA epitope tag at their N-terminal end. All plasmids were verified by sequencing.

mRNA synthesis

Capped mRNAs encoding HA tagged full length Bicc1, HA tagged Bicc1 variants and Firefly luciferase reporter mRNAs contained the TCE of the *cripto1* mRNA 3'UTR or the 3'UTR of the *cyclinb1* mRNA were synthesized as described (Fritz and Sheets 2001; Zhang et al. 2009; Sheets et al. 1994; Zhang 2013; Zhang et al. 2013).

Luciferase reporter mRNAs

Reporter mRNAs were diluted to a concentration of 2.5 nM. 5 nanoliters (12.5 amol) was injected into embryonic cells. When injected embryos reached the appropriate stage, extracts were prepared and analyzed for luciferase activity (Fritz and Sheets 2001; Zhang et al. 2009; Sheets et al. 1994).

Immunoblotting

The analysis of proteins by immunoblotting was done as described (Zhang et al. 2013) using monoclonal anti-HA, anti-actin and polyclonal anti-Bicc1 antibodies (Park et al. 2016).

qRT-PCR

Quantitative RT-PCR to analyze reporter mRNAs and endogenous mRNAs associated with Bicc1 was performed as described (Park et al. 2011).

Immunoprecipitations

Embryos were injected with mRNA encoding Bicc1 variants fused at the N-terminus with an HA epitope tag (Ha-Bicc1) (Zhang et al. 2013). When injected embryos reached the appropriate Blastula stage (st.7) injected embryos were lysed in 100ul of TNMEN-150 buffer (Cooke, Prigge, and Wickens 2010). The lysate was centrifuged (4^0C, 10min at 5000rpm) and the supernatant incubated with anti-HA antibody coupled to protein-G agarose (2hrs, 4^0C). The beads were collected (1min, 3,000rpm) and washed

4x in 1 ml TNMEN 150 buffer. For each wash, the beads were incubated in buffer at 4^0C for 5min, spun at 3000 rpm 1 min and supernatant removed. RNA was isolated from the washed beads for analysis by qRT-PCR (Park et al. 2016).

Bicc1 protein expression and purification

Bicc1 variants were cloned into pET28b Bacterial expression vectors as N-terminal fusions with a His-6 SUMO tag (Malakhov et al. 2004). Cultures of *E.coli* cells containing each plasmid were grown to an OD600 of 0.6 and induced with 1mM IPTG at 25 degrees C overnight. The cells were collected and lysed in B-PER reagent, 1/2X TBS, 2.5mM MgCl2, 2.5% glycerol, 1mM BME, 10mM Imidazole, 200mM NaCl, 1mM ATP, and protease inhibitors. The soluble lysate was applied to a nickel chromatography resin and allowed to bind overnight. The resin was then washed 5 times with 1xTBS, 5mM MgCl2, 5% glycerol, 2mM BME, 20mM Imidazole, 400mM NaCl, 1mM ATP, followed by a wash of the same conditions with no ATP. The proteins were then eluted with 450mM imidazole and dialyzed in 1x TBS.

EMSA

Recombinant SUMO-Bicc1 N-terminal fusion proteins were expressed and purified as described above. The Cripto1 and CyclinB1 3'-fluorescein-labeled RNA substrates were purchased from IDT. Binding reactions (50µl) contained SUMO-Bicc1 protein, 10mM Hepes, pH 7.5, 1mM EDTA, 50 mM KCl, 0.02% Tween 20, 0.1mg/ml yeast tRNA, 100µg/ml BSA, 2mM DTT, and 10nM fluorescent RNA. Reaction products were analyzed on 7.5% (1× TBE) native polyacrylamide gels (Dowdle et al. 2017). The gels were then scanned at 473 nm using a fluorimager.

Fluorescent Polarization assays for RNA binding

Binding reactions as described above were assembled into individual wells of a 96-well black round bottom plate. The reactions were scanned using a plate reader with an excitation wavelength at 485nm and an emission wavelength of 528nm in the parallel and perpendicular direction (Pagano, Clingman, and Ryder 2011). The data was analyzed using Gen5 software.

Contributions

S.P. M.D. and M.S. were involved in all phases of the work, including planning, executing and analyzing most of the experiments, and writing the manuscript. C.A.F was involved in discussion of experimental results and writing the manuscript. D.W.H. planned and performed the maternal loss of function experiments and assisted in writing relevant sections of the manuscript. S.B. performed some experiments and assisted in the writing of the manuscript. All of the authors have read the manuscript and agree with its content.

Acknowledgements

We thank Andy Chung and Nithin Charlly for assistance with initial gel shift analysis and Laura Vanderploeg for preparing the figures. This work was supported by National Science Foundation Grant 1050395 (M.D.S.), NIH R21HD076828 and R01HD091921, (M.D.S.), the University of Wisconsin Graduate School (M.D.S.), and by a grant from the NIH, R01GM083999 (D.W.H.). Megan Dowdle was supported by a SciMed GRS Advanced Opportunity Fellowship through University of Wisconsin-Madison Graduate

School and Biotechnology Training Program through the National Institute of General Medical Sciences of the National Institutes of Health (T32GM008349).

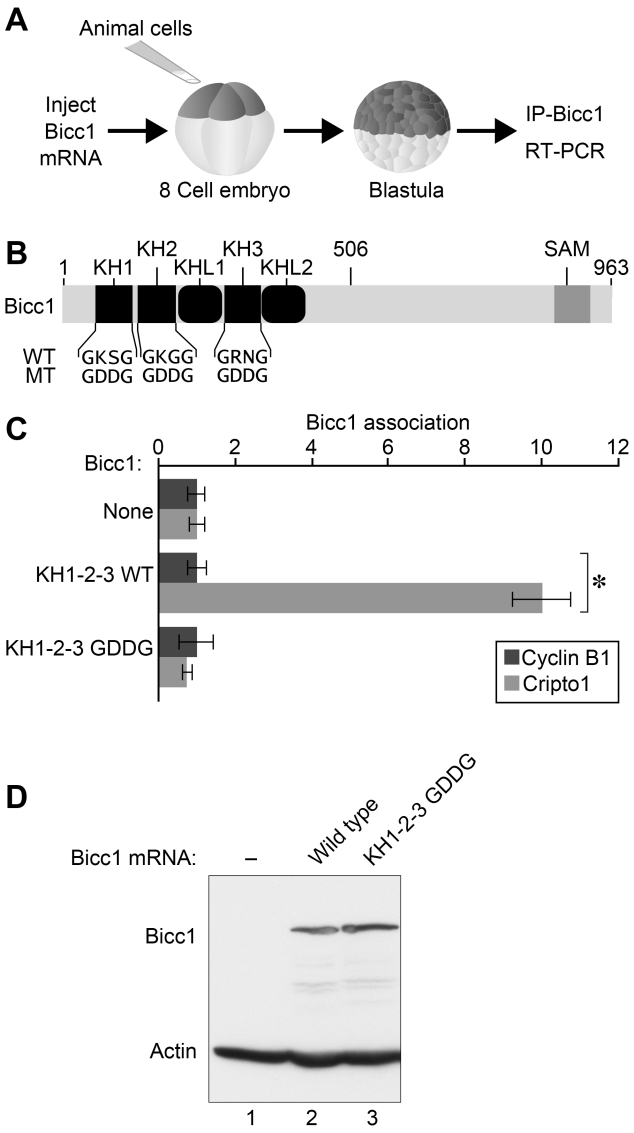

Figure 1. The Bicc1 KH domains are required for RNA binding.
A. Animal cell assay for *in vivo* Bicc1 binding. Animal cells of eight-cell *Xenopus* embryos were injected with mRNA encoding HA tagged Bicc1. Some injected samples included luciferase reporter mRNAs. When embryos reached stage 9, Bicc1 was immunoprecipitated with an HA antibody and the associated RNA isolated for analysis (Park et al. 2016; Zhang et al. 2013). RNA samples were reverse transcribed, and the cDNA used as template for q-PCR. ***B.*** Diagram of Bicc1 protein showing the three KH domains (KH1, KH2 and KH3) with the wild-type (WT) GXXG motif and the GDDG substitutions (MT) created and analyzed. ***C.*** The Bicc1 protein containing GDDG substitutions (KH1-2-3 GDDG) was expressed in *Xenopus* embryos and analyzed for binding to endogenous mRNAs (see Fig.1A). The KH1-2-3 GDDG protein was defective for RNA binding in comparison to the Bicc1 WT protein. *Error bars* represent the standard error from three separate experiments. *T-test P-value <0.05 ***D.*** Immunoblot analysis with an anti-HA antibody was used to monitor the expression of the Bicc1 proteins expressed in embryos for RNA binding assays.

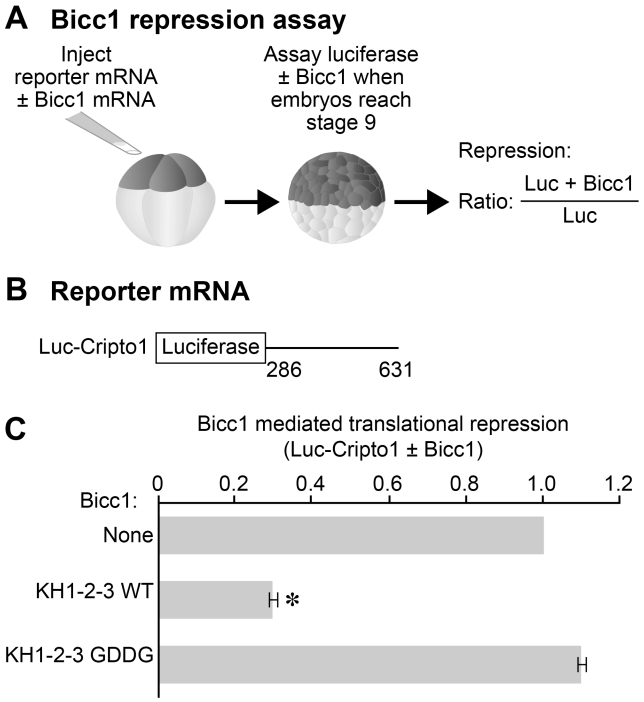

Figure 2. The Bicc1 KH domains are required for translational repression.
A. Animal cell assay for Bicc1 translational repression. Animal cells of eight-cell *Xenopus* embryos were injected with luciferase reporter mRNAs. Some of the embryos were given a second injection of mRNA encoding full-length *Xenopus* Bicc1 (Zhang et al. 2013). When embryos reached stage 9-10, luciferase assays were performed. Repression, as measured by the ratio of luciferase exhibited by a reporter mRNA with and without Bicc1 expression, was calculated and plotted. ***B.*** Diagram of Cripto1 3'UTR fragment incorporated into luciferase reporter mRNAs used to analyze translational repression. ***C.*** The Bicc1 protein containing GDDG substitutions (KH1-2-3 GDDG) was defective for repressing the Luc-Cripto1 reporter mRNA while the Bicc1 wild-type protein repressed the reporter efficiently. *Error bars* represent the standard error from three separate experiments.

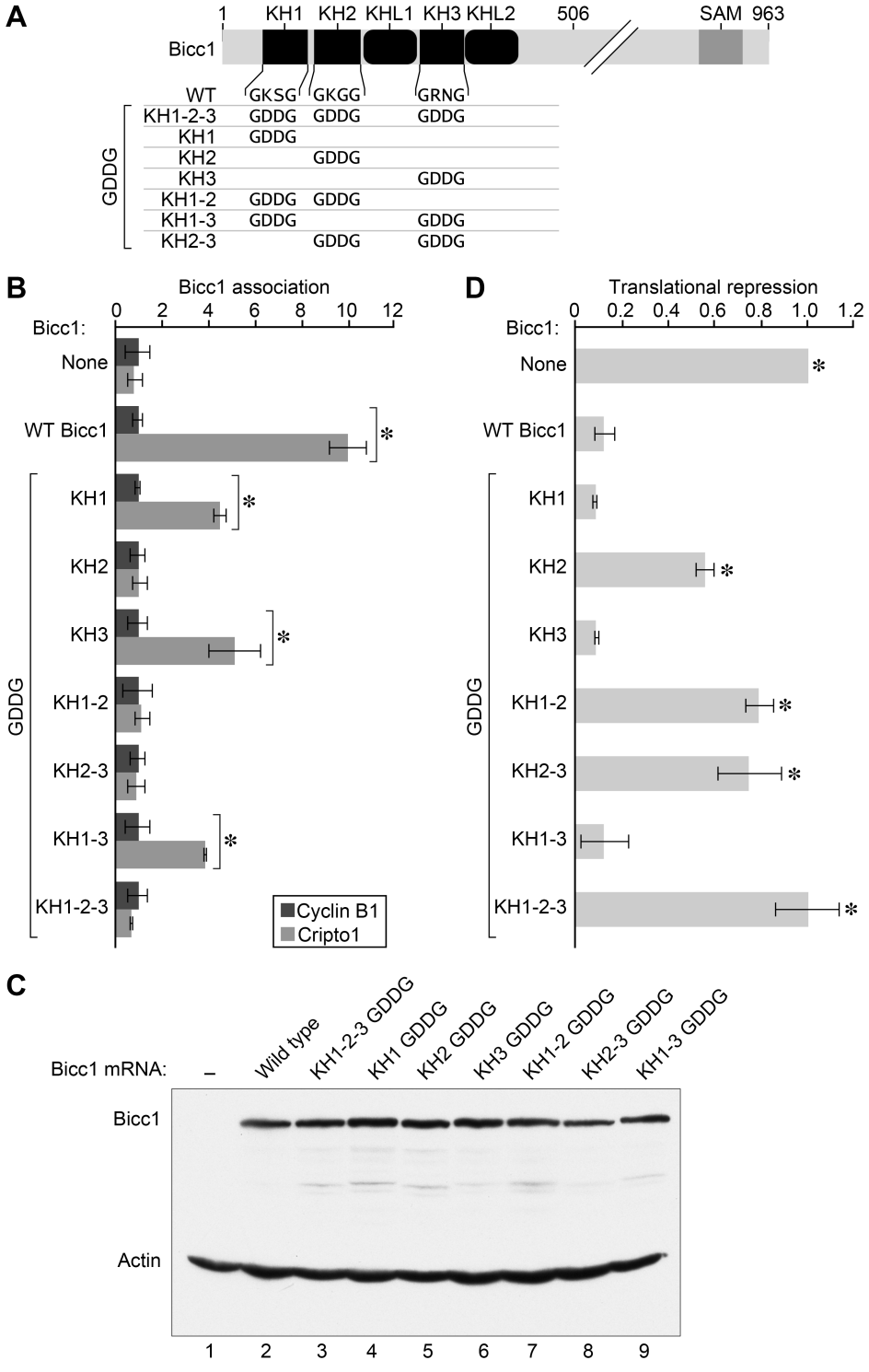
A
1
KH1
KH2
KHL1
KH3
KHL2
506
SAM
963
Bicc1
WT
KH1-2-3
KH1
KH2
KH3
KH1-2
KH1-3
KH2-3
GDDG
GKSG
GKGG
GRNG
GDDG
B
Bicc1 association
Bicc1:
None
WT Bicc1
KH1
KH2
KH3
KH1-2
KH2-3
KH1-3
KH1-2-3
GDDG
Cyclin B1
Cripto1
D
Translational repression
C
Bicc1 mRNA:
–
Wild type
KH1-2-3 GDDG
KH1 GDDG
KH2 GDDG
KH3 GDDG
KH1-2 GDDG
KH2-3 GDDG
KH1-3 GDDG
Bicc1
Actin
1
2
3
4
5
6
7
8
9

Figure 3. The KH2 domain is a major determinant of Bicc1 RNA binding and translational repression.

A*.** Diagram of Bicc1 protein showing the three KH domains (KH1, KH2 and KH3) with the wild-type (WT) GXXG motif and the single and double GDDG substitutions (MT) created and analyzed. ***B. The Bicc1 protein variants containing GDDG substitutions were expressed in *Xenopus* embryos and analyzed for binding to endogenous mRNAs (see Fig.1A). The KH2 GDDG, KH1-2 GDDG, KH2-3 GDDG and KH1-2-3 GDDG proteins were defective for RNA binding in comparison to the Bicc1 WT protein. *Error bars* represent the standard error from three separate experiments. *T-test P-value of <0.05, **T-Test P-Value of <0.0005 when compared to Bicc1 WT protein binding to Cripto1 RNA. ***C.*** Immunoblot analysis with an anti-HA antibody was used to monitor the expression of the different protein variants used in RNA binding assays. ***D.*** The Bicc1 protein variants containing GDDG substitutions were expressed in *Xenopus* embryos and analyzed for translational repression using the Luc-Cripto1 reporter mRNA (see Fig.3A). The KH2 GDDG, KH1-2 GDDG, KH2-3 GDDG and KH1-2-3 GDDG proteins were defective for translational repression in comparison to the Bicc1 WT protein. *Error bars* represent the standard error from three separate experiments. *T-test P-value of <0.05 compared to No Protein control. ‡T-test P-value of <0.05 compared to KH2.

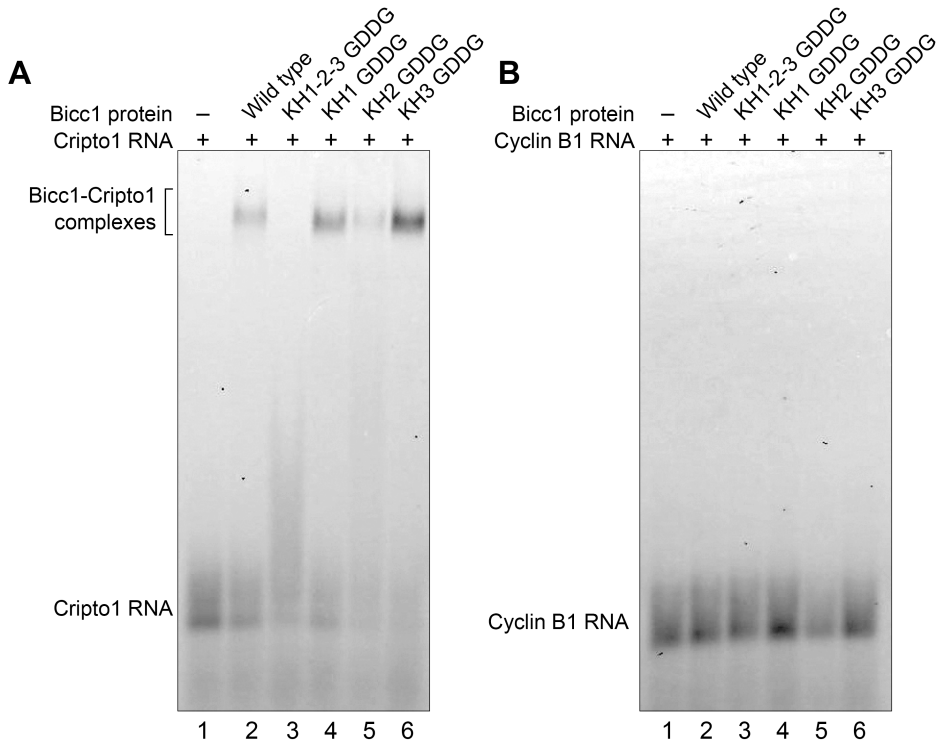

Figure 4. The Bicc1 KH2 domain is required for direct interaction with RNA targets ***A.*** Electromobility shift assays of Bicc1 wild-type and protein variants and RNA substrate that is the Bicc1 binding site from the 3'UTR of the *Xenopus* Cripto1 mRNA. Recombinant proteins consisting of the Bicc1 N-terminus (aa 1-506) were expressed and purified from *E.coli*. Wild-type protein along with KH1 GDDG, KH2 GDDG, KH3 GDDG and KH1-2-3 GDDG were generated for analysis. Binding reactions consisting of the different proteins mixed with a fluorescently labeled 32nt RNA representing a well characterized Bicc1 binding site derived from the 3'UTR of the *Xenopus* Cripto1 mRNA were analyzed on native polyacrylamide gels electrophoresed horizontally (Dowdle et al. 2017). The KH2 GDDG and KH1-2-3 GDDG proteins were defective for complex formation compared to wild-type Bicc1 and other Bicc1 variants. ***B.*** Electromobility shift assays of Bicc1 wild-type and protein variants and RNA substrate derived from the 3'UTR of the *Xenopus* Cyclin B1 mRNA. The Cyclin B1 mRNA is not a Bicc1 target and represents a negative control for binding.

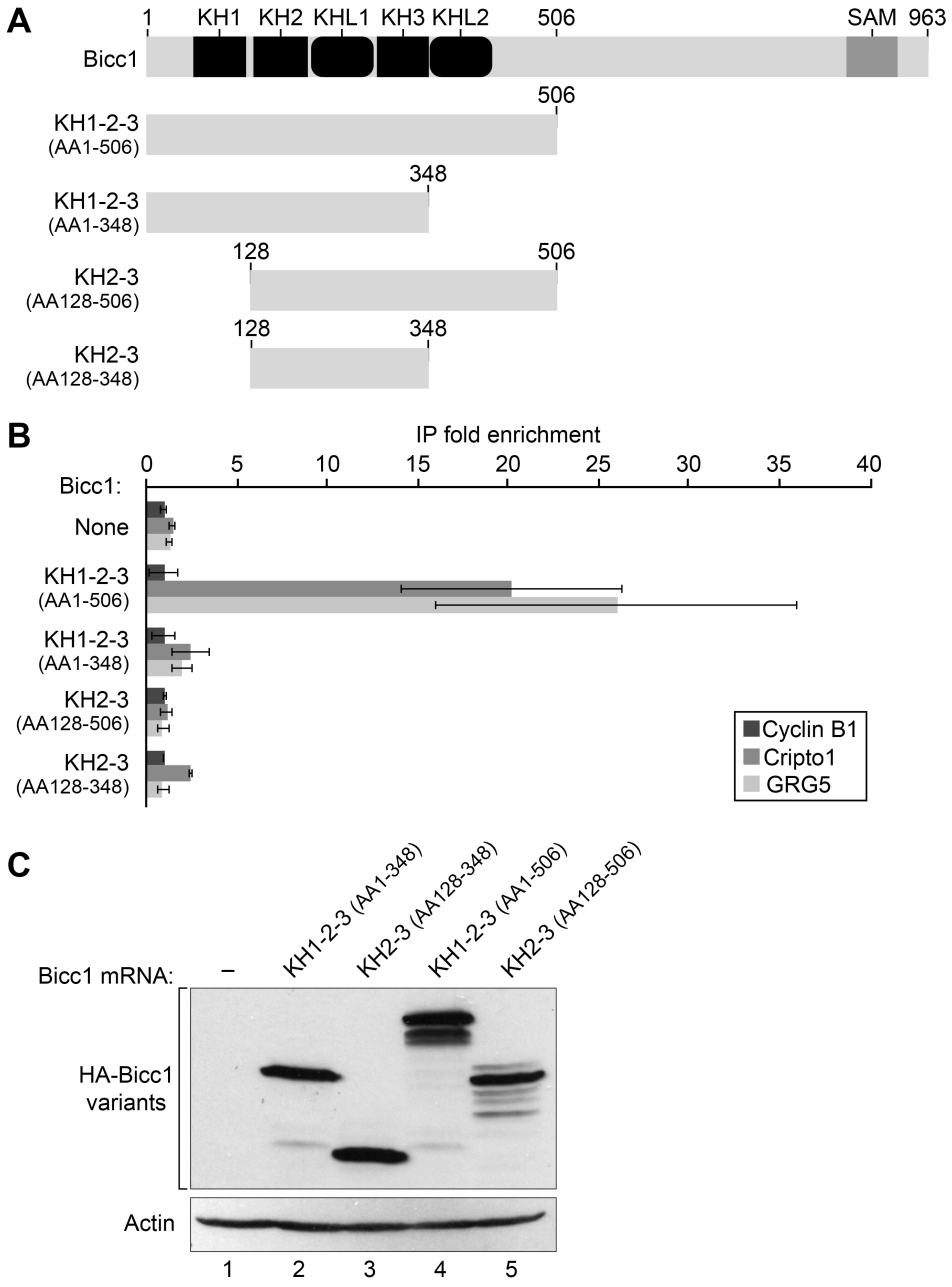

Figure 5. The entire N-terminal region of Bicc1 is required for RNA binding *in vivo*. ***A.*** Diagram of the intact Bicc1 N-terminal region (KH1-2-3 AA 1-506) and the different derivatives that lack KHL2 (KH1-2-3 AA1-348), lack KH1 (KH2-3 AA 128-506) or KH1 and KHL2 (KH2-3 AA 128-348). ***B.*** The endogenous *Xenopus* maternal Cripto1 and GRG5 mRNAs were bound by the intact Bicc1 N-terminus, but none of the derivatives lacking different regions bound to these mRNAs. None of the proteins bound to the Cyclin B1 mRNA, a negative control for this experiment as the Cyclin B1 is not a Bicc1 target. *Error bars* represent the standard error from three separate experiments. ***C.*** Immunoblot

analysis with an anti-HA antibody was used to monitor the expression of the different protein variants used in RNA binding assays.

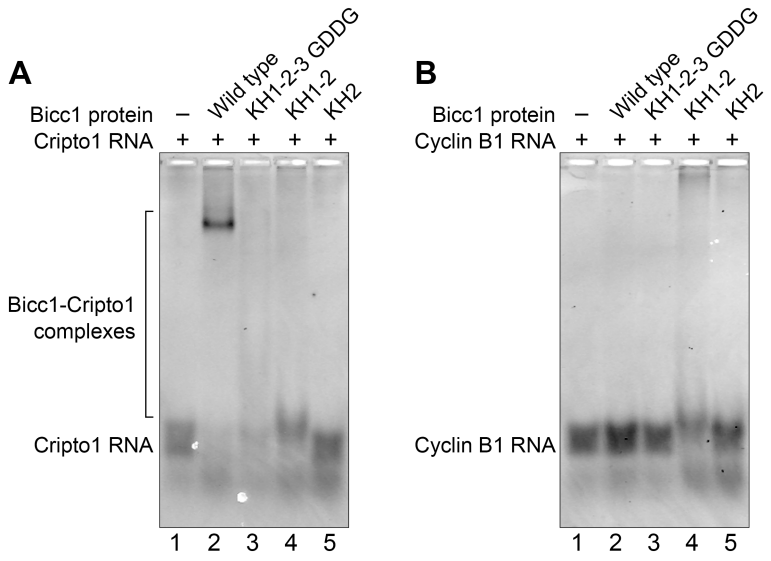

Figure 6. The KH2 domain is not sufficient for RNA binding.
A. Electromobility shift assays of Bicc1 wild-type and protein variants. Recombinant proteins consisting of the Bicc1 wild-type N-terminus (aa 1-506), the KH1-2-3 GDDG protein (aa 1-506), KH1-2 protein (aa 41-201) and KH2 protein (aa 126 – 201) were expressed and purified from *E.coli*. Binding reactions consisting of the different proteins mixed with a fluorescently labeled 32nt RNA representing a well characterized Bicc1 binding site derived from the 3'UTR of the *Xenopus* Cripto1 mRNA were analyzed on native polyacrylamide gels electrophoresed horizontally (Dowdle et al. 2017). ***B.*** Electromobility shift assays of Bicc1 wild-type and protein variants with an RNA substrate derived from the 3'UTR of the *Xenopus* Cyclin B1 mRNA. The Cyclin B1 mRNA is not a Bicc1 target and represents a negative control for binding.

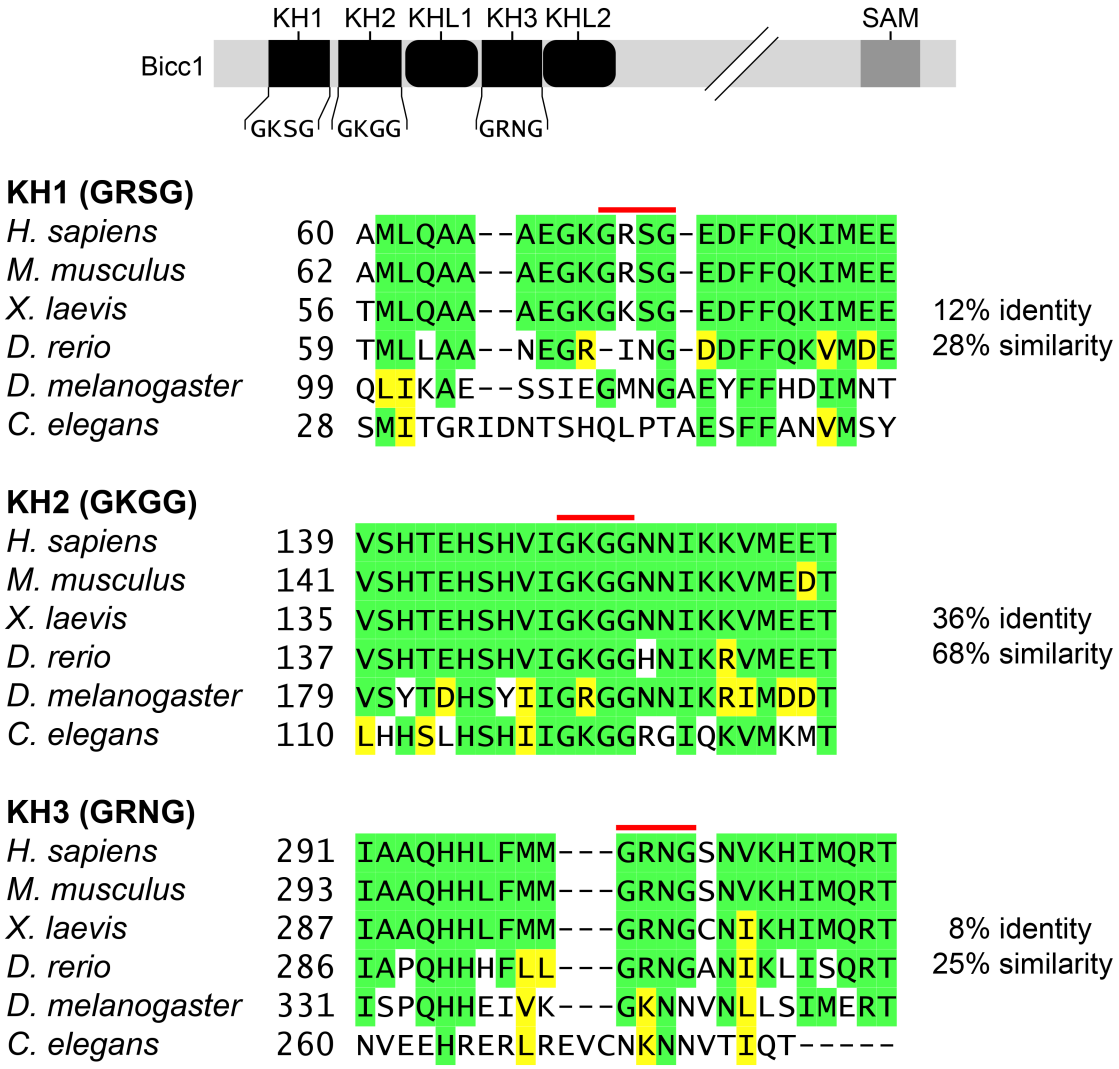

Figure 7. The KH2 domain is an evolutionary conserved feature of Bicc1 proteins. Amino acid sequences from vertebrate and invertebrate Bicc1 proteins were analyzed with Clustal Omega. The regions surrounding the GXXG motif of each KH domain are shown. Residues identical to human Bicc1 are highlighted in green, while similar residues are highlighted in yellow. The comparison of full-length Bicc1 proteins is presented in **Fig.S4**.

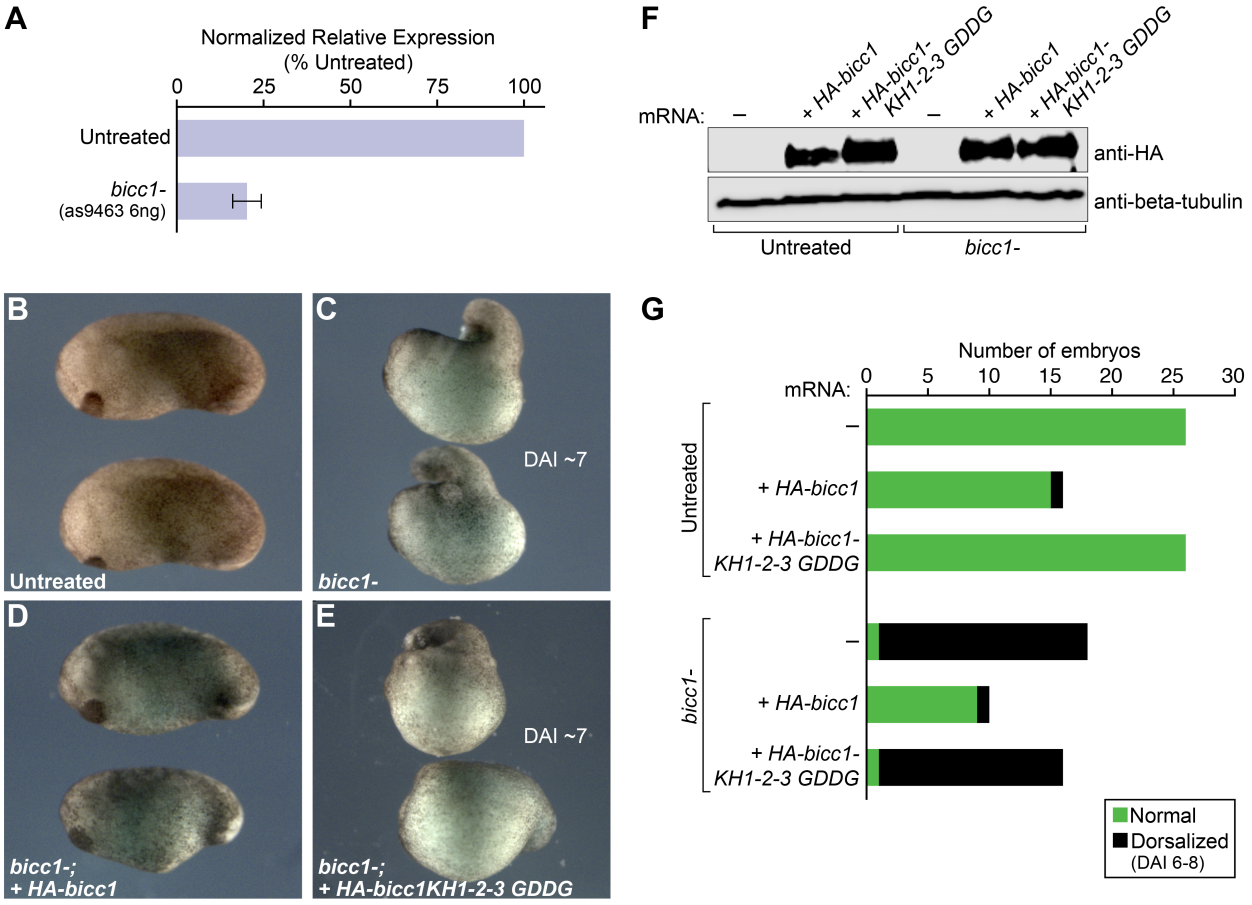

Figure 8. The Bicc1 KH domains are required for Bicc1's function in embryonic patterning.
Validated Bicc1 antisense phosphorothioate oligonucleotide (oligo 9463) was injected into oocytes and the oocytes matured overnight. Matured oocytes were treated with vital dyes, transferred to an ovulating host female, and the laid eggs from manipulated oocytes were fertilized. ***A-D.*** Phenotypes of control and sibling *experimental Xenopus* embryos. ***A.*** Control embryos (stage 30) ***B.*** Stage 30 embryos depleted of maternal *bicc1* mRNA. The maternal knock-out embryos develop with expanded dorsal-anterior structures (DAI 7). ***C.*** The defects from depleting embryos of *bicc1* were rescued by wild-type *bicc1* mRNA. Embryos depleted of maternal *bicc1* mRNA were injected at the vegetal pole with wild-type HA-*bicc1* mRNA (20pg). ***D.*** The defects from depleting embryos of *bicc1* were not rescued by KH1-2-3 GDDG *bicc1* mRNA. Embryos depleted of maternal *bicc1* mRNA were injected at the vegetal pole with KH1-2-3 GDDG *bicc1* mRNA (20pg). ***E.*** The wild-type and KH1-2-3 GDDG *bicc1* proteins were expressed at comparable levels. Proteins from maternal knock-out embryos injected with the different mRNAs were analyzed by immunblotting and probing with an HA antibody.

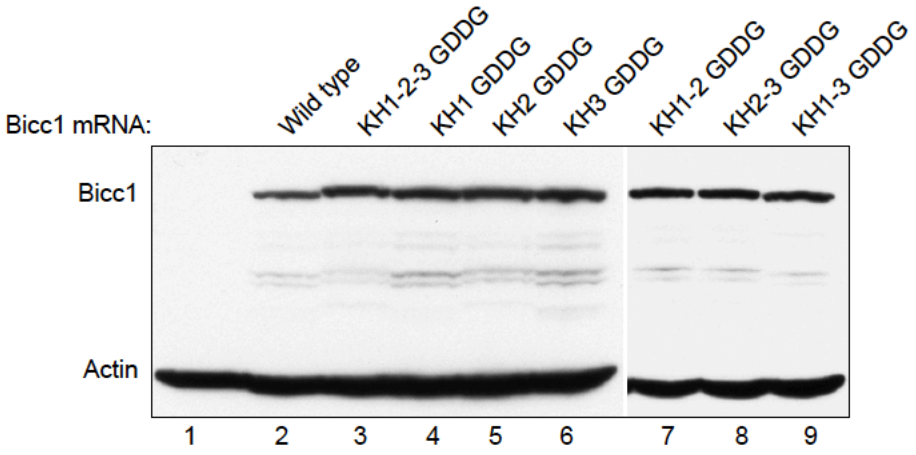

Figure. S1. **Expression of Bicc1 protein variants**.
The Bicc1 protein variants containing GDDG substitutions were expressed in *Xenopus* embryos and analyzed for binding to endogenous mRNAs (see **Fig.3**). Immunoblot analysis with an anti-HA antibody was used to monitor the expression of the different protein variants used in RNA binding assays. The blot was probed with an antibody to cytoskeletal actin to serve as a control.

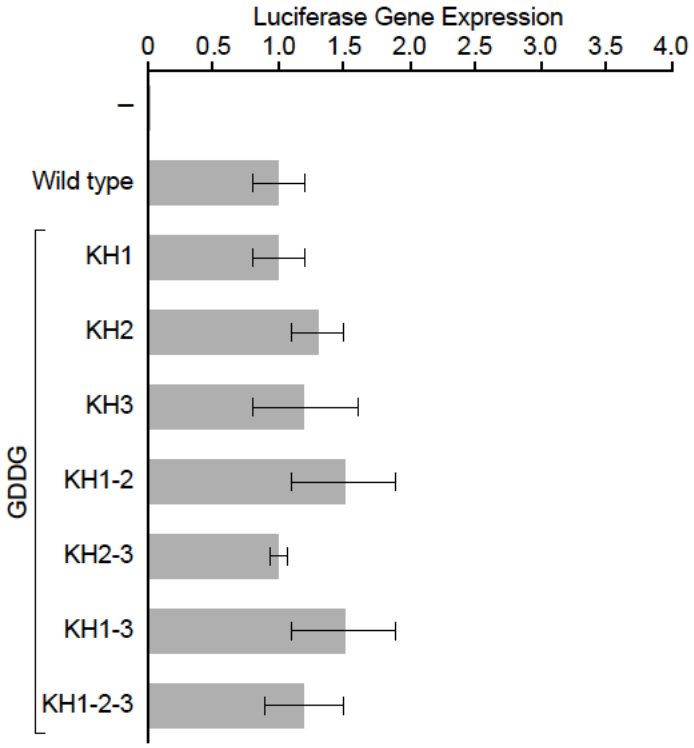

Figure. S2. Reporter mRNA stability.
RT-PCR was used to quantitate the luciferase-Cripto mRNA (Luc-Cripto1) used to analyze Bicc1 variants for translational repression. *Error bars* represent the standard error from three separate experiments.

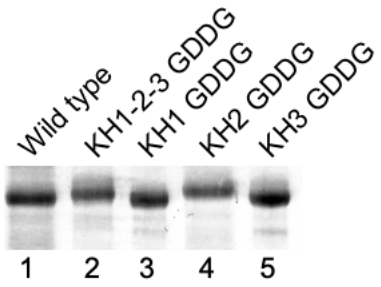

Figure. S3. Bicc1 protein variants used for RNA binding assays. The different Bicc1 protein variants (aa 1-506) purified from E.coli and used for RNA binding assays were analyzed by SDS polyacrylamide gel electrophoresis.

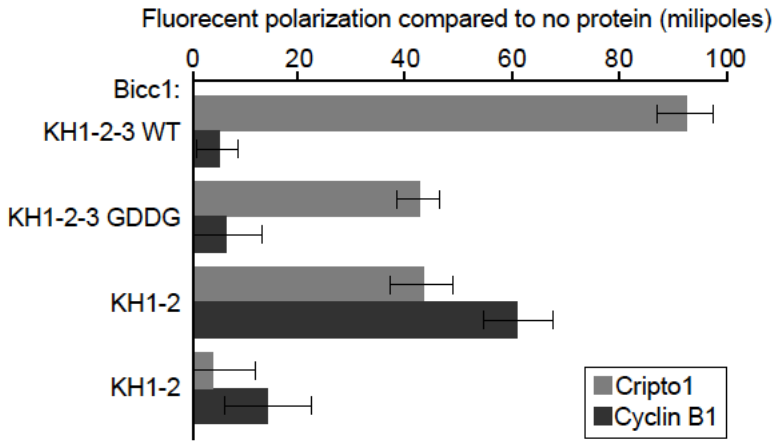

Figure. S4. RNA binding of wild-type and variant Bicc1 proteins by fluorescence polarization.
RNA binding assays were assembled using different Bicc1 proteins and either the fluorescently labeled Cripto1 RNA or the negative control cyclinB1 RNA. The binding of proteins to the RNAs was analyzed using a fluorescent plate reader and values were plotted in comparison to control reactions containing no protein.

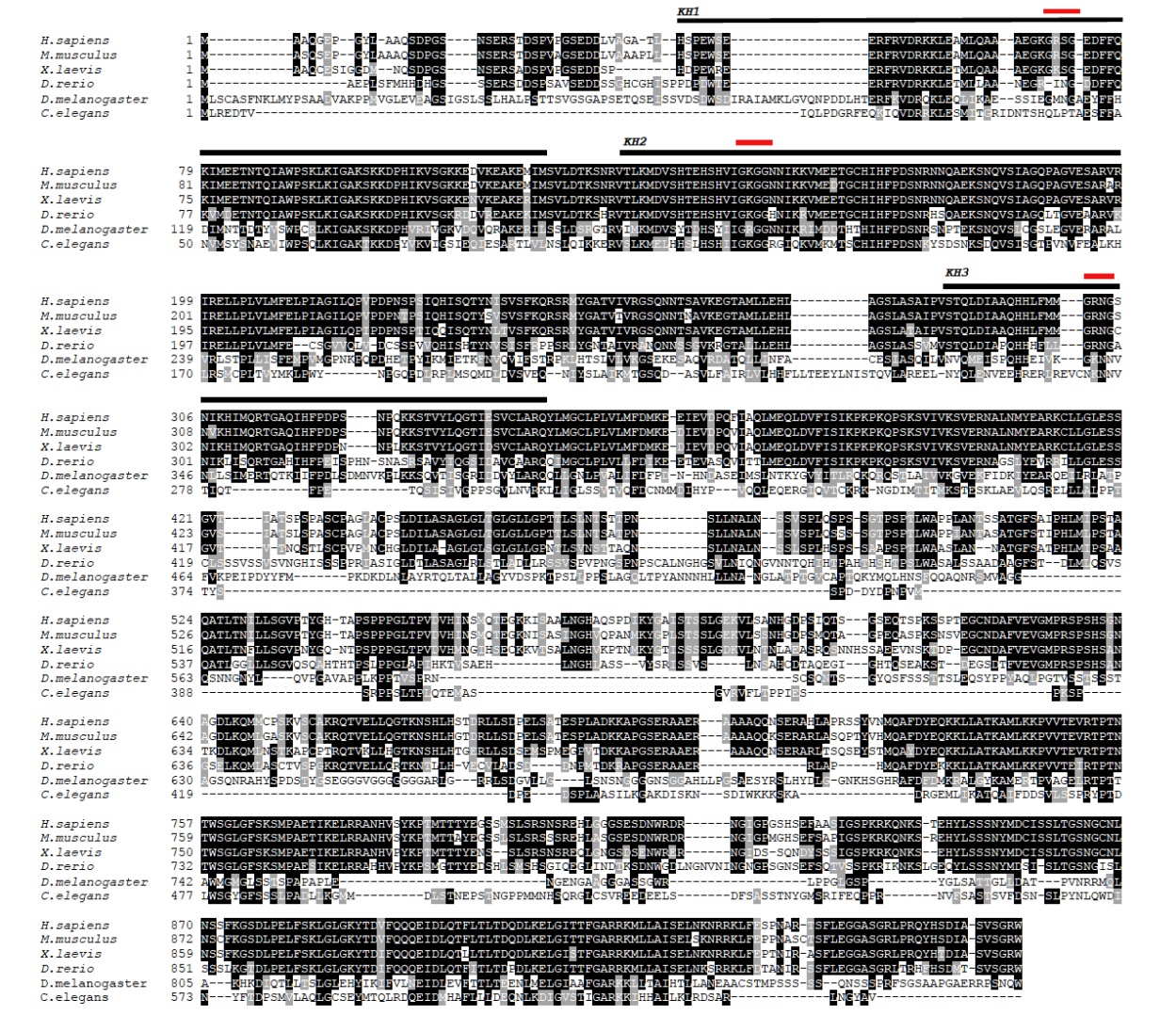

Figure. S5. The evolutionarily conserved features of Bicc1 proteins.
Amino acid sequences from vertebrate and invertebrate Bicc1 proteins were analyzed with Clustal Omega. Residues identical to human Bicc1 are highlighted in black, while similar residues are highlighted in grey. Each of the KH domains and the associated GXXG motifs are indicated.

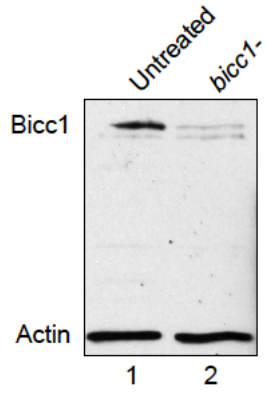

Figure. S6. Antisense oligos block accumulation of the Bicc1 protein during oocyte maturation. Oocytes were injected with the 9463 oligonucleotide and incubated for two hours. The injected oocytes were matured overnight with progesterone then the proteins analyzed by immunoblotting and probing with the Bicc1 antibody.

References:

Bull, Alice Louise (1966) Bicaudal, a genetic factor which affects the polarity of the embryo in Drosophila melanogaster. J Exp Zool B Mol Dev Evol, 161: 221-41.

Chao, J. A., Y. Patskovsky, V. Patel, M. Levy, S. C. Almo, and R. H. Singer (2010) ZBP1 recognition of beta-actin zipcode induces RNA looping. Genes Dev, 24: 148-58.

Cheng, M. H., and R. P. Jansen (2017) A jack of all trades: the RNA-binding protein vigilin. Wiley interdisciplinary reviews. RNA, 8.

Cooke, A., A. Prigge, and M. Wickens (2010) Translational repression by deadenylases. J Biol Chem, 285: 28506-13.

Dominguez, D., P. Freese, M. S. Alexis, A. Su, M. Hochman, T. Palden, C. Bazile, N. J. Lambert, E. L. Van Nostrand, G. A. Pratt, G. W. Yeo, B. R. Graveley, and C. B. Burge (2018) Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Mol Cell, 70: 854-67 e9.

Dowdle, M. E., S. B. Imboden, S. Park, S. P. Ryder, and M. D. Sheets (2017) Horizontal Gel Electrophoresis for Enhanced Detection of Protein-RNA Complexes. Journal of visualized experiments : JoVE.

Du, Z., J. K. Lee, S. Fenn, R. Tjhen, R. M. Stroud, and T. L. James (2007) X-ray crystallographic and NMR studies of protein-protein and protein-nucleic acid interactions involving the KH domains from human poly(C)-binding protein-2. Rna, 13: 1043-51.

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

Fritz, B. R., and M. D. Sheets (2001) Regulation of the mRNAs encoding proteins of the BMP signaling pathway during the maternal stages of Xenopus development. Dev Biol, 236: 230-43.

Gamberi, C., and P. Lasko (2012) The Bic-C family of developmental translational regulators. Comparative and functional genomics, 2012: 141386.

Gerstberger, S., M. Hafner, and T. Tuschl (2014) A census of human RNA-binding proteins. Nat Rev Genet, 15: 829-45.

Heasman, J., S. Holwill, and C. C. Wylie (1991) Fertilization of cultured Xenopus oocytes and use in studies of maternallyA inherited molecules. Methods Cell Biol, 36: 213-30.

Hentze, M. W., A. Castello, T. Schwarzl, and T. Preiss (2018) A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol, 19: 327-41.

Hollingworth, D., A. M. Candel, G. Nicastro, S. R. Martin, P. Briata, R. Gherzi, and A. Ramos (2012) KH domains with impaired nucleic acid binding as a tool for functional analysis. Nucleic Acids Res, 40: 6873-86.

Kraus, M. R., S. Clauin, Y. Pfister, M. Di Maio, T. Ulinski, D. Constam, C. Bellanne-Chantelot, and A. Grapin-Botton (2012) Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. Human mutation, 33: 86-90.

Lemaire, L. A., J. Goulley, Y. H. Kim, S. Carat, P. Jacquemin, J. Rougemont, D. B. Constam, and A. Grapin-Botton (2015) Bicaudal C1 promotes pancreatic NEUROG3+ endocrine progenitor differentiation and ductal morphogenesis. Development, 142: 858-70.

Mahone, M., E. E. Saffman, and P. F. Lasko (1995) Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. Embo J, 14: 2043-55.

Maisonneuve, C., I. Guilleret, P. Vick, T. Weber, P. Andre, T. Beyer, M. Blum, and D. B. Constam (2009) Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. Development, 136: 3019-30.

Malakhov, M. P., M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks, and T. R. Butt (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. Journal of structural and functional genomics, 5: 75-86.

Nakel, K., F. Bonneau, C. R. Eckmann, and E. Conti (2015) Structural basis for the activation of the C. elegans noncanonical cytoplasmic poly(A)-polymerase GLD-2 by GLD-3. Proc Natl Acad Sci U S A, 112: 8614-9.

Nicastro, G., A. M. Candel, M. Uhl, A. Oregioni, D. Hollingworth, R. Backofen, S. R. Martin, and A. Ramos (2017) Mechanism of beta-actin mRNA Recognition by ZBP1. Cell reports, 18: 1187-99.

Nicastro, G., I. A. Taylor, and A. Ramos (2015) KH-RNA interactions: back in the groove. Current opinion in structural biology, 30: 63-70.

Pagano, J. M., C. C. Clingman, and S. P. Ryder (2011) Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. Rna, 17: 14-20.

Park, S., S. Blaser, M. A. Marchal, D. W. Houston, and M. D. Sheets (2016) A gradient of maternal Bicaudal-C controls vertebrate embryogenesis via translational repression of mRNAs encoding cell fate regulators. Development, 143: 864-71.

Park, S., E. E. Patterson, J. Cobb, A. Audhya, M. R. Gartenberg, and C. A. Fox (2011) Palmitoylation controls the dynamics of budding-yeast heterochromatin via the telomere-binding protein Rif1. Proc Natl Acad Sci U S A, 108: 14572-7.

Piazzon, N., C. Maisonneuve, I. Guilleret, S. Rotman, and D. B. Constam (2012) Bicc1 links the regulation of cAMP signaling in polycystic kidneys to microRNA-induced gene silencing. Journal of molecular cell biology, 4: 398-408.

Saffman, E. E., S. Styhler, K. Rother, W. Li, S. Richard, and P. Lasko (1998) Premature translation of oskar in oocytes lacking the RNA-binding protein bicaudal-C. Mol Cell Biol, 18: 4855-62.

Sheets, M. D., C. A. Fox, T. Hunt, G. Vande Woude, and M. Wickens (1994) The 3'-untranslated regions of c-mos and cyclin mRNAs stimulate translation by regulating cytoplasmic polyadenylation. Genes Dev, 8: 926-38.

Teplova, M., L. Malinina, J. C. Darnell, J. Song, M. Lu, R. Abagyan, K. Musunuru, A. Teplov, S. K. Burley, R. B. Darnell, and D. J. Patel (2011) Protein-RNA and protein-protein recognition by dual KH1/2 domains of the neuronal splicing factor Nova-1. Structure, 19: 930-44.

Valverde, R., L. Edwards, and L. Regan (2008) Structure and function of KH domains. Febs J, 275: 2712-26.

Yaguchi, S., J. Yaguchi, and K. Inaba (2014) bicaudal-C is required for the formation of anterior neurogenic ectoderm in the sea urchin embryo. Scientific reports, 4: 6852.

Zhang, Y., A. Cooke, S. Park, C. N. Dewey, M. Wickens, and M. D. Sheets (2013) Bicaudal-C spatially controls translation of vertebrate maternal mRNAs. Rna, 19: 1575-82.

Zhang, Y., K. D. Forinash, J. McGivern, B. Fritz, K. Dorey, and M. D. Sheets (2009) Spatially restricted translation of the xCR1 mRNA in Xenopus embryos. Mol Cell Biol, 29: 3791-802.

Zhang, Y., S. Park, S. Blaser, and M. D. Sheets (2014) Determinants of RNA binding and translational repression by the Bicaudal-C regulatory protein. J Biol Chem, 289: 7497-504.

Zheng, S., P. Lan, X. Liu, and K. Ye (2014) Interaction between ribosome assembly factors Krr1 and Faf1 is essential for formation of small ribosomal subunit in yeast. J Biol Chem, 289: 22692-703.

Chapter 4:

Defining critical RNA features of the Bicc1-RNA interface

Megan E. Dowdle and Michael D. Sheets

I designed and performed all the experiments and collected all the data. I wrote the manuscript with Michael Sheets.

ABSTRACT

Bicaudal-C (Bicc1) is a conserved RNA binding protein and developmental regulator that functions in metazoans to selectively repress the translation of specific mRNAs. This repression is essential for controlling the synthesis of proteins that regulate developmental decisions. Despite Bicc1's biological relevance, there is a lack of experimental structure-function data that addresses how this protein binds to specific mRNA targets. Bicc1 binds to a 32-nucleotide element in the 3'UTR of the CR1 mRNA. However, it is unknown the exact sequence motif(s) that Bicc1 recognizes. Previous studies have shown that mutations in the leader and loop sequences of the element adversely affect Bicc1's ability to stably bind the CR1 mRNA. In order to determine the exact sequence motif that Bicc1 binds to in this 32-nucleotide structure, we have systemically changed the nucleotides in both the leader and the loop and monitored the binding via electrophoretic mobility shift assays (EMSAs) and fluorescent polarization. Our results have indicated that Bicc1 binding to the CR1 target is complex. Sequences in both the leader and the loop are crucial for mRNA binding indicating a possible bipartite motif. Additionally, while several RNA mutants impair Bicc1 binding, no single mutation abolishes Bicc1 binding to the CR1 mRNA target. Results from these experiments are critical for understanding the mechanisms by which Bicc1 influences embryonic development as well as understanding how RNA-binding domains function together to identify specific target RNA sequences.

INTRODUCTION:

Bicaudal-C (Bicc1) is a conserved RNA binding protein and developmental regulator that functions in metazoans to selectively repress the translation of specific mRNAs (Gamberi and Lasko, 2012). This repression is essential for controlling the synthesis of proteins that in turn regulate key developmental decisions. In early *Drosophila* and *Xenopus* embryos, defects in Bicc1 function lead to aberrant anterior-posterior patterning (Saffman et al., 1998; Wessely and De Robertis, 2000; Mohler and Wieschaus, 1986; Park et al., 2016). In later development, Bicc1 influences proper organogenesis (Maisonneuve et al., 2009; Li et al., 2015). Additionally, defects in Bicc1 have been linked to human diseases such as polycystic kidney disease (Piazzon et al., 2012; Tran et al., 2010; Gamberi et al., 2017; Rothé et al., 2019), cancer (Li et al., 2020), and major depressive disorder (MDD) (Bermingham et al., 2012; Ota et al., 2015; Wang et al., 2019). However, despite its biological importance, it is not well understood how Bicc1 selectively binds and represses its specific mRNA targets.

Work from the Sheets lab using early *Xenopus* embryos demonstrated that Bicc1's ability to bind mRNA is critical for its biological role in translational repression (Dowdle et al., 2019). When a mutant Bicc1 that cannot bind mRNA is expressed in *Xenopus* embryos depleted of wild-type Bicc1 the embryos develop exaggerated anterior structures at the expense of the posterior (Dowdle et al., 2019) (**Chapter 3**). Studies in *Xenopus* indicate that the N-terminal region of the Bicc1 protein, which contains three canonical KH-domains and two KH-like (KHL) domains, is both necessary and sufficient for mRNA binding (Zhang et al., 2014). KH domains are conserved RNA binding features first identified in the hnRNP K protein (Grishin, 2001;

Valverde et al., 2008). Canonical KH domains are known to bind to RNA through their evolutionarily conserved GXXG motifs (Nicastro et al., 2015; Nicastro et al., 2017), and through studies of the Nova-1 protein it has been proposed that individual KH domains recognize tetranucleotide motifs in their specific RNA substrates (Jensen et al., 2000). It has also been suggested that proteins with multiple RNA binding domains function on a modular basis (Lunde et al., 2007; Hentze et al., 2018). For example, RNA binding proteins like hnRNP K (Paziewska et al., 2004), ZBP1 (Nicastro et al., 2017), and KSRP (Díaz-Moreno et al., 2010) contain multiple RNA binding domains that function modularly and combinatorially. Thus, an emerging theme is that RNA-binding proteins only utilize a certain subset of their domains to bind RNA. In depth *in vitro* and *in vivo* analysis revealed that for the *Xenopus* Bicc1 protein, the second KH domain (KH2) is critical for Bicc1's mRNA binding activity, while the KH1 and KH3 domains make minimal contributions to binding (Dowdle et al., 2019) (**Chapter 3**). However, it was also shown that though KH2 is necessary for binding, it is not sufficient for specific binding to Bicc1 target mRNAs (Dowdle et al., 2019).

With regards to the RNA features that define Bicc1 binding sites, the most extensive information has come from studies of the *Xenopus* Cripto1 mRNA and the 32-nucleotide Bicc1 binding site that resides within its 3'UTR (Zhang et al., 2014). The 32-nucleotide Bicc1 binding site forms a putative stem-loop structure with a 12-nucleotide leader sequence preceding the stem-loop (**Fig.1A**). Through analysis of this RNA, it was found that both sequence and structural features are required for Bicc1 binding. It was determined that the stem structure was important, but the exact sequence of the stem was not critical (Zhang et al., 2014). However, the stem-loop was not sufficient for

binding suggesting that other regions of the 32-nucleotide RNA contribute important information necessary recognition by Bicc1 (Zhang et al., 2014).

In this study, I have used biochemical assays to provide new information about the RNA binding features of the Bicc1 protein. By making mutations in the 32-nucleotide binding site of Cripto-1 I analyzed the binding affinity of Bicc1 to the various mutant sequences through electrophoretic mobility shift assays (EMSA) and fluorescent polarization (FP) analysis. I first showed that Bicc1 binds specifically to the Cripto-1 binding site with high affinity and that this binding is reliant on sequences within the loop and leader. This indicates that Bicc1 recognizes a non-contiguous or bipartite binding site in its RNA targets. By mutating and analyzing individual nucleotides in the loop and leader, I was able to determine important sub-regions within the loop and the leader and identify preferred nucleotides at specific positions for efficient Bicc1 binding. Additionally, I have purified a recombinant N-terminal version of the human Bicc1 protein. Using this protein, I determined that the human protein binds to the same 32-nucleotide target as *Xenopus* Bicc1. This suggests that the binding features between the human and *Xenopus* protein are conserved.

RESULTS:

RNA binding analysis of *Xenopus* Bicc1. To analyze Bicc1 RNA binding functions a recombinant protein consisting of the N-terminal region (aa 1-506) of *Xenopus laevis* Bicc1 was expressed and purified from *E.coli* (Dowdle et al., 2017) (**Chapter 2**). This protein was used in EMSA (**Fig.1B**) and FP binding (**Fig.1C**) assays with fluorescein labeled RNA substrates. The well-characterized 32-nucleotide binding

site from the 3'UTR of the Cripto-1 mRNA was used as the substrate (**Fig.1A**) while a 32-nucleotide RNA from the 3'UTR of the CyclinB1 mRNA was used as the negative control (Dowdle et al., 2017) (**Chapter 2**). Reactions containing increasing amounts of Bicc1 demonstrate that the protein has an apparent K_d of 176nM for its specific substrate (**Figs.1A and 1B**). This binding was very specific as the protein binds very inefficiently to the negative control RNA from the 3'UTR of the Cyclin B1 mRNA.

Binding analysis of human Bicc1. The functions of Bicc1 are defective in several human disease states. For example, Bicc1 mutations predicted to cause loss of function are associated with kidney disease (Kraus et al., 2012; Piazzon et al., 2012; Leal-Esteban et al., 2018; Rothé et al., 2019), while DNA translocations that result in a fusion between the fibroblast growth factor receptor 2 (FGFR2) and Bicc1 genes (FGFR2-Bicc1) cause cancer of the bile duct called cholangiocarcinoma (Arai et al., 2014; Ying et al., 2019; Li et al., 2020). Studies in the Sheets lab have focused on analyzing *Xenopus* Bicc1 and its binding to known RNA targets from *Xenopus* embryos (Zhang et al., 2013; Park et al., 2016). However, I wanted to exploit our expertise and approaches to begin analyzing human Bicc1 to potentially provide insights into Bicc1 associated disease states. To begin, I specifically wanted to determine if human Bicc1 would recognize *Xenopus* RNA substrates. This experiment required production and purification of the human Bicc1 N-terminal domain protein. Despite sharing a high degree of amino acid similarity with the *Xenopus* protein (Dowdle et al., 2019), the human protein was insoluble when expressed in *E.coli* (S. Park unpublished). To address this issue, we took two approaches. First, we used a transient expression system in HEK293 human cells grown in suspension and were able to

obtain human Bicc1 protein that was soluble. However, low yields with this approach suggested that it would require significant effort to purify sufficient quantities for experiments (S. Moffett unpublished). In a second approach, we created a modified plasmid for expressing human Bicc1 in *E.coli*. In this plasmid the human Bicc1 N-terminal region was codon optimized for *E.coli* expression. This strategy was successful, and I was able to express and purify soluble human Bicc1 N-terminal protein from *E.coli* (**Fig. 2A)**. This is the first time the human protein has been successfully generated in sufficient quantities for biochemical analysis.

The binding activity of the human Bicc1 N-terminal protein was analyzed in both EMSA (**Fig. 2B)** and FP assays (**Fig. 2C**) with the 32-nucleotide binding site of *Xenopus*Cripto-1 (**Fig.1A**) and Cyclin B1 RNA was used as the negative control. Reactions containing increasing amounts of protein show that human Bicc1 binds very efficiently with an apparent Kd of 31nM for the *Xenopus* Cripto-1 substrate (**Figs. 2B and 2C**) while binding very inefficiently to the negative control CyclinB1. The human Bicc1 shows a higher affinity for the Cripto-1 substrate compared to the *Xenopus* protein and also displays a higher affinity for Cyclin B1, though still very inefficient compared to Cripto-1 binding. This could be indicative that there may be slight preferences in affinity between the human and *Xenopus* Bicc1 for the Cripto-1 RNA due to slight differences in their KH domains, particularly KH1 and KH3. This data shows for the first time that the binding between the RNA binding activities for *Xenopus* Bicc1 and human Bicc1 are conserved.

The loop and leader regions of the 32nt Cripto1 RNA are required for Bicc1 binding. The 32-nucleotide binding site of Cripto-1 is comprised of a leader, loop, and

stem sequences (**Fig. 3A top**). Previous studies of the 32 nucleotide Bicc1 binding site from the *Xenopus* Cripto1 mRNA determined that structure of the stem rather than its sequence was an important determinant of binding (Zhang et al., 2014). These studies also suggested that efficient recognition by Bicc1 relied upon features of the RNA in addition to the stem (Zhang et al., 2014). Therefore, to begin examining the necessity of specific sequences beyond the stem, variant 32nt RNAs were generated in which the sequences of the leader and the loop were changed (**Fig. 3A middle and bottom**). These mutant leader and loop RNAs were then examined for Bicc1 binding via EMSA with a fixed amount of Bicc1 protein (**Fig. 3B**) and FP assays with increasing amounts of protein (**Fig. 3C**). I observed that multiple nucleotide changes of the leader or the loop sequences significantly decreased Bicc1 binding to the Cripto-1 substrate (**Fig. 3B and 3C**). This demonstrates that the leader and loop are important features for Bicc1 binding. In addition, because efficient binding requires both wild-type sequences at the loop and the leader this indicates that the binding site is complex and the required sequence elements are non-continuous. Similar non-continuous and complex sequence elements have recently been described for other RNA binding proteins and in some cases when there are two adjacent sequence elements they are referred to as bipartite (Dominguez et al., 2018; Loughlin et al., 2019).

Defining nucleotides in the loop of the 32nt Cripto1 RNA required for Bicc1 binding. The Cripto1-Loop RNA containing multiple nucleotide substitutions was defective for recognition and binding by the Bicc1 protein (**Fig.3B**). To determine the contribution and importance of individual nucleotides of the loop for Bicc1 Binding, additional variants were generated in which each variant contained a single nucleotide

change of a specific position within the loop (**Fig.4A**). Each RNA variant was analyzed using EMSA (**Fig.4Bi, ii, and iii**) and FP (**Fig.4C**) assays for binding to the Bicc1 N-terminal protein.

These experiments revealed that while there are significant impacts on binding, no single nucleotide abolished binding to the level of the multiple substitution loop mutant (**Figs.3A and 3B**). Still, some nucleotide preferences were observed. For positions 18-20 (**Fig. 4A**), there was a stronger preference for the A and C nucleotides for Bicc1 binding (**Figs. 4B and C**). Position 21 was only location where only the wild-type nucleotide (a U) was optimal for binding (**Figs. 4B and C**). At position 22 either pyrimidine (U or C) was sufficient for binding, while position 23 A or C nucleotides were sufficient.

These results suggest that the nucleotides positions 21-23 are the most critical for Bicc1 binding and hence the least tolerable to changes. However, the data also suggest that while there are preferences for particular nucleotides at specific positions, a strict consensus sequence of the loop for optimal binding is difficult to define. I observed some differences between the two binding assays with the same RNA substrates. These slight differences between the more qualitative EMSAs and the more quantitative FP (**compare Figs. 4B and C**) could be due to the fact that the fluorescent polarization assay samples a dynamic population measuring both stable and unstable RNA-protein complexes at the same time. As a result, subtle changes might not be as easily captured in this assay.

Defining nucleotides in the Leader of the 32nt Cripto1 RNA required for Bicc1 binding. The Cripto1 32nt RNA containing ten nucleotide substitutions of the

Leader was defective for binding by Bicc1 protein (**Figs.3B and 3C**). To define the specific nucleotides of the leader important for Bicc1 binding, addition RNA variants were generated in which two or three nucleotides were changed a specific position within the leader (**Fig.5A**). Each RNA variant was analyzed using EMSA (**Fig.5B**) and FP (**Fig.5C**) binding assays with Bicc1 protein. RNA variants in which contained substitutions of the 2-4 and 11-12 positions (2,4CG, 2-4ACG, 2-4GAC and 11-12UG) were bound by Bicc1 with an efficiency similar to the binding of the wild-type Cripto1 RNA (binding values of individual RNAs greater than 0.8 where binding of the WT RNA is 1.0) (**Figs.5B and 5C**). In contrast, the 5-6UU, 8-10GAC and 8-10CUG substitutions diminished Bicc1 binding (binding values of individual RNAs between 0.3 and 0.6,**Figs.5B and 5C**). Interestingly, the 5-7GAC, 8-10ACA and 12G variants bound to Bicc1 with a higher efficiency than wild-type RNA. Together these results suggest that nucleotides at positions, 5-7 and 8-10 are important determinants of the leader for Bicc1 binding while nucleotides at positions 2-4 and 11-12 minimally impact binding.

To begin to define individual nucleotides of the leader important for Bicc1 binding, separate RNA variants containing each possible nucleotide change at the 8, 9 and 10 positions of the 32nt Cripto1 RNA (Fig.6A) were analyzed using EMSA (**Fig.6B**) and FP (**Fig.6C**) binding assays. All the RNA variants which contained substitutions of position 8 (8U, 8C or 8A) were bound by Bicc1 with an efficiency close to the binding to wild-type RNA. Similar results were observed with single nucleotide variants of position 10, except for the 10A substitution. The 10A substitution reduced binding to 52% of the wild-type RNA. In contrast, all nucleotide substitutions at position 9 diminished Bicc1 binding (binding values of individual RNAs between 0.35 and 0.5) (**Fig.6B and 6C**).

DISCUSSION AND FUTURE STUDIES:

Previous studies of Bicc1 identified a 32-nt binding site from the 3'UTR of the *Xenopus* Cripto1 mRNA (Zhang et al., 2014). This work demonstrated that Bicc1 binding required both structural and sequence features present in the 32-nt RNA. These results suggested that Bicc1 recognition of RNA substrates was complex and not the binding of a protein to a simple sequence element. To address this complexity, I performed in-depth studies of the sequences within the 32nt RNA that are required for efficient and specific Bicc1 binding.

My results suggest that Bicc1 interacts with the Cripto-1 mRNA via a bipartite binding site. Multinucleotide substitutions of either the loop or the leader decreased Bicc1 binding to the 32nt RNA. The loop and leader sequences are separated by the stem of the stem-loop structure. Previous studies indicated that base-pairing to form the double-stranded stem was important for binding, but the exact sequence of the stem could be changed extensively without affecting binding (Zhang et al 2014). Thus, my results indicate that the loop and leader sequences important for Bicc1 binding are non-contiguous and together form a complex element referred to as bipartite binding site.

Recent studies have surprisingly revealed the presence of bipartite binding sites in the RNA targets of several RNA binding proteins (Dominguez et al., 2018; Loughlin et al., 2019; Biswas et al., 2019). A broad study of 78 RNA-binding proteins in humans found that several of the proteins analyzed preferred to bind to short bipartite RNA motifs over longer 6-mer motifs (Dominguez, et al. 2018). Furthermore, studies conducted with members of the IGF2 family of RNA-binding proteins found that these proteins recognize a bipartite motif in their targets (Nicastro, et al., 2017; Biswas et al.,

2019), though these studies were conducted with shorter fragments of RNA and thus did not look into structure of the RNA substrate. Another protein that binds to a bipartite motif is fused in sarcoma (FUS). In a recent study, the authors found that FUS requires both sequence elements and a stem-loop structure similar to Bicc1 with the recognition of specific sequences of RNA through both its Zinc Finger (ZnF) domain and its RRM domain, additionally the RRM recognized a stem-loop structure as and that in certain cases these two domains work synergistically to enhance RNA binding (Loughlin et al., 2019). Consequently, the similarities between the substrate of FUS and Cripto-1 may provide a blueprint for how Bicc1 might bind to its RNA targets.

My data support a model where the individual elements of the Bicc1 bipartite binding site function together to promote efficient binding. Multinucleotide substitutions of the loop or the leader significantly reduced Bicc1 binding, indicating that in the absence of one sequence the other cannot compensate and promote efficient binding. In addition, RNAs consisting of just the stem-loop (Zhang et al 2014) or just the leader (Dowdle et al data not shown) were not efficiently bound by Bicc1. These results suggest that the loop and leader sequences function together to promote binding. For example, the two elements of the bipartite binding site may form a specific RNA structure that is optimal for Bicc1 recognition (Dominguez et al., 2018; Loughlin et al., 2019). Alternatively, the two elements of the bipartite site could represent separate regions of interaction with the Bicc1 protein and such an interaction promotes efficient binding (Loughlin et al., 2019). Another possibility is that the two elements of the bipartite site serve to multimerize Bicc1 and this promotes binding. Distinguishing between these possibilities will require additional biochemical experiments, structural

analysis of the Bicc1-RNA complex and the identification and characterization of Bicc1 binding sites from other target mRNAs.

The analysis of individual nucleotides in the loop or regions in leader did not result in a complete loss of binding. This suggests that individual nucleotide changes may be compensated for by the presence of two distinct binding sites in the RNA. Recent studies on mouse Bicc1 have shown that in extracts from HEK293T cells Bicc1 preferentially binds to RNAs containing a GAC sequence (Minegishi et al., 2020). However, this study lacked quantitative binding analysis of this motif and did not analyze the importance of this sequence within the context of an authentic 3'UTR. While mutational analysis of the leader suggests a similar preference (**Fig. 5B** and **Fig. 5C**), this cannot be the only determinant of Bicc1 binding. In particular, while the loop is important for Bicc1 binding, it lacks a GAC sequence. Furthermore, changes of leader nucleotides at positions 8 or 10 did not abolish Bicc1 binding despite removing the natural GAC sequence. However, in other contexts this sequence does seem to have the ability to enhance binding as the mutant that most strongly enhanced binding (**5-7GAC, Fig. 5B and 5C**) created two adjacent GAC sequences in the leader.

My results revealed that the nucleotide requirements for Bicc1 binding of the loop and leader sequences were complex and did not readily lend themselves to parsing out consensus binding motifs. However, K-mer analysis may be able to harness the subtleties in the binding data that are not immediately obvious in order to create consensus sequences (Audoux et al., 2017). If such consensus motifs can be derived, they might prove useful for analyzing other Bicc1 regulated 3'UTRs for similar binding sites.

The experiments I have described analyzed Bicc1 binding using *in vitro* biochemical assays. An important set of future studies will use the advantages of *Xenopus* embryos to analyze binding *in vivo* (Park et al 2016). It will also be important to use luciferase reporter assays to test the ability of Bicc1 to repress reporter mRNAs containing variants of the 32nt binding site (Zhang et al 2014). These *in vivo* assays are crucial to building an understanding of the biological relevance of the mutations in the Cripto-1 target. Previous studies in our lab have shown that in some cases mild effects in *in vitro* binding assays can have a more severe impact *in vivo* (Dowdle et al., 2019). Therefore, *in vivo* assays may provide additional information to understand and facilitate identifying a consensus sequence motif optimal for Bicc1 binding.

Finally, I have demonstrated that human Bicc1 can be purified for biochemical assays and that this protein binds with similar affinity and specificity as *Xenopus* Bicc1. This is the first time that human Bicc1 has been generated and used in biochemical studies. This protein can now be used to further identify potential human targets as well as analyze the role that disease-causing mutations impact the ability for human Bicc1 to bind to its mRNA targets.

This work advances our understanding in how Bicc1 binds to its RNA targets. I found that Bicc1 binds to at least one of its target RNAs through a bipartite binding site. This information coupled with the nucleotide preferences I identified could be useful as we seek to identify and characterize the Bicc1 binding sites in additional target mRNAs. Furthermore, this work adds to the field of understanding how multiple RNA-binding domain containing proteins identify and bind to unique targets with high specificity despite recognizing minimal, short RNA sequence motifs.

METHODS

Bicc1 protein expression and purification

Bicc1 proteins were cloned into pET28b Bacterial expression vectors as N-terminal fusions with a His-6, SUMO tag (Malakhov et al., 2004). Cultures of *E.coli* cells containing each plasmid were grown to an OD600 of 0.6 and induced with 1mM IPTG at 25 degrees C overnight. The cells were collected and lysed in B-PER reagent, 1/2X TBS, 2.5mM MgCl2, 2.5% glycerol, 1mM BME, 10mM Imidazole, 200mM NaCl, 1mM ATP, and protease inhibitors. The soluble lysate was applied to a HisTrap FF chromatography column. The column was washed 10 times with 1xTBS, 5mM MgCl2, 5% glycerol, 2mM BME, 20mM Imidazole, 400mM NaCl and the proteins were then eluted with 450mM imidazole and dialyzed in 1x TBS.

EMSA

Recombinant SUMO-Bicc1 N-terminal proteins were expressed and purified as described above. The 32-nucleotide Cripto1 and CyclinB1 3'-fluorescein-labeled RNA substrates were purchased from IDT. Binding reactions (50µl) contained SUMO-Bicc1 protein, 10mM Hepes, pH 7.5, 1mM EDTA, 50 mM KCl, 0.02% Tween 20, 0.1mg/ml yeast tRNA, 100µg/ml BSA, 2mM DTT, and 10nM RNA. Reaction products were analyzed on 7.5% (1× TBE) native polyacrylamide gels (Dowdle et al., 2017). The gels were then scanned at 473 nm using a fluorimager.

Fluorescent Polarization assays for RNA binding

Binding reactions as described above were assembled into individual wells of a 96-well black round bottom plate. The reactions were scanned using a plate reader with an excitation wavelength at 485nm and an emission wavelength of 528nm in the parallel

and perpendicular direction (Pagano et al., 2011). The data were analyzed using Gen5 software to generate fluorescent polarization binding graphs.

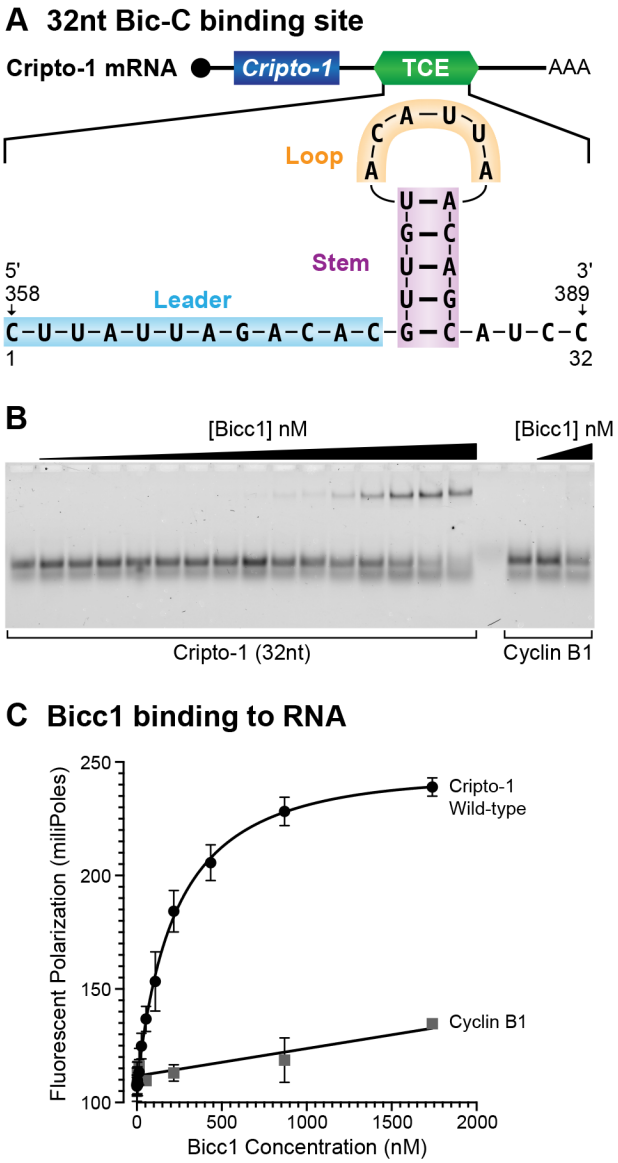

Figure 1. Bicc1 binds to a 32-nucleotide binding site in Cripto-1.
A. Diagram of the Cripto-1 mRNA with the translational control element (TCE) highlighted and expanded. The sequence of the 32-nucleotide binding site with its predicted stem-loop structure is shown with leader, stem, and loop highlighted. ***B.*** EMSA experiment analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1. The free RNA is shown at the bottom with the bound complex highlighted at the top. ***C.*** Fluorescent polarization assay analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1. The $K_{d,app}$ was determined by analyzing total binding of Cripto-1 compared to Cyclin B1 with Prism software and is the average of at least three experiments. Error shown is standard deviation. Little or no binding was observed for Cyclin B1 so the $K_{d,app}$ is not reported.

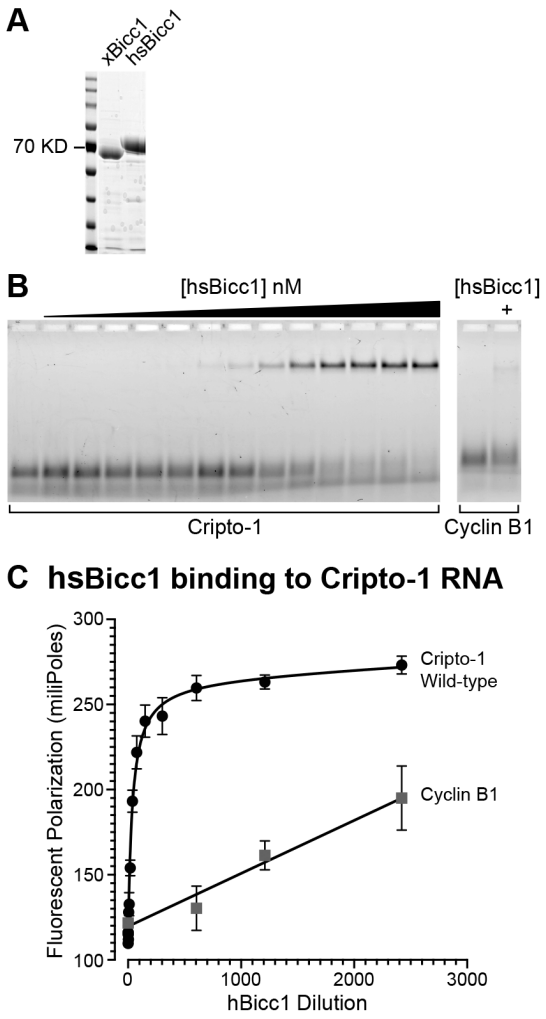

Figure 2. Human Bicc1 binds to the 32-nucleotide *Xenopus* Cripto-1 binding site. ***A.*** The human protein purified from *E. coli* and used for RNA binding assays was analyzed with SDS polyacrylamide gel electrophoresis. ***B.*** An EMSA experiment analyzing the affinity of human Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1. The free RNA is shown at the bottom with the bound complex highlighted at the top. ***C.*** Fluorescent polarization assays analyzing the affinity of human Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1. The $K_{d,app}$ was determined by analyzing total binding of Cripto-1 compared to Cyclin B1 with Prism software and is the average of at least three experiments. Error shown is standard deviation. Inefficient binding was observed for Cyclin B1 so the $K_{d,app}$ was not reported.

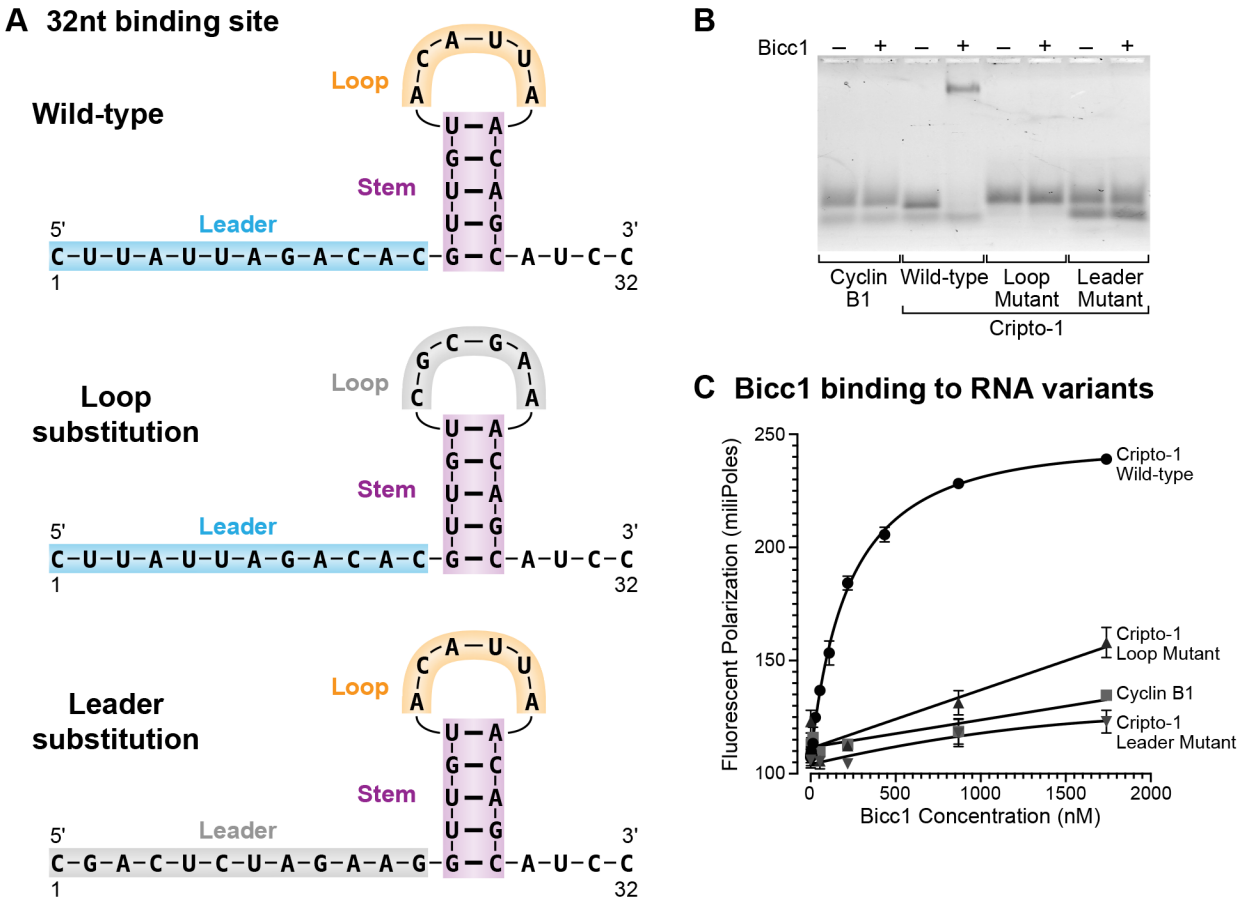

Figure 3. Bicc1 requires sequences in both the leader and the loop to bind RNA. ***A.*** Diagrams of the wild-type Cripto-1 32-nucleotide binding site (top), the Cripto-1 loop mutant sequence highlighted in grey (middle), and the Cripto-1 leader mutant sequence highlighted in grey (bottom). ***B.*** EMSA experiments comparing the binding of Bicc1 to the Cripto-1 wild type, Cripto-1 loop mutant, Cripto-1 leader mutant, and Cyclin B1. The assays were run with 350nM of Bicc1 protein to 10nM of fluorescein labeled RNA. The free RNA is shown at the bottom of the gel while bound complexes are shown at the top of the gel. ***C.*** Fluorescent polarization assays analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to the Cripto-1 loop mutant, Cripto-1 leader mutant, and a 32-nucleotide site in Cyclin B1. The $K_{d,app}$ was determined by analyzing total binding of Cripto-1 compared to Cyclin B1 with Prism software and is the average of at least three experiments. Error shown is standard deviation. Inefficient binding was observed for the Cripto-1 loop mutant, Cripto-1 leader mutant, and Cyclin B1 so the $K_{d,app}$ was not reported.

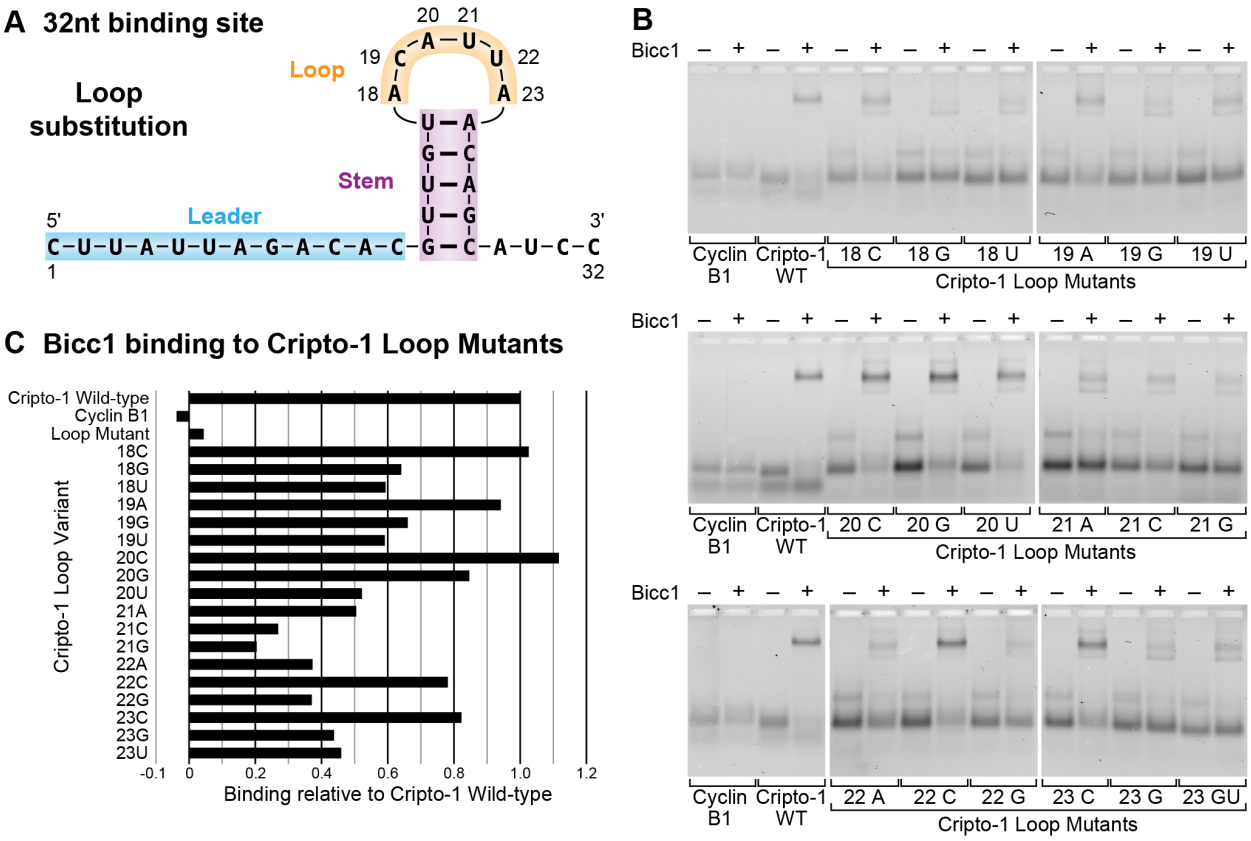

Figure 4. Analysis of Bicc1 binding to individual nucleotide mutations in the loop. ***A.*** Diagram of the Cripto-1 32-nucleotide binding site with the loop nucleotides numbered. ***B.*** EMSA experiments analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to Cyclin B1 and individual loop mutants. Nucleotides 18 and 19 were assayed at in the top panel, 20 and 21 in the middle panel, and 22 and 23 in the bottom panel. The assays were run with 350nM of Bicc1 protein to 10nM of fluorescein labeled RNA. The free RNA is shown at the bottom with the bound complex highlighted at the top. ***C.*** Fluorescent polarization experiments analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1 and individual loop mutants. The assays were run with 350nM of Bicc1 protein. The results are reported as the average of at least three experiments.

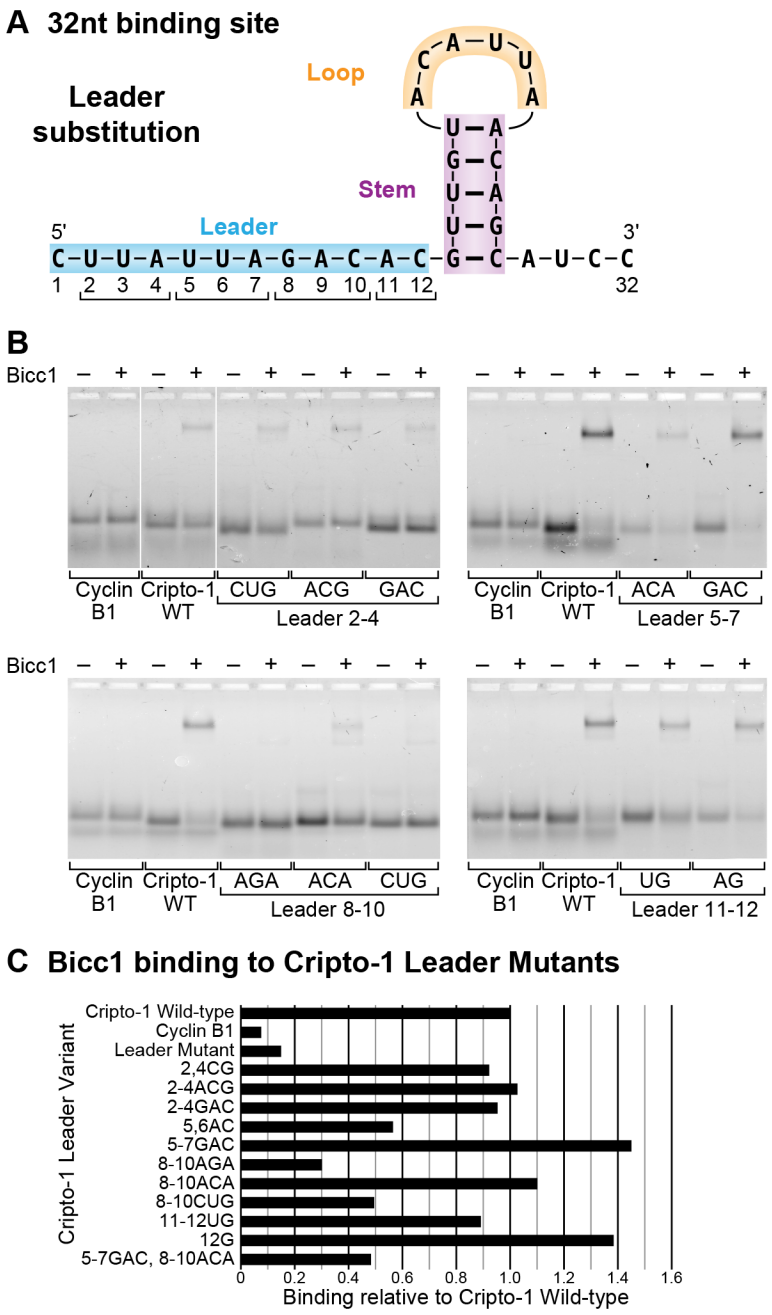

Figure 5. Analysis of Bicc1 binding to mutations in the leader.
A. Diagram of the Cripto-1 32-nucleotide binding site with the leader nucleotides numbered. ***B.*** EMSA experiments analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to Cyclin B1 and leader mutants. Nucleotides 2-4 were assayed at in the top left panel, 5-7 in the top right panel, 8-10 in the bottom left panel, and 11-12 in the bottom right panel. The assays were run with 350nM of Bicc1 protein to 10nM of fluorescein labeled RNA. The free RNA is shown at the bottom with the bound complex highlighted at the top. ***C.*** Fluorescent polarization experiments analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1 and leader mutants. The assays were run with

350nM of Bicc1 protein. The results are reported as the average of at least three experiments.

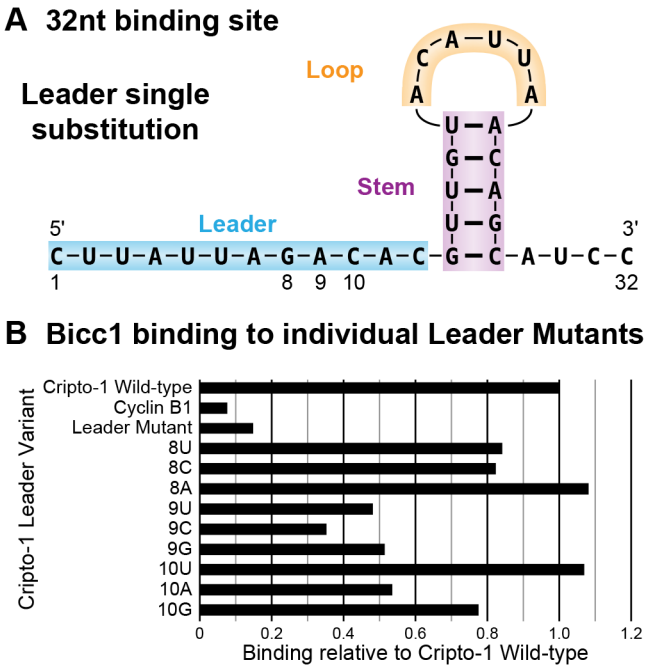

Figure 6. Analysis of Bicc1 binding to individual mutations in the leader nucleotides 8-10.
A. Diagram of the Cripto-1 32-nucleotide binding site with the leader nucleotides numbered. ***B.*** Fluorescent polarization experiments analyzing the affinity of Bicc1 for a 3'-fluorescein labeled Cripto-1 binding site compared to a 32-nucleotide site in Cyclin B1 and individual leader mutants. The assays were run with 350nM of Bicc1 protein. The results are reported as the average of at least three experiments.

REFERENCES

Arai, Y., Totoki, Y., Hosoda, F., Shirota, T., Hama, N., Nakamura, H., Ojima, H., Furuta, K., Shimada, K., Okusaka, T., et al. (2014). Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology* 59, 1427–34.

Audoux, J., Philippe, N., Chikhi, R., Salson, M., Gallopin, M., Gabriel, M., Le Coz, J., Drouineau, E., Commes, T., & Gautheret, D. (2017). DE-kupl: exhaustive capture of biological variation in RNA-seq data through k-mer decomposition. *Genome biology*, *18*(1), 243.

Bermingham, R., Carballedo, A., Lisiecka, D., Fagan, A., Morris, D., Fahey, C., Donohoe, G., Meaney, J., Gill, M. and Frodl, T. (2012). Effect of genetic variant in BICC1 on functional and structural brain changes in depression. *Neuropsychopharmacology* 37, 2855–62.

Dominguez, D., Freese, P., Alexis, M. S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N. J., Van, N. E. L., Pratt, G. A., et al. (2018). Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol Cell* 70, 854–867.e9.

Dowdle, M. E., Imboden, S. B., Park, S., Ryder, S. P. and Sheets, M. D. (2017). Horizontal Gel Electrophoresis for Enhanced Detection of Protein-RNA Complexes. *J Vis Exp*.

Dowdle, M. E., Park, S., Blaser, I. S., Fox, C. A., Houston, D. W. and Sheets, M. D. (2019). A single KH domain in Bicaudal-C links mRNA binding and translational repression functions to maternal development. *Development* 146 (10).

Díaz-Moreno, I., Hollingworth, D., Kelly, G., Martin, S., García-Mayoral, M., Briata, P., Gherzi, R. and Ramos, A. (2010). Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. *Nucleic Acids Res* 38, 5193–205.

Gamberi, C., Hipfner, D. R., Trudel, M. and Lubell, W. D. (2017). Bicaudal C mutation causes myc and TOR pathway up-regulation and polycystic kidney disease-like phenotypes in Drosophila. *PLoS Genet* 13, e1006694.

Gamberi, C. and Lasko, P. (2012). The Bic-C family of developmental translational regulators. *Comp Funct Genomics* 2012, 141386.

Grishin, N. V. (2001). KH domain: one motif, two folds. *Nucleic Acids Res* 29, 638–43.

Hentze, M. W., Castello, A., Schwarzl, T. and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* 19, 327–341.

Jensen, K. B., Musunuru, K., Lewis, H. A., Burley, S. K. and Darnell, R. B. (2000). The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proc Natl Acad Sci U S A* 97, 5740–5.

Kraus, M. R., Clauin, S., Pfister, Y., Di, M. M., Ulinski, T., Constam, D., Bellanné-Chantelot, C. and Grapin-Botton, A. (2012). Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. *Hum Mutat* 33, 86–90.

Leal-Esteban, L. C., Rothé, B., Fortier, S., Isenschmid, M. and Constam, D. B. (2018). Role of Bicaudal C1 in renal gluconeogenesis and its novel interaction with the CTLH complex. *PLoS Genet* 14, e1007487.

Li Y, Klena NT, Gabriel GC, Liu X, Kim AJ, Lemke K, Chen Y, Chatterjee B, Devine W, Damerla RR, Chang C, Yagi H, San Agustin JT, Thahir M, Anderton S, Lawhead C, Vescovi A, Pratt H, Morgan J, Haynes L, Smith CL, Eppig JT, Reinholdt L, Francis R, Leatherbury L, Ganapathiraju MK, Tobita K, Pazour GJ, Lo CW. (2015) Global genetic analysis in mice unveils central role for cilia in congenital heart disease. *Nature* 521, 520-4.

Li, F., Peiris, M. N. and Donoghue, D. J. (2020). Functions of FGFR2 corrupted by translocations in intrahepatic cholangiocarcinoma. *Cytokine Growth Factor Rev* 52, 56–67.

Loughlin, F. E., Lukavsky, P. J., Kazeeva, T., Reber, S., Hock, E. M., Colombo, M., Von, S. C., Pauli, P., Cléry, A., Mühlemann, O., et al. (2019). The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. *Mol Cell* 73, 490–504.e6.

Lunde, B. M., Moore, C. and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* 8, 479–90.

Maisonneuve, C., Guilleret, I., Vick, P., Weber, T., Andre, P., Beyer, T., Blum, M. and Constam, D. B. (2009). Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. *Development* 136, 3019–30.

Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M., Weeks, S. D. and Butt, T. R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics* 5, 75–86.

Minegishi, K., Rothé, B., Komatsu, K. R., Ono, H., Ikawa, Y., Nishimura, H., Miyashita, E., Takaoka, K., Bando, K., Kiyonari, H., et al. (2020). Fluid flow-induced left-right asymmetric decay of Dand5 mRNA in the mouse embryo requires Bicc1-Ccr4 RNA degradation complex. *bioRxiv*.

Mohler, J. and Wieschaus, E. F. (1986). Dominant maternal-effect mutations of Drosophila melanogaster causing the production of double-abdomen embryos. *Genetics* 112, 803–22.

Nicastro, G., Candel, A. M., Uhl, M., Oregioni, A., Hollingworth, D., Backofen, R., Martin, S. R. and Ramos, A. (2017). Mechanism of β-actin mRNA Recognition by ZBP1. *Cell Rep* 18, 1187–1199.

Nicastro, G., Taylor, I. A. and Ramos, A. (2015). KH-RNA interactions: back in the groove. *Curr Opin Struct Biol* 30, 63–70.

Ota, K. T., Andres, W., Lewis, D. A., Stockmeier, C. A. and Duman, R. S. (2015). BICC1 expression is elevated in depressed subjects and contributes to depressive behavior in rodents. *Neuropsychopharmacology* 40, 711–8.

Pagano, J. M., Clingman, C. C. and Ryder, S. P. (2011). Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* 17, 14–20.

Park, S., Blaser, S., Marchal, M. A., Houston, D. W. and Sheets, M. D. (2016). A gradient of maternal Bicaudal-C controls vertebrate embryogenesis via translational repression of mRNAs encoding cell fate regulators. *Development* 143, 864–71.

Paziewska, A., Wyrwicz, L. S., Bujnicki, J. M., Bomsztyk, K. and Ostrowski, J. (2004). Cooperative binding of the hnRNP K three KH domains to mRNA targets. *FEBS Lett* 577, 134–40.

Piazzon, N., Maisonneuve, C., Guilleret, I., Rotman, S. and Constam, D. B. (2012). Bicc1 links the regulation of cAMP signaling in polycystic kidneys to microRNA-induced gene silencing. *J Mol Cell Biol* 4, 398–408.

Rothé, B., Gagnieux, C., Leal-Esteban, L. C. and Constam, D. B. (2019). Role of the RNA-binding protein Bicaudal-C1 and interacting factors in cystic kidney diseases. *Cell Signal* 1094-99.

Saffman, E. E., Styhler, S., Rother, K., Li, W., Richard, S. and Lasko, P. (1998). Premature translation of oskar in oocytes lacking the RNA-binding protein bicaudal-C. *Mol Cell Biol* 18, 4855–62.

Tran, U., Zakin, L., Schweickert, A., Agrawal, R., Döger, R., Blum, M., De Robertis, E. M. and Wessely, O. (2010). The RNA-binding protein bicaudal C regulates polycystin 2 in the kidney by antagonizing miR-17 activity. *Development* 137, 1107–16.

Valverde, R., Edwards, L. and Regan, L. (2008). Structure and function of KH domains. *FEBS J* 275, 2712–26.

Wang, Z., Zhou, D., Li, S., Zhang, Y. and Wang, C. (2019). Underlying mechanisms of recombinant adeno-associated virus-mediated bicaudal C homolog 1 overexpression in the medial prefrontal cortex of mice with induced depressive-like behaviors. *Brain Res Bull* 150, 35–41.

Wessely, O. and De Robertis, E. M. (2000). The Xenopus homologue of Bicaudal-C is a localized maternal mRNA that can induce endoderm formation. *Development* 127, 2053–62.

Ying, X., Tu, J., Wang, W., Li, X., Xu, C. and Ji, J. (2019). FGFR2-BICC1: A Subtype Of FGFR2 Oncogenic Fusion Variant In Cholangiocarcinoma And The Response To Sorafenib. *Onco Targets Ther* 12, 9303–9307.

Zhang, Y., Cooke, A., Park, S., Dewey, C. N., Wickens, M. and Sheets, M. D. (2013). Bicaudal-C spatially controls translation of vertebrate maternal mRNAs. *RNA* 19, 1575–82.

Zhang, Y., Park, S., Blaser, S. and Sheets, M. D. (2014). Determinants of RNA binding and translational repression by the Bicaudal-C regulatory protein. *J Biol Chem* 289, 7497–504.

Chapter 5:

Conclusions, Perspectives, and Future Directions

Summary

My work has focused on analyzing the mechanisms of Bicc1 binding to its target mRNAs from both the protein and mRNA perspective. Chapter 2 describes the horizontal EMSA technique I used to analyze the protein and RNA variants throughout my thesis (Dowdle et al., 2017). In Chapter 3 results from both *in vitro* and *in vivo* experiments, demonstrated that the KH2 domain in the N-terminal of Bicc1 is a critical feature for RNA binding. This domain is evolutionarily conserved in both invertebrate and vertebrate species (Dowdle et al., 2019) (Chapter 3). However, while KH2 was critical for binding, it was not sufficient for specificity. Results presented in Chapter 3 also show that disrupting Bicc1's ability to bind to RNA disrupts its ability to repress RNA translation, leading to defects in development that cause excess anterior structures to form in the *Xenopus* embryo. My results presented in Chapter 4 suggests that Bicc1 binds to the Cripto-1 RNA target through the use of a non-contiguous bipartite binding site. The best characterized Bicc1 binding site is a 32-nucleotide region from the 3'UTR of the *Xenopus* Cripto-1 mRNA (Zhang et al., 2013; Zhang et al., 2014). I performed an in-depth analysis of the nucleotides required for Bicc1 binding and my results indicate that sequences in both the leader and the loop regions of the 32-nucleotide RNA were necessary for efficient Bicc1 binding. However, my results also indicate that Bicc1 binding is complex and no single nucleotide change abolished binding as effectively as changing multiple consecutive residues. Despite this challenge, I was still able to determine regions that are more important for binding; the 8-10 and 21-23 nucleotide positions in the leader and loop respectively. In addition, I successfully generated for the first time a recombinant human Bicc1 protein for RNA binding studies.

With this protein I observed that the human Bicc1 exhibits very similar binding properties for the *Xenopus* 32-nucleotide Cripto-1 as the *Xenopus* protein. These similarities suggest that my in-depth analysis of *Xenopus* Bicc1 is directly relevant to ultimately defining the RNA binding properties of the human Bicc1 protein and its potential roles in specific disease states.

Future Directions

Determine how Bicc1 specifically recognizes RNA targets

The results presented in Chapter 2 demonstrate that the KH2 domain was a structural feature of the Bicc1 protein critical for its RNA binding. However, while the GXXG motifs in KH domains are necessary for RNA binding, they do not confer specificity. NMR studies conducted with KH domain containing proteins including Nova1, KSRP, and ZBP1, demonstrated that the GXXG motif makes contact with the phosphate backbone of the RNA target while other surface exposed amino acid residues make contact with particular RNA bases to provide specificity (Hollingworth et al., 2012; Nicastro et al., 2017). Additionally, structural studies of KH domains bound to their RNA target indicate that KH and other domains often act in a combinatorial fashion to mediate their binding to specific RNA targets (Nicastro et al., 2017; Biswas et al., 2019; Yang et al., 2017). As Bicc1 contains three canonical KH domains, this combinatorial binding could explain why KH2 is critical but not sufficient for efficient or specific RNA binding.

1.1 Mutational analysis of Bicaudal-C KH domains

A mutational approach will be used to determine which residues in the three canonical KH domains are required for specific RNA binding. Surface exposed amino

acids that are likely important for specific mRNA binding can be predicted by comparing the Bicc1 KH2 domain to structural studies conducted on other KH2 domains through sequence comparison (Biswas et al., 2019; Yang et al., 2017; Du et al., 2007; Nicastro et al., 2017; Chao et al., 2010; Lewis et al., 2000) (**Fig. 1**). Additionally, because KH2 is highly conserved (Dowdle et al., 2019) (Chapter 3) the amino acids that are conserved from invertebrates to vertebrates can be selectively mutated to alanine (**Fig. 2**). Finally, an alanine scanning approach can be taken to mutate every amino acid in KH2 to alanine. The effects the mutations have on RNA binding can be observed through the *in vitro* assays used extensively throughout in my thesis. The top candidates for potential binding defects can then be assayed *in vivo* through the RNA-IPs and luciferase repression assays outlined in Chapter 3 of my thesis. After analyzing KH2, the same approach can be taken with KH1 and KH3 to determine the impact these domains may have on specific and efficient RNA binding.

1.2 Structural studies of Bicc1 binding to RNA

Studies have revealed that RNA binding proteins often bind their substrates in specific conformations. Despite the information that biochemical assays can reveal to us, these specific conformations can only be confirmed through structural studies. As Bicc1 is soluble, it is theoretically a good candidate for structural studies such as X-Ray Crystallography; however, crystallography studies to determine the structure have not been successful (Collaboration with Traci Hall, NIHEHS, unpublished). Also truncating the N-terminal to isolate individual domains for NMR has only yielded many insoluble proteins (Dowdle, M.E. unpublished). Despite these challenges, I have made some progress in purifying shortened KH1-5 and truncated KH2-5 proteins that could

potentially be used for NMR. Additionally, with the advances of Cryo-EM on the University of Wisconsin campus, it may be possible to purify larger Bicc1 constructs for structural analysis as well. These studies could reveal both the specific points of contact, confirming any motifs discovered through the *in vitro* RNA binding studies in Chapter 4, and show the role that KH1 and KH3 play in binding to RNA targets. Additionally, it would be beneficial to the field in understanding the role that the KH domains play not just in recognizing RNA targets but also in how they may play a role in stabilizing the structure of the N-terminal domain.

1.3 Determine minimal binding sites in other Bicc1 targets

In order to fully understand how Bicc1 recognizes its targets on both a sequence and structural basis there needs to be further characterization of other Bicc1 targets. The Sheets lab has previously identified 62 mRNA targets of Bicc1 in *Xenopus* (Zhang et al., 2013). However, only the Cripto-1 target has been extensively characterized. The approaches I have used for my thesis work can be applied to identify minimal binding sequences in Bicc1's other targets. The ultimate goal is to define a holistic consensus binding site that can then be used to help determine potential targets of Bicc1 in other species beyond *Xenopus*. Altogether, these potential studies would help to understand how Bicc1 binds to its many specific targets.

Identify human mRNA targets of Bicc1

Though Bicc1's biological relevance in humans is clear, there are no confirmed human Bicc1 targets. In order to understand the role that Bicc1 plays in human development, organogenesis, and disease it is necessary to determine the mRNA targets that Bicc1 regulates.

2.1 RNA-immunoprecipitation and sequencing (RIP-Seq) experiments in human cell lines

Defects in Bicc1 have been linked to heart (Katoh, 2016; Zhang et al., 2005), kidney (Rothé et al., 2020; Leal-Esteban et al., 2018; Kraus et al., 2012), and nervous system (Ota et al., 2015; Lewis et al., 2010) defects in humans. All three of these cell types can be differentiated from induced pluripotent stem cells (IPSCs) (Zhang et al., 2009; Takasato et al., 2015; Karumbayaram et al., 2009). In the Sheets lab we have been able to successfully differentiate IPSCs to both neurons and cardiomyocytes and have shown that Bicc1 is expressed in both cell types despite not being highly expressed in the IPSCs. With these differentiated stem cell lines, we are ideally positioned to identify human targets of Bicc1. This can be done through RNA-immunoprecipitation followed by deep sequencing (RIP-Seq) of the pulldown targets. The pulldowns can be analyzed through bioinformatics to categorize the identified RNAs into functional categories. In addition to RIP-seq, there are multiple methods of crosslinking immunoprecipitation (CLIP) (Van et al., 2016; Huppertz et al., 2014) can be used to comprehensively identify potential human targets of Bicc1.

2.2 Validate Human targets *in vitro*

The most promising and highly enriched targets identified *through in vivo* RIP-Seq and CLIP assays can be verified as direct targets through the *in vitro* experiments using the soluble, recombinant version of the N-terminal human Bicc1 I generated.. This N-terminal protein can also be used for footprinting experiments to identify a minimal binding site in these targets as well. By identifying and validating human mRNA targets,

we can potentially understand the role that mis-regulated Bicc1 targets play in human disease states.

Determine the mechanism of repression of Bicc1

While my work has focused on the mechanism of RNA binding, and work done by the Bowie and Constam labs have focused on the role of the SAM domain (Rothé et al., 2015; Rothé et al., 2018) there is still much to learn about how Bicc1 represses its mRNA targets. Through work done by the Sheets lab, it was determined that Bicc1 repressed the translation of RNA targets through its C-terminal region independent of the SAM domain (Zhang et al., 2014). This domain has been characterized as being intrinsically disordered and therefore not much can be predicted from structural and sequence comparisons. Repression was found to be cap and initiation factor dependent but independent of the Poly(A) tail (Zhang et al., 2014). Studies in *Drosophila* have shown that the *Drosophila* Bicc1 recruits members of the CCR4-NOT complex to deadenylate and thus repress its targets (Chicoine et al., 2007). Deadenylation has also been suggested to be the potential mode of repression in vertebrates (Minegishi et al., 2020) though biochemical studies have not verified this mechanism.

3.1 Determine which proteins interact with Bicc1

One path to analyzing the repression domain and its mode of action is to define specific protein partners that may facilitate its functions. There are many traditional methods to determine which proteins interact with Bicc1 such as yeast two-hybrid assays and affinity pulldowns with the C-terminal repression domain. However, previous experiments conducted with these techniques have not yielded useful results. Recently there have been new methods of spatially tagging interacting protein partners *in vivo*

with biotin using modified BirA enzymes from *E.coli* (Fairhead and Howarth, 2015). One of the most promising methods using these enzymes is Turbo-ID (Branon et al., 2018). While biotinylation through BirA has been shown to be slow and inefficient, the Turbo-ID enzyme has been engineered to work quickly and much more efficiently (Branon et al., 2018). We can create a recombinant Bicc1 C-terminal fused to the Turbo-ID enzyme to biotinylate proteins in close proximity with Bicc1. The biotinylated proteins can be purified with streptavidin and identified via mass spectrometry. Once potential protein partners have been identified, we can potentially then return to more traditional techniques such as yeast two-hybrids and co-immunoprecipitations to verify that they directly bind to Bicc1. By determining which proteins interact with Bicc1 we will be able to the mechanisms by which Bicc1 represses its mRNA targets. For example, Bicc1 could interact with proteins that compete with initiation factors to bind to the 5' cap of the mRNA or it could compete with poly(A) binding protein for binding to the poly(A) tail to prevent translation initiation (Hentze et al., 2018, Tahmasebi et al., 2019). If Bicc1 does bind to members of the CCR4-NOT complex, this would also serve as biochemical confirmation of the mechanism of repression that is seen in *Drosophila*.

Significance and Concluding Remarks

Results from my own studies (Dowdle 2019) (Chapter 3), other studies from the Sheets lab (Park et al., 2016), and studies conducted in both invertebrates (Saffman et al., 1998; Chicoine et al., 2007) and vertebrates (Maisonneuve et al., 2009; Leal-Esteban et al., 2018; Tran et al., 2010) demonstrate that Bicc1 is a biologically relevant RNA binding protein. The results of my thesis are significant as this is the first time the

mechanism of the RNA binding of Bicc1 has been determined. As this protein is evolutionarily conserved and my work with the human Bicc1 has shown specificity for the *Xenopus* target this will be useful in identifying targets of Bicc1 in humans. Furthermore, as I have purified a recombinant form of the human protein, this protein can then be used in biochemical experiments used to determine how Bicc1 functions in diseases caused by mutant Bicc1. My thesis has focused on determining the mechanism on which Bicc1 binds to its RNA targets. The results presented in my thesis have broad implications for understanding not just Bicc1 but also understanding how KH domain containing proteins in general may recognize and bind their RNA targets. And, while my studies have provided new insights into this mechanism, there are still many questions which remain to be answered regarding to how Bicc1 binds to its targets and the potential roles of the KH1 and KH3 domains. Additionally, while it seems likely that the *Xenopus* and human Bicc1 proteins recognize and bind targets in a similar manner (Chapter 4), there are still few human targets known. Finally, it is critical to determine how Bicc1 represses its mRNA targets. The results of my current work combined with results from proposed future experiments, will together provide a new perspective into how Bicc1 impacts development and human disease.

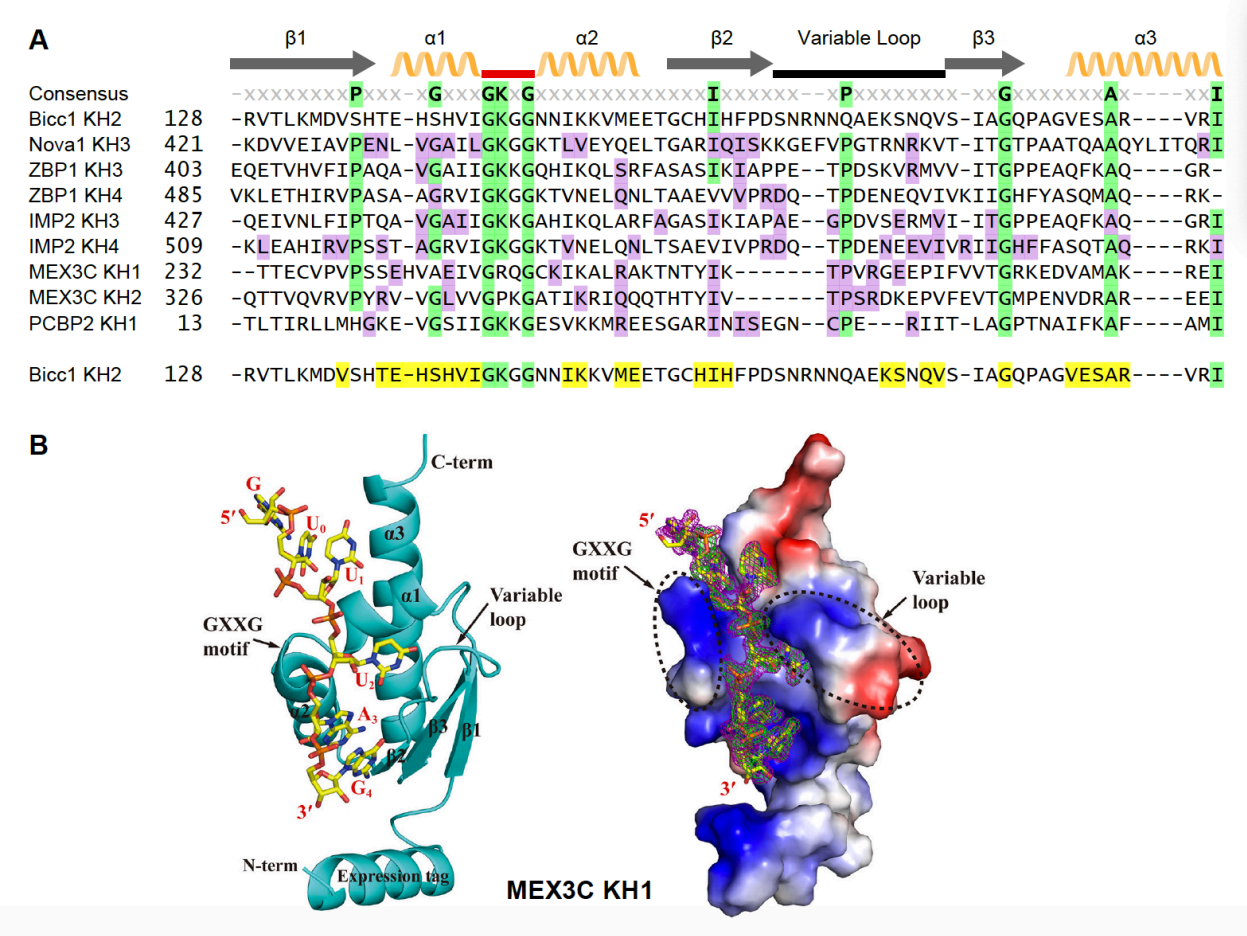

Figure 1. RNA-binding KH domains show conserved features and interactions through the KH domain.
A. Amino acid sequences of RNA-binding KH domains (Biswas et al., 2019; Yang et al., 2017; Du et al., 2007; Nicastro et al., 2017; Chao et al., 2010; Lewis et al., 2000) were aligned using MUSCLE alignment software with a consensus sequence derived from residues that were 75% conserved across sequences. The conserved KH fold (Hollingworth et al., 2012) is shown above the consensus sequence while conserved residues are highlighted in green. Residues that are known to create specific contacts with RNA bases are highlighted in pink. Designed mutations in the Bicc1 KH2 domain are highlighted in yellow. **B.** The structure of MEX3C KH1 bound to RNA (Yang et al., 2017, PDB 5WWW) with structural features labeled and the variable loop and GXXG highlighted.

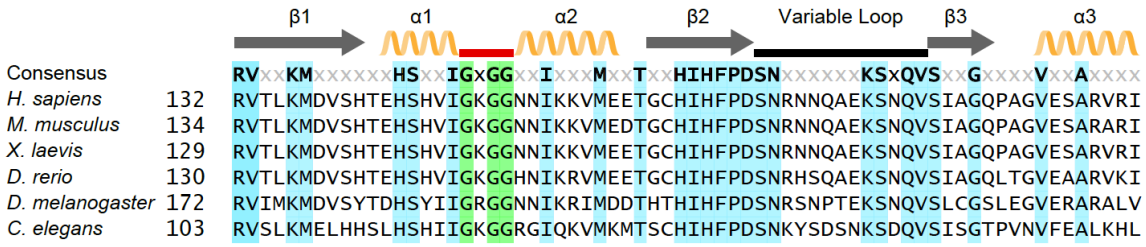

Figure 2. The KH2 domain is an evolutionarily conserved feature of Bicc1 proteins Amino acid sequences of the Bicc1 KH2 domain were analyzed with MUSCLE alignment software with amino acids identical to the human protein highlighted in cyan with the GxxG motif highlighted in green. A consensus sequence was derived from amino acids that were conserved in all sequences. The conserved KH fold is shown above the consensus sequence with the GXXG sequence shown in red and the variable loop region shown in black.

References

Biswas, J., Patel, V. L., Bhaskar, V., Chao, J. A., Singer, R. H. and Eliscovich, C. (2019). The structural basis for RNA selectivity by the IMP family of RNA-binding proteins. *Nat Commun* 10, 4440.

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N. and Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* 36, 880–887.

Chao, J. A., Patskovsky, Y., Patel, V., Levy, M., Almo, S. C. and Singer, R. H. (2010). ZBP1 recognition of beta-actin zipcode induces RNA looping. *Genes Dev* 24, 148–58.

Chicoine, J., Benoit, P., Gamberi, C., Paliouras, M., Simonelig, M. and Lasko, P. (2007). Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev Cell* 13, 691–704.

Dowdle, M. E., Imboden, S. B., Park, S., Ryder, S. P. and Sheets, M. D. (2017). Horizontal Gel Electrophoresis for Enhanced Detection of Protein-RNA Complexes. *J Vis Exp*.

Dowdle, M. E., Park, S., Blaser, I. S., Fox, C. A., Houston, D. W. and Sheets, M. D. (2019). A single KH domain in Bicaudal-C links mRNA binding and translational repression functions to maternal development. *Development* 146 (10).

Du, Z., Lee, J. K., Fenn, S., Tjhen, R., Stroud, R. M. and James, T. L. (2007). X-ray crystallographic and NMR studies of protein-protein and protein-nucleic acid interactions involving the KH domains from human poly(C)-binding protein-2. *RNA* 13, 1043–51.

Fairhead, M. and Howarth, M. (2015). Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol* 1266, 171–84.

Hentze, M. W., Castello, A., Schwarzl, T. and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* 19, 327–341.

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, 6873–86.

Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L. E., Sibley, C. R., Sugimoto, Y., Tajnik, M., König, J. and Ule, J. (2014). iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* 65, 274–87.

Karumbayaram, S., Novitch, B. G., Patterson, M., Umbach, J. A., Richter, L., Lindgren, A., Conway, A. E., Clark, A. T., Goldman, S. A., Plath, K., et al. (2009). Directed

differentiation of human-induced pluripotent stem cells generates active motor neurons. *Stem Cells* 27, 806–11.

Katoh, M. (2016). FGFR inhibitors: Effects on cancer cells, tumor microenvironment and whole-body homeostasis (Review). *Int J Mol Med* 38, 3–15.

Kraus, M. R., Clauin, S., Pfister, Y., Di, M. M., Ulinski, T., Constam, D., Bellanné-Chantelot, C. and Grapin-Botton, A. (2012). Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. *Hum Mutat* 33, 86–90.

Leal-Esteban, L. C., Rothé, B., Fortier, S., Isenschmid, M. and Constam, D. B. (2018). Role of Bicaudal C1 in renal gluconeogenesis and its novel interaction with the CTLH complex. *PLoS Genet* 14, e1007487.

Lewis, C. M., Ng, M. Y., Butler, A. W., Cohen-Woods, S., Uher, R., Pirlo, K., Weale, M. E., Schosser, A., Paredes, U. M., Rivera, M., et al. (2010). Genome-wide association study of major recurrent depression in the U.K. population. *Am J Psychiatry* 167, 949–57.

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, 323–32.

Minegishi, K., Rothé, B., Komatsu, K. R., Ono, H., Ikawa, Y., Nishimura, H., Miyashita, E., Takaoka, K., Bando, K., Kiyonari, H., et al. (2020). Fluid flow-induced left-right asymmetric decay of Dand5 mRNA in the mouse embryo requires Bicc1-Ccr4 RNA degradation complex. *bioRxiv*.

Maisonneuve, C., Guilleret, I., Vick, P., Weber, T., Andre, P., Beyer, T., Blum, M. and Constam, D. B. (2009). Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. *Development* 136, 3019–30.

Nicastro, G., Candel, A. M., Uhl, M., Oregioni, A., Hollingworth, D., Backofen, R., Martin, S. R. and Ramos, A. (2017). Mechanism of β-actin mRNA Recognition by ZBP1. *Cell Rep* 18, 1187–1199.

Ota, K. T., Andres, W., Lewis, D. A., Stockmeier, C. A. and Duman, R. S. (2015). BICC1 expression is elevated in depressed subjects and contributes to depressive behavior in rodents. *Neuropsychopharmacology* 40, 711–8.

Park, S., Blaser, S., Marchal, M. A., Houston, D. W. and Sheets, M. D. (2016). A gradient of maternal Bicaudal-C controls vertebrate embryogenesis via translational repression of mRNAs encoding cell fate regulators. *Development* 143, 864–71.

Rothé, B., Gagnieux, C., Leal-Esteban, L. C. and Constam, D. B. (2020). Role of the RNA-binding protein Bicaudal-C1 and interacting factors in cystic kidney diseases. *Cell Signal* 68, 109499.

Rothé, B., Leal-Esteban, L., Bernet, F., Urfer, S., Doerr, N., Weimbs, T., Iwaszkiewicz, J. and Constam, D. B. (2015). Bicc1 Polymerization Regulates the Localization and Silencing of Bound mRNA. *Mol Cell Biol* 35, 3339–53.

Rothé, B., Leettola, C. N., Leal-Esteban, L., Cascio, D., Fortier, S., Isenschmid, M., Bowie, J. U. and Constam, D. B. (2018). Crystal Structure of Bicc1 SAM Polymer and Mapping of Interactions between the Ciliopathy-Associated Proteins Bicc1, ANKS3, and ANKS6. *Structure* 26, 209–224.e6.

Saffman, E. E., Styhler, S., Rother, K., Li, W., Richard, S. and Lasko, P. (1998). Premature translation of oskar in oocytes lacking the RNA-binding protein bicaudal-C. *Mol Cell Biol* 18, 4855–62.

Tahmasebi, S., Sonenberg, N., Hershey, J. W. B. and Mathews, M. B. (2019). Protein Synthesis and Translational Control: A Historical Perspective. *Cold Spring Harb Perspect Biol* 11,

Takasato, M., Er, P. X., Chiu, H. S., Maier, B., Baillie, G. J., Ferguson, C., Parton, R. G., Wolvetang, E. J., Roost, M. S., Chuva, de S. L. S. M., et al. (2015). Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* 526, 564–8.

Tran, U., Zakin, L., Schweickert, A., Agrawal, R., Döger, R., Blum, M., De, R. E. M. and Wessely, O. (2010). The RNA-binding protein bicaudal C regulates polycystin 2 in the kidney by antagonizing miR-17 activity. *Development* 137, 1107–16.

Van, N. E. L., Pratt, G. A., Shishkin, A. A., Gelboin-Burkhart, C., Fang, M. Y., Sundararaman, B., Blue, S. M., Nguyen, T. B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* 13, 508–14.

Yang, L., Wang, C., Li, F., Zhang, J., Nayab, A., Wu, J., Shi, Y. and Gong, Q. (2017). The human RNA-binding protein and E3 ligase MEX-3C binds the MEX-3-recognition element (MRE) motif with high affinity. *J Biol Chem* 292, 16221–16234.

Zhang, J., Wilson, G. F., Soerens, A. G., Koonce, C. H., Yu, J., Palecek, S. P., Thomson, J. A. and Kamp, T. J. (2009). Functional cardiomyocytes derived from human induced pluripotent stem cells. *Circ Res* 104, e30–41.

Zhang, S. X., Garcia-Gras, E., Wycuff, D. R., Marriot, S. J., Kadeer, N., Yu, W., Olson, E. N., Garry, D. J., Parmacek, M. S. and Schwartz, R. J. (2005). Identification of direct

serum-response factor gene targets during Me2SO-induced P19 cardiac cell differentiation. *J Biol Chem* 280, 19115–26.

Zhang, Y., Cooke, A., Park, S., Dewey, C. N., Wickens, M. and Sheets, M. D. (2013). Bicaudal-C spatially controls translation of vertebrate maternal mRNAs. *RNA* 19, 1575–82.

Zhang, Y., Park, S., Blaser, S. and Sheets, M. D. (2014). Determinants of RNA binding and translational repression by the Bicaudal-C regulatory protein. *J Biol Chem* 289, 7497–504.