

**Investigating the regulation of the production of the long non-coding RNA
BC200 and its truncated isoform BC120**

By: Alyssa Cohen

A thesis submitted to Faculty of Graduate and Postdoctoral Studies of The
University of Manitoba

in partial fulfillment of the requirements of the degree of

Master of Science

Department of Chemistry

Faculty of Science

University of Manitoba

Winnipeg, Manitoba, Canada

Copyright © 2026 by Alyssa Cohen

Copyright Permissions

Figure 1.1. Panel A adapted from ⁽¹⁾. Panels B and C adapted from ⁽²⁾.

Figure 1.2. Adapted from (27).

Figure 1.3. Adapted from ⁽³⁾.

Figure 1.5. Adapted from ⁽²⁾.

Figure 1.6. Panel A adapted from ⁽⁴⁾. Panel B adapted with permission from ⁽⁵⁾.

Figure 1.7. Adapted from ⁽²⁾.

Figure 2.2. Adapted from ⁽³⁾.

Abstract

BC200 is a long non-coding RNA closely related to the 7SL, a component of the signal recognition particle and the ancestor to all Alu-like RNAs. Consisting of an Alu domain, followed by an adenosine-rich region, unique sequence and poly-cytosine tract, BC200 has many protein binding partners. Interacting with the SRP9/14 heterodimer, BC200 has been shown to act as a regulator of transcription and translation. It has also been shown that BC200 serves as a positive regulator of translation in tumour cells, along with a promoter of tumour cell migration. Recent sequencing studies done by the McKenna lab led to the discovery a novel non-coding RNA almost identical to the first ~120 nucleotides of BC120. This RNA, known as BC120, is expressed endogenously in a myriad of cell lines and types, and is highly expressed in ovarian tumours. The similarity of this RNA to BC200 has led the McKenna group to question whether this RNA arises from BC200 processing events related to the poly-cytosine tract of the RNA, or from non-canonical secondary structures. Overexpression plasmids containing truncations of BC200 with varying cytosine tract lengths were developed and transfected into MCF-7 cells prior to extracting the RNA and probing for BC200 and BC120 via Northern Blot. This showed an increase in cellular BC200 as the cytosine-tract length in the truncation increased. Additionally, *in vitro* transcribed and biotinylated BC200 truncations were used in pull-down assays prior to Western Blotting to probe for poly-C binding protein 2, showing that at least 5 cytosines are required for poly-C binding protein 2 to interact BC200. *In vitro* transcription conditions relating to monovalent cation salt concentration, incubation temperature, and nucleoside triphosphate concentration were altered to assess which conditions, if any, altered the amounts of BC120 produced during transcription. The addition of potassium chloride promotes the production of BC120 during *in vitro* transcription of BC200, as well as a non cytosine-tract

truncation, BC185. Detectable changes were not observed in the reduced incubation temperature or nucleoside triphosphate conditions. Overall, this project aimed to characterize the processes that contribute to the unique expression and transcriptional patterns observed in BC120.

Acknowledgements

I would like to begin by thanking my supervisor Dr. Sean McKenna. Dr. McKenna has been my supervisor since my time as an undergraduate student. Aside from providing invaluable guidance throughout my time in his research group, and providing feedback and edits to make to my thesis, Dr. McKenna always made sure to remind me that even though I am my own worst critic, I have been making progress. Without him, I do not think I would have had the ability to persevere through all the curveballs this research has thrown at me.

Throughout my time in the McKenna lab, there have been lab members who also helped me navigate the trials and tribulations associated with research. Firstly, I would like to thank Daniel Gussakovsky who, while supervising me as an undergraduate, helped me hone many of the techniques that I used almost every day during my graduate studies. Secondly, I want to acknowledge Dr. Evan Booy for always answering questions I had about my project. Thirdly, Mira Brown, who helped me navigate the process of protein purification, an area I had little experience in. Next, I want to acknowledge Liam Zamora who, although I only got to really know during the last half of my graduate studies had become not just a colleague, but a good friend. Lastly, I want to thank the two students who I directly supervised during my graduate studies, Bruce Barrientos and Khoi Ho who contributed to the experiments contained within this thesis. I also want to acknowledge the Faculty of Science, Department of Chemistry, Graduate Enhancement of Tri-Agency Stipends (GETS) and the Faculty of Graduate and Post-Doctoral Studies for research funding.

Finally, I want to thank my friends and loved ones for their unwavering support in my pursuit of graduate studies.

Dedication

This thesis is dedicated to my support network, which, despite being small, was strong enough to keep me going when I was ready to give up.

Table of Contents

Copyright Permissions.....	i
Abstract.....	ii
Acknowledgements	iv
Dedication	v
List of Tables.....	viii
List of Figures.....	ix
List of Abbreviations.....	xi
Chapter 1: Introduction	1
1.1 Non-coding RNA	1
1.1.1. Coding versus non-coding RNA.	1
1.1.2 lncRNAs	2
1.1.2. Alu RNAs	2
1.2 The Signal Recognition Particle	5
1.2.1 7SL	5
1.2.2 The signal recognition particle (SRP)	7
1.2. BC200 and BC120	9
1.3.1 BC200.....	9
1.3.2 BC120.....	12
1.4 Thesis Outline.....	18
Chapter 2: Materials and Methods	19
2.1 Reagents	19
2.2 Cell culture.....	19
2.2.1 Cell plating and passaging	19
2.2.2 Freezing down cells, storage and culture recovery	20
2.3 Molecular cloning.....	21
2.3.1 Plasmid construction and validation.....	21
2.3.2 Plasmid transformation, production and storage of glycerol stocks	23
2.4. Assessing the relative levels of BC200 and BC120 RNAs in cells.....	25
2.4.1 Cell transfection, collection and RNA purification.....	25
2.4.2. Quantification of transfection efficiency	26
2.4.3. Northern Blotting	27

2.4.4. BC120-discriminating RT-qPCR.....	29
2.5. The production and purification of a Ribonuclease Inhibitor.....	31
2.5.1 RNH1 production and purification.....	31
2.5.2 Assessing purified RNH1 efficacy	35
2.6. Assessing the protein binding partners of BC200 truncations	35
2.6.1 <i>In vitro</i> transcription plasmid preparation	35
2.6.2 T7 RNA Polymerase production and purification	38
2.6.3 T7 polymerase testing and transcription optimization	40
2.6.4. Large-scale transcriptions and RNA purifications.....	41
2.6.5 RNA-biotin ligations and anti-biotin magnetic bead preparation.	42
2.6.6 Pull-down assays to assess truncation-protein binding partner interactions.....	43
2.6.7 Western blotting against BC200 truncation binding partners.	43
2.7. Assessing how transcription conditions impact the production of BC120.....	44
2.7.1 Temperature-dependent transcriptions	45
2.7.2 Salt conditions	45
2.7.3 Varied NTP concentrations	46
Chapter 3: Results.....	47
3.1. Acknowledgements	47
3.2. 5 cytosines in the poly-C region of BC200 truncations results in increased levels of cellular BC200.	47
3.3. Poly-C binding protein 2 requires approximately 5 cytosines to engage in RNA-protein interactions with BC200.	53
3.4. BC120 is formed during BC200 <i>in vitro</i> transcription reactions in the presence of KCl, a salt that stabilizes G-quadruplexes, while lower transcription temperatures or NTP concentrations lead to inconclusive results.	55
3.4.1. BC120 formation increases in the presence of KCl	55
3.4.2. The reduction of NTP concentrations and changes in <i>in vitro</i> transcription temperature do not appear to impact BC120 production <i>in vitro</i>	57
Chapter 4: Discussion and Future Directions	60
4.1. Discussion	60
4.2. Future Directions.....	64
Works cited	67

List of Tables

Table 2.1. Primers used in the production of inserts used to make BC200 truncation overexpression vectors.	22
Table 2.2. The composition and concentrations of each component of the buffers used to purify RNH-1.	32
Table 2.3. The primers used to generate the inserts required to make <i>in vitro</i> transcription plasmids of different BC200 poly-C tract truncations.	36
Table 2.4. The composition of each buffer used in the purification of T7 RNA polymerase.	38

List of Figures

Figure 1.1. All Alu RNAs can be traced back to an ancestral lncRNA progenitor.	4
Figure 1.2. The signal recognition particle (SRP) is a ribonucleoprotein complex implicated in the regulation of protein synthesis and cellular protein localization.	7
Figure 1.3. BC200 can be grouped into distinct regions based on its secondary structure and sequence.	10
Figure 1.4. Many proteins have been confirmed to interact with distinct regions of BC200.	12
Figure 1.5. The presence of BC120, an Alu monomer with a sequence and secondary structure virtually identical to BC200 was identified	14
Figure 1.6. The presence of non-canonical secondary structures formed in DNA and/or RNA may modulate the production of BC120 transcripts from the genomic BC200 sequence.	16
Figure 1.7. BC120 may arise from cellular post-transcriptional processing events of full-length BC200.	17
Figure 2.1. Re-digestion of BC200 truncation overexpression constructs confirms a drop-out of the expected size if cloning were successful.	24
Figure 2.2. Schematic illustrating the principle behind the BC120-discriminating RT-qPCR assay.	30
Figure 2.3. Size-exclusion chromatography is a highly effective secondary step to achieve highly pure RNH-1.	34
Figure 3.1. The probe targeting nucleotides 86-107 in BC200 non-specifically binds to 7SL and other Alus in the cell.	48
Figure 3.2. The length of cytosines in the poly-C tract of the overexpression plasmid impacts the cellular levels of BC200 and BC120 RNA.	49
Figure 3.3. The quantification of GFP mRNA via RT-qPCR serves as a means to assess transfection efficiency.	51
Figure 3.4. Attempts to quantify BC200 and BC120 via a probe-based RT-qPCR assay yielded inconsistent results.	52
Figure 3.5. Approximately 5 cytosines are necessary for noticeable interactions with the poly-C tract of BC200.	54
Figure 3.6. The presence of KCl salt during the <i>in vitro</i> transcription of BC200 and BC185 promotes the production of BC120.	56

Figure 3.7. Reduced NTP concentrations (A) and reduced incubation temperatures (B) do not have a significant impact on the amount of BC120 produced during *in vitro* transcription. 59

List of Abbreviations

A	Adenosine
ATP	Adenosine triphosphate
BC1	Brain cytoplasmic 1 RNA
BC120	Brain cytoplasmic RNA 120
BC200	Brain cytoplasmic RNA 200
bU	Bromouridine
C	Cytosine
cDNA	Complementary DNA
C-hairpin	Cytosine hairpin
CSDE1	Cold Shock Domain Containing E1
Ct	Cycle threshold
CTP	Cytosine triphosphate
CV	Column volume
Da	Dalton
DHX36	DEAH-box Helicase 36
DIG	Digoxigenin
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
dsDNA	Double-stranded DNA
dT	Deoxythymidine
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-mesenchymal transition
ER	Endoplasmic reticulum
EV	Empty vector
FASTA	Fast-All

FBS	Fetal bovine serum
Ffs	Fifty-four homologue
FLAM	Free left Alu monomer
FPLC	Fast protein liquid chromatography
FRAM	Free right Alu monomer
G	Guanosine
G4	G-quadruplex
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
GTPase	Guanosine triphosphatase
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HNRNPK	Heterogenous nuclear ribonucleoprotein K
HRP	Horseradish peroxidase
IP	Immunoprecipitation
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IVT	<i>In vitro</i> transcription
K	Potassium
KCl	Potassium chloride
kDa	Kilodalton
KH	K homology
LB	Luria Bertani
Li	Lithium
LiCl	Lithium chloride
LINE-1	Long Interspersed Nuclear Element-1
LNA	Locked nucleic acid
lncRNA	Long non-coding RNA

MCF-7	Michigan Cancer Foundation-7
MEM	Minimal essential medium
MgCl ₂	Magnesium chloride
miRNA	MicroRNA
mRNA	Messenger RNA
MΩ · cm	Mega ohm-cm
Na	Sodium
NaCl	Sodium chloride
ncRNA	Non-coding RNA
NP-40	Nonyl phenol-40
OD ₆₀₀	Optical density at 600 nm
Oligo	Oligonucleotide
<i>OPA1</i>	Optic atrophy 1
PABPN1	Poly(A)-binding protein nuclear 1
PABPN2	Poly(A)-binding protein nuclear 2
PCBP	Poly-(C)-binding protein
PCBP2	Poly(C)-binding protein 2
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEI	Polyethyleneimine hydrochloride
PMSF	Phenylmethylsulfonyl fluoride
Ppi	Pyrophosphate
PVDF	Polyvinylidene fluoride
RHAU	RNA Helicase associated with AU-rich element; also known as DHX36
RNAP	RNA polymerase
Rpm	Rotations per minute
rRNA	Ribosomal RNA

RT-qPCR	Reverse transcriptase-quantitative polymerase chain reaction
saRNA	Self-amplifying RNA
SCC	Sterile sodium citrate
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SINE	Small interspersed nuclear element
siRNA	Small interfering RNA
SRP	Signal recognition particle
SRP14	Signal recognition particle 14
SRP19	Signal recognition particle 19
SRP54	Signal recognition particle 54
SRP68	Signal recognition particle 68
SRP68/72	Signal recognition particle 68/72 heterodimer
SRP72	Signal recognition particle 72
SRP9	Signal recognition particle 9
SRP9/14	Signal recognition particle 9/14 heterodimer
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
STRAP	Serine-threonine kinase receptor-associated protein
T	Thymidine
TAE	Tris-acetate-EDTA
TBE	Tris-borate- EDTA
TBS-T	Tris-buffered saline with Tween 20
TRIS	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
UTP	Uridine triphosphate

ΔC_t

Delta Ct

 $\Delta\Delta C_t$

Delta-delta Ct

Chapter 1: Introduction

1.1 Non-coding RNA

1.1.1. Coding versus non-coding RNA.

A concept that is foundational to our understanding of gene expression is the central dogma of molecular biology (6). According to the central dogma, deoxyribonucleic acid, or DNA, acts as the form of genetic material stored by the cell. Its contents are then transcribed into ribonucleic acid, or RNA. This RNA is then translated by the ribosome into proteins, the cellular machinery. For a very long time, it was thought that all DNA in the cell got transcribed into RNA that acted as the template for translation. This idea was solidified by the results of the Human Genome Project, which found that only 1-2% of the human genome actually contained protein-coding stretches of DNA (7). The rest of this DNA began being referred to as ‘junk DNA’ as it was believed to not carry any biological and functional significance (8). But as time went on, it was found that these pieces of junk DNA weren’t actually ‘junk’ at all. These stretches of DNA act as the template for non-coding RNAs (ncRNAs), which, as time went on, were found to carry functional significance.

ncRNA acts as a broad term for any RNA in the cell that does not act as the template for protein synthesis (9). ncRNAs are further characterized into two groups known as housekeeping and regulatory ncRNAs. Housekeeping RNAs are expressed in all cell types and carry functions pertaining to gene expression. Examples of housekeeping ncRNAs include ribosomal and transfer RNAs (rRNA and tRNA, respectively), both of which play pivotal roles in protein translation. Regulatory ncRNAs on the other hand, function to regulate epigenetic processes,

transcription, and post-transcriptional processes. Small-interfering RNAs (siRNAs) and microRNAs (miRNAs) are pertinent and well-known types of regulatory ncRNAs.

1.1.2 lncRNAs

Another type of regulatory ncRNA of interest to researchers is the group of RNAs called long non-coding RNAs (lncRNAs) (9). lncRNA is a broad term that encompasses many RNAs that are at least 200 nucleotide bases long, and, like other regulatory RNAs play many roles in modulating gene expression and protein production in the cell (10). Broadly, lncRNAs have been found to participate in processes related to transcriptional regulation, translational regulation, as well as post-transcriptional processing, among others. lncRNAs can be grouped based on their location in the genome as well as how they arise in said locations, also known as their context (10). A pertinent type of ncRNA gaining increasing interest in the scientific community is the Alu RNA.

1.1.2. Alu RNAs

Alu RNAs are long ncRNAs (lncRNAs) that are derived from DNA sequences known as Alu elements (11). Comprising about 11% of the human genome, these elements gained their pervasiveness through retrotransposition. Alu elements are a type of short-interspersed element (SINE), meaning that they are non-autonomous retro-transposable elements, dependent on the function of a trans-activating factor called long interspersed element-1 (LINE-1) (1,11,12). LINE-1 associates with a complex containing the signal recognition particle 9/14 heterodimer (SRP9/14), and Alu RNA. Alu complementary DNA (cDNA) is generated through the reverse transcriptase activity carried by LINE-1. This cDNA is then brought back to the genomic DNA, which is cut via LINE-1 endonuclease activity. Finally, the cDNA is inserted into the genome via

the integrase activity of LINE-1. Aside from sharing the process of LINE-1 mediated retrotransposition, Alu elements share sequence similarities which, at the RNA level, incur secondary structural similarities (1,11,12). At the DNA level, Alu elements contain an internal RNA polymerase (RNAP) III promoter region containing A and B box promoter elements (1,11). A poly-adenosine (poly-A) tract separates the left arm from the right arm, and this right arm is followed by another terminal poly-adenosine tract.

Despite characteristics that are common amongst all Alu elements and RNAs, there are sequence differences that allow Alus to be further characterized into different families. These families include the AluY, AluS, and AluJ families, ordered from evolutionarily youngest to oldest (11,13,14). The evolution of these Alu families is dictated through the retrotransposition events they undergo, which, naturally, lead to the accumulation of mutations (11). Due to their integration into the genome, there are very few, if any deletion events that occur. As such, the evolution and accumulation of Alu families in various primate species can be looked at, even being used as a tool to determine speciation. It is important to recall that, despite being divided into different subfamilies, the transcription products of all Alu elements (Alu RNAs) share a secondary structure similar to that of their ancestral RNA, 7SL (1,11).

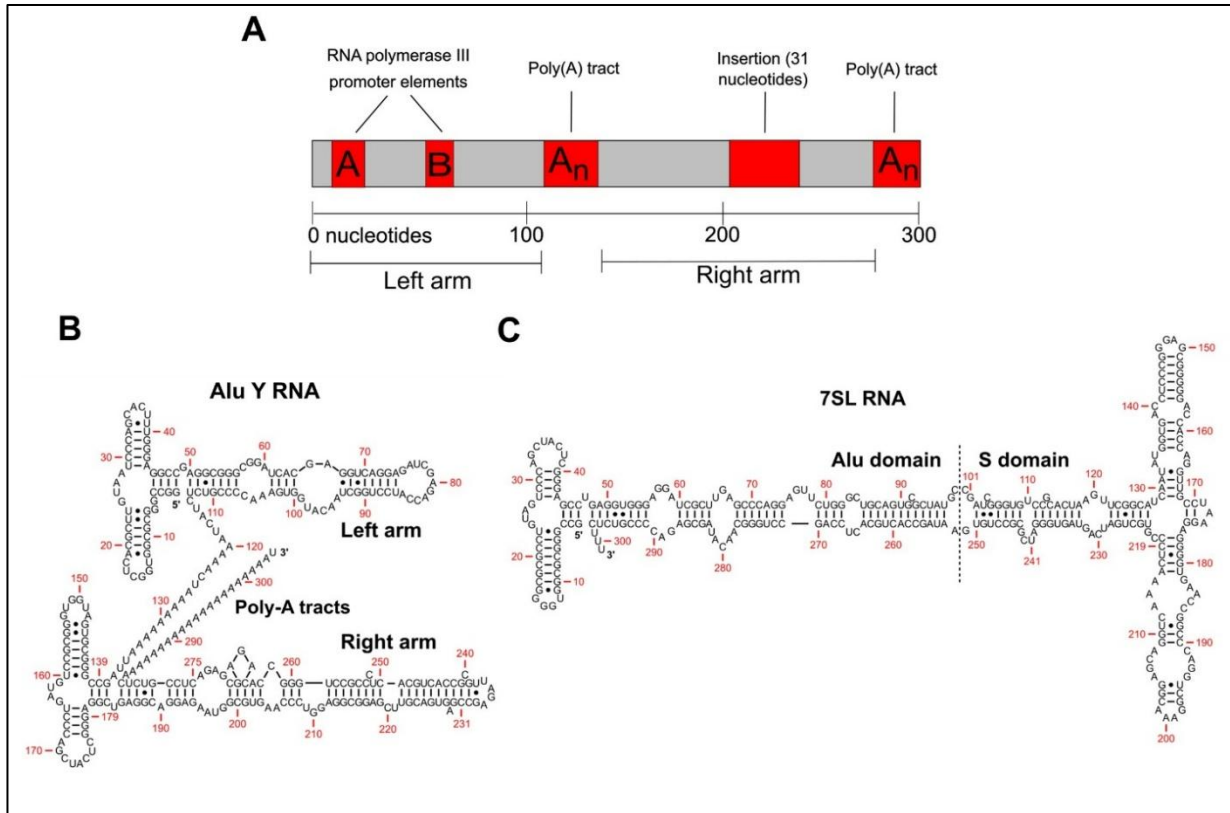


Figure 1.1. All Alu RNAs can be traced back to an ancestral lncRNA progenitor. (A) The genomic sequence of a canonical Alu RNA, containing internal RNAP III promoter elements, two poly-A tracts, a unique 31 nucleotide insertion and a transcription termination sequence. (B) The sequence and secondary structure of a canonical Alu Y RNA. The left and right arms of the RNA are depicted as being above and below each other respectively, separated by the internal poly-adenosine tract. The RNA is flanked by another poly-adenosine tract. (C) The sequence and secondary structure of the 7SL RNA.

Alu elements are of mounting interest to researchers due to their implications in a myriad of human disease states (1). In the context of disease, Alu elements, along with their transcription products are contributors to disease progression and severity. An example of an Alu insertion associated with disease development is that of the integration of an Alu repeat into the intron of the optic atrophy 1 (*OPA1*) gene (15). In a study done on a large family, all presenting with autosomal dominant optic atrophy, this insertion was found in intron 7 of the *OPA1* sequence,

causing an in-frame deletion in the eighth exon. The authors of this study postulated that the result of this in-frame deletion resulted in OPA1 losing one of its important enzymatic functions, guanosine triphosphatase (GTPase) activity, disrupting important regulatory networks and causing disease.

Not only are Alu element insertions into a gene disruptive to important cellular and regulatory processes, so are their transcriptional products. A pertinent example of this is seen in a paper by Ruocco *et. al*, detailing how Alu RNAs are implicated in cancer progression (16). In the context of colorectal cancer, it has been found that a high abundance of Alu RNAs interact with a microRNA (miRNA) called miR-566, which, when downregulated, is associated with colorectal cancer progression (16,17). In this paper, it is found that upon Alu RNA transfection, miR-566 levels in the cell decrease relative to non-transfected conditions (16). Further, the high abundance of Alu RNAs in colorectal cancer is due to a deficit in the DICER1 protein, a phenomenon the authors also found in ovarian, renal and breast cancers. The findings in this paper led to the conclusion that an overabundance of Alu RNAs in cancerous cells deplete miR-566, inducing the epithelial-mesenchymal transition (EMT), responsible for the development of malignancy. Thus, at both the DNA and RNA level, Alu elements play a pivotal role in many disease states in the body, making them an increasingly interesting avenue of biomedical study.

1.2 The Signal Recognition Particle

1.2.1 7SL

The secondary structural similarities between all Alu RNAs points to the fact that all Alu elements are derived from an ancestral ncRNA known as 7SL (1,11). Research has shown that Alu elements arose from an inefficient retrotransposition event of a 7SL deletion occurred before

the evolutionary divergence of primates and rodents (11). In rodents, this event resulted in the evolution of B1 repeats. In primates, this event gave rise to the free left Alu monomer (FLAM) and free right Alu monomer (FRAM) sequences. Eventually, these two elements dimerized in primates, giving rise to the highly retrotransposable Alu elements (18,19).

7SL, like other Alu elements, is functionally significant in the cell. This ncRNA acts as a component of the signal recognition particle (SRP) (20). Namely, it interacts with the signal recognition particle 9/4 (SRP9/14) heterodimer. This ribonucleoprotein complex functions to pause protein synthesis through interactions with the ribosome (20,21). In addition, 7SL RNA, along with the rest of the SRP play a role in protein localization, directing newly synthesized proteins to and from the endoplasmic reticulum (ER) for further processing and localization events. It has also been found that outside of protein localization, the SRP, namely SRP9/14 play a crucial role in regulating the transcription of Alu RNAs, including 7SL and one of its important descendant RNAs, Brain Cytoplasmic RNA 200 (BC200) (2).

The structure of 7SL is that of two main regions, being the Alu domain, and the S domain (22). Both of the 7SL domains play an important functional role in regulating protein synthesis. For context, the process of protein translation consists of three main steps. In order of occurrence, those steps are initiation, elongation, and termination (23). In the initiation step of translation, a meticulous stepwise process results in the formation of the complete 80S initiation complex, consisting of various elongation factors, the 40S subunit, 60S subunit, poly-A binding proteins, and messenger RNA (mRNA). In the elongation step, amino acids are brought to the ribosome and are catalytically joined together. The order of amino acids in a peptide chain is determined by the codons in the mRNA, which pertain to specific amino acids. When a newly synthesized peptide chain emerges from the ribosome, the Alu domain of the 7SL RNA binds to

a signal sequence on the newly synthesized peptide (22). The Alu domain on the other hand, functions to pause elongation.

1.2.2 The signal recognition particle (SRP)

The signal recognition particle (SRP) is a ribonucleoprotein particle that is essential to cellular function. Being a heterogenous ribonucleoprotein, the protein component of the mammalian SRP consists of six distinct peptides, namely SRP9, 14, 19, 54, 68 and 72 as characterized by Walter and Blöbel and cited by Saraogi and Shan (24). As summarized by Kellogg et al., these six proteins, along with the 7SL RNA can be divided into two distinct domains, the Alu and S domains (25). The SRP9/14 heterodimer binds to the Alu segment of 7SL to form the Alu domain, while SRP 19, 54, 68/72 heterodimer bind to the S segment, forming the S domain (22,25–28). SRP 54 is a heavily conserved component of the SRP, with a homologue (fifty-four homologue, ffs) being the sole protein that makes up the eubacterial SRP (29).

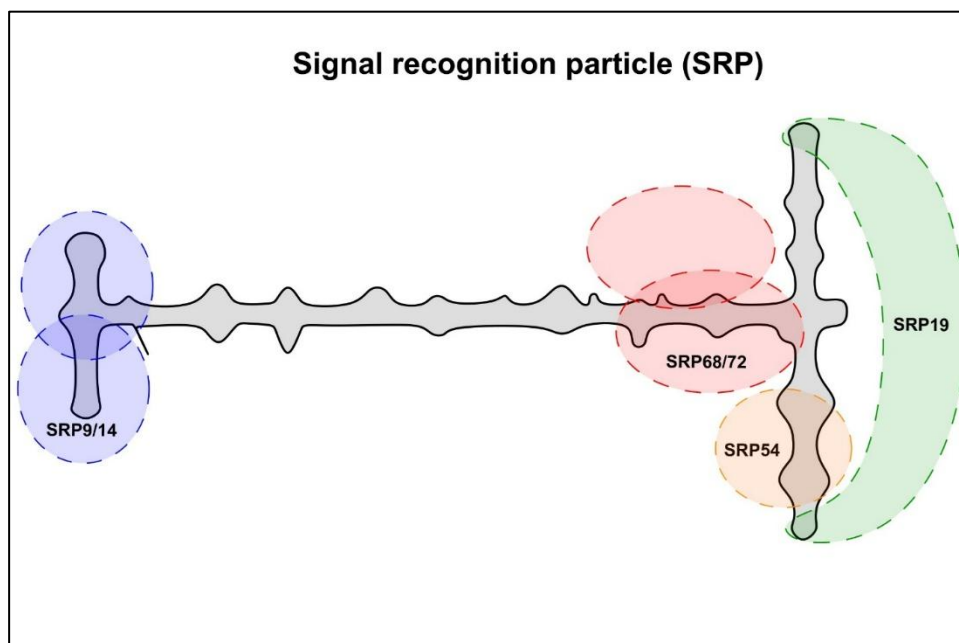


Figure 1.2. The signal recognition particle (SRP) is a ribonucleoprotein complex implicated in the regulation of protein synthesis and cellular protein localization. The SRP9/14 heterodimer

bind to the Alu region of 7SL, implicated in translational arrest (17,18). SRP19, SRP54, and the SRP68/72 heterodimer bind to different regions of the S domain (25). Without SRP19 binding and causing a conformation change in 7SL, SRP54 is unable to interact with the S domain, which docks ribosomes to the endoplasmic reticulum.

The understanding of the SRP in the cell has progressed, but much is still to be learned about it in completion. In relation to the Alu domain, it has been found that neither SRP9, nor SRP14 can bind to 7SL in isolation; they must be dimerized in order to interact with the Alu domain of 7SL (30). Relating to the S domain, the binding of SRP54 is SRP19-dependent; in that SRP19 causes a conformational change in the secondary structure of 7SL, facilitating SRP54 binding (31–33). The assembly of the SRP follows a meticulous group of steps referred to as the SRP cycle. As summarized by Bovia and Strub, the SRP recognizes a signal sequence on the newly synthesized peptide emerging from the ribosome (34). It is this binding which causes the elongation step of translation to be paused. The SRP redirects this peptide-ribosome complex to the SRP receptor on the endoplasmic reticulum (ER) membrane via a guanosine triphosphate (GTP) hydrolysis-mediated mechanism. This complex comes together with the translocon channel protein in the ER membrane, inserting the polypeptide. From here, another signal sequence is recognized, causing the SRP to dissociate from the polypeptide-ribosomal complex, thus allowing translation to continue.

While 7SL acts as part of the SRP, it has been determined that the SRP9/14 heterodimer, alongside its functions in protein translocation, plays a role in modulating the transcription of 7SL, along with its descendent RNA, BC200 (2). A paper by Gussakovsky *et. al* proposed a mechanism by which the SRP9/14 heterodimer recruits SRP9/14 to the 7SL/BC200 genetic locus, beginning transcription (2). SRP9/14 interacts with the beginnings of BC200 or 7SL RNA,

before the full RNAs get transcribed by RNA polymerase III. When SRP9/14 is not bound to this RNA, the subsequent transcriptional products get degraded. While this mechanism was proposed for the regulation of 7SL and BC200 transcription, it is likely that the transcription of other Alu RNAs follows similar mechanisms.

1.2. BC200 and BC120

1.3.1 BC200

BC200 is a long non-coding RNA closely related to the Alu family. It is comprised of one Alu domain, and a tail containing an adenosine-rich region and a cytosine-rich region, separated by a unique sequence (35). In other words, BC200 is an Alu monomer as opposed to an Alu dimer. Sequencing studies comparing the Alu domain of BC200 with that of the consensus sequences of other Alu families has shown that BC200 carries many similarities to the AluJ subfamily, suggesting a close evolutionary relationship with canonical AluJ elements (36). Primate-specific BC200 and its rodent-specific analog brain cytoplasmic 1 (BC1) RNA exhibit somewhat similar expression patterns, with both being heavily localized to neuronal dendrites in the brain (37–39). In the brain, BC200 acts primarily as a regulator of protein synthesis, interacting with SRP9/14 heterodimer (40). However, this RNA has also been implicated in the progression of neurodegenerative diseases such as Alzheimer's, where its upregulated expression is implicated in the degeneration of neuronal cells (41).

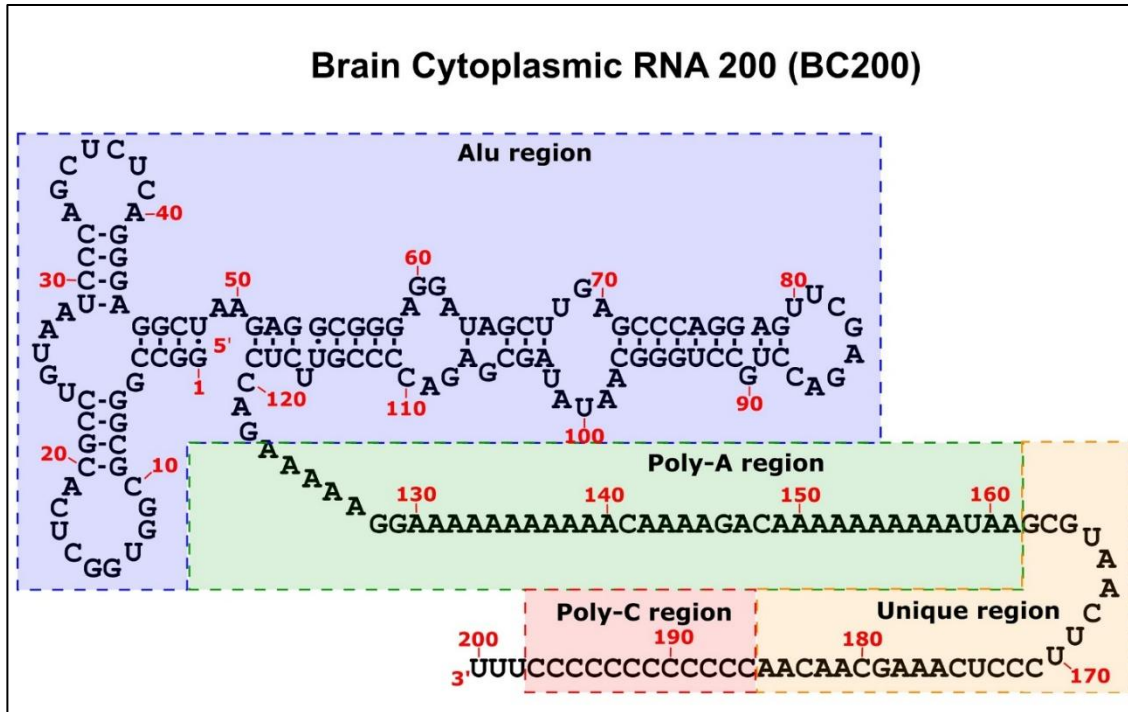


Figure 1.3. BC200 can be grouped into distinct regions based on its secondary structure and sequence.

Although BC200 was initially found to be localized in dendrites, subsequent research has shown that BC200 is present at varying levels in certain cell types, namely being present at high amounts in certain cancers (23–27). Increasingly, BC200 is being considered in the medical science community as a cancer biomarker, indicative of disease progression and prognosis (1,42–45). The extent to which BC200 is involved in cancer development and progression is vast, being involved in a myriad of processes. In the context of non-small cell lung cancer for instance, BC200 promotes chemotherapy drug resistance, along with cancer metastasis. (46) This is further supported by other studies demonstrating that in many cancers, BC200 functions to facilitate tumour cell proliferation and migration, contributing to metastases (47–55).

The capacity of BC200 to regulate protein synthesis is of major interest to researchers as a means to explain how BC200 functions in a cancer context. This type of work has been done

through polysome profiling, which makes use of a sucrose density gradient to separate mRNA with varying degrees of active translation occurring; in other words, mRNAs with more ribosomes attached to it actively undergoing translation are more dense and travel to the bottom of the gradient, while mRNAs with less active translation occurring are less dense and stay higher up in the gradient (⁵⁶). In a polysome profile study upon cells exhibiting BC200 overexpression relative to a control or knock-down condition, it has been found that BC200 promotes protein synthesis in cancer cells, which positively contributes to their viability (57). These results contradict work which previously found that BC200 was a translation inhibitor (58–61).

The primary and secondary structural properties of BC200 allow it to bind to an array of protein binding partners. Insights into the RNA-protein interactions that BC200 is a part of can act as a means to understanding its function in both cancerous and non-cancerous cells (62). Proteomic studies have shown that BC200 binds to many proteins, including but not limited to the SRP9/14 heterodimer, RNA Helicase associated with A/U-rich element (RHAU, also known as DHX36), cold shock domain-containing E1 (CSDE1), serine/threonine kinase receptor associated protein (STRAP), heterogeneous nuclear ribonucleoprotein K (HNRNPK), poly-Adenosine binding protein 2 (PABP2) and poly-Cytosine binding protein 2 (PCBP2) (62–64). These proteins bind to distinct regions in the structure of BC200, with SRP9/14 binding to the Alu domain, RHAU, CSDE1 and STRAP interacting with the A-rich region towards the 3' end of the molecule, and HNRNPK and PCBP2 interacting with the 3' cytosine (C)-rich region (62,63).

Of the proteins mentioned, those of interest include HNRNPK, PCBP2 and PABPN1. Work by Booy *et. al* led to the postulation that CSDE1 promotes proper folding of BC200 (62,65). Moreover, it is probable that CSDE1 binding to BC200 acts to stabilize the RNA,

preventing it from being processed down or degraded in the cell (62). It has also been found that HNRNPK, along with PCBP2, PABPN2, along with other poly-(A) binding proteins have been shown to interact with various ncRNAs to modulate other cellular processes, or contribute to RNA stabilization in cancerous, and non-cancerous contexts (66,67,66,68–70). Thus, the idea that these proteins, which have already been shown to interact with BC200, is an avenue to be considered, which could provide more information pertaining to its cellular roles (62).

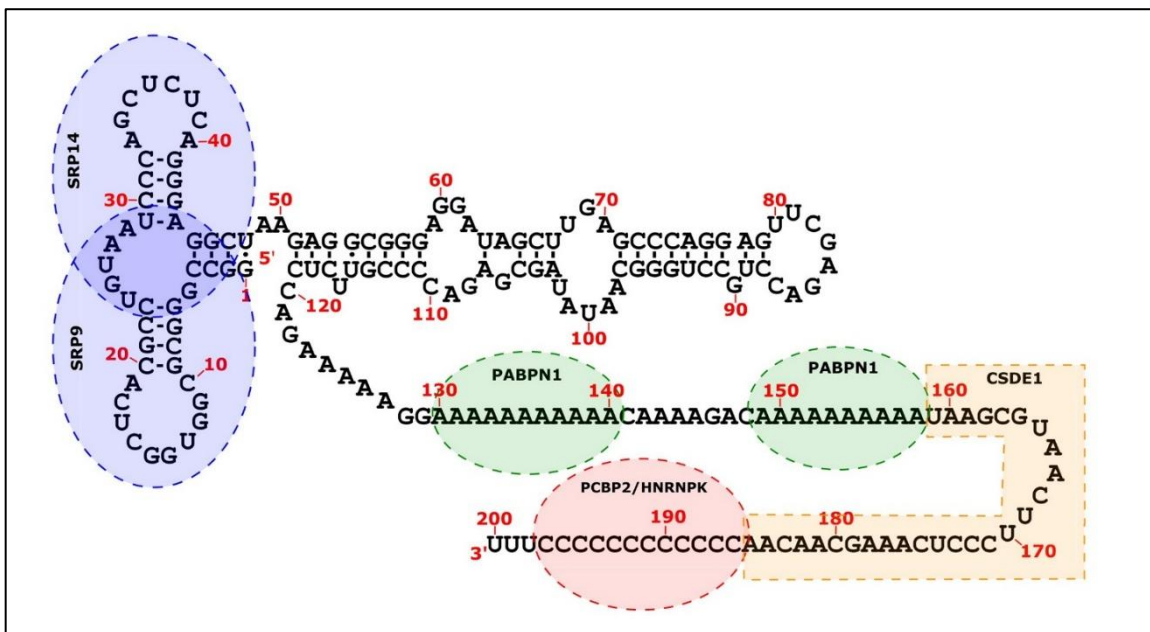


Figure 1.4 Many proteins have been confirmed to interact with distinct regions of BC200 ⁽⁶²⁾. With an Alu region similar to that of 7SL, BC200 interacts with the SRP9/14 heterodimer. PABPN1, CSDE1, PCBP2/HNRNPK have been confirmed as interactors with BC200 via proteomic studies involving immunoprecipitation (IP) assays to confirm what regions of BC200 they bind to (62).

1.3.2 BC120

In a recent paper published by Booy *et al.*, the presence of Brain Cytoplasmic RNA 120 (BC120), which displays sequence similarity to BC200 is detailed (3). What is intriguing is that

this BC120 is expressed at higher levels in certain tissues relative to BC200, and both BC200 and BC120 are found in high abundance in ovarian cancer tumor samples. These findings are indicative of a complex means of regulating the abundance of BC200 and BC120 in cells, which is not yet understood. While BC200 itself is used as a cancer biomarker and an indicator of disease prognosis, the presence of BC120 presents an interesting avenue not yet investigated when looking to understand the implications of ncRNA transcripts, the regulation of their expression, and how they function in the context of cancer (1,3,42–45).

BC120 is an Alu-related monomer, with a sequence similar to the first 118 nucleotides of BC200, with the addition of two uracil's at the 3' end (3). As seen in Figure 1.5, the last two nucleotides of BC120 are highlighted in red. Upon observation of this BC200 truncation, Booy *et al.* conducted sequencing on RNA from Michigan Cancer Foundation-7 (MCF-7) cells and normal brain tissues (3). Interestingly, it was found that this BC200 truncation displayed heterogeneity at the 3' end. At the 3' end, modifications observed including the addition of 1-4 uracil's, or various combinations of adenines and uracils. In brain tissue, this truncation terminated at position 118 of BC200, with the addition of two uracil's at the 3' end. In MCF-7 cells, transcript lengths varied but demonstrated an average length of 120 nucleotides, with most of the transcripts being 119 nucleotides long. With the most common forms of this transcript being 120 nucleotides long, this ncRNA was thereafter referred to as BC120, with the knowledge that 3' end heterogeneity was present within clones of the same cell type, as well as between cell types.

Half-life experiments done to assess the stability of BC120 compared to BC200 find that it is much more stable than its longer counterpart. In these experiments, actinomycin D, which stalls the activity of RNA polymerase III, was added to cells after transfection with a BC200

overexpression plasmid. Levels of BC200 and BC120 at different time points was probed via Northern Blot. At the end of these experiments, it was found that after 12 hours, BC120 levels remained relatively stable in the cell, while BC200 levels diminished significantly. This further supports the indication that the expression of BC120 is likely driven partly by regulatory events in the cell.

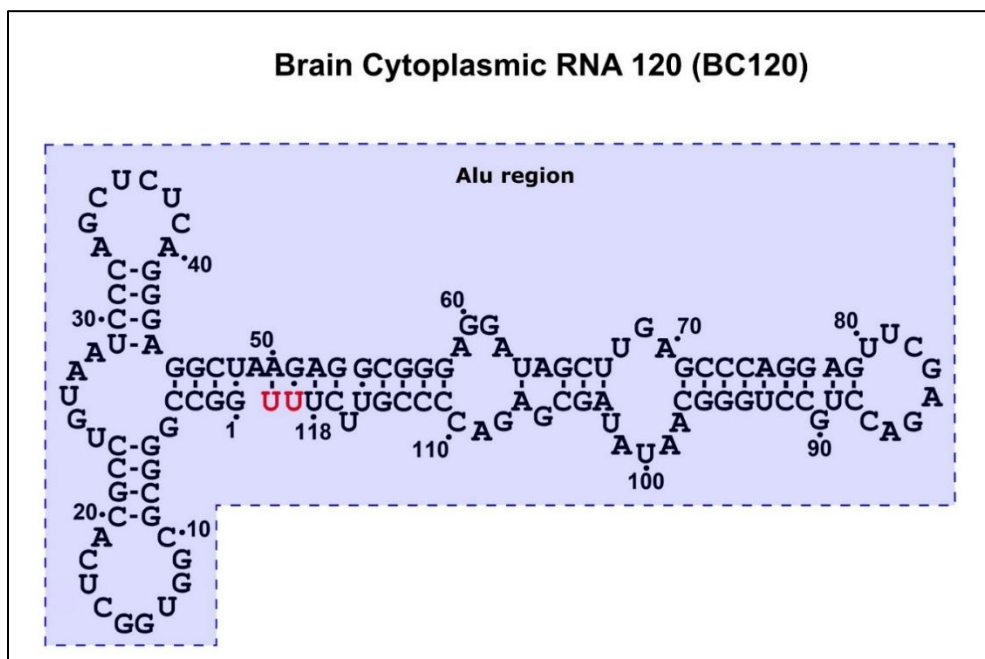


Figure 1.5. The presence of BC120, an Alu monomer with a sequence and secondary structure virtually identical to BC200 was identified (³). BC120 is comprised of an Alu domain without a Poly-A, unique, or Poly-C region. This monomer is found to be more stable than BC200 in half-life studies, displaying distinct expression patterns from its longer counterpart.

DNA, like RNA, has can fold into complex secondary structures when single-stranded (71,72). At the DNA level, the sequence of BC200 contains a long stretch of guanosines (Gs), which, once transcribed, result in a poly-cytosine (C) tract at the RNA level. These stretches of Gs can potentially form a secondary structure known as a G-quadruplex (G4), which are tetrads of Gs stabilized by a monovalent cation, which undergo base-stacking (73,74). Certain salts have

been found to favor or disfavor the stabilization of G-quadruplexes (75). For instance, potassium cations (K^+) have shown a propensity to stabilize G-quadruplexes, while lithium cations (Li^+) destabilize them. In the literature, G-quadruplexes in DNA have been shown to regulate transcription (76). In addition, G-quadruplex helicases have been found to regulate transcription and other cellular processes (77,78). This makes it possible that one, or multiple G-quadruplex helicases may serve to regulate the length of BC200 produced by winding or unwinding the quadruplex that forms in its genomic sequence.

Another complex nucleic acid structure that can play a role in modulating transcription is that of the R-loop. R-loops are formed during transcription, occurring when newly transcribed RNA, referred to as the nascent strand, re-enters the transcription bubble (5). When this occurs, the non-template DNA is displaced by the RNA hybridizing with the template strand. The formation of R-loops is seen to occur behind the actively transcribing RNAP II, and is often found to inhibit or arrest transcription. This is likely due to changes in nucleic acid topology introducing energetic barriers, or destabilization of the nucleic acid-RNAP II complex that is traditionally formed during transcription. Interestingly, R-loops are commonly formed in regions including G-rich DNA. This is intriguing considering that it has been found that R-loop-associated G-quadruplexes can modulate transcriptional activity (79). Thus, it is possible that the production of BC120 is partly regulated at the transcriptional level due to secondary structure formation on the genomic sequence of BC200.

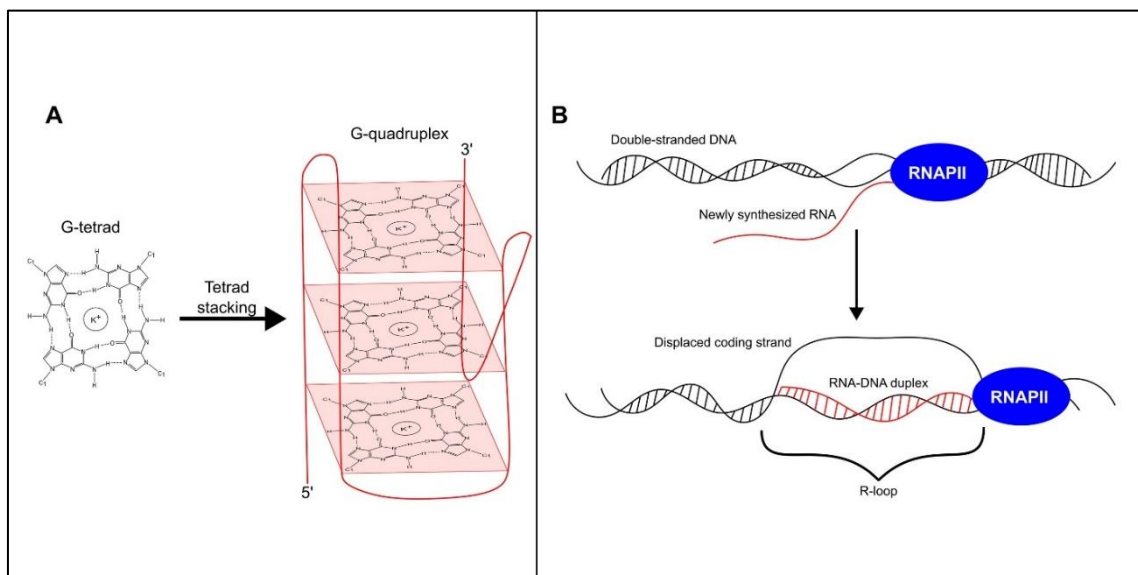


Figure 1.6. The presence of non-canonical secondary structures formed in DNA and/or RNA may modulate the production of BC120 transcripts from the genomic BC200 sequence. (A) G-quadruplexes form when G-tetrads, which are stabilized by the presence of monovalent cations (commonly potassium ions) stack (73,74). (B) R-loops are formed during transcription when the newly synthesized (nascent) RNA strand anneals with the template strand, displacing the coding strand of DNA (5). Conditions that slow transcription, such as GC-rich regions have been shown to favor R-loop formation.

In the previously mentioned study, it was postulated that BC120 possibly arises from BC200 processing events (3). This is further supported when examining some of the pertinent proteins that interact with BC200, the binding of which may protect BC200 from processing to BC120. The proteins in question being PCBP2 and HNRNPK, both of which, as mentioned prior, bind to the C-rich region towards the 3' end of BC200 (62). It is this binding at the 3' end which, hypothetically, may block exonucleases from acting on this region of the molecule. It is also highly probable that it is not just the poly-C tract of BC200 that allows for its transcription to be regulated.

Poly-A binding protein nuclear 1 (PABPN1) is a confirmed binding partner of BC200, binding within the A-rich stretch of the molecule (62). The binding of this protein to the A-rich region of BC200 may serve as yet another protein that shields the lncRNA from being post-transcriptionally processed into BC200. This idea is further supported when examining previous RNA-probing experiments done by our research group. In the paper in which the discovery of BC120 is detailed, Northern blotting experiments were done (3). Briefly, Northern blotting allows us to probe for the presence of specific RNAs via the use of probes specific to certain RNA sequences. In these experiments, not only was BC120 identified, but the presence of a slightly larger lncRNA intermediate as well. This RNA was named BC161, due to its approximate length of 161 nucleotides, confirmed via overexpression and further Northern blotting assays. It is the presence of this intermediate which makes it likely that poly-A binding protein binding, along with that of poly-C binding proteins serve to shield BC200 from post-transcriptional processing down to BC120.

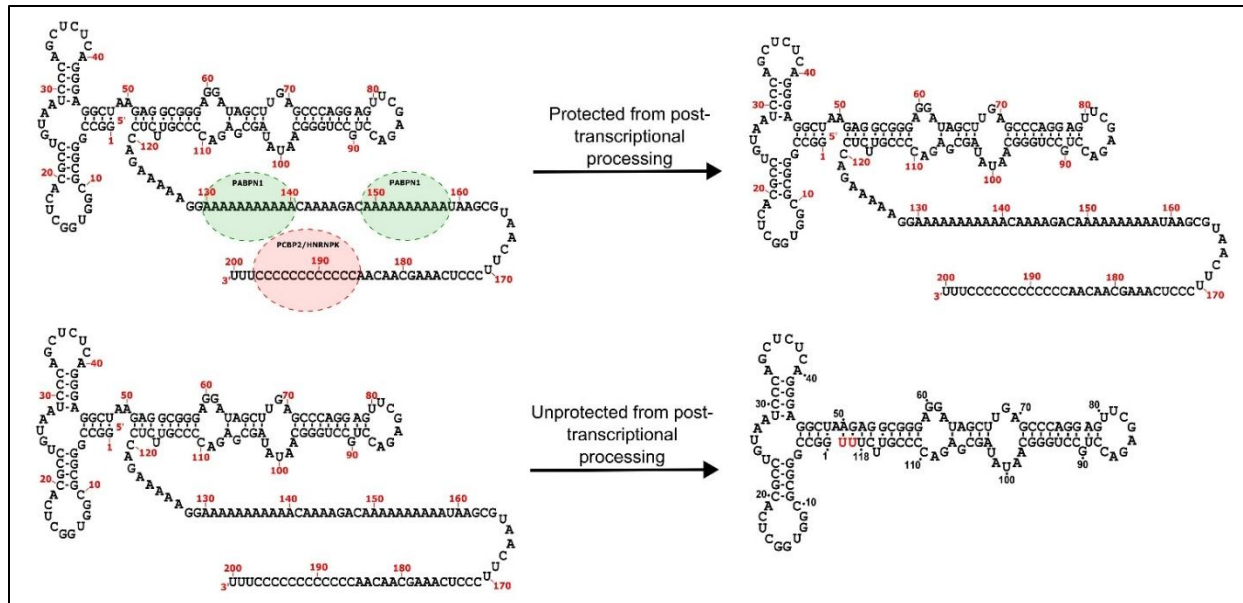


Figure 1.7. BC120 may arise from cellular post-transcriptional processing events of full-length BC200. Confirmed binding partners PCBP2, HNRNPK and PABPN1 all bind to different portions of the BC200 tail (62). It is hypothesized that when these proteins are not bound to the

tail of BC200, it is unprotected from processing to BC120. The binding behaviour of these proteins may play a significant role in modulating BC120 production.

1.4 Thesis Outline

Currently, the role of BC120 is understudied, possibly being ignored in prior studies investigating BC200 function. However, how the abundance of this molecule and its longer counterpart is regulated in the cell is also yet to be determined. In this thesis, I will describe the insights gleaned into the factors and processes that favor and disfavor BC120 production *in vitro* and *in cellulo*. Chapter 2 will include descriptions of the methods used to carry out my experiments. In Chapter 3, I will analyze the results of these experiments, including objective explanations of the conditions that favor the production of BC120. In the concluding chapter of this thesis, I will interpret these findings, and end with an elaboration on the future directions of this research.

Chapter 2: Materials and Methods

2.1 Reagents

Chemicals and reagents used over the duration of my work were acquired from Thermo Fisher Scientific (Ottawa, Canada) or New England Biolabs (Whitby, Canada). These standard solutions were all molecular biology grade. The specific preparation of buffers and other solutions are detailed later in the chapter. All buffer solutions were prepared using Milli-Q[®] - purified ultra pure double deionized water, purified at a resistance of 18.2 mega ohm-cm ($M\Omega \cdot cm$).

2.2 Cell culture

The human embryonic kidney (HEK)293T cell line was gifted by Dr. Dr. Thomas Klonisch, while the MCF-7 cell line was gifted from Dr. Spencer Gibson. Cell cultures were maintained at a temperature of 37°C and an atmospheric carbon dioxide (CO₂) gas concentration of 5% in a humidified incubator. Cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific) with a supplementation of 10% Fetal Bovine Serum (FBS) (Thermo Fisher Scientific).

2.2.1 Cell plating and passaging

In order to ensure that cells did not get overcrowded, die, and that ideal levels of RNA and proteins could be produced for downstream assays, cell plating densities were optimized based on the behaviour of the cell, as well as incubation time. The cell densities used for MCF-7 cells with different incubation times in units of cells/mL are as follows: 500×10^3 for 24 hours,

325×10^3 for 48 hours, and 225×10^3 for 72 hours. Cells were grown in either 15 cm plates or 6-well plates (Thermo Fisher Scientific), at plating volumes of 24 mL and 2 mL respectively.

When cells were solely being maintained, they were passaged two to three times a week. 0.05% trypsin-ethylenediaminetetraacetic acid (EDTA) (Thermo Fisher Scientific) was used to enzymatically disrupt the adherence of MCF-7 cells. An equal volume of DMEM (Thermo Fisher Scientific) supplemented with 10% FBS (Thermo Fisher Scientific) was used to neutralize the trypsin. Both HEK293T and MCF-7 cells were spun down in a swinging bucket rotor centrifuge (Thermo Fisher Scientific) at 5000 rpm for five minutes. Cells were resuspended in cell culture media before being plated.

2.2.2 Freezing down cells, storage and culture recovery

In order to store cells for prolonged periods of time, cells were harvested following the steps detailed in section 2.3.1 Cell plating and passaging. Following centrifugation and aspiration of the supernatant, the cells were resuspended in 3 mL of freezing media. This media is comprised of 80% FBS, 10% DMEM and 10% dimethyl sulfoxide (DMSO). The cell suspension was aliquoted 1 mL at a time in 1.5 mL cryovials (Thermo Fisher Scientific). These vials were stored overnight in a Bel-Art Cryo-Safe Freeze Controller (Thermo Fisher Scientific) filled with isopropanol, which was further placed in a -80 °C freezer. Following this overnight storage in the freezer, the cryovials were transferred to liquid nitrogen.

Frozen cell cultures were recovered via thawing in a 37 °C water bath. Cells were then quickly transferred to a 150 mm plate (Thermo Fisher Scientific) with 25 ml of cell culture media. The plate was gently rocked back and forth to distribute cells evenly on the plate. Cells were passaged and maintained as described in section 2.3.1 cell plating and passaging.

2.3 Molecular cloning

A central component of my research involved the construction of overexpression and *in vitro* transcription (IVT) vectors containing truncations of BC200. This process is referred to as molecular cloning. The process of generating these constructs, validating their sequences, as well as storing transformants containing these plasmids is detailed in the subsequent sections.

2.3.1 Plasmid construction and validation

The generation of overexpression constructs is based off the production of an insert sequence, digesting the vector and insert with identical restriction endonucleases, and ligating them together. This process begins with the purification of the plasmid backbone. For overexpression constructs, the commercial plasmid backbone pcDNA3.1 (+) (Thermo Fisher Scientific) was used, while the puC57 backbone (Thermo Fisher Scientific) was used for *in vitro* transcription constructs. The pcDNA 3.1 backbone contains the coding sequence for green fluorescent protein (GFP), which acts as a reporter of transfection efficiency. The production of inserts was done via polymerase chain reaction (PCR) off a plasmid containing the full-length BC200 sequence, with primers that both add a restriction endonuclease cut site to the produced fragment and modify the number of cytosines produced. The sequences for these primers can be found in Table 2.1.

Table 2.1. Primers used in the production of inserts used to make BC200 truncation overexpression vectors.

Primer¹	Sequence (5'-3')
BglII BC200 Forward	GTAGCAAGATCTGGCCGGGCGCGGT
MluI BC185 + C Reverse	TGCGTAACGCGTAAAAGTT
MluI BC185 + 2Cs Reverse	TGCGTAACGCGTAAAAGGTT
MluI BC185 + 3Cs Reverse	TGCGTAACGCGTAAAAGGGTT
MluI BC185 + 4Cs Reverse	TGCGTAACGCGTAAAAGGGGTT
MluI BC185 + 5Cs Reverse	TGCGTAACGCGTAAAAGGGGGTT
MluI BC185 + 6Cs Reverse	TGCGTAACGCGTAAAAGGGGGGTT
MluI BC185 + 7Cs Reverse	TGCGTAACGCGTAAAAGGGGGGGTT
MluI BC185 + 8Cs Reverse	TGCGTAACGCGTAAAAGGGGGGGGTT
MluI BC185 + 9Cs Reverse	TGCGTAACGCGTAAAAGGGGGGGGGTT
MluI BC185 + 10Cs Reverse	TGCGTAACGCGTAAAAGGGGGGGGGGTT
MluI BC185 + 11Cs Reverse	TGCGTAACGCGTAAAAGGGGGGGGGGGTT

¹Prior to beginning this project, overexpression plasmids containing the sequences for full-length BC200 (12Cs), along with 6Cs and 11Cs were already produced.

Following the production of these inserts, both the plasmid and the fragment were cut with the restriction endonucleases BglII and MluI (Thermo Fisher Scientific). Digestion reactions were prepared according to manufacturer instructions. Following digestion, samples were run on a 2% agarose gel (Thermo Fisher Scientific) prepared in Tris-acetate-EDTA (TAE) buffer at 100 volts for 1 hour. Gels were stained with SYBR[®]-Safe in-gel DNA stain (Thermo Fisher Scientific) and imaged using the FluorChem[®] Q imaging system from Alpha Innotech

(now ProteinSimple) (San Jose, California). Fragments containing linearized plasmid or desired insert were excised from the gel and purified using the GeneJET gel extraction kit (Thermo Fisher Scientific). Digested inserts and vectors were ligated with T4 DNA ligase from New England Biosciences according to manufacturer instructions.

2.3.2 Plasmid transformation, production and storage of glycerol stocks

To generate large quantities of these constructs, transformation into a competent cell line was required. NEB® 5-alpha Competent *E. coli*, a DH5 α -derived strain (New England Biolabs) was transformed using the 5 minute transformation protocol provided by the manufacturer, before spread plating onto Luria-Bertani (LB) agar (Thermo Fisher Scientific) supplemented with 100 μ g/ml ampicillin for selection of transformants. Plates were incubated in a 37 °C incubator and left to incubate overnight. The following day, single colonies were picked from the plate and used to inoculate multiple 5 mL cultures of LB broth (Thermo Fisher Scientific) supplemented with 100 μ g/ml ampicillin. Cultures were grown overnight in round bottom test tubes placed inside of a shaking incubator at a temperature of 37 °C and 200-250 rpm. Constructs were purified from these liquid cultures the following day using the GeneJet Plasmid Miniprep kit (Thermo Fisher Scientific). Plasmids were sent off for Sanger sequencing by Genewiz (Azenta Life Sciences) (Montreal, Quebec) in pre-mixed reactions. Fast-All (FASTA) files obtained from sequencing were input into SnapGene for comparison to the sequence that should have been obtained if the upstream cloning steps were successful.

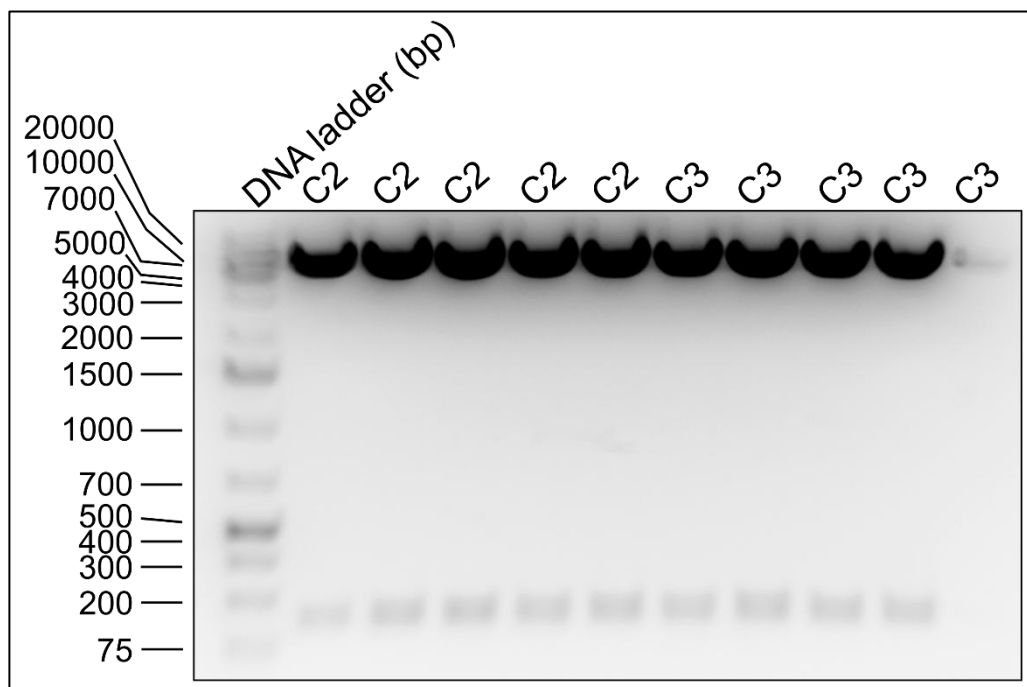


Figure 2.1. Re-digestion of BC200 truncation overexpression constructs confirms a drop-out of the expected size if cloning were successful. Each construct was purified on a small-scale via mini-prep, with each bacterial suspension grown from a distinct colony from a spread-plate. This was done just in case one of the colonies contained an incorrectly cloned plasmid. The 2Cs and 3Cs constructs were run on this gel serve as an example of what was repeated after each cloning experiment. These re-digestions were ran on a 2% agarose gel stained with SYBR[®] Safe (Thermo Fisher Scientific) and detected using the FluorChem[®] Q imaging system (ProteinSimple) using the Cy2 excitation and emission filters.

Once the transformations were confirmed to be correct, a large-scale overnight culture was inoculated using the small-scale culture used for the miniprep. Growth conditions for the large-scale culture were as previously described, aside from growth in a bacterial cell culture flask. Larger amounts of plasmid were purified using the GeneJet Plasmid Maxiprep kit (Thermo Fisher Scientific). For final confirmation of successful cloning, sequencing was carried out as previously described. Upon final confirmation that the cloning was successful, glycerol stocks for the successful transformants were created. Glycerol stocks were prepared by combining 500 μ L of 50% glycerol (Thermo Fisher Scientific) with 500 μ L of the liquid bacterial culture in a 2

mL cryovial (Sigma Aldrich). Prepared glycerol stocks were stored long-term in a -80 °C freezer. Cultures were recovered from glycerol stocks by scraping off a bit of the stock with a micropipette tip and injecting it into LB broth supplemented with the appropriate antibiotic.

2.4. Assessing the relative levels of BC200 and BC120 RNAs in cells.

2.4.1 Cell transfection, collection and RNA purification

MCF-7 cells were collected, spun down and diluted to the seeding density as detailed in section 2.2.1 cell plating and passaging. The transfection mixture includes the following components added in the order written: 4 µg of plasmid, 6 µL of Polyethylenimine (PEI) Max 40000 daltons (Da) from Polysciences Inc. (Warrington, Pennsylvania) diluted to a concentration of 1 mg/ml and 400 µL of Opti-MEM™ (minimal essential medium) reduced serum medium (Thermo Fisher Scientific) in a microcentrifuge tube (Thermo Fisher Scientific). This mixture was mixed via vortexing before being left to incubate at room temperature in the biosafety cabinet for at least 15 minutes before use. Following incubation, the transfection mixture was added to the center of each well in a 6-well plate (Thermo Fisher Scientific), before 2 mL of cell suspension was added on top of the mixture. Plates were gently rocked back and forth to ensure that cells were distributed evenly over the surface of the well before incubation as detailed in section 2.2.1 cell plating and passaging.

24 hours later, the cell culture media was aspirated and fresh cell culture media was added. Following another 24 hours of incubation, cells were observed using the EVOS FL Auto Imaging System microscope on the GFP fluorescence channel to check for successful transfection. The cell culture media in each well of the plate was aspirated, before each well was rinsed with trypsin to remove any residual media. After rinsing with trypsin, 750 µL of trypsin

was added to each well of the plate before incubation as detailed in section 2.2.1. An equal amount of cell culture media was added to each well to neutralize trypsinization and rinse the cells off the bottom of the plate. The suspension was transferred into a microcentrifuge tube. Cells were spun down in a tabletop microcentrifuge at 500 x g for 5 minutes. Total cellular RNA was extracted using the Monarch Spin RNA Isolation Kit (Mini) (New England Biolabs) according to manufacturer instructions. RNA concentration and purity were measured via NanoDrop 200C Spectrophotometer (Thermo Fisher Scientific).

2.4.2. Quantification of transfection efficiency

While observation of GFP via fluorescence microscope allowed us to assess that transfection occurred, it only gave a qualitative assessment. As such, transfection efficiency was further assessed via reverse-transcriptase/real-time quantitative PCR (RT-qPCR). Transfection efficiency was assessed to ensure that any changes seen in BC200 and BC120 abundance in experimental conditions was not due to inconsistencies in plasmid transfection for each condition. The TaqMan probing system was used, consisting of forward and reverse primers, along with a short oligo containing a 5' end reporter, and a 3' end quencher. The sequence of each component is as follows: GFP forward primer, GTCCGCCCTGAGCAAAG; GFP reverse primer, TTTTTTTTTTTTTTTTTTTTTTTTTTTT oligo deoxythymidine₂₄ (dT₂₄); GFP probe, AGGACCATGTGATCGCGCTTCTC. RT-qPCR was performed using an Applied Biosystems StepOnePlus instrument (Thermo Fisher Scientific). The cycle threshold (Ct) values from the transfected conditions were compared to a non-transfected condition using the delta-delta Ct ($\Delta\Delta Ct$) method to determine the fold-change in GFP mRNA abundance.

2.4.3. Northern Blotting

To probe for the relative abundance of BC200 and BC120 in different experimental conditions, a Northern Blotting methodology was employed. The RNA collected according to section 2.4.1 cell transfection and collection was normalized to the lowest concentration obtained for a particular biological replicate of the experiment. An equal volume of 2X denaturing RNA load dye was mixed with the RNA, before heating each sample to 95 °C for five minutes using an Eppendorf ThermoMixer C (Thermo Fisher Scientific). The load dye contained 95% formamide, 0.025% sodium dodecyl sulfate (SDS), 0.025% bromophenol blue, 0.025% xylene cyanol and 0.5 mM EDTA. Samples were separated by size via urea-denaturing acrylamide gel electrophoresis using an 8% gel. 1.5 mm thick gels were cast using the BioRad Mini-PROTEAN® Tetra handcast system (Mississauga, Ontario). Each gel was pre-run in 1X Tris-Borate-EDTA (TBE) buffer at 13 W for one hour. Samples were run at the same wattage for 30 minutes.

The RNA was transferred to a positively-charged BrightStar-Plus nylon membrane (Thermo Fisher Scientific) using the Bio-Rad Mini Trans-Blot® tank transfer system in 0.5X TBE buffer. Buffers were run at 100V for one hour in a large ice-bath with an ice pack placed in the transfer system to prevent excess heat generation. The RNA was crosslinked to the membrane using the Spectro-linker XL-1000 UV crosslinker (Thermo Fisher Scientific) at 240 mJ/cm². A pre-hybridization step was performed on the membrane using a solution containing the UltrHyb Oligo Hybridization Buffer (Thermo Fisher Scientific) along with 2X blocking reagent (Roche Life Science). This step was performed at a temperature of 60 °C for 30 minutes with shaking. An overnight locked nucleic acid (LNA) probe hybridization step was performed with shaking at 54 °C with the probes at a concentration of 3 nM. Probes were digoxigenin (DIG)-labeled at both

the 5' and 3' ends. Probes targeting nucleotides 33-52 and 86-107 were tested when optimizing this experiment. The probe sequences are as follows: BC200 33-52, TCT+TAGCCTCCC+T+G+A+G +AGCT; BC200 86-107, TC+GC+TAT+ATTG+CCC+AGGC+AGGT. The LNA probes were custom synthesized and obtained from Qiagen (synthesized in Västerås, Västmanland). The “+” symbol to the left of a nucleotide indicates that it is an LNA.

The following day, the blots were washed with 2x, 0.5x and 0.1x saline sodium citrate (SSC) buffer (Thermo Fisher Scientific) with 0.1% SDS (Thermo Fisher Scientific) twice each for 20 minutes at a time at 60 °C. The blots were blocked with a solution containing 100 mM maleic acid, 150 mM sodium chloride (NaCl) pH 7.5 plus 2x blocking reagent for 30 minutes with shaking at room temperature. A following blocking step was performed under the same exact conditions with EveryBlot blocking buffer (Bio-Rad). The primary antibody (anti-DIG) (Jackson ImmunoResearch) was added to the blocking buffer at a 1:2000 dilution, before being left to incubate overnight with shaking at 4 °C. The next day, blots were washed three times with tris-buffered saline plus 0.1% Tween 20 (TBS-T). The blot was then incubated in EveryBlot blocking buffer (Bio-Rad) containing a secondary antibody (goat, anti-mouse) conjugated with horseradish peroxidase (HRP) (Thermo Fisher Scientific) diluted 1:10,000. The blots were incubated for two hours with shaking at room temperature before three five-minute washes with TBS-T. The blots were visualized by chemiluminescence using the Immobilon Forte HRP substrate (Thermo Fisher Scientific). Blots were imaged using the FluorChem[®] Q imaging system (ProteinSimple) with the chemi emission channel.

2.4.4. BC120-discriminating RT-qPCR

Northern Blots are a semi-quantitative method, meaning that supplementing our work with a quantitative method was another way to observe changes in BC200 and BC120 abundance. Because BC120 shares an identical resemblance to the first 120 nucleotides of BC200 at the sequence level, a method of discriminating BC120 from BC200 was developed via RT-qPCR (3). Briefly, cellular RNA was purified following expression plasmid transfection as described in section 2.4.1 Cell transfection, collection and RNA purification. The RNA was then ligated with a 5' phosphorylated linker at the 3' end. The sequence of the linker is as follows: ACTGTAGGCACCATCAAT. Ligations were prepared at a total volume of 50 uL and consisted of the following: 1x T4 RNA Ligase buffer (New England Biolabs), 1mM adenosine triphosphate (ATP), 1 μ M linker, 100 units of T4 RNA ligase, 10% DMSO, 10% polyethylene glycol (PEG) 8000, 40 units of RiboLock ribonuclease (RNase) inhibitor and 1.25 μ g of RNA. The ligation reactions were cleaned using the Monarch RNA Cleanup Kit (10 μ g) (New England Biolabs). RT-qPCR to detect linker-ligated BC120 was done using the following primers and probes: BC120 forward, AGCTCTCAGGGAGGCTAAGA; linker reverse complement, ATTGATGGTGCCTACAGT; BC120 probe, ATATTGCCCAGGCAGGTCTCGAAC. To the RT-qPCR reaction, an LNA probe labeled with DIG at the 5' and 3' ends of the oligo was added to a concentration of 250 nM. The sequence of this oligonucleotide (oligo) is complimentary to nucleotides 157-177 of BC200, inhibiting its amplification (Exiqon design ID 535069-2, Qiagen).

For quantifying BC200 linker-ligated RNA, a standard RT-qPCR run was set up using the following primers and probes: BC200 forward (nucleotides 158-180), ATAAGCGTAACTTCCCTCAAAGC, and the linker reverse complement. RT-qPCR for GFP in each sample was also conducted for normalization purposes using the methods outlined in section 2.4.2. Quantification of transfection efficiency. Both the BC200, BC120 and GFP Ct values obtained for each transfection condition were normalized to the expression of each RNA in an empty vector condition. After normalizing for GFP, the difference between the levels of BC200 detected, as well as the levels of BC120 detected after the discriminating RT-qPCR assay were assessed.

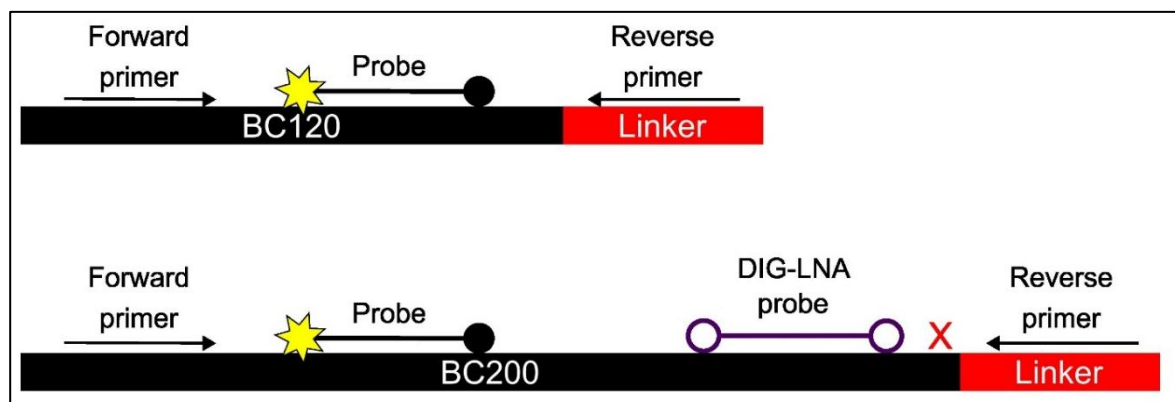


Figure 2.2. Schematic illustrating the principle behind the BC120-discriminating RT-qPCR assay. Pictured is cellular RNA that has been ligated with a linker sequence that the reverse primer can bind to, and a reverse primer that binds to the same region of BC200 and BC120. A DIG-labelled LNA probe binds to the tail region of BC200. When RT-qPCR is performed, the probe inhibits elongation and reducing the levels of BC200 detected by the assay. This data can be compared to RT-qPCR data in the absence of the probe to determine relative BC200 and BC120 levels in the sample

2.5. The production and purification of a Ribonuclease Inhibitor.

RNA is very susceptible to degradation by ribonuclease (RNase) enzymes in solution. As such, for all downstream experiments involving the use and manipulation of RNA, the addition of an RNase inhibitor was essential. We produce and purify this RNase inhibitor in house.

2.5.1 RNH1 production and purification

Prior to producing RNH1, all of the buffers for the protein purification were produced and filter sterilized using a bottle-top filter (Thermo Fisher Scientific). Buffers were prepared according to the table below (Table 2.2). Production then begins by transforming chemically competent BL21 cells (New England Biolabs) with a protein expression plasmid containing the sequence for RNH1. Transformation was performed according to manufacturer instructions, and transformants were selected for on LB agar plates with 50 µg/mL kanamycin. Plates were incubated overnight at 37 °C. A colony was used to inoculate a 10 mL overnight liquid culture in LB (Thermo Fisher Scientific) + 50 µg/mL kanamycin, grown overnight with shaking at 37 °C. Half of the liquid culture was used the following day to inoculate a 1L LB + kanamycin broth culture. Once the culture reached an optical density at 600 nm (OD₆₀₀) between 0.4-0.6, high-yield protein expression was induced using a 1M stock of Isopropyl-β-D-thiogalactopyranoside (IPTG) to make a final concentration in the broth of 1 mM. The liquid culture was grown overnight in a shaking incubator at 18 °C.

Table 2.2. The composition and concentrations of each component of the buffers used to purify RNH1.

Buffer component	Concentration			
	Lysis	Wash	Elution	SEC
HEPES (pH 7.6)^a	20 mM	20 mM	20 mM	20 mM
NaCl	300 mM	1 M	300 mM	100 mM
DTT^b	1 mM	1 mM	1 mM	1 mM
Triton X-100	0.03%	0.03%	0.03%	0.03%
Glycerol	5%	5%	5%	5%
Imidazole	10 mM	25 mM	500 mM	-
PMSF^c	1 mM	-	-	-

^a 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)

^b Dithiothreitol (DTT)

^c Phenylmethylsulfonyl fluoride (PMSF)

The following morning, the BL21 culture was spun down at 5000 x g for 10 minutes before being resuspended in 30 mL of lysis buffer per 1 L of liquid culture. The cell suspension was sonicated using a 150 Sonic Dismembrator Sonicator (Thermo Fisher Scientific) for 5 minutes at 50% amplification for 10 seconds at a time, with 30 seconds in between each pulse. Following sonication, the culture was spun down at 33,000 x g for 30 minutes. After centrifugation, the supernatant, which contains protein, was transferred to a syringe equipped with a 0.45 μ M polyvinylidene difluoride (PVDF) filter and filtered into a 50 mL Falcon tube.

The ÄKTA Go (Cytiva) system was used to purify our protein via affinity chromatography, along with the UNICORN software to define run parameters. All lines in use

were rinsed with filtered Milli-Q[®] water before priming. A HisTrap HP 1 mL nickel column was hooked up to the ÄKTA GO and cleaned with 5 column volumes (CVs) of filtered Milli-Q[®] water. All of the lines that were to be in use for the protein purification were primed with the appropriate buffers: lysis buffer for the sample line, wash buffer for line A and elution buffer for line B. Following priming, the nickel column was equilibrated with 5 CV of wash buffer. The sample was loaded onto the column and rinsed with 20 CV of wash buffer, followed by elution with x mL of elution buffer. 4.5 mL fractions were collected using the ÄKTA F9-R fraction-collection system (Cytiva). Protein containing fractions were assessed via Bradford assay in a 96-well plate based on the peak observed via the chromatography trace, and 15% sodium-dodecyl sulfate polyacrylamide gel (SDS-PAGE) electrophoresis was run at 100V for 1 hour to confirm RNH1-containing fractions. The SDS-PAGE was stained with a Coomassie Brilliant Blue solution containing 50% methanol, 10% glacial acetic acid, 40% water, 0.1% Coomassie Brilliant Blue R-250 with rocking for 30 minutes. Destaining was done using a solution composed of the following: 50% methanol, 10% glacial acetic acid, 40% water with rocking for ~1 hour, or until the gel could be visualized with the naked eye. Following staining, gels were scanned.

RNH1-containing fractions were pooled together for size-exclusion chromatography (SEC). A HiLoad Superdex 75 c26/600 column connected to the ÄKTA Purifier 10 system (Cytiva) was equilibrated with 1 CV of SEC buffer. The RNH1-containing pooled fraction was manually loaded into a 10 mL sample loop and injected onto the column. The column was then rinsed with 1 CV of SEC buffer at a flow rate of 1.5 mL/min, and 5 mL fractions were collected using the Frac-950 fraction collection system (Cytiva). RNH1 containing fractions were collected based on the chromatography trace, and once again, a Bradford assay and SDS-PAGE

was run to confirm the presence of RNH1. These fractions were pooled together and concentrated using a Millipore Amicon® Ultra (Thermo Fisher Scientific) spin column protein concentrator that was equilibrated with elution buffer. DTT concentrations in the buffer were brought up to 8 mM and an equal volume of glycerol was added. The protein was stored at -20 °C.

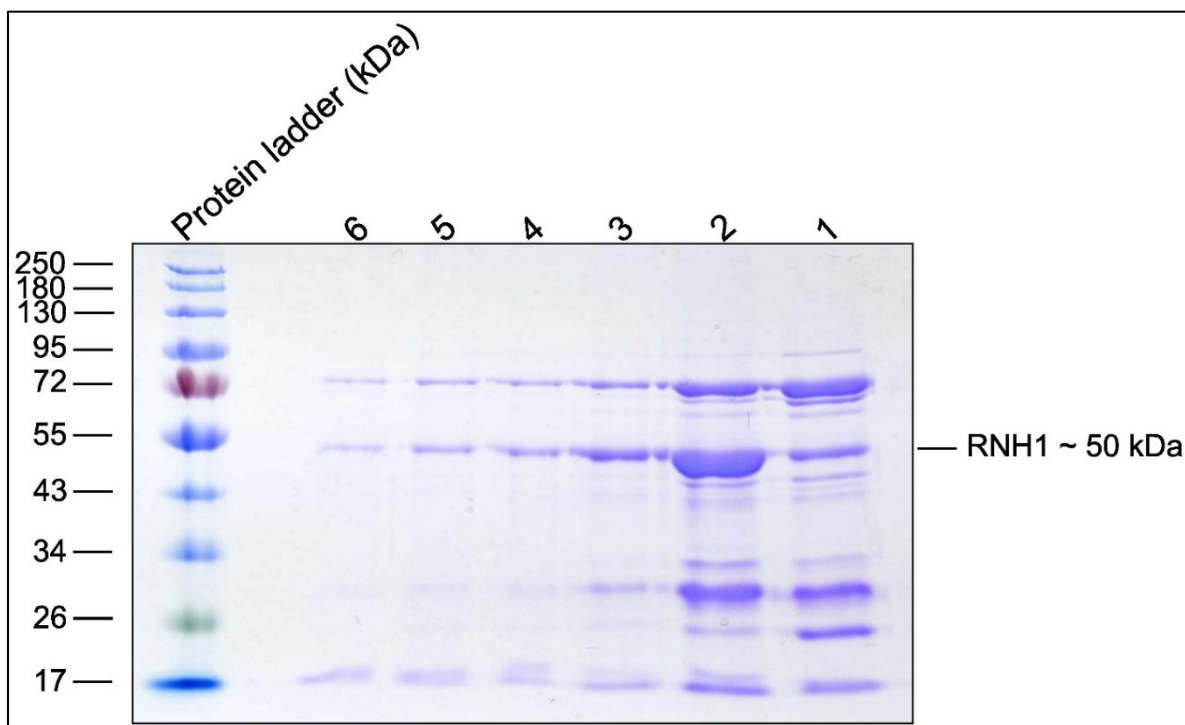


Figure 2.3. Size-exclusion chromatography is a highly effective secondary step to achieve highly pure RNH1. Following affinity chromatography, SEC is performed to desalt the protein. Fractions were collected and ran on a 15% SDS-PAGE at 100V for one hour to determine which fractions contained the highest protein yield. The numbers on the gel (1-6) indicate the six fractions within the highest point of the chromatography trace peak and as such thought to contain the most RNH1. Those fractions were pooled together and concentrated. Glycerol was added to the protein at an equal volume before it was stored and tested.

2.5.2 Assessing purified RNH1 efficacy

Following purification, the efficacy of RNH1 was tested. A native TBE-polyacrylamide gel was used to run the following experimental reactions: RNA with no RNase, RNA with RNase at 1 ng/ μ L + commercial RNase inhibitor (New England Biolabs), RNA + RNase at 1 ng/ μ L + increasing volumes of RNH1. The gel was run for 30 minutes at 13W and stained with 0.1% toluidine blue staining solution (1g of toluidine blue dye, 10 mL glacial acetic acid, topped off to 1 L with Milli-Q[®] water). The optimal amount of RNH1 used per reaction was determined to be the lowest volume that provides adequate protection against RNA degradation from RNases. For all downstream experiments, that volume of RNH1 was added when required in experiments.

2.6. Assessing the protein binding partners of BC200 truncations

As discussed in the introduction to this thesis, we hypothesized that the production of BC120 may be modulated by proteins interacting with the poly-C tract of BC200, protecting it from degradation. We also hypothesized that these interactions are dependent on the number of cytosines in this tract. Knowing some of the putative and confirmed BC200 binding partners (62–64), we wanted to determine the minimal C-tract length required for poly-C binding proteins to interact with BC200.

2.6.1 *In vitro* transcription plasmid preparation

The construction of IVT plasmids was conducted similarly as the method detailed in section 2.3.1 Plasmid construction and validation. Following validation, plasmids were digested according to the enzyme at their linearization site, being BbsI or DraI. According to

manufacturers instructions (Thermo Fisher, New England Biolabs), it was assumed that 1 μ L of enzyme cut 1 μ g of DNA every 5 minutes for 2 hours. As such, reactions were set up with appropriate volumes of 10x reaction buffer, water plasmid and restriction enzyme before being left to incubate at 37 °C for two hours to overnight. Following enzymatic digestion, plasmids were purified via phenol extraction and ethanol precipitation. Briefly, restriction digest reactions were transferred to microcentrifuge tubes containing vacuum grease and mixed with an equal volume of phenol from phenol:chloroform:isoamyl alcohol saturated in TRIS buffer (Thermo Fisher Scientific). After vigorously mixing, the tubes were spun in a Legend Micro 17 tabletop centrifuge (Thermo Fisher Scientific) at 3000 x g for five minutes. The protein separates into the organic phase which is denser and thus sinks to the bottom of the tube, while the DNA stays in the aqueous phase, which is less dense. The two phases are separated by the vacuum grease in the microcentrifuge tube.

Table 2.3. The primers used to generate the inserts required to make IVT plasmids of different BC200 poly-C tract truncations.

Primer	Sequence
T7 BC200 Forward	CCCGGGATCCGATTAATACGACTCACTATAGG
T7 BC185 + 1C Reverse	AATGCATCTAGAGAAGACTTGTTGTTGCTTTGAGGGAAGTTA
T7 BC185 + 2Cs Reverse	AATGCATCTAGAGAAGACTTGTTGTTGCTTTGAGGGAAGTTA
T7 BC185 + 3Cs Reverse	AATGCATCTAGAGAAGACTTGGGTTGTTGCTTTGAGGGAAGTTA

T7 BC185 + 4Cs Reverse	AATGCATCTAGAGAAGACTTGGGGTTGTTGCTTTGAGGGAAGTTA
T7 BC185 + 5Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGTTGTTGCTTTGAGGGAAGTTA
T7 BC185 + 6Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGGTTGTTGCTTTGAGGGAAGTTA
T7 BC185 + 7Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGGGTTGTTGCTTTGAGGGAAGTT A
T7 BC185 + 8Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGGGGTTGTTGCTTTGAGGGAAG TTA
T7 BC185 + 9Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGGGGGTTGTTGCTTTGAGGGAA GTТА
T7 BC185 + 10Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGGGGGGTTGTTGCTTTGAGGGA AGTTA
T7 BC185 + 11 Cs Reverse	AATGCATCTAGAGAAGACTTGGGGGGGGGGGTTGTTGCTTTGAGGG AAGTTA

After pipetting off the aqueous phase into a fresh tube, 3 M sodium acetate at a pH of 5.2 was added to the tube containing the DNA, at a volume 1/10th of that of the DNA in the tube. After pipetting to mix, cold 95% ethanol was added in a 3-fold excess to the tube and mixed via vortexing. Tubes were spun on a Legend Micro 21 R tabletop centrifuge (Thermo Fisher Scientific) at maximum speed (21,100 x g) at 4 °C for 15 minutes to pellet the precipitated plasmid. After the supernatant was removed, three reaction volumes worth of 70% ethanol was used to wash the pellet, before another spin under the same conditions for 5-10 minutes to re-

pellet the plasmid. The DNA was resuspended in molecular biology-grade water (Thermo Fisher Scientific) and its concentration was determined using a Nanodrop 2000C spectrophotometer (Thermo Fisher Scientific). The plasmid was diluted to a concentration of 500 ng/ μ L.

2.6.2 T7 RNA Polymerase production and purification

In order to transcribe BC200 truncations, the purification of a highly efficient DNA-dependent RNA polymerase was required. T7 polymerase was chosen due to its high efficiency (80). T7 protein purification buffers were prepared according to the table below (Table 2.4). A 200 mL overnight culture was inoculated by scraping off the top of a premade glycerol stock of BL21s (New England Biolabs) in LB broth with 100 μ g/ml and grown using the same conditions outlined in sections 2.3.1 and 2.3.2. This culture was grown overnight and used to inoculate a large 2 L culture. The liquid culture was grown to an optical density at 600 nm (OD_{600}) of 0.6 before protein production was induced by adding isopropyl- β -D-thiogalactopyranoside (IPTG) to the culture to a final concentration of 1 mM. The culture was grown for another 3 hours before being harvested by centrifugation at 5000 xg for 10 min using a Sorvall Lynx 6000 centrifuge (Thermo Fisher Scientific).

Table 2.4. The composition of each buffer used in the purification of T7 RNA polymerase.

	Concentration			
Component	Lysis buffer	Wash buffer	Elution buffer	SEC buffer
Tris-HCl (pH 8)	50 mM	50 mM	50 mM	
HEPES (pH 7.6)	-	-	-	20 mM
NaCl	100 mM	300 mM	100 mM	100 mM

β-mercaptoethanol	5 mM	5 mM	20 mM	-
Glycerol	5%	5%	20%	5%
Imidazole	1 mM	25 mM	200 mM	-
PMSF^a	1 mM	1 mM	1 mM	-
EDTA	-	-	1 mM	-
Triton X-100	-	-	0.1%	0.03%
DTT	-	-	-	1 mM

Following centrifugation, the supernatant was discarded, and the bacteria was resuspended using 20 mL of lysis buffer per litre of culture. Lysate was transferred to a metal sonication cup, and the lysate was sonicated for 3 minutes and 20 seconds total, in 20 second intervals with 30 second breaks in-between at a 50% amplitude using a 150 Sonic Dismembrator Sonicator (Thermo Fisher Scientific). The lysate was then spun down at 14500 x g for 30 minutes at 4 °C. The supernatant was filtered using a 0.45 µm PVDF filter before affinity purification using a HisTrap HP 1 mL column (Cytiva) and the ÄKTA Go (Cytiva) fast protein liquid chromatography (FPLC) system. After sample loading, the column was rinsed with 20 column volumes (CV) of wash buffer, and then eluted for 30 CV using a linear gradient elution from the wash buffer to a 100% concentration of the elution buffer. Eluant was collected in 3 mL fractions. Both a Bradford assay and 15% SDS-PAGE electrophoresis were run to confirm T7-containing fractions as described in section 2.5.1 RNH1 production and purification. The SDS-PAGE was stained and destained as previously described.

Target protein-containing fractions were pooled together and further purified using size exclusion chromatography using the ÄKTA Purifier 10 system (Cytiva). Prior to SEC, a HiLoad Superdex 75 26/600 column (Cytiva) was equilibrated with 1 CV of SEC buffer, and the sample loop was washed with filtered Milli-Q water and primed with buffer before injecting the sample. SEC was run using 1 CV of SEC buffer, and 5 mL fractions were collected. Once again, target protein-containing fractions were confirmed via Bradford assay and SDS-PAGE. Fractions containing the target protein were pooled together and concentrated and stored as described in sections 2.5.1 and 2.5.2.

2.6.3 T7 polymerase testing and transcription optimization

Following the purification of T7 RNA polymerase, the optimal conditions for the transcription of a given plasmid, along with the efficacy of the purified T7 were tested. Before this, 10X *in vitro* transcription buffer containing 400 mM Tris-HCl (pH 8.1 at 37 °C), 10 mM spermidine (Thermo Fisher Scientific), 0.01% Triton X-100 (Thermo Fisher Scientific), 100 mM DTT in Milli-Q water. A stock mix of nucleotide triphosphates (NTPs) were generated by diluting each to a stock concentration of 100 mM, adjusting the pH to 7.0 and combining all four and diluting to a final concentration of 50 mM (12.5 mM of each NTP), and storing it at -20 °C. According to the protocol detailed by Booy *et. al* (81), a small-scale, 50 µL reaction is set up with the following components: 5 µL 10x transcription buffer, 5 µL of linearized template at a concentration of 500 ng/µL, 8 µL 50 mM NTP (8 mM final concentration), 4-20 mM magnesium chloride (MgCl₂) from a 100 mM working stock, and 0.5-2 µL of T7 RNA polymerase, brought up to 50 µL with molecular biology grade water (Thermo Fisher Scientific).

To test T7 RNA polymerase efficacy, an *in vitro* transcription reaction is set up for a plasmid that has its ideal transcription conditions known; i.e, what the ideal concentration of MgCl_2 is for transcription off of that plasmid. Reactions were set up from 0.5-2 μL of T7 in 0.5 μL increments. Reactions were incubated for 1 hour at 37 °C in an Eppendorf ThermoMixer C (Thermo Fisher Scientific), before denaturing urea-TBE gel electrophoresis was conducted as described in section 2.4.3. Following electrophoresis, gels were stained with rocking for 5-10 minutes with 0.1% toluidine blue staining solution (1g of toluidine blue dye, 10 mL glacial acetic acid, topped off to 1 L with Milli-Q® water). Gels were destained with Milli-Q® water and imaged to check for which volume of T7 offered the best transcription.

To determine the optimum MgCl_2 concentration for the transcription of each plasmid, transcription reactions were set up as described prior, with equal volumes of T7 in each reaction, with MgCl_2 concentrations adjusted from 4-20 mM. Samples were prepared, run, and the gel was stained as described. The optimal MgCl_2 concentration was the one in which, for a given volume of T7, produced the most RNA off of a given plasmid. For further downstream experiments. MgCl_2 concentrations, as well as the volume of T7 used in a 50 μL reaction were kept the same for consistency. This allowed other experimental conditions to be manipulated, knowing that these two factors were not confounding results.

2.6.4. Large-scale transcriptions and RNA purifications

To purify large amounts of RNA, transcription conditions, determined in 2.5.2 could be upscaled to a volume of 10 mL. Any pyrophosphate (Ppi) precipitate that was generated during the transcription was pelleted via centrifugation at 3000 x g for 5 min and the supernatant transferred to a clean tube. To quench the reaction, EDTA was added to the reaction to a final

concentration of 50 mM. A phenol extraction, similar to as described in section 2.5.1 without an ethanol precipitation was performed to remove the T7 RNA polymerase enzyme. Tubes were spun for 10 minutes at 3000 x g.

To desalt the freshly transcribed RNA, an RNA buffer consisting of 10 mM sodium phosphate pH 6.6 and 100 mM NaCl was prepared and used to equilibrate an Econo-Pac® 10-DG desalting column (BioRad). Samples were added to the column starting with a 3mL volume, which was allowed to flow through completely, before another mL of sample was applied to the column. 5 mL of RNA buffer was applied to the column to elute the RNA. For larger samples, 25 mL of buffer was applied to the column and allowed to flow-through before this process was repeated.

RNA was separated from its template via size-exclusion chromatography using a HiLoad Superdex 200 PR 26/60 column (Cytiva) connected to the ÄKTA Purifier 10 system (Cytiva). Both the RNA sample and the RNA buffer were filtered using a 0.22 µm filter. The SEC column was equilibrated with 1 mL of RNA buffer, before the sample was injected into the FPLC. RNA was eluted by running RNA buffer on the column, while fractions were collected using the Frac-950 fraction collector (Cytiva). Denaturing urea-TBE acrylamide gel electrophoresis was used to determine RNA-containing fractions.

2.6.5 RNA-biotin ligations and anti-biotin magnetic bead preparation.

The truncated RNAs produced and purified as described in section 2.5.3 were subsequently labeled to be used in further pull-down assays. RNAs were labeled via biotinylation using the Pierce 3' end RNA biotinylation kit (Thermo Fisher scientific) according to the

specifications of the manufacturer, followed by purification using the Monarch Spin RNA Cleanup Kit (50 µg) from New England Biolabs.

2.6.6 Pull-down assays to assess truncation-protein binding partner interactions

MCF-7 cells were passaged and maintained before being collected as described in section 2.2.1. Cells were grown to about 80% confluency prior to immunoprecipitation. Cells were scraped in cold PBS (Thermo Fisher Scientific) and spun down at 500xg for 5 minutes. This was done twice to ensure that cells were rid of any residual transfection reagent and cell culture media. Cells were lysed via resuspension in radioimmunoprecipitation assay (RIPA) buffer composed of 50 mM Tris pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% nonident P-40 (NP-40), 0.5% sodium deoxycholate, 0.1% SDS along with the addition of Halt protease inhibitor cocktail (Thermo Fisher Scientific) and RNH1 (made in house). Cells were incubated at 4 °C for 10 minutes with mixing. Buffer composition was adjusted to the following conditions for IP buffer: 25 mM HEPES (pH 7.9), 5 mM KCl, 0.5 mM MgCl₂, and 0.25% v/v NP-40. Tubes were spun in a tabletop microcentrifuge at 5000 rpm and 4 °C for 5 minutes. The Bradford assay was used to determine protein concentrations the lysates before normalizing them to a concentration of 5 mg/mL. Following normalization, biotin-labeled RNA was added to 500 µL of lysate to a final concentration of 250 nM and incubated at 4 °C for one hour with end over end mixing. Following incubation, 50 µL of streptavidin-agarose beads were added to the lysate, which was left to incubate with end-over-end mixing at 4 °C for 30 minutes. Beads were spun down for 1 minute at 3000 rpm and followed by three washes in RIPA buffer, and then three washes in PBS.

2.6.7 Western blotting against BC200 truncation binding partners.

Biotin-streptavidin pull-down assays were conducted as described in section 2.6.6. After being washed, 50 μ L of 2X SDS loading dye (0.31 M Tris-HCl (pH 6.8), 50% glycerol, 10% SDS, 0.05% bromophenol blue) was added to resuspend the beads before boiling them for 5 minutes at 95 °C. The resuspension was run on SDS-PAGE as described in Section 2.5.1.

Proteins were transferred to PVDF membranes (Thermo Fisher Scientific) using the TransBlot Turbo transfer system (BioRad). Following transfer to the membrane, the blot was blocked with 5% skimmed milk powder in TBS-T for 30 minutes at room temperature. Primary antibody targeting PCBP2 (Rabbit, Abcam) was diluted in 5% milk in TBS-T and allowed to incubate overnight with rocking at 4 °C. After the overnight incubation, blots were washed three times in TBS-T, before incubation with the secondary antibody conjugated with HRP for 1 hour in 5% skimmed milk in TBS-T with rocking at room temperature. The blot was imaged using Immobilon Forte HRP substrate (Thermo Fisher Scientific) and the FluorChem Q[®] imaging system (ProteinSmart) with the chemi emission channel.

2.7. Assessing how transcription conditions impact the production of BC120.

Previously, members of my research group have produced BC200 by *in vitro* transcription for a myriad of different experimental procedures. Upon running small-scale trial transcriptions and running them using denaturing urea-TBE polyacrylamide gel electrophoresis, the presence of a smaller band, correlating with the size of the newly characterized BC120 RNA, was observed upon staining (3). With this observation in mind, a component of my research involved investigating the experimental conditions which can influence the presence, or absence of BC120 upon *in vitro* transcription of BC200.

2.7.1 Temperature-dependent transcriptions

Temperature is believed to play a role in the outcomes of *in vitro* transcription, as increasing or lowering the temperature has been shown to have pronounced outcomes on the products produced. As such, standard *in vitro* transcription reactions were set up as detailed in section 2.6.3 at the standard temperature, along with the following experimental temperature conditions: 20, 24, 28, 30 and 34 °C. Once again, samples were run on a denaturing urea-TBE polyacrylamide gel, stained, imaged, and quantified as described in section 2.6.3.

2.7.2 Salt conditions

Prior to *in vitro* transcription, stock solutions of sodium chloride (NaCl), potassium chloride (KCl) and lithium chloride (LiCl) were prepared at a 1M concentration with Milli-Q water. *In vitro* transcription plasmids containing the sequences for BC200, as well as BC185 (a truncation that lacks the 3' poly-C tract, or the poly-G tract at the DNA level), were purified as described in section 2.3.2, and linearized as described in section 2.6.1. Following linearization, plasmids were treated without any additional salts, or in the presence of KCl, NaCl, and LiCl. To each plasmid, stock monovalent cations were added to a concentration of 100 mM in a microcentrifuge tube before heating to 95 °C for 5 minutes on an Eppendorf ThermoMixer C (Thermo Fisher Scientific), and then left on a microcentrifuge tube rack at room temperature to cool. After cooling, plasmids were stored at 4 °C for long-term preservation. Plasmid without any salt added was also treated this way for the sake of consistency. The heating of the plasmid and the cooling was done to allow any possible secondary structures at the DNA level to form prior to use as template in an *in vitro* transcription reaction. 50 µL *in vitro* transcription reactions were

set up and ran according to the method detailed in section 2.6.3, with the addition of monovalent cation stock to bring the concentration back up to 100 mM in the final reaction. Gels were run and stained as described in section 2.6.3.

2.7.3 Varied NTP concentrations

One of the main factors impacting the outcomes of *in vitro* transcription is that of the NTP concentrations used. To assess this, *in vitro* transcription reactions were set up as previously described in section 2.6.3, with the concentration of NTPs in the reaction cut in half; as such, the concentration of adenosine (A), uridine (U), cytosine (C) and guanosine (G) nucleotides in the solution was 1 mM each, adding up to a final concentration of 4 mM. In addition, test transcriptions were conducted at differing concentrations of MgCl_2 to optimize the experiment, as suggested by McKenna *et al.* when experimental conditions are changed (82) Gels were run as described in section 2.6.3. However, SYBR Gold stain diluted in 1:10,000 (Thermo Fisher Scientific) was used in place of staining with 0.1% toluidine blue solution for some experimental trials and the gel was destained with Milli-Q[®] water. Gels stained with SYBR Gold were imaged using the FluorChem Q[®] imaging system (ProteinSmart) on the Cy3 excitation and emission filters. Other gels were stained as described in section 2.6.3.

Chapter 3: Results

3.1. Acknowledgements

While the experiments conducted to collect data for this project were developed by me, I received assistance in collecting some data. The data presented in Figure 3.5 was collected by Bruce Barrientos. The data presented in Figures 3.7 and 3.8 were collected by Khoi Ho.

3.2. 5 cytosines in the poly-C region of BC200 truncations results in increased levels of cellular BC200.

To assess how the levels of cellular BC200 and BC120 are affected by poly-C tract length, cells were transfected and incubated with overexpression plasmids containing sequences corresponding to BC200 truncations with 1 C in the tract all the way up to full-length BC200 (12 Cs). Transfection efficiency was checked via RT-qPCR to quantify GFP mRNA. As displayed in Figure 3.3, levels of GFP mRNA demonstrate some degree of variability between conditions relative to the empty vector (EV) condition, which is to be expected. However, the variability seen in the transfection efficiency does not appear to impact the trends seen BC200 and BC120 abundance via Northern blotting (Figure 3.1) Under these conditions, we could reasonably assume that the changes we saw in the relative levels of cellular BC200 and BC120 were due to the overexpression of a BC200 truncation, and not inconsistencies in transfection efficiency.

Subsequently, Northern blotting was conducted to probe for changes in the abundance of BC200 and BC120 upon overexpression of BC200 truncations with increasing C-tract lengths. Initially, the probes targeting nucleotides 33-52 and 86-107 of BC200 were used together. Using this method yielded noisy Northern Blots not suitable for quantification. This is because 86-107

probe was specific to the Alu region, making it bind strongly with 7SL in the cell and very little else. A combination of this Alu-specific probe, along with another probe that bound to a sequence more specific to BC200 and its other isoforms was tested. This combination yielded Northern blots with significant background noise, rendering them unsuitable for quantification.

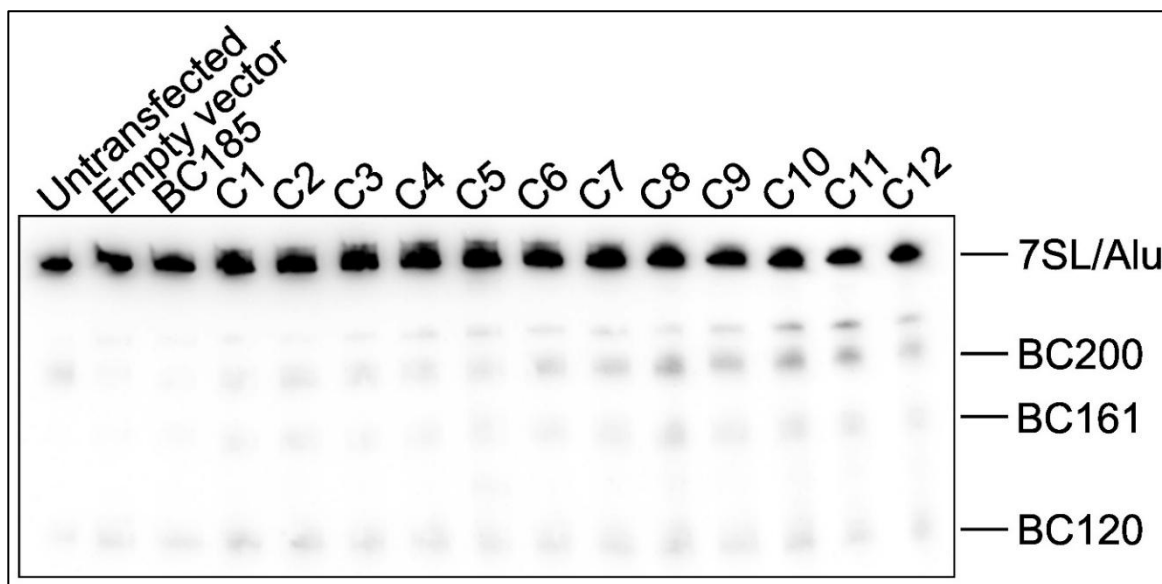


Figure 3.1. The probe targeting nucleotides 86-107 in BC200 non-specifically binds to 7SL and other Alus in the cell. RNA samples were separated by size via urea-denaturing acrylamide gel electrophoresis using an 8% gel at 13W for 30 minutes. RNA was transferred to a nylon membrane and crosslinked before hybridization with a DIG-labeled LNA probe complimentary to part of the Alu region of BC200. The blot was incubated with a primary antibody against DIG prior to incubation with a secondary antibody conjugated to HRP. RNA was detected via chemiluminescence. C1 through to C11 refer to the number of Cs in the BC200 truncation overexpressed in each condition.

Lastly, a probe targeting nucleotides 33-52 in BC200 was used on its own. Despite still exhibiting a small degree of binding to 7SL, this probe was found to be much more specific to BC200 and its truncations. As illustrated by Figure 3.2, as the length of cytosines encoded by each overexpression constructs increases, so too does the level of cellular BC200, while levels of BC120 remain relatively constant over time. Examination of BC200 and BC120 levels via

densitometry further demonstrates this trend. In the ‘Untransfected’ condition, an intense band at the BC200 position is present, along with very faint BC161 and BC120 band. These bands correspond to endogenous levels of each RNA species, showing that without overexpression, endogenous levels of BC200 are noticeably higher than BC120, differing from the truncation overexpression conditions. According to Figure 3.2, panel B, there is a degree of variability present with the densitometry data.

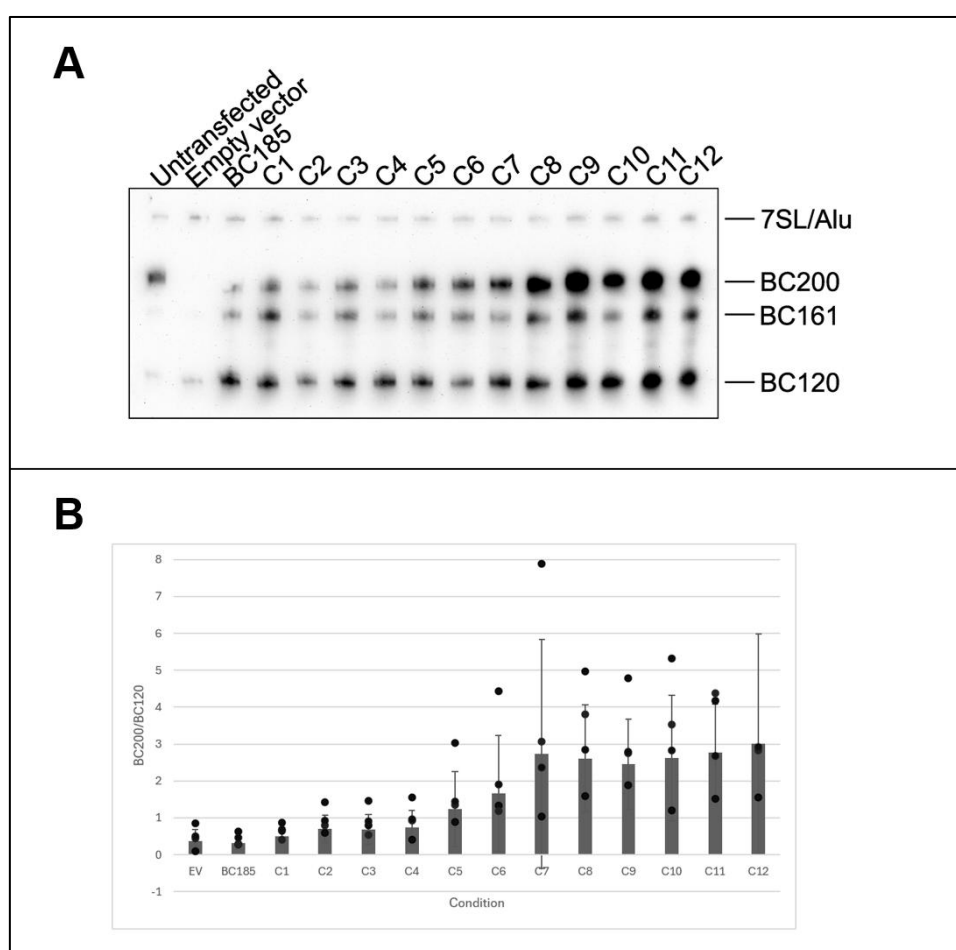


Figure 3.2. The length of cytosines in the poly-C tract of the overexpression plasmid impacts the cellular levels of BC200 and BC120 RNA. (A). An exemplary Northern blot showing this trend, which becomes noticeable by the 5C condition. RNA samples were separated by size via urea-denaturing acrylamide gel electrophoresis using an 8% gel at 13W for 30 minutes. RNA was transferred to a nylon membrane and crosslinked before hybridization with a DIG-labeled LNA probe complimentary to part of the Alu region of BC200. The blot was incubated with a primary

antibody against DIG prior to incubation with a secondary antibody conjugated to HRP. RNA was detected via chemiluminescence (B). Quantified densitometry data showing the ratio of BC200 and BC120 in the cell, indicated by BC200/BC120 on the y-axis. C12 refers to full-length BC200. Empty Vector refers to a GFP-containing overexpression vector without the insertion of BC200 or one of its truncations. n=4.

When performing transfection-based experiments, measuring transfection efficiency is imperative to ensure that the trends observed are true trends observed due to experimental conditions, and not due to inconsistencies in plasmid transfection. Because the overexpression plasmids used in these experiments contained the coding sequence for GFP, GFP mRNA was measured via RT-qPCR. Data was quantified via delta-Ct (Δ Ct) relative to the empty vector (EV) condition). These results are displayed in Figure 3.3. According to this figure, there is some variation in the fold-change of GFP levels between each condition relative to EV. Ideally, the fold-change would have a value of 1 for each condition, but the variation is within an anticipated range. For every transfection experiment, GFP levels were tested to examine consistency between biological replicates.

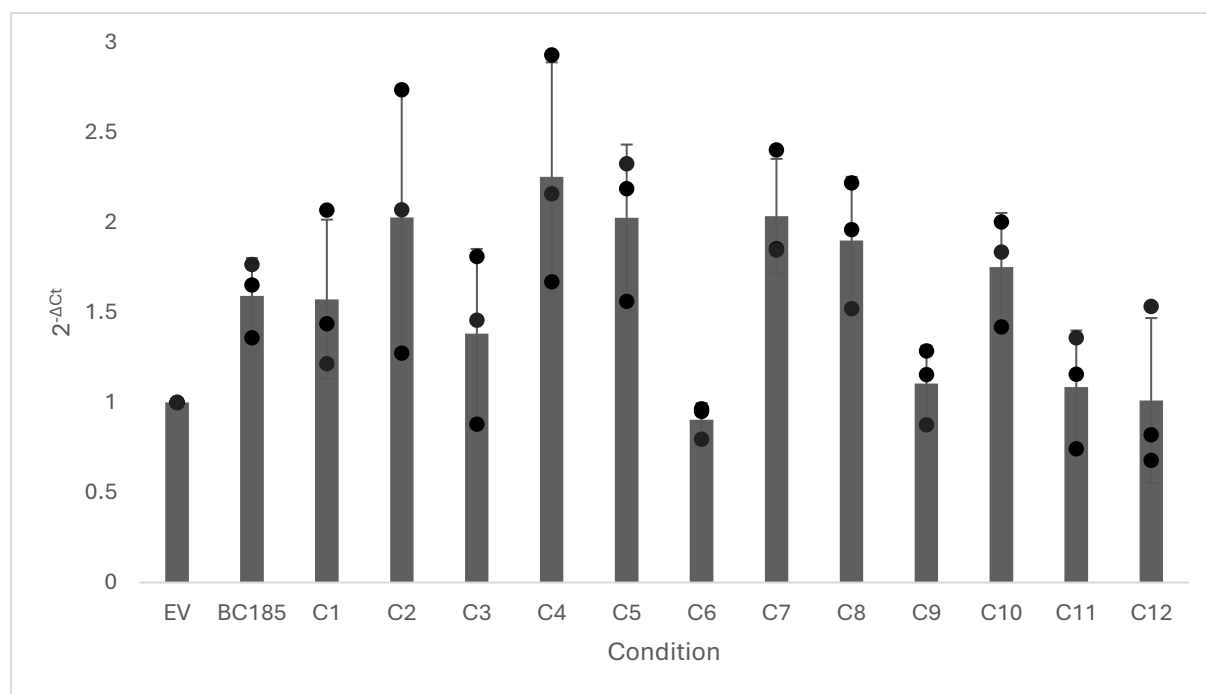


Figure 3.3. The quantification of GFP mRNA via RT-qPCR serves as a means to assess transfection efficiency. The $\Delta\Delta C_t$ method was used to compare the amplification of GFP mRNA in all experimental conditions, relative to the EV condition. Three biological replicates were used for quantification. Each biological replicate of a sample used in quantification is the average of three technical replicates, aside from one biological sample set.

In addition to Northern Blotting, a BC120-discriminating RT-qPCR assay was attempted as a way to quantitatively assess the relative abundance of BC200 and BC120 for each condition. This method was developed by the McKenna research group as a means to detect BC120 exclusively, since BC200 and BC120 share their first ~120 nucleotides (3). The results of these assays are displayed in Figure 3.4. This experiment was attempted four times, with $\Delta\Delta C_t$ data not necessarily aligning with trends seen in the Northern Blot experiments (Figure 3.2). While the relative amplification of full-length BC200 trends towards increasing as the length of the C-tract in the overexpression vector increases, it is not as robust of a trend as seen in the Northern blots. In addition, for some of the higher C-tract length conditions, especially C8, the variability in the data is especially large. With how difficult it was to get consistency between biological

replicates, it was concluded that this approach to investigating the hypothesis associated with this study would take lots of troubleshooting relative to the timeframe allotted to complete this project. As such, the decision was made to no longer further pursue this approach.

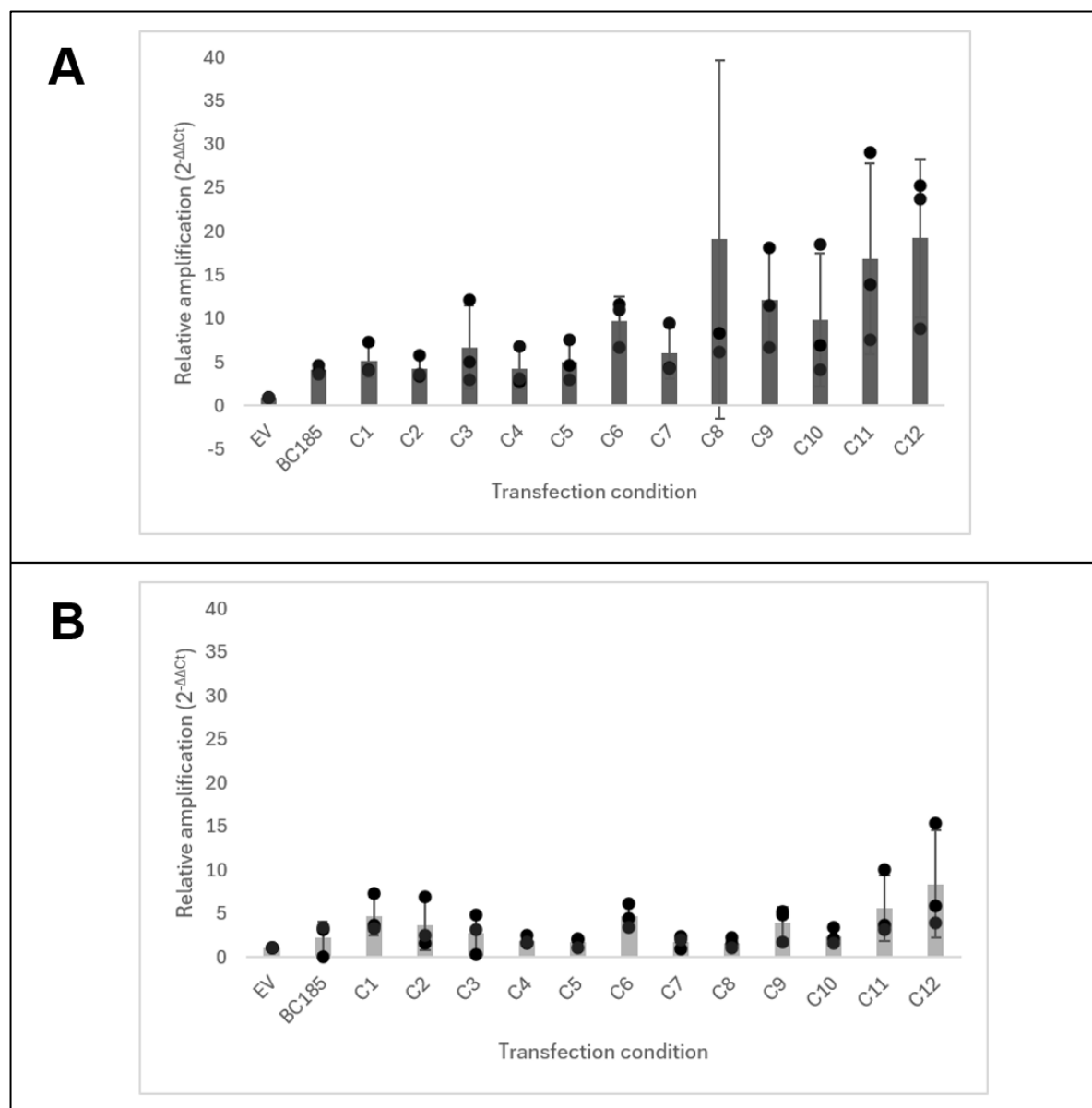


Figure 3.4. The combined Ct (A) and BC120-alone Ct values (B) yielded variation. GFP was used as a housekeeping gene to normalize the data. The $\Delta\Delta Ct$ analyses are the culmination of 3 biological replicates. All experimental conditions are compared relative to the EV condition. Briefly, RT-qPCR to detect BC120 was conducted using a primer and probe set with a forward primer complimentary to BC120, with a reverse primer complimentary to the linker, with a probe binding to BC120 between the two. A 5' and 3' DIG-labeled probe complimentary to nucleotides

157-177 was added to the reaction. RT-qPCR for the BC200 was done with a forward primer complimentary to BC200, along with the same reverse primer used to quantify BC120.

3.3. Poly-C binding protein 2 requires approximately 5 cytosines to engage in RNA-protein interactions with BC200.

It was found that generally, as the number of cytosines in the BC200 truncation increased in the overexpression construct, so too did cellular levels of BC200. As illustrated in Figure 3.2, panel B, significant increases in the ratio of BC200 to BC120 are not seen until the C5 truncation condition. Intrigued by this trend, an investigation into the RNA-protein interactions between *in vitro* transcribed truncations and known BC200 interactors was undertaken. *In vitro* transcription was used to synthesize BC200 truncations with 1 cytosine in its C tract, all the way up to full-length BC200 with 12 C's in its C-tract. This RNA was ligated with biotin, which, in conjunction with streptavidin coated magnetic beads can be used to precipitate or pull-down the biotinylated RNAs, along with their protein binding partners. These beads were mixed with loading dye and heated before separating bound proteins by size via SDS-PAGE. Proteins were transferred to a membrane which was blocked prior to incubation with a primary antibody against PCBP2. After washing, the blot was incubated with a secondary antibody conjugated to HRP such that proteins could be detected via chemiluminescence.

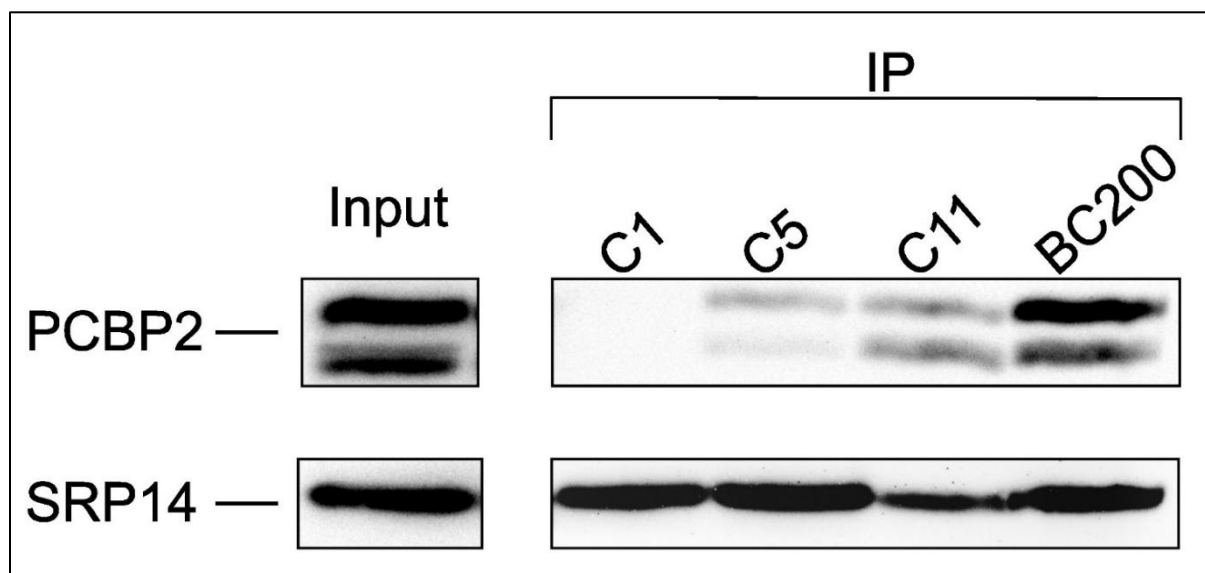


Figure 3.5. Approximately 5 cytosines are necessary for noticeable interactions with the poly-C tract of BC200. Probing for SRP14 was used as a loading control, as its abundance in lysate should not be impacted by the number of cytosines in the C-tract of BC200. “Input” refers to total cell lysate before a pull-down assay, while “IP” designates the PCBP2 pulled-down with the magnetic streptavidin-coated beads. Bound protein was separated by size via 15% SDS-PAGE at 100V for 1 hour before transfer to PVDF membrane which was blocked with 5% skim milk in TBS-T. The blot was incubated with a primary antibody against PCBP2, a secondary antibody conjugated to HRP. The blot was imaged using chemiluminescence.

As illustrated by Western Blot, four *in vitro* transcribed and biotinylated RNA pulldowns were done, testing four different poly-C tract lengths, including full-length BC200. SRP14 was used as a loading control. Referring to Figure 1.4, the Alu domain of BC200, nearest to the 5' end of the molecule is the region where SRP 14 interacts with BC200 as part of the SRP 9/14 heterodimer (62–64). PCBP2 is confirmed to interact with BC200 only in the poly-C region closest to the 3' end of the molecule. With the Alu domain of BC200 being unaffected by truncations at the 3' end, neither is SRP 14 binding to BC200. There is little to no PCBP2 detected in the pull-down for C1, with a slight increase for the C5 condition. There is more for C11, and finally, the most intense bands for the BC200 pull-down. This is indicative of the

affinity for PCBP2 to BC200 being greater when the poly-C tract length is at its full length. Due to time constraints, multiple successful biological replicates were not obtained. However, the trend seen in this trial is promising, providing future avenues of investigation for the continuation of this project.

3.4. BC120 is formed during BC200 *in vitro* transcription reactions in the presence of KCl, a salt that stabilizes G-quadruplexes, while lower transcription temperatures or NTP concentrations lead to inconclusive results.

3.4.1. BC120 formation increases in the presence of KCl

Literature has shown that structure formation at the DNA level can impact what products are formed during transcription (5,76). It is also known that the presence of monovalent cations such as K^+ can stabilize the formation of G-quadruplexes in nucleic acids (73–75). With these points in mind, it was imperative to investigate how the presence of certain monovalent cations impact BC200 transcription *in vitro*. To do so, *in vitro* transcription vectors containing the genomic sequence of BC200, along with a truncation of BC200 lacking the poly-C tract (BC185) were purified, linearized and repurified via phenol extraction and ethanol precipitation. Plasmids were treated with 100 mM KCl, NaCl, LiCl, or no additional salt before heating and cooling. Each plasmid was used in a small-scale *in vitro* transcription reaction and separated via urea-TBE acrylamide gel electrophoresis to assess their impacts on transcription. These results are illustrated in Figure 3.6.

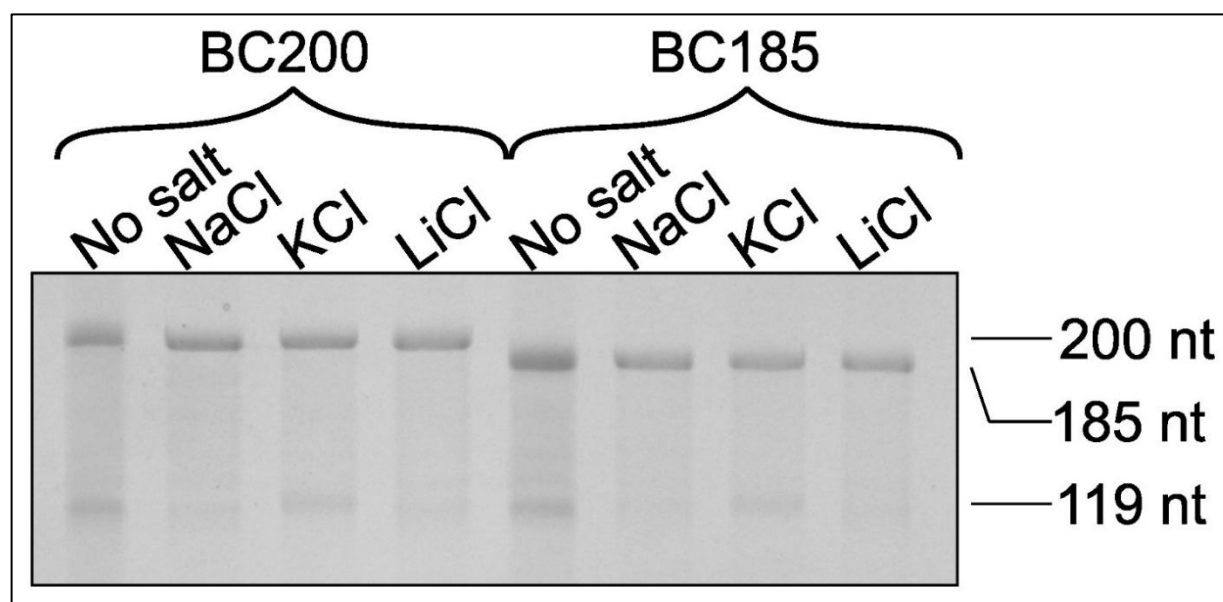


Figure 3.6. The presence of KCl salt during the *in vitro* transcription of BC200 and BC185 promotes the production of BC120. A distinct band that is ~119 nucleotides (nt) long is seen in both the BC185 KCl transcription, along with the BC185 transcription. The formation of a band at this length is negligible in the NaCl and LiCl conditions for both BC200 and BC185. BC200 and BC185 *in vitro* transcription plasmids were heated in the presence of 100 mM salt prior to *in vitro* transcription. Following transcription, reactions were mixed with load dye and ran on an 8% urea-denaturing acrylamide gel for 30 minutes at 13W. Gels were stained with a Coomassie Brilliant Blue staining solution and destained before imaging.

Across all experimental replicates was the trend of seeing the efficiency of transcription being slightly hampered by the presence of an extra 100 mM of salt in the *in vitro* transcription reaction mix. In Figure 3.6, 119 nucleotide bands are noticeable in the ‘No salt’ *in vitro* transcriptions of BC200 and BC185. This may be due to the heating and cooling steps of the transcription plasmids as detailed in section 2.7.3, leading to the formation of secondary structures in the linearized plasmid DNA during the reannealing process. In the NaCl conditions for both BC200 and BC185, the presence of the BC120 band is incredibly faint. In Figure 3.6, this band is labelled as “119 nt.” This discrepancy is in place due to the 3’ end heterogeneity of BC120 displayed within different MCF-7 clones, and between cell types (3). This indicates that

NaCl has little impact on the formation of a secondary structure in the DNA level that could lead to BC120 production. Interestingly, in the KCl conditions for both the BC200 and BC185 *in vitro* transcriptions, a higher presence of a smaller secondary band is seen. These results support what was originally hypothesized, which is that KCl may form G-quadruplexes at the DNA level that prematurely stall or terminate transcription, thus resulting in the transcription of BC120. The presence of BC120 in the *in vitro* transcription of BC185 with 100 mM KCl is intriguing due to the fact that BC185 was transcribed alongside BC200 in these experiments as a control due to its lack of a poly-C tract; in other words, its inability to form G-quadruplexes at the DNA level. As expected, the LiCl condition for both plasmids led to a lower yield of a BC120 band, indicating that there was no secondary structure stabilization at the DNA level in those transcriptions.

3.4.2. The reduction of NTP concentrations and changes in *in vitro* transcription temperature do not appear to impact BC120 production *in vitro*

The main objective of these experiments was to further explore how the formation of non-canonical structures could influence the production of BC120 during *in vitro* transcription of BC200. In the literature, reduced NTP concentrations, as well as temperature have been found to slow transcription (83,84). It is also known that decreased transcription rates can lead to the formation of non-canonical secondary structures, as reviewed by Zhang and Landick (85). With this in mind, an exploration into how altering these two conditions during *in vitro* transcription of BC200 impacts BC120 production was pursued. For the reduced NTP experiments, continuous optimization of MgCl₂ concentrations was conducted, since the efficiency of *in vitro* is sensitive to MgCl₂ concentrations, of which the ideal concentration depends on the plasmid and other reaction conditions (82). As shown in Figure 3.7, panel A, a MgCl₂ concentration gradient was

conducted to determine the ideal salt concentration for *in vitro* transcription, similar to when a standard transcription reaction is being optimized. Following optimization, the replicates of this experiment that were conducted were largely inconclusive. As such, the variable of NTP concentration was no longer pursued for the duration of this project. This is not to say that a reduction in NTP concentration does not impact BC200 production *in vitro*, however. Comparing transcription *in cellulo* and *in vitro* is difficult to do, especially since the cell is a dynamic environment.

It was believed that deviating from standard *in vitro* transcription conditions may have had an impact on the abundance of RNA transcribed, meaning that more sensitive detection methods may be necessary to examine the effects of reduced NTP concentrations. As such, experimental trials using SYBR[®] Gold staining as opposed to toluidine blue were employed. SYBR[®] Gold is a widely used commercial gel stain that is incredibly sensitive to double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), and single-stranded RNA (ssRNA), with a fluorescence enhancement about 1000-fold greater than ethidium bromide (EtBr) staining (86). This makes it a valuable stain to use when imaging gels with low levels of nucleic acids loaded. To this, *in vitro* transcriptions with reduced NTP concentrations were conducted with varying concentrations of MgCl₂ as described for the standard staining experiments. Transcriptions were ran using denaturing urea-TBE polyacrylamide gel electrophoresis and stained with SYBR[®] Gold diluted 1:10,000 and destained before imaging using the Cy3 excitation and emission channels. SYBR Gold staining, despite its increased sensitivity, does not facilitate the observation of any distinct changes in BC120 production with reduced NTP concentrations (86). To prevent redundancy, the results are not shown.

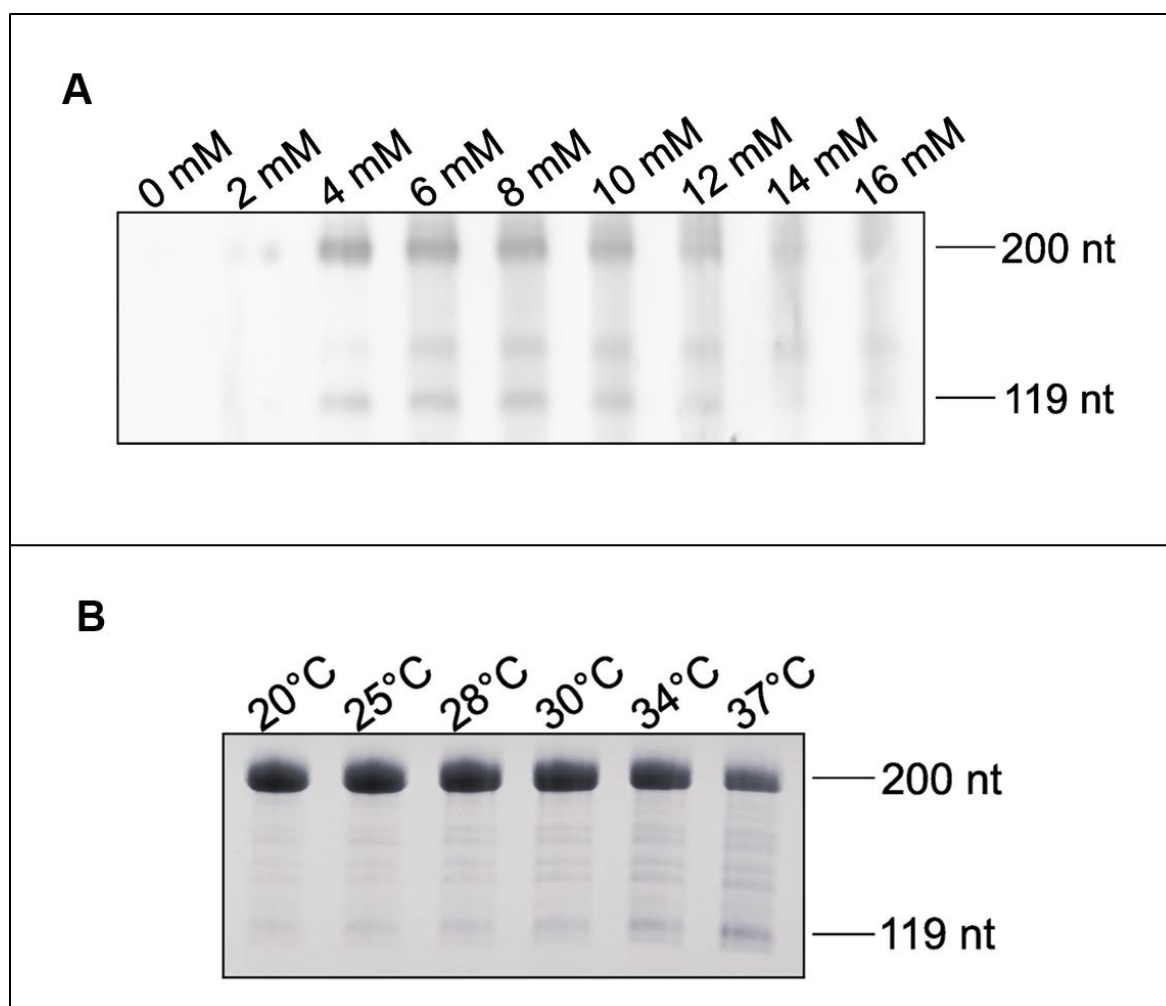


Figure 3.7. Reduced NTP concentrations (A) and reduced incubation temperatures (B) do not have a significant impact on the amount of BC120 produced during *in vitro* transcription. (A). Multiple transcriptions at reduced NTP concentrations with different concentrations of MgCl_2 were done for optimization purposes, yielding inconclusive results. (B). Reducing the incubation temperature during an IVT reaction relative to standard conditions leads to inconclusive results in relation to the relative abundances of BC200 and BC120. Transcription reactions were mixed with load dye and ran on an 8% urea-denaturing acrylamide gel for 30 minutes at 13W. Gels were stained with a toluidine blue staining solution and destained before imaging.

To assess how the incubation temperature of an *in vitro* transcription reaction may influence the production of BC120, standard transcription reactions of BC200 were set up and incubated for an hour at six different temperatures: 20, 25, 28, 30, 34 and 37 °C. A representative acrylamide gel is displayed in Figure 3.7, panel B. It can be seen that changes in the incubation

temperature of *in vitro* transcription reactions does not cause any significant shift towards producing any more or less BC120 in comparison to standard reaction conditions at 37 °C. With these results in mind, along with the results of the reduced NTP experiments, the conclusion that neither reduced NTP concentrations, nor reduced or *in vitro* transcription reaction temperatures cause any shifts at the transcriptional level that significantly impact the levels of BC120 seen during *in vitro* transcription.

Chapter 4: Discussion and Future Directions

4.1. Discussion

In this thesis, there are two main objectives that were investigated; that of how the length of the poly-C tract of BC200 may play a role in modulating its processing into BC120, possibly through RNA-protein interactions, and how the formation of non-canonical secondary structures at the DNA level and during transcription in general could influence BC120 production. These two areas were addressed through a combination of both *in cellulo* and *in vitro* experimental methods. While some conclusions were able to be drawn from the data collected relating to each of these areas, questions were raised regarding how differing experimental methodologies could approach these research questions differently, possibly leading to more robust and defined conclusions.

BC200 is a long non-coding RNA implicated cancer cell viability (57). In various ovarian cancer cell lines, a shortened version of BC200, named BC120, was found in high amounts (3). This discovery is what led to questions about how BC120 is produced in the cell; whether it be through the processing of BC200, or through transcriptional mechanisms. Beginning with investigating the BC200 processing hypothesis, Northern blotting RNA purified from MCF-7

cells transfected with different BC200 truncations pointed towards the trend that as the length of cytosines in the poly-C tract is increased, so too does the amount of full-length BC200 that is present in the cell. However, this increase is not very prominent in the data until the 5-cytosine BC200 truncation transfection condition, as seen in Figure 3.2, panel A. This data, in conjunction with the pull-down assay results displayed in Figure 3.5 supports the idea that interactions between nucleotides in the poly-C tract of BC200 and poly-C binding proteins such as PCBP2 protect BC200 from processing down into BC120.

PCBP2, along with the other PCBP2s within the family it belongs to, contain hnRNP K homology (KH) domains; three of them to be more precise (87). Structurally, in eukaryotes, KH domains consist of a beta sheet, followed by two alpha-helices, two more beta sheets and an alpha helix (88). According to crystal structures resolved by Du *et al.*, each KH domain of PCBP2 can interact with a C-rich sequence containing 3 C's (89,90). Tying this back to the pull-down assay results, it is seen that there is no between PCBP2 and the C1 BC200 truncation (Figure 3.5). However, there is a noticeable increase in the amount of PCBP2 pulled-down in the assay starting with the C5 condition. As such, with the number of cytosines in the C-tract of BC200, the known binding behaviour of PCBP2 KH domains to short cytosine-rich sequences, and the Northern blotting results obtained, it is probable that through interactions with proteins such as PCBP2, BC200 is shielded from, or prevented from events that would lead to cleavage of its tail region. Thus, increasing cellular levels of BC120. With further verification and experimental replicates to support the observations obtained in the pull-down assay, a new area relating to ncRNA regulation can be explored.

To explore ways in which the regulation of BC120 production may occur at the transcriptional level, IVT plasmids containing the sequences of BC200 and BC185 were

transcribed off of in conditions that deviate from standard IVT reaction conditions. At the beginning of this project, it was thought that at the DNA level, the genomic sequence for the poly-C tract of BC200 (which, at the DNA level is a stretch of G's) may form a secondary structure such as a G-quadruplex that could contribute to the premature termination of BC200 transcription, thus providing a possible explanation as to how BC120 arises. The first deviation from standard conditions was treating the plasmids with different salts that dissociate into monovalent cations (K, Li, Na). In the transcription reaction for BC200 containing KCl, a noticeable BC120 band appears alongside the BC200 band. However, this band is also seen in the BC185 transcription reaction with the same salt. At the DNA level, the sequence for BC185 does not contain the G-rich region that the BC200 sequence does. This leads to the conclusion that the formation of G-quadruplexes may not be implicated in BC120 transcription *in vitro*.

While K^+ ions are implicated in the stabilization of G-quadruplexes, they are also known to coordinate various non-secondary structures at both the DNA and RNA. For instance, it has been found that a structure known as a cytosine-hairpin (C-hairpin) commonly forms in single-stranded DNA (ssDNA) in the presence of K^+ ions (91). These structures are prone to forming in C-rich repetitive sequences such as the fragile X triplet repeat $(GCC)_n \bullet (GCC)_n$ (92). In BC200, at positions 119 and 120, there are two cytosine nucleotides, while at positions 128 and 129, there are two guanosine nucleotides (Figures 1.3, 1.4). While out of the scope of this project, it is probable that through the addition of K^+ ions, a C-hairpin structure is formed, which can act as a transcriptional modulator, thus resulting in the transcription of BC120. Investigating such a possibility could provide an interesting avenue for future studies into this specific area, further contributing to the literature surrounding the secondary structural modulation of gene expression (93).

Regarding the *in vitro* transcription experiments involving reduced NTP concentrations, the results are inconclusive. Any experiments conducted *in vitro* are difficult to compare with how similar conditions would manifest in a living cell. The impact of reduced NTP concentrations on transcription *in cellulo* is not widely studied in the literature. This makes drawing any conclusions and expanding on this area difficult to do, without first optimizing and repeating these experiments to draw on possibly more conclusive results.

In relating to the effects of temperature on the transcription of BC200 and BC120, there were no clear conclusions that could be drawn from these experiments. It appears as if there are no notable impacts on the levels of BC120 that are produced during *in vitro* transcription at incubation temperatures that deviate from the standard (37 °C). In the literature relating to decreased incubation temperatures for *in vitro* transcription reactions, it is found that for self-amplifying RNA (saRNA), lowering the incubation temperature for the reaction to temperatures as low as 21 °C, albeit with a longer incubation time of 2-4 hours, the yield of RNA is decreased, while the quality is increased (94). In this same paper, the ideal conditions for high quality mRNA are much higher, at the standard temperature of 37 °C, up to even 42 °C. There is little to nothing in the literature relating to how altering the temperature of IVT incubation for ncRNA products, presenting a gap that needs to be investigated more.

In a cellular context, lots of work in the literature has been done relating to how temperature changes change the transcriptional profile of the cell; in other words, what types of RNAs are produced. As expected, the transcriptomic profile in situations relating to fluctuations in temperature are those relating to heat-shock and maintaining homeostasis (95,96). There is nothing in the literature relating to how a shift in temperature may impact the specific

mechanisms of transcription, such as the formation of secondary structures that can hinder the transcription of a full-length transcription product, including the transcription of ncRNAs.

The development of malignant cells in eukaryotes is often triggered by aberrant gene expression and unregulated cell cycle processes, amongst other factors (97–99). It is also known that ncRNAs, such as Alu and Alu-like RNAs such as BC200 are implicated in cancer and other disease states (1,43,47,48,51,55). Because of this, studies into BC200 and its functional capabilities are rapidly growing in popularity. The discovery of BC120, a truncated version of BC200 with high levels of expression in ovarian cancers opens up many questions that need to be addressed (³). While an understanding that secondary structure formation may play a role in the production of BC120 *in vitro*, lots of work still needs to be done come close to understanding how BC120 is formed. An idea of how BC120 and BC200 are regulated is essential to being able to comprehend how they function in the context of both malignant and non-malignant cell lines.

4.2. Future Directions

This project has laid the foundation for future investigations into BC200 and its newly discovered truncation BC120. To fully understand how the poly-C tract of BC200 plays a role in regulating its levels in the cell, further refinement of the cell transfection process prior to Northern blotting and RT-qPCR needs to be developed. Since writing this thesis, members of the McKenna lab have picked up where this project has left off, troubleshooting the experimental pipeline, including scrutinizing the MaxiPrep protocol used to purify the overexpression plasmids. In addition to the Northern Blots, experimental replicates of the BC200 truncation pull-down assays involving more of the *in vitro* transcribed truncations must be done to further establish the trend seen between poly-C tract length and PCBP binding. In addition to probing

for PCBP2, probing for HNRNPK would be beneficial, as an interplay between these proteins in regulating cellular levels of BC200 and BC120 is possible.

Experiments relating to the stability of cellular levels of BC200 and BC120 upon transfection with the overexpression plasmids created for this project is also an important avenue to explore. A lab member has already pursued this investigation, conducting bromouridine (bU) chase experiments, similar to those previously published (2). In short, bU experiments involve treating cells with 5'-bU, followed by treatment with regular cell culture medium after incubation. Following this treatment, a pull-down assay, similar to that described to probe for BC200 C-tract interactors is done, and the levels of RNA at different timepoints is assessed to determine stability. Experiments such as these would provide insight into how the poly-C tract of BC200 impacts the stability of BC200 in the cell, relating back to how protein interactors influence the cellular levels of this transcript and BC120.

There are more conditions that can be tested *in vitro* to assess changes in the production of BC200 and BC120 during *in vitro* transcription. Spermidine is a common reagent included in many IVT buffers, playing a role in allowing T7 to effectively dissociate from the template DNA, amongst others (94,100). The standard buffer used for these experiments has a final spermidine concentration of 1 mM. Altering the spermidine concentration in the IVT buffer and assessing how the levels of BC200 and BC120 produced during transcription is an intriguing aspect to explore. The IVT buffer used for these experiments also contains PEG 8000, a molecular crowding agent, enhancing transcription relative to buffer solutions without it (101). As such, one could make the inference that increasing PEG 8000 levels in the buffer solution may increase the ability of T7 to interact with the template. The inverse would also likely be true. Assessing how both spermidine and PEG 8000 impact the relative levels of BC200 and BC120

produced during IVT, while difficult to translate to the regulation of these molecules in a cellular context, are interesting to explore from a biochemical perspective.

When all is said and done, lots has yet to be learned about BC200. Even more needs to be learned about BC120. It is known that BC200 functions to positively regulate translation and promote tumor cell viability, while what BC120 does is yet to be studied (47,57). Investigating the effects of BC120 on translation in cancer cells, as well as its impacts on cell malignancy are imperative to establishing how these two molecules may or may not interplay to impact cancer prognosis. Understanding the function of BC120, along with what conditions alter its production are essential in establishing a knowledge base that can be tapped into in the future. BC200 is an established biomarker for certain cancers (43,48,51,55). Depending on the information gleaned about BC120 in the future, there is the possibility that this molecule may also serve as a marker of cancer prognosis. The behaviour of BC200 in a cancer context, combined with the expression patterns observed for BC120 are a testament to how impactful ncRNAs are; not only in relation to malignant cells, but also in healthy ones. In short, an understanding of Alu RNAs, the factors influencing their expression, and the protein interaction networks they are a part of are crucial steps towards tapping into these RNAs, leading to the development of possibly life-saving treatments, improving the quality of many lives.

Works cited:

- (1) Gussakovsky, D.; McKenna, S. A. Alu RNA and Their Roles in Human Disease States. *RNA Biol.* **2021**, *18* (sup2), 574–585. <https://doi.org/10.1080/15476286.2021.1989201>.
- (2) Gussakovsky, D.; Booy, E. P.; Brown, M. J. F.; McKenna, S. A. Nuclear SRP9/SRP14 Heterodimer Transcriptionally Regulates 7SL and BC200 RNA Expression. *RNA* **2023**, *29* (8), 1185–1200. <https://doi.org/10.1261/rna.079649.123>.
- (3) Booy, E. P.; Gussakovsky, D.; Brown, M.; Shwaluk, R.; Nachtigal, M. W.; McKenna, S. A. lncRNA BC200 Is Processed into a Stable Alu Monomer. *RNA* **2024**, *30* (11), 1477–1494. <https://doi.org/10.1261/rna.080152.124>.
- (4) Brázda, V.; Luo, Y.; Bartas, M.; Kaura, P.; Porubiaková, O.; Šťastný, J.; Pečinka, P.; Verga, D.; Da Cunha, V.; Takahashi, T. S.; Forterre, P.; Myllykallio, H.; Fojta, M.; Mergny, J.-L. G-Quadruplexes in the Archaea Domain. *Biomolecules* **2020**, *10* (9), 1349. <https://doi.org/10.3390/biom10091349>.
- (5) Belotserkovskii, B. P.; Tornaletti, S.; D'Souza, A. D.; Hanawalt, P. C. R-Loop Generation during Transcription: Formation, Processing and Cellular Outcomes. *DNA Repair* **2018**, *71*, 69–81. <https://doi.org/10.1016/j.dnarep.2018.08.009>.
- (6) Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561–563.
- (7) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Miklos, G. L. G.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Francesco, V. D.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R.-R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z. Y.; Wang, A.; Wang, X.; Wang, J.; Wei, M.-H.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S. C.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y.-H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigo, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.;

Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y.-H.; Coyne, M.; Dahlke, C.; Mays, A. D.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. The Sequence of the Human Genome. *Sci. Am. Assoc. Adv. Sci.* **2001**, *291* (5507), 1304–1351.

(8) Makałowski, W. Genomic Scrap Yard: How Genomes Utilize All That Junk k. *Gene* **2000**, *259* (1–2), 61–67. [https://doi.org/10.1016/s0378-1119\(00\)00436-4](https://doi.org/10.1016/s0378-1119(00)00436-4).

(9) Zhang, P.; Wu, W.; Chen, Q.; Chen, M. Non-Coding RNAs and Their Integrated Networks. *J. Integr. Bioinforma.* **2019**, *16* (3), 20190027. <https://doi.org/10.1515/jib-2019-0027>.

(10) Ma, L.; Bajic, V. B.; Zhang, Z. On the Classification of Long Non-Coding RNAs. *RNA Biol.* **2013**, *10* (6), 924–933. <https://doi.org/10.4161/rna.24604>.

(11) Deininger, P. Alu Elements: Know the SINES. *Genome Biol.* **2011**, *12* (12), 236. <https://doi.org/10.1186/gb-2011-12-12-236>.

(12) Dewannieux, M.; Esnault, C.; Heidmann, T. LINE-Mediated Retrotransposition of Marked Alu Sequences. *Nat. Genet.* **2003**, *35* (1), 41–48. <https://doi.org/10.1038/ng1223>.

(13) Lin, L.; Shen, S.; Tye, A.; Cai, J. J.; Jiang, P.; Davidson, B. L.; Xing, Y. Diverse Splicing Patterns of Exonized Alu Elements in Human Tissues. *PLoS Genet.* **2008**, *4* (10), e1000225. <https://doi.org/10.1371/journal.pgen.1000225>.

(14) Batzer, M. A.; Deininger, P. L. Alu Repeats and Human Genomic Diversity. *Nat. Rev. Genet.* **2002**, *3* (5), 370–379. <https://doi.org/10.1038/nrg798>.

(15) Gallus, G. N.; Cardaioli, E.; Rufa, A.; Pozzo, P. D.; Bianchi, S.; D'Eramo, C.; Collura, M.; Tumino, M.; Pavone, L.; Federico, A. Alu-Element Insertion in an OPA1 Intron Sequence Associated with Autosomal Dominant Optic Atrophy.

(16) Di Ruocco, F.; Basso, V.; Rivoire, M.; Mehlen, P.; Ambati, J.; De Falco, S.; Tarallo, V. Alu RNA Accumulation Induces Epithelial-to-Mesenchymal Transition by Modulating miR-566 and Is Associated with Cancer Progression. *Oncogene* **2018**, *37* (5), 627–637. <https://doi.org/10.1038/onc.2017.369>.

(17) Drusco, A.; Nuovo, G. J.; Zanesi, N.; Di Leva, G.; Pichiorri, F.; Volinia, S.; Fernandez, C.; Antenucci, A.; Costinean, S.; Bottoni, A.; Rosito, I. A.; Liu, C.-G.; Burch, A.; Acunzo, M.; Pekarsky, Y.; Alder, H.; Ciardi, A.; Croce, C. M. MicroRNA Profiles Discriminate among Colon Cancer Metastasis. *PLoS ONE* **2014**, *9* (6), e96670. <https://doi.org/10.1371/journal.pone.0096670>.

(18) Quentin, Y. Fusion of a Free Left Alu Monomer and a Free Right Alu Monomer at the Origin of the Alu Family in the Primate Genomes. *Nucleic Acids Res.* **1992**, *20* (3), 487–493. <https://doi.org/10.1093/nar/20.3.487>.

(19) Kriegs, J. O.; Churakov, G.; Jurka, J.; Brosius, J.; Schmitz, J. Evolutionary History of 7SL RNA-Derived SINES in Supraprimates. *Trends Genet.* **2007**, *23* (4), 158–161. <https://doi.org/10.1016/j.tig.2007.02.002>.

- (20) Walter, P.; Blobel, G. Signal Recognition Particle Contains a 7S RNA Essential for Protein Translocation across the Endoplasmic Reticulum. *Nature* **1982**, *299* (5885), 691–698. <https://doi.org/10.1038/299691a0>.
- (21) Peter Walter; Vishwanath, L. Mechanism of Protein Translocation Across the Endoplasmic Reticulum Membrane. *Annu. Rev. Cell Dev. Biol.* **1986**, *2*, 499–516. <https://doi.org/10.1146/annurev.cb.02.110186.002435>.
- (22) Hasler, J. Alu RNP and Alu RNA Regulate Translation Initiation in Vitro. *Nucleic Acids Res.* **2006**, *34* (8), 2374–2385. <https://doi.org/10.1093/nar/gkl246>.
- (23) Blanchet, S.; Ranjan, N. Translation Phases in Eukaryotes. In *Ribosome Biogenesis*; Entian, K.-D., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2022; Vol. 2533, pp 217–228. https://doi.org/10.1007/978-1-0716-2501-9_13.
- (24) Saraogi, I.; Shan, S. Molecular Mechanism of Co-translational Protein Targeting by the Signal Recognition Particle. *Traffic* **2011**, *12* (5), 535–542. <https://doi.org/10.1111/j.1600-0854.2011.01171.x>.
- (25) Kellogg, M. K.; Tikhonova, E. B.; Karamyshev, A. L. Signal Recognition Particle in Human Diseases. *Front. Genet.* **2022**, *13*, 898083. <https://doi.org/10.3389/fgene.2022.898083>.
- (26) Walter, P. SIGNAL SEQUENCE RECOGNITION AND PROTEIN TARGETING TO THE ENDOPLASMIC RETICULUM MEMBRANE. **2026**.
- (27) Siegel, V.; Walter, P. Removal of the Alu Structural Domain from Signal Recognition Particle Leaves Its Protein Translocation Activity Intact. *Nature* **1986**, *320* (6057), 81–84. <https://doi.org/10.1038/320081a0>.
- (28) Sauer-Eriksson, A. E.; Hainzl, T. S-Domain Assembly of the Signal Recognition Particle. *Curr. Opin. Struct. Biol.* **2003**, *13* (1), 64–70. [https://doi.org/10.1016/S0959-440X\(02\)00010-6](https://doi.org/10.1016/S0959-440X(02)00010-6).
- (29) Luijck, J. SRP-Mediated Protein Targeting: Structure and Function Revisited. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2004**. <https://doi.org/10.1016/j.bbamer.2004.03.013>.
- (30) Strub, K.; Walter, P. Assembly of the Alu Domain of the Signal Recognition Particle (SRP): Dimerization of the Two Protein Components Is Required for Efficient Binding to SRP RNA. *Mol. Cell. Biol.* **1990**, *10* (2).
- (31) Walter, P.; Blobel, G. Disassembly and Reconstitution of Signal Recognition Particle. *Cell* **1983**, *34* (2), 525–533. [https://doi.org/10.1016/0092-8674\(83\)90385-9](https://doi.org/10.1016/0092-8674(83)90385-9).
- (32) Politz, J. C.; Yarovoi, S.; Kilroy, S. M.; Gowda, K.; Zwieb, C.; Pederson, T. Signal Recognition Particle Components in the Nucleolus. *Proc. Natl. Acad. Sci.* **2000**, *97* (1), 55–60. <https://doi.org/10.1073/pnas.97.1.55>.
- (33) Romisch, K.; Ribes, V.; High, S.; Lutcke, H.; Tollervey, D.; Dobberstein, B. Structure and Function of Signal Recognition Particle (SRP).
- (34) Bovia, F.; Strub, K. The Signal Recognition Particle and Related Small Cytoplasmic Ribonucleoprotein Particles. *J. Cell Sci.* **1996**, *109* (11), 2601–2608. <https://doi.org/10.1242/jcs.109.11.2601>.

- (35) Labuda, D.; Zitekiewicz, E. Evolution of Secondary Structure in the Family of 7SL-like RNAs. *J. Mol. Evol.* **1994**, *39* (5). <https://doi.org/10.1007/BF00173420>.
- (36) Jurka, J.; Milosavljevic, A. Reconstruction and Analysis of Human Alu Genes. *J. Mol. Evol.* **1991**, *32* (2), 105–121. <https://doi.org/10.1007/BF02515383>.
- (37) Tiedge, H.; Chen, W.; Brosius, J. Primary Structure, Neural-Specific Expression, and Dendritic Location of Human BC200 RNA. *J. Neurosci.* **1993**, *13* (6), 2382–2390. <https://doi.org/10.1523/JNEUROSCI.13-06-02382.1993>.
- (38) Martignetti, J. A.; Brosius, J. BC200 RNA: A Neural RNA Polymerase III Product Encoded by a Monomeric Alu Element. *Proc. Natl. Acad. Sci.* **1993**, *90* (24), 11563–11567. <https://doi.org/10.1073/pnas.90.24.11563>.
- (39) Tiedge, H.; Freneau, R. T.; Weinstock, P. H.; Arancio, O.; Brosius, J. Dendritic Location of Neural BC1 RNA. *Proc. Natl. Acad. Sci.* **1991**, *88* (6), 2093–2097. <https://doi.org/10.1073/pnas.88.6.2093>.
- (40) Kremerskothen, J.; Zopf, D.; Walter, P.; Cheng, J.-G.; Nettermann, M.; Niewerth, U.; Maraia, R. J.; Brosius, J. Heterodimer SRP9/14 Is an Integral Part of the Neural BC200 RNP in Primate Brain. *Neurosci. Lett.* **1998**, *245* (3), 123–126. [https://doi.org/10.1016/S0304-3940\(98\)00215-8](https://doi.org/10.1016/S0304-3940(98)00215-8).
- (41) Mus, E.; Hof, P. R.; Tiedge, H. Dendritic BC200 RNA in Aging and in Alzheimer's Disease. *Proc. Natl. Acad. Sci.* **2007**, *104* (25), 10679–10684.
- (42) Shin, H.; Kim, Y.; Kim, M.; Lee, Y. BC200 RNA: An Emerging Therapeutic Target and Diagnostic Marker for Human Cancer.
- (43) Samson, J.; Cronin, S.; Dean, K. BC200 (BCYRN1) – The Shortest, Long, Non-Coding RNA Associated with Cancer. *Non-Coding RNA Res.* **2018**, *3* (3), 131–143. <https://doi.org/10.1016/j.ncrna.2018.05.003>.
- (44) Non Coding RNA Molecules as Potential Biomarkers in Breast Cancer. In *Advances in Cancer Biomarkers: From biochemistry to clinic for a critical revision*; Scatena, R., Ed.; Advances in Experimental Medicine and Biology; Springer Netherlands: Dordrecht, 2015; Vol. 867, pp 263–275. <https://doi.org/10.1007/978-94-017-7215-0>.
- (45) Cao, X.-G.; Zhao, R.; Zhu, C.; Li, X.; Cao, W.; Zong, H.; Hu, H. BC200 LncRNA a Potential Predictive Marker of Poor Prognosis in Esophageal Squamous Cell Carcinoma Patients. *OncoTargets Ther.* **2016**, 2221. <https://doi.org/10.2147/OTT.S99401>.
- (46) LncRNA BC200 Regulates the Cell Proliferation and Cisplatin Resistance in Non-Small Cell Lung Cancer via PI3K/AKT Pathway.
- (47) Booy, E. P.; McRae, E. K.; Koul, A.; Lin, F.; McKenna, S. A. The Long Non-Coding RNA BC200 (BCYRN1) Is Critical for Cancer Cell Survival and Proliferation. *Mol. Cancer* **2017**, *16* (1), 109. <https://doi.org/10.1186/s12943-017-0679-7>.
- (48) Chen, X.; Zhao, Y.; Wang, D.; Lin, Y.; Hou, J.; Xu, X.; Wu, J.; Zhong, L.; Zhou, Y.; Shen, J.; Zhang, W.; Cao, H.; Hong, X.; Hu, T.; Zhan, Y. The HNF4 α -BC200-FMR1–Positive Feedback Loop Promotes Growth and Metastasis in Invasive Mucinous Lung Adenocarcinoma. *Cancer Res.* **2021**, *81* (23), 5904–5918. <https://doi.org/10.1158/0008-5472.CAN-21-0980>.

- (49) Zheng, H.; Chen, C.; Luo, Y.; Yu, M.; He, W.; An, M.; Gao, B.; Kong, Y.; Ya, Y.; Lin, Y.; Li, Y.; Xie, K.; Huang, J.; Lin, T. Tumor-derived Exosomal *BCYRN1* Activates WNT5A/VEGF-C/VEGFR3 Feedforward Loop to Drive Lymphatic Metastasis of Bladder Cancer. *Clin. Transl. Med.* **2021**, *11* (7), e497. <https://doi.org/10.1002/ctm2.497>.
- (50) Huo, W.; Qi, F.; Wang, K. Long Non-coding RNA BCYRN1 Promotes Prostate Cancer Progression via Elevation of HDAC11. *Oncol. Rep.* **2020**, *44* (3), 1233–1245. <https://doi.org/10.3892/or.2020.7680>.
- (51) Zhao, R.; Cao, X.; Jin, S.; Li, R.; Zhong, Q.; Jiang, M.; Han, J.; Guo, C.; Zong, H. LncRNA BC200 Promotes Esophageal Squamous Cell Cancer Migration and Invasion and Can Regulate ATF4 Expression. *Front. Oncol.* **2020**, *10*, 1392. <https://doi.org/10.3389/fonc.2020.01392>.
- (52) Mu, M.; Niu, W.; Zhang, X.; Hu, S.; Niu, C. LncRNA BCYRN1 Inhibits Glioma Tumorigenesis by Competitively Binding with miR-619-5p to Regulate CUEDC2 Expression and the PTEN/AKT/P21 Pathway. *Oncogene* **2020**, *39* (45), 6879–6892. <https://doi.org/10.1038/s41388-020-01466-x>.
- (53) Song, S.; Bian, W.-G.; Qin, Z.; Zeng, D.; Xu, J.-J.; Tang, H.-C. LncRNA BCYRN1 Promotes Cell Migration and Invasion of Non-Small Cell Lung Cancer via the miR-30b-3p/ROCK1 Axis. *Neoplasma* **2022**, *69* (03), 583–593. https://doi.org/10.4149/neo_2022_210726N1045.
- (54) Arima, J.; Yoshino, H.; Fukumoto, W.; Kawahara, I.; Saito, S.; Li, G.; Fukuda, I.; Iizasa, S.; Mitsuke, A.; Sakaguchi, T.; Inoguchi, S.; Matsushita, R.; Nakagawa, M.; Tatarano, S.; Yamada, Y.; Enokida, H. LncRNA BCYRN1 as a Potential Therapeutic Target and Diagnostic Marker in Serum Exosomes in Bladder Cancer. *Int. J. Mol. Sci.* **2024**, *25* (11), 5955. <https://doi.org/10.3390/ijms25115955>.
- (55) Zhang, S.; Liu, N.; Cao, P.; Qin, Q.; Li, J.; Yang, L.; Xin, Y.; Jiang, M.; Zhang, S.; Yang, J.; Lu, J. LncRNA BC200 Promotes the Development of EBV-Associated Nasopharyngeal Carcinoma by Competitively Binding to miR-6834-5p to Upregulate TYMS Expression. *Int. J. Biol. Macromol.* **2024**, *278*, 134837. <https://doi.org/10.1016/j.ijbiomac.2024.134837>.
- (56) Rodriguez-Martinez, A.; Young-Baird, S. K. Polysome Profiling Is an Extensible Tool for the Analysis of Bulk Protein Synthesis, Ribosome Biogenesis, and the Specific Steps in Translation. *Mol. Biol. Cell* **2025**, *36* (4), mr2. <https://doi.org/10.1091/mbc.E24-08-0341>.
- (57) Booy, E. P.; Gussakovsky, D.; Choi, T.; McKenna, S. A. The Noncoding RNA BC200 Associates with Polysomes to Positively Regulate mRNA Translation in Tumor Cells. *J. Biol. Chem.* **2021**, *296*, 100036. <https://doi.org/10.1074/jbc.RA120.015775>.
- (58) Jang, S.; Shin, H.; Lee, J.; Kim, Y.; Bak, G.; Lee, Y. Regulation of BC 200 RNA -mediated Translation Inhibition by Hn RNP E1 and E2. *FEBS Lett.* **2017**, *591* (2), 393–405. <https://doi.org/10.1002/1873-3468.12544>.
- (59) Kondrashov, A. V.; Kieffmann, M.; Ebnet, K.; Khanam, T.; Muddashetty, R. S.; Brosius, J. Inhibitory Effect of Naked Neural BC1 RNA or BC200 RNA on Eukaryotic in Vitro Translation Systems Is Reversed by Poly(A)-Binding Protein (PABP). *J. Mol. Biol.* **2005**, *353* (1), 88–103. <https://doi.org/10.1016/j.jmb.2005.07.049>.
- (60) Jang, S.; Shin, H.; Lee, Y. Functional Analysis of RNA Motifs Essential for BC200 RNA-Mediated Translational Regulation. *BMB Rep.* **2020**, *53* (2), 94–99. <https://doi.org/10.5483/BMBRep.2020.53.2.153>.

- (61) Lin, D.; Pestova, T. V.; Hellen, C. U. T.; Tiedge, H. Translational Control by a Small RNA: Dendritic BC1 RNA Targets the Eukaryotic Initiation Factor 4A Helicase Mechanism. *Mol. Cell. Biol.* **2008**, *28* (9), 3008–3019. <https://doi.org/10.1128/MCB.01800-07>.
- (62) Booy, E. P.; McRae, E. K.; Ezzati, P.; Choi, T.; Gussakovsky, D.; McKenna, S. A. Comprehensive Analysis of the BC200 Ribonucleoprotein Reveals a Reciprocal Regulatory Function with CSDE1/UNR. *Nucleic Acids Res.* **2018**, *46* (21), 11575–11591. <https://doi.org/10.1093/nar/gky860>.
- (63) Booy, E. P.; McRae, E. K. S.; Howard, R.; Deo, S. R.; Ariyo, E. O.; Dzananovic, E.; Meier, M.; Stetefeld, J.; McKenna, S. A. RNA Helicase Associated with AU-Rich Element (RHAU/DHX36) Interacts with the 3'-Tail of the Long Non-Coding RNA BC200 (BCYRN1). *J. Biol. Chem.* **2016**, *291* (10), 5355–5372. <https://doi.org/10.1074/jbc.M115.711499>.
- (64) Bovia, F. The SRP9/14 Subunit of the Human Signal Recognition Particle Binds to a Variety of Alu-like RNAs and with Higher Affinity than Its Mouse Homolog. *Nucleic Acids Res.* **1997**, *25* (2), 318–326. <https://doi.org/10.1093/nar/25.2.318>.
- (65) Ray, S.; Anderson, E. C. Stimulation of Translation by Human Unr Requires Cold Shock Domains 2 and 4, and Correlates with Poly(A) Binding Protein Interaction. *Sci. Rep.* **2016**, *6* (1), 22461. <https://doi.org/10.1038/srep22461>.
- (66) Xu, Y.; Wu, W.; Han, Q.; Wang, Y.; Li, C.; Zhang, P.; Xu, H. New Insights into the Interplay between Non-Coding RNAs and RNA-Binding Protein HnRNPK in Regulating Cellular Functions. *Cells* **2019**, *8* (1), 62. <https://doi.org/10.3390/cells8010062>.
- (67) Fabbri, M.; Girnita, L.; Varani, G.; Calin, G. A. Decrypting Noncoding RNA Interactions, Structures, and Functional Networks. *Genome Res.* **2019**, *29* (9), 1377–1388. <https://doi.org/10.1101/gr.247239.118>.
- (68) Li, Z.; Zhang, T.; Yang, X.; Peng, Y. Role of Noncoding RNA and Protein Interaction in Pancreatic Cancer. *Chin. Med. J. (Engl.)* **2025**, *138* (9), 1019–1036. <https://doi.org/10.1097/CM9.00000000000003587>.
- (69) Nicholson-Shaw, A. L.; Kofman, E. R.; Yeo, G. W.; Pasquinelli, A. E. Nuclear and Cytoplasmic Poly(A) Binding Proteins (PABPs) Favor Distinct Transcripts and Isoforms. *Nucleic Acids Res.* **2022**, *50* (8), 4685–4702. <https://doi.org/10.1093/nar/gkac263>.
- (70) Wang, H.; Chen, Y.; Liu, Y.; Li, Q.; Luo, J.; Wang, L.; Chen, Y.; Sang, C.; Zhang, W.; Ge, X.; Yao, Z.; Miao, L.; Liu, X. The lncRNA ZFAS1 Regulates Lipogenesis in Colorectal Cancer by Binding Polyadenylate-Binding Protein 2 to Stabilize SREBP1 mRNA. *Mol. Ther. Nucleic Acids* **2022**, *27*, 363–374. <https://doi.org/10.1016/j.omtn.2021.12.010>.
- (71) Cox, R. A.; Littauer, U. Z. Secondary Structure of Ribonucleic Acid in Solution. *Nature* **1959**, *184*, 818–819. <https://doi.org/https://doi-org.uml.idm.oclc.org/10.1038/184818a0>.
- (72) Ts'O, P. O. P.; Helmkamp, K.; Sander, C. Secondary Structures of Nucleic Acids in Organic Solvents: IV. Effects of Electrolytes.
- (73) Sen, D.; Gilbert, W. Formation of Parallel Four-Stranded Complexes by Guanine-Rich Motifs in DNA and Its Implications for Meiosis. *Nature* **1988**, *334* (6180), 364–366. <https://doi.org/10.1038/334364a0>.

- (74) Wang, Y.; Patel, D. J. Guanine Residues in d(T2AG3) and d(T2G4) Form Parallel-Stranded Potassium Cation Stabilized G-Quadruplexes with Anti Glycosidic Torsion Angles in Solution. *Biochemistry* **1992**, *31* (35), 8112–8119. <https://doi.org/10.1021/bi00150a002>.
- (75) Bhattacharyya, D.; Mirihana Arachchilage, G.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, *4*. <https://doi.org/10.3389/fchem.2016.00038>.
- (76) Kim, N. The Interplay between G-Quadruplex and Transcription. *Curr. Med. Chem.* **2019**, *26* (16), 2898–2917. <https://doi.org/10.2174/0929867325666171229132619>.
- (77) Schult, P.; Paeschke, K. The DEAH Helicase *DHX36* and Its Role in G-Quadruplex-Dependent Processes. *Biol. Chem.* **2021**, *402* (5), 581–591. <https://doi.org/10.1515/hsz-2020-0292>.
- (78) Wu, G.; Xing, Z.; Tran, E. J.; Yang, D. DDX5 Helicase Resolves G-Quadruplex and Is Involved in *MYC* Gene Transcriptional Activation. *Proc. Natl. Acad. Sci.* **2019**, *116* (41), 20453–20461. <https://doi.org/10.1073/pnas.1909047116>.
- (79) Yang, L.; Lee, C.-Y.; Paul, T.; Myong, S. Opposing Roles of Two R-Loop Associated G-Quadruplexes in Tuning Transcription Activity. *Nucleic Acids Res.* **2025**, *53* (18), gkaf930. <https://doi.org/10.1093/nar/gkaf930>.
- (80) Conrad, T.; Plumbom, I.; Alcobendas, M.; Vidal, R.; Sauer, S. Maximizing Transcription of Nucleic Acids with Efficient T7 Promoters. *Commun. Biol.* **2020**, *3* (1), 439. <https://doi.org/10.1038/s42003-020-01167-x>.
- (81) Booy, E. P.; Meng, H.; McKenna, S. A. Native RNA Purification by Gel Filtration Chromatography. In *Recombinant and In Vitro RNA Synthesis: Methods and Protocols*; Conn, G. L., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; Vol. 941. <https://doi.org/10.1007/978-1-62703-113-4>.
- (82) McKenna, S. A.; Kim, I.; Puglisi, E. V.; Lindhout, D. A.; Aitken, C. E.; Marshall, R. A.; Puglisi, J. D. Purification and Characterization of Transcribed RNAs Using Gel Filtration Chromatography. *Nat. Protoc.* **2007**, *2* (12), 3270–3277. <https://doi.org/10.1038/nprot.2007.480>.
- (83) McDaniel, B. A. M.; Grundy, F. J.; Artsimovitch, I.; Henkin, T. M. Transcription Termination Control of the S Box System: Direct Measurement of S-Adenosylmethionine by the Leader RNA. *Proc. Natl. Acad. Sci.* **2003**, *100* (6), 3083–3088. <https://doi.org/10.1073/pnas.0630422100>.
- (84) Lewicki, B. T. U.; Margus, T.; Remme, J.; Nierhaus, K. H. Coupling of rRNA Transcription and Ribosomal Assembly in Vivo. *J. Mol. Biol.* **1993**, *231* (3), 581–593. <https://doi.org/10.1006/jmbi.1993.1311>.
- (85) Zhang, J.; Landick, R. A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. *Trends Biochem. Sci.* **2016**, *41* (4), 293–310. <https://doi.org/10.1016/j.tibs.2015.12.009>.
- (86) Tuma, R. S.; Beaudet, M. P.; Jin, X.; Jones, L. J.; Cheung, C.-Y.; Yue, S.; Singer, V. L. Characterization of SYBR Gold Nucleic Acid Gel Stain: A Dye Optimized for Use with 300-Nm Ultraviolet Transilluminators. *Anal. Biochem.* **1999**, *268* (2), 278–288. <https://doi.org/10.1006/abio.1998.3067>.

- (87) Makeyev, A. V.; Liebhaber, S. A. The Poly(C)-Binding Proteins: A Multiplicity of Functions and a Search for Mechanisms. *RNA* **2002**, *8* (3), 265–278. <https://doi.org/10.1017/S1355838202024627>.
- (88) Grishin, N. V. KH Domain: One Motif, Two Folds. *Nucleic Acids Res.* **2001**, *29* (3), 638–643. <https://doi.org/10.1093/nar/29.3.638>.
- (89) Du, Z.; Lee, J. K.; Tjhen, R.; Li, S.; Pan, H.; Stroud, R. M.; James, T. L. Crystal Structure of the First KH Domain of Human Poly(C)-Binding Protein-2 in Complex with a C-Rich Strand of Human Telomeric DNA at 1.7 Å. *J. Biol. Chem.* **2005**, *280* (46), 38823–38830. <https://doi.org/10.1074/jbc.M508183200>.
- (90) Du, Z.; Lee, J. K.; Fenn, S.; Tjhen, R.; Stroud, R. M.; James, T. L. 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* **2007**, *13* (7), 1043–1051. <https://doi.org/10.1261/rna.410107>.
- (91) Gao, B.; Hou, X.-M. Opposite Effects of Potassium Ions on the Thermal Stability of I-Motif DNA in Different Buffer Systems. *ACS Omega* **2021**, *6* (13), 8976–8985. <https://doi.org/10.1021/acsomega.0c06350>.
- (92) Chen, X.; Mariappan, S. V.; Catasti, P.; Ratliff, R.; Moyzis, R. K.; Laayoun, A.; Smith, S. S.; Bradbury, E. M.; Gupta, G. Hairpins Are Formed by the Single DNA Strands of the Fragile X Triplet Repeats: Structure and Biological Implications. *Proc. Natl. Acad. Sci.* **1995**, *92* (11), 5199–5203. <https://doi.org/10.1073/pnas.92.11.5199>.
- (93) Wadkins, R. Targeting DNA Secondary Structures. *Curr. Med. Chem.* **2000**, *7* (1), 1–15. <https://doi.org/10.2174/0929867003375461>.
- (94) Popova, P. G.; Lagace, M. A.; Tang, G.; Blakney, A. K. Effect of in Vitro Transcription Conditions on Yield of High Quality Messenger and Self-Amplifying RNA. *Eur. J. Pharm. Biopharm.* **2024**, *198*, 114247. <https://doi.org/10.1016/j.ejpb.2024.114247>.
- (95) Pessa, J. C.; Joutsen, J.; Sistonen, L. Transcriptional Reprogramming at the Intersection of the Heat Shock Response and Proteostasis. *Mol. Cell* **2024**, *84* (1), 80–93. <https://doi.org/10.1016/j.molcel.2023.11.024>.
- (96) Place, R. F.; Noonan, E. J. Non-Coding RNAs Turn up the Heat: An Emerging Layer of Novel Regulators in the Mammalian Heat Shock Response. *Cell Stress Chaperones* **2014**, *19* (2), 159–172. <https://doi.org/10.1007/s12192-013-0456-5>.
- (97) Sinkala, M. Mutational Landscape of Cancer-Driver Genes across Human Cancers. *Sci. Rep.* **2023**, *13* (1), 12742. <https://doi.org/10.1038/s41598-023-39608-2>.
- (98) Baylin, S. B. Aberrant Patterns of DNA Methylation, Chromatin Formation and Gene Expression in Cancer. *Hum. Mol. Genet.* **2001**, *10* (7), 687–692. <https://doi.org/10.1093/hmg/10.7.687>.
- (99) Collins, K.; Jacks, T.; Pavletich, N. P. The Cell Cycle and Cancer.
- (100) Nuss, L.; Herbst, J. Stimulation of in Vitro Transcription of T4 DNA by the Pobamine Spermidine.
- (101) Chung, S.; Lerner, E.; Jin, Y.; Kim, S.; Alhadid, Y.; Grimaud, L. W.; Zhang, I. X.; Knobler, C. M.; Gelbart, W. M.; Weiss, S. The Effect of Macromolecular Crowding on Single-Round Transcription by

Escherichia Coli RNA Polymerase. *Nucleic Acids Res.* **2019**, *47* (3), 1440–1450.
<https://doi.org/10.1093/nar/gky1277>.