ABSTRACT

Defining the molecular determinants required for RNA binding functions of the Bicaudal-C (Bicc1) translational repressor protein

Bicaudal-C (Bicc1) is an RNA binding protein that functions as translational repressor in all animal species. Its repression activities are required for a multitude of biological processes such as guiding the earliest steps of embryonic development and organ homeostasis. For example, defects in Bicc1 expression or its activities can give rise to polycystic kidney disease in humans, a disorder that is characterized by an abnormal collection of cysts in the kidney. In addition, Bicc1 functions in vertebrate embryos to control the translation of specific maternal mRNAs that encode key developmental regulatory proteins. During the early stages of vertebrate development defects in Bicc1 trigger cell fate changes that lead to embryonic malformations similar to some human birth defects. A general feature is that in all of its biological contexts Bicc1 binds to specific mRNAs and it is through this binding that it selects specific mRNAs for translational repression. My project is focused on defining the molecular determinants of the Bicc1 protein required for that specific mRNA binding. Recent results from the Sheets lab identified a 70 amino acid region of the Bicc1 protein, the KH2 domain that is required for mRNA binding. However, the particular molecular features within the KH2 domain important for RNA binding activity are unknown. To address this issue, I am generating mutations of the KH2 domain that change evolutionarily conserved amino acid residues in all Bicc1 proteins. These mutant proteins will be analyzed with gel shift RNA binding assays in vitro as well as in vivo RNA binding experiments performed in Xenopus embryos. My goal is to define the amino acid residues in the KH2 domain that are responsible for providing specificity for Bicc1 binding to specific mRNAs, guide its repression activities and mediate its biological functions.

Author Name/Major

Mentor Name/Department

Author Signature

Mentor Signature

Date
COVER SHEET

TITLE: Defining the determinants for RNA binding by Bicaudal-C (BicC1) mutagenesis

AUTHOR'S NAME: Sonam Dolma

MAJOR: Biochemistry

DEPARTMENT: Biochemistry

MENTOR: Michael Sheets

DEPARTMENT: Biomolecular Chemistry

YEAR: 2020

The author hereby grants to University of Wisconsin-Madison the permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.
Acknowledgements

I would first and foremost like to thank my parents for their unconditional love and their countless sacrifices for me and my sister’s future. If today I am able to pursue higher education whole-heartedly and have the opportunity to actualize my dreams, it is due solely to their unceasing kindness and support.

I sincerely would like to thank Dr. Michael Sheets for giving me the opportunity of being a part of the Sheets Lab and his continued support of my endeavors. It has been an incredibly enriching experience in lab for the past few years.

I would also like to thank my graduate student mentor, Megan Dowdle, for guiding me and helping me throughout my research years.
Defining the molecular determinants required for RNA binding functions of the Bicaudal-C (Bicc1) translational repressor protein

By: Sonam Dolma
Department of Biochemistry
Mentor: Dr. Michael Sheets, Department of Biomolecular Chemistry

ABSTRACT

Bicaudal-C (Bicc1) is an RNA binding protein that functions as translational repressor in all animal species. Its repression activities are required for a multitude of biological processes such as guiding the earliest steps of embryonic development and organ homeostasis. For example, defects in Bicc1 expression or its activities can give rise to polycystic kidney disease in humans, a disorder that is characterized by an abnormal collection of cysts in the kidney. In addition, Bicc1 functions in vertebrate embryos to control the translation of specific maternal mRNAs that encode key developmental regulatory proteins. During the early stages of vertebrate development defects in Bicc1 trigger cell fate changes that lead to embryonic malformations similar to some human birth defects. A general feature is that in all of its biological contexts Bicc1 binds to specific mRNAs and it is through this binding that it selects specific mRNAs for translational repression. My project is focused on defining the molecular determinants of the Bicc1 protein required for that specific mRNA binding. Recent results from the Sheets lab identified a 70 amino acid region of the Bicc1 protein, the KH2 domain that is required for mRNA binding. However, the particular molecular features within the KH2 domain important for RNA binding activity are unknown. To address this issue, I am generating mutations of the KH2 domain that change evolutionarily conserved amino acid residues in all Bicc1 proteins. These mutant proteins will be analyzed with gel shift RNA binding assays in vitro as well as in vivo RNA binding experiments performed in Xenopus embryos. My goal is to define the amino acid residues in the KH2 domain that are responsible for providing specificity for Bicc1 binding to specific mRNAs, guide its repression activities and mediate its biological functions.
Introduction

Background and Significance

The Bicaudal-C (Bicc1) protein is a protein that binds specific target mRNAs and represses their translation—this is essential for translational modulation of maternal mRNAs in developing embryos as it dictates where, when, and how much protein is expressed from each target mRNA\(^1\). In developing embryos, translational regulation of maternal mRNAs is an important form of gene regulation and dictates the proteome until the developing zygote begins to synthesize its own DNAs and RNAs\(^2,3\). Due to this, Bicc1 has a great influence on many of the proteins that perform important functions in cells such as the kidneys, pancreas and heart\(^4\); therefore, it is imperative in the functionality and development of tissues and, consequently, proper organ formation and function.

Post-transcriptional modulation by Bicaudal-C

The governance of gene expression is an essential part of all biological processes, as it affects its proteome and its functions. The dysregulation of genes can lead to a plethora of abnormalities due to either reduced or over expression of the given gene. Gene regulations occur at different levels, beginning from gene transcription to post-transcriptional gene translation. In early animal development, the regulation of post-transcriptional events is especially significant when regulatory events occur after RNA synthesis\(^11\). This focus of this thesis is on the translational repressor Bicaudal-C RNA-binding protein and its post-transcriptional regulation of mRNA translation.
RNA binding proteins and mRNA regulation

There are roughly 1,500 human proteins known to bind to RNAs. Bicaudal-C is an RNA binding protein. RNA binding proteins play an important role in post-transcriptional regulation. Visualization of the basic structure of eukaryotic mRNAs is critical in understanding the interface between mRNAs and RNA-binding proteins (Figure 1). A mature eukaryotic mRNA consists of a 5’ cap, then a 5’ untranslated region (UTR), an open reading frame (ORF), 3’ UTR and a polyadenylated tail. The ORF is the coding sequence which encodes the protein. The untranslated regions (5’ and 3’ UTR) hold regulatory sequences which are responsible for regulation of mRNA translation and localization. These regions of regulatory sequences provide the basis for interactions with specific RNA binding proteins which dictate an mRNA’s behavior.

Figure 1. Structure of a mature eukaryotic mRNA.

The influence of mRNA translation on the proteome of the cell requires post-transcriptional events to be highly regulated, as it can greatly alter the cell’s identity and functionality. mRNA translation is most tightly regulated during the initiation, elongation, and termination phases of translation.

RNA binding proteins play a crucial part in the translational regulation of specific mRNAs. As depicted in Figure 2, RNA-binding proteins can as activators (Green) by recruiting translation factors such as the 4E initiation complex, the 40S and 60S ribosomal subunits, or the
polyadenylated binding protein (PABP). They can also act as translation inhibitors (Red) by repressing the aforementioned factors. Bicaudal-C protein is an example of a translational repressor.

**Figure 2.** RNA binding proteins can either activate (Green) or inhibit (Red) translation.

*Translational repression by RNA binding proteins*

As stated above, the initiation phase of mRNA translation is tightly regulated in eukaryotes as part of the regulation process. Translational repressors inhibit translation by impeding the process of initiation, which inhibits mRNA protein synthesis. Although most mRNA translation is unaltered, there are specific sets of mRNAs that are repressed translationally. For these specific mRNAs, there are specific RNA binding proteins that denote specificity for the inhibition. The *Drosophila* protein Bruno is a model example of a translational repressor protein. Bruno identifies specific binding sites on the oskar mRNA in *Drosophila* embryos to inhibit translation. Bruno does this by interfacing with the Cup protein, which then binds directly to eIF4E, preventing the initiation phase, and leading to translational repression.
Bicaudal-C: translational repressor protein

Bicaudal-C (Bicc1) is a conserved RNA binding protein that functions as a translational repressor in all animal species. Bicaudal-C was first discovered in *Drosophila* embryos during mutation analyses that gave rise to defects such as double-abdomen or “bicaudal” phenotype in embryonic anterior-posterior patterning\(^\text{17}\) (Figure 3).

![Diagram](image)

**Figure 3.** Translational repression mediated by Bicc1 has a critical impact on proper development of the Xenopus embryo.

Subsequent experiments and analyses of protein sequencing disclosed that Bicaudal-C was a mRNA specific RNA-binding protein with several putative binding domains\(^\text{18}\). Further experiments on the Bicaudal-C protein revealed its function as a translational repressor, when *Drosophila* embryos lacking the Bicaudal-C protein overexpressed the translation of the *oskar*
mRNA\(^{19}\) (Figure 4). Homologs of Bicaudal-C have since then been discovered throughout the animal kingdom.

![mRNA recognition vs. translational repression diagram](image)

**Figure 4.** *Bicaudal-C: translational repressor protein*

The importance of Bicc1 on proper adult organ function is illustrated in kidney cells where Bicc1 binds to specific mRNA targets to modulate their translation\(^5\). If Bicc1 is abnormal, this mode of regulation is jeopardized and thus the unrepressed mRNAs are erroneously translated, resulting in an accumulation of the proteins encoded by the target mRNAs in kidney cells\(^6\). The defect in Bicc1 induces irregular build-up of proteins and causes the cells to form abnormal kidney tubules, which leads to polycystic kidney disease\(^7\). One of the main issues lies in how Bicc1 recognizes and binds to specific mRNAs to regulate their translation.

The Sheets Lab studies Bicc1 in the context of vertebrate *Xenopus laevis*. In embryos of the frog *Xenopus laevis*, the maternally deposited Bicc1 is confined to the vegetal hemisphere of the oocyte and embryo\(^{20}\). The *Xenopus* protein is advantageous for studying Bicc1’s functions, as the protein is highly conserved between humans and *Xenopus* (Figure 5 and 6)\(^8\). The Xenopus embryos are a good model due to their similarities and their ease of manipulation.
Using the Xenopus embryos, the Sheets Lab has made substantial identifications and characterizations of target mRNAs of the Bicc1 protein. Of the 62 mRNA targets discovered, the Cripto-1 (CR1) mRNA is the most characterized target of Bicaudal-C. Our lab has previously shown that the N-terminal region of Bicc1 is responsible for the RNA-binding activity of the protein. In all Bicc1 proteins, the N-terminal region contains three canonical K Homology (KH) domains.
domains. Canonical KH domains have a GXXG motif that is integral for RNA-binding activity (REF, REF)\textsuperscript{10}. With this knowledge, our lab was able to individually examine each KH domain by mutating their GXXG motifs. Through this method, we established that the KH2 domain was the main determinant of Bicc1’s RNA-binding activity, while KH1 and KH3 contributed more minimally. However, while the GXXG motif is integral for RNA-binding, it does not confer RNA target specificity.

Therefore, my research is focused on the mechanism of Bicaudal-C’s RNA binding activities and elucidating what factors confer Bicc1 RNA-binding specificity. I hypothesized that specificity for the Bicc1–RNA interaction is conferred by the amino acid residues in the regions surrounding the $GXXG$ motif in the KH2 domain.

**Methods**

*Methodology Overview*

In order to test this hypothesis, I made specific amino acid mutations in the KH2 region of Bicc1 and observe the RNA-binding activity of the variant Bicc1 proteins \textit{in vivo}. The KH domains contain multiple regions that are present on the surface of the protein which are available to contact the RNA. I am targeting six of these predicted nucleic acid binding regions in Bicaudal-C KH\textsubscript{2} domain: $HIH161-163$, $ME155-156$, $IK153-154$, $I144$, $SH141-142$, and $TEH138-140$. The goal is to successfully create mutations and targeted sites by mutating the wild-type amino acids to alanine. Generating Bicc1 mutant DNA involves four steps: cutting the mRNA expression plasmid with enzymes for cloning; PCR amplification of the Bicc1 fragments
containing the desired mutations; cloning the PCR fragments with Gibson Assembly enzyme mix; and identifying the correct BicC1 mutants with DNA sequencing.

**Figure 7.** Predicted nucleic acid binding regions in Bic-C KH2
Running PCR and purifying the PCR product

**DNA Polymerase PCR**

A DNA polymerase chain reaction (PCR) was run with the interest of amplifying the insert DNA. Each of the six template DNA used different forward and reverse primers to align the fragments together. The 50uL PCR reactions contained 10uL 5x Q5 Reaction Buffer, 1uL of 10 mM dNTPs, 2.5 uL of 10 uM Forward Primer, 2.5 uL of 10 uM Reverse Primer, 1uL of template DNA, 0.5 uL of Q5 Hi-Fidelity DNA Polymerase, and 32.5 uL of Nuclease-Free water. The reactions were placed in the thermocycler and the 2x Phusion protocol was selected with the appropriate pre-set annealing temperature.
Isolating the insert and vector by gel purification

The PCR reactions were then run on a 1% agarose gel for 35 minutes. The DNA fragments viewed under ultraviolet (UV) lights with crisp bands were cut out from the gel using a razor before continuing with gel extraction protocol. The cut gel slices were placed in a 1.5 ml microcentrifuge tube with the addition of 900 uL of binding buffer to dissolve the agarose and denature the proteins to promote DNA binding. The gel mixture with the binding buffer was incubated at 50 degrees Celsius for 15 minutes to dissolve the gel. After ensuring the gel was completely dissolved and at an optimal yellow color, up to 800 uL of the solubilized gel mixture was transferred to a purification column and centrifuged for one minute, followed by discarding the flow-through. Approximately 700 uL of wash buffer pre-diluted with ethanol was added to the purification column and centrifuged twice to ensure complete removal of residual buffer. The purification column was then transferred to another clean 1.5 mL microcentrifuge tube and eluted with 25 uL of elution buffer and centrifuged for 1 minute. A NanoDrop Spectrophotometer was used to determine recovered DNA concentration from the gel extraction. The purified DNA was then stored at -20 degrees Celsius.

Ligating the inserts into vectors

To facilitate the fusion of the insert into the recipient plasmid, ligation reactions were set up. The 10 uL ligation reactions were composed of 2 uL of the vector Afe1779 ssp1, 3 uL of each of the two segments, and 2 uL of infusion mix. The ligation reactions were then incubated for 15 minutes on the 50 degrees hotplate.
**Bacterial transformations**

Bacterial transformations were then used to introduce the DNA into a bacterial cell for the ease of storing and replicating plasmids. Agar plates with the AMP-antibiotic and the DH5-alpha competent cells were warmed up to room temperature. The 10 uL ligated reactions were then mixed with 75 uL of the competent cells and incubated on ice for 25 minutes. After heat shocking each transformation and placing back on ice for 2 minutes, 900 uL of SOC media were added to the bacteria and grown in the shaking incubator for at least 45 minutes at 37 degrees Celsius. The transformations were then plated onto the agar plate and incubated overnight at 37 degrees Celsius.

**Plasmid miniprep and sequencing**

Individual bacterial colonies were picked from the streaked plate to inoculate 5 mL of LB medium and 5 uL of AMP-antibiotic in a glass tube. The bacterial culture was then placed in the shaking incubator overnight at 37 degrees Celsius. The bacterial pellet were then collected the next day by the centrifugation of the bacterial culture and removing all supernatant. The pelleted cells were completely resuspended in 250 uL of the resuspension solution by pipetting up and down. Then 250 uL of lysis solution were added, followed up 350 uL of the neutralization solution and inverted multiple times. The cells were centrifuged for 5 minutes and the supernatant was transferred to the spin column and centrifuged for 1 minute. Wash solution (500 uL) was applied to the spin column and centrifuged 1 minute twice. The spin column was then transferred to a clean 1.5 mL microcentrifuge tube and eluted with 50 uL of the elution buffer and centrifuged for 2 minutes. The purified plasmid DNA was stored at -20 degrees Celsius. The final product was then sequenced at the UW Biotech Center to verify the plasmid.
**Bicc1 protein expression and purification**

In order to express Bicc1 proteins, pET28b bacterial expression vectors as N-terminal fusions with a His-6, SUMO tag was used for the cloning of Bicc1 proteins. Each plasmid was grown in cultures of *E. coli* cells to 600 OD of 0.6 and induced with 1mM IPTG at 25 degrees Celsius overnight. Afterwards, the cells were collected and lysed in 2.5 mM MgCl₂, 2.5% glycerol, 0.5x TBS, B-PER reagent, 10 mM Imidazole, 1mM BME, 1mM ATMP, 200 mM NaCl, and protease inhibitors. The resulting soluble lysate was then applied to a HisTrap FF chromatography column. The column was then washed ten times with 1x TBS, 5 mM MgCl₂, 5% glycerol, 2mM BME, 20mM Imidazole, 400 mM NaCl. Finally, the proteins were eluted with 450 mM of imidazole and completed by dialyzing in 1x TBS solution.

**Luciferase reporter mRNAs assay**

Reporter mRNAs would be diluted to a concentration of 2.5 nM and 5 nanoliters would be injected into the embryonic cells. After the injected embryos reach a suitable period, extracts of the embryonic cells would be prepared and further analyzed for luciferase activity to assay binding activity.
Results & Future Directions

[Disclaimer: Unable to complete experiments and get adequate results due to the COVID-19 pandemic]

Figure 8. Successful cloning and sequencing of the IK1530154 mutant.
Figure 9. Successful cloning and sequencing of the SH141-142 mutant.
Figure 10. Successful cloning and sequencing of the TEH138-140 mutant.

The region highlighted in purple is the Bicc1 KH2 domain where we have the proposed mutations that flank the GXXG motif within the region. The circled region shows the successful mutations in the KH2 domain of Bicc1.
Figure 11. Successful expression and purification of Bicc1 protein. The control wild-type Bicc1 weighs approximately 70 kilodaltons, and the purified Bicc1 protein measures up to the same.

The cloning of all the DNA plasmids into the RNA vectors and protein expression vectors were successful, however further experiments need to be conducted in order to examine how the mutants affect the translational repressor capabilities of the Bicc1 protein. For future directions, it would be crucial to conduct luciferase repression assays to analyze mutant Bicc1 binding activity. The next steps will be to express the Bicc1 variants in embryos along with reporter mRNA containing the luciferase open reading frame (ORF), followed by the known Bicc1 target: Cripto1 mRNA and non-target: cyclin, to analyze their ability to repress translation. These in vivo assays are essential for constructing an understanding of the biological relevance of the proposed mutations in the Cripto1 mRNA target. Previous studies published in the Sheets Lab have revealed that in vitro binding assays can have a different impact than in vivo assays, thus conducting in vivo assays in the future would be important for providing additional information in understanding and ultimately analyzing the mutants for defects in Bicc1 binding to mRNA.
References


