

**Characterization and Validation of Specific Nucleocytoplasmic RNA-Binding
Proteins**

MASTER THESIS

In partial fulfillment to get the academic degree

„Master of Science in Engineering“

Master program

„Biotechnology“

Management Center Innsbruck

Supervisor

Prof. Dr. Thomas Tuschl

The Rockefeller University, New York

Internal supervisor

Prof. Dr. Katrin Bach

Author

B.Sc. Kathrin Freisl

1510352003

Declaration

I hereby declare, under oath, that this master thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extent of the original quotes is indicated.

The paper has not been submitted for evaluation to another examination authority or has been published in this form or another.

Innsbruck, 29/09/2017

Acknowledgement

At this point I want to thank everyone who supported and guided me throughout the time I worked on this thesis.

First of all I would like to express my gratitude to Prof. Dr. Thomas Tuschl for giving me the opportunity to work in his laboratory and to make so many new and great experiences. I am thankful for his guidance, invaluable constructive criticism and friendly advice during my work.

I want to express my warmest thanks to Dr. Cindy Meyer who supported me during the whole time with an enormous kindness and sympathy. I truly appreciate her sharing all her experience with me, spending so much time on helping me and giving me the confidence to work on her project. Thank you so much for the great supervision.

Furthermore I want to thank the whole Tuschl Lab for integrating me to be part of this productive research environment. My special thanks go to Aitor Garcia and Marlene Quintas Pinho for all their help, support and kindness in the daily lab work.

I would also like to thank Prof. Dr. Bach for supervising this thesis and her support to finish my master's degree in New York.

Finally I am sincerely grateful to my family and Martin who continuously supported and believed in me.

Abstract

RNA-binding proteins (RBP) function in every aspect of RNA metabolism and determine the fate of an RNA molecule from its synthesis to its decay. 1,542 RBP are encoded by the human genome that dynamically interact with their specific RNA targets to control post-transcriptional gene regulation (PTGR) at many levels including mRNA translation, modification and decay. About 700 RBPs have been suggested to bind mRNAs and are grouped into 400 families of mRBPs with up to 10 family members. In order to understand the function of an mRBP, it is important to characterize the function of all family members since functional overlap and compensation can be assumed. The development of powerful methods, such as label-free protein quantification or RNA sequencing, to detect and quantify expressed proteins in the cell as well as to determine binding sites of target mRNAs at nucleotide resolution, allow scientists to gain a comprehensive understanding of the function of RBPs in target RNA metabolism. This is very important since dysfunction of several RBPs causes defects in PTGR pathways or leads to diseases, such as neurodegenerative diseases or cancer. Although a large variety of RBPs are subject of past and current research, sufficient understanding of the function and target specificities of the majority of mRBP families is often lacking.

The G3BP1 mRBP family is one of those and comprises two members, namely G3BP1 and G3BP2. The aim of this thesis was to identify the molecular targets of G3BP1 proteins as well as their function in target RNA metabolism. In order to do so, we first generated CRISPR/Cas9-mediated single G3BP1 or G3BP2 knockout HEK293 cells as well as G3BP1 and G3BP2 double knockout HEK293 cells. We observed defects in global protein translation and cellular growth if G3BP1 and G3BP2 function was lost. Furthermore, western blotting analysis revealed the activation of the eukaryotic translation initiation factor 2 (EIF2S1) indicating the induction of a cellular stress response. To identify RNA targets, we performed Photoactivatable Ribonucleoside Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) for G3BP1. G3BP1 preferentially targeted mRNAs across CDS and 3' UTRs by its RNA-binding domain. The ability to bind target RNAs was further essential for G3BP1 to localize to cytoplasmic stress granules (SGs). G3BP1 and G3BP2 double knockout cells lost the capability of SG assembly when exposed to severe environmental stress, such as heat shock or sodium arsenite treatment.

For future studies, it will now be of high interest to elucidate the detailed mechanism how G3BP1 proteins modulate the translation of target mRNAs and whether the phosphorylation of EIF2S1 contributes to detected translational issues.

Kurzfassung

RNA-bindende Proteine (RBP) sind an allen Aktionen des RNA Metabolismus beteiligt und bestimmen so das Schicksal der RNA Moleküle von der Synthese bis hin zu deren Abbau. Das menschliche Genom codiert 1.542 RBPs, welche durch die dynamische Interaktion mit spezifischen RNA-Fragmenten post-transkriptionale Genregulation (PTGR) auf vielen Ebenen wie der mRNA Translation, der Modifikation oder dem Abbau regulieren. Es wird angenommen, dass etwa 700 dieser RBP spezifisch an mRNA binden, welche wiederum in 400 mRBP Familien untergliedert sind und bis zu 10 Familienmitglieder haben können. Um die Funktion eines einzelnen mRBP zu verstehen ist es wichtig, alle Familienmitglieder zu charakterisieren, da eine gegenseitige funktionelle Bekräftigung oder Kompensation möglich ist. Die Entwicklung leistungsstarker Methoden, wie Label-freie Protein-Quantifizierung oder RNA-Sequenzierung zur Detektion und Quantifizierung von Proteinen ermöglicht ein umfassendes funktionelles Verständnis von RBPs.

Dies ist deshalb so wichtig, da Fehlfunktionen von RBPs zu Defekten in der PTGR führen und neurodegenerative Erkrankungen oder Krebs auslösen können. Obwohl eine Vielzahl von RBPs Gegenstand vergangener und aktueller Forschung sind, fehlt bei der Mehrheit der mRBP-Familien ein hinreichendes Verständnis. Die G3BP1-mRBP Familie ist eine der unzureichend Beschriebenen. Sie besteht aus zwei Mitgliedern: G3BP1 und G3BP2.

Ziel dieser Arbeit war es, die molekularen Substrate von G3BP1 sowie deren Funktion im mRNA-Metabolismus zu identifizieren. Um dies zu erreichen, wurden mittels CRISPR/Cas9 Technologie G3BP1- oder G3BP2 Einzelknockout sowie G3BP1- und G3BP2-Doppel-Knockout-HEK293-Zellen erzeugt. G3BP1 und G3BP2 defiziente Zellen zeigten verminderte Proteintranslation und zellulären Wachstum. Darüber hinaus konnte durch Western Blot Analyse die Phosphorylierung des eukaryotischen Translationsinitiationsfaktors 2 (EIF2S1) nachgewiesen werden, was die Aktivierung einer zellulären Stressantwort anzeigt. Um RNA-Substrate zu identifizieren, wurde Photoactivatable Ribonucleoside Enhanced Crosslinking und Immunoprecipitation (PAR-CLIP) angewandt. Dieser zeigt, dass G3BP1 durch sein RNA bindendes Motiv (RRM) mRNAs bevorzugt in der CDS und 3'UTRs bindet. Auch für die Anlagerung von G3BP1 in cytoplasmatischen Stressgranulen (SG) war die Bindung zu mRNAs essentiell. Für zukünftige Studien ist es nun von großem Interesse, den Mechanismus zu beschreiben, auf dem G3BP1 mRNA Translation beeinflusst und welche Rolle die Aktivierung von EIF2S1 dabei spielt.

List of content

1	Introduction.....	1
1.1	RNA-binding proteins in post-transcriptional gene regulation.....	1
1.2	The Ras-GTPase-activating protein SH3-domain-binding protein 1 (G3BP1) RNA-binding protein family	2
1.2.1	Structure and localization of the G3BP1 protein family	2
1.2.2	Suggested functions of the G3BP1 protein family	2
1.2.3	G3BP1 protein family in human diseases.....	4
1.3	Specific aims of this thesis	4
2	Material.....	5
3	Methods.....	11
3.1	Cell cultivation	11
3.2	Generation of stable Flp-In T-REx HEK293 cell lines and Plasmids	11
3.3	CRISPR/Cas9	12
3.4	Growth Curves	12
3.5	SDS-PAGE and Western Blot	12
3.6	Total RNA extraction	13
3.7	PAR-CLIP	13
3.8	Fluorescence microscopy.....	14
3.9	Flag-IP and Mass Spectrometry.....	15
3.10	Polysome profiling	15
3.11	Ribosome profiling.....	16
4	Results.....	17
4.1	G3BP1 protein binds mRNA in CDS and 3'UTR.....	17
4.2	CAPRIN1 and USP10 are direct interactors of G3BP1.....	20
4.3	Double knockout of G3BP1 and G3BP2 causes cell growth defects.....	22
4.4	G3BP1 proteins are essential for stress granule formation	23
4.5	Double knockout of G3BP1 and G3BP2 causes translational repression	25
4.6	Double knockout of G3BP1 and G3BP2 causes EIF2S1-mediated stress response	26
4.7	RRM domain is crucial for RNA-binding and functionality of G3BP1.....	28

5	Discussion	30
5.1	G3BP1 protein binds mRNA in CDS and 3'UTR	30
5.2	CAPRIN1 and USP10 are direct interactors of G3BP1	30
5.3	G3BP1 proteins are essential for stress granule formation	32
5.4	Double knockout of G3BP1 and G3BP2 causes translational repression	32
5.5	Double knockout of G3BP1 and G3BP2 causes EIF2S1-mediated stress response	34
5.6	RRM domain is crucial for RNA-binding and functionality of G3BP1	35
5.7	Conclusion.....	35
6	Outlook	36
	List of figures	IX
	List of tables	IX
	List of references.....	X

1 Introduction

1.1 RNA-binding proteins in post-transcriptional gene regulation

Mechanisms of post-transcriptional gene regulation (PTGR) control the expression of genes at many levels, including RNA splicing, RNA transport, mRNA stability, translation, modification and decay ¹. RNA-binding proteins (RBP) are important regulators of the RNA metabolism addressing different processes from its synthesis to its decay. RBPs often bind thousands of RNAs in cells at specific target sites what is important for their functionality. In total there are 1,542 human RBPs encoded in the genome that act in the most diverse intracellular processes by interacting with all categories of RNAs ². According to their RNA binding sites, RBPs are subdivided in groups interacting with the same RNA category, such as rRNAs, mRNAs, tRNAs, small nuclear RNAs or microRNAs. This is a useful tool, since RBPs of the same subclass often show similar functions in terms of PTGR pathways or lead to similar phenotypes in human diseases like cancer, viral infections or neurodegeneration ^{2,3}. The most abundant group across all tissues, are mRNA-binding proteins (mRBPs) predominantly regulating mRNA maturation, localization, and degradation. The about 700 individual mRBPs are assigned to 400 families ².

One of them is the Ras-GTPase-activating protein SH3-domain-binding protein 1 (G3BP1) family that has two family members: G3BP1 and G3BP2 ⁴.

RBPs contain at least one of the approximately 600 different RNA-binding domains (RBDs) known for their abilities to interact with RNA ². According to their RBDs, RBPs are further subdivided as the domain can augur some insights towards RNA targets and binding preference². By far the most abundant RBD in mRBP is the RNA recognition motif (RRM) found in 405 mRBPs out of 692 ². Further frequently occurring RBDs are the K homology (KH) domain, the DEAD motif, double-stranded RNA-binding motif (DSRM) or zinc-finger domain, ⁵.The RRM domain comprises approximately 80-90 amino acids and recognizes single-stranded target RNAs in a size window of about 4-8 nucleotides. Through stacking interactions as well as electrostatic and hydrogen bonds at the surface of β -sheet RRM domains interact with their target RNA ⁵. As the RNA-binding ability of a single RBD is limited by their short recognition sequence, most RBP contain more than one RBD. Other than possibly expected there are relatively few RBDs that cover an enormous functional range of RBPs and their RNA-binding targets by a variety of their combination ³.

1.2 The Ras-GTPase-activating protein SH3-domain-binding protein 1 (G3BP1) RNA-binding protein family

1.2.1 Structure and localization of the G3BP1 protein family

The Ras-GTPase-activating protein SH3-domain-binding protein 1 (G3BP1) family of mRBPs comprises two family members: G3BP1 and G3BP2⁴. G3BP1 is 466 amino acids (aa) long, G3BP2 482 aa. Both proteins contain two domains, one N-terminal nuclear transport factor 2 (NTF2) and one C-terminal RRM. The NTF2 domain is also found in few other RBP including NTF2 itself, the highly-related export receptor TAP and its co-factor p15 which regulate nuclear export of mRNA⁶. Although both G3BP1 family members mainly localize to the cytoplasm^{4,7}, the NTF2 domain allows G3BP1 proteins to shuttle between nucleus and cytoplasm⁸. Therefore, G3BP1 and G3BP2 are suggested to be mRNA shuttling proteins.

1.2.2 Suggested functions of the G3BP1 protein family

When G3BP1 protein family was first described, it was suggested to function in Ras signal transduction pathways by its interaction with RasGAP⁴. This function however was revisited by several studies including experiments on Rasputin (Rin), the *Drosophila melanogaster* homologue of G3BP1^{9,10}.

In the early 2000, G3BP1 was thought to comprise intrinsic endonuclease activity and to regulate mRNA stability (Tourriere 2001). Furthermore, various additional functions of G3BP1 proteins have been described, such as its DNA/RNA helicase activity¹¹, the regulation of viral infection through the interaction with CAPRIN1¹², as well as its role in protein ubiquitination through its association with ubiquitin specific peptidase 10 (USP10)¹³. Protein ubiquitination is a common strategy of post-translational modification to regulate the turnover of specific proteins by changing their stability, localization or activity¹⁴. Therefore specific ubiquitin-conjugated enzymes bind single or poly ubiquitin nucleotides covalently to lysine residues of the targeted protein to label them for further treatment. One of the best described pathways addresses the addition of an ubiquitin residue as a signal for proteasome digest like it is shown for transcription factor NF- κ B, a translation inhibitor¹⁵. But also target signaling to control diverse physiological and pathological processes¹⁶ or the transient inactivation of specific proteins are common mechanism triggered by ubiquitinating enzymes¹⁴. Contrary to mentioned ubiquitinating enzymes, there are also deubiquitinating enzymes that catalyze the removal of ubiquitin residues from tagged proteins by hydrolysis^{17,18}.

In case of G3BP1 being a direct interactor to USP10, a regulatory effect of G3BP1 protein family on protein turnover is likely, albeit in an indirect manner.

Another study by Tourriere and co-workers suggested G3BP1 to be involved in mRNA degradation as it functions as phosphorylation-dependent endoribonuclease that cleaves sequence specific at the 3' UTR of c-myc mRNAs⁸. As G3BP1 phosphorylation and activity was shown to be cell cycle dependent, G3BP1 was assigned to function as growth factor sensor effecting mRNA stability and translation⁹. Additionally, G3BP1 was shown to directly interact with the mRNA encoding the mitochondrial H⁺-ATP synthase subunit β (β -F1-ATPase)¹⁹ and inhibit the translational initiation of β -F1-ATPase suggesting the G3BP1 proteins also play a role in regulation of protein translation.

The best and most-detailed described function of G3BP1 proteins is their role in stress granule (SG) formation²⁰⁻²². SG formation describes a mechanism to regulate protein synthesis when translation initiation needs to be stalled as a consequence of environmental stress like viral infections, heat stress or nutrition starvation^{23,24}. SG formation is mostly regulated by the phosphorylation of serine residue 51 at the α -subunit of eukaryotic translation initiation factor (EIF2S1/eIF2 α). EIF2S1 in turn is regulated by four stress-sensing serine/threonine kinases that are activated by various stress conditions: EIF2AK1/HRI is phosphorylated by oxidative stress or heat shock, EIF2AK2/PKR is activated according to dsRNA, EIF2AK3/PERK senses ER stress and EIF2AK4/GCN2 is triggered by amino acid starvation or UV light²⁵. Phosphorylated EIF2S1 prevents the assembly of the EIF2/GTP/tRNAⁱMet ternary complex by inhibiting efficient GDP-GTP exchange. Initiator tRNAⁱMet can no longer bind to the 40S ribosomal subunit causing stalled translation inhibition²⁶. In a dynamical process connected with the disassembly of translating polysomes, that run off of mRNA, untranslated mRNAs, translation initiation components, the 40S ribosomal subunit and many mRBPs like G3BP1 accumulate to a cytoplasmic RNA-RBP complex, namely SG²⁴. Besides the common pathway of SG formation triggered by the phosphorylation of EIF2S1, there are also EIF2S1 independent pathways triggered by the control of mechanistic target of rapamycin (mTOR) or by the interaction with the eIF4F complex consisting of eIF4E, eIF4A and eIF4G²⁶.

The G3BP1 protein family has already been described rudimentarily in various biological mechanisms like mRNA metabolism, signal transduction and cell cycle regulation. However, next to the role in SG formation, the actual function of G3BP1s is still poorly understood wherefore the characterization of this protein family is of great interest.

1.2.3 G3BP1 protein family in human diseases

G3BP1 protein family is assumed to be implicated in several pathways including cell proliferation²⁷, signal transduction²⁸ and mRNA metabolism⁸ that are known to be crucial in cancer biology²⁹, for what reason G3BP1 protein family may state a promising target for cancer therapy. G3BP1 was shown to be overexpressed in various cancer types, including breast, colon and pancreas^{27,30,31}. Further G3BP1 is described to supported proliferation by specifically interaction with peripheral myelin protein 22 (PMP22). Also G3BP1 mediated increased S phase entry was measured in breast cancer cells²⁷. Further studies in mouse embryonic fibroblasts and during development revealed altered expression of essential growth factors dependent on G3BP1 depletion. In consequence G3BP1 knockout in mice resulted in embryonic lethality and growth retardation³².

Increasing evidence based on advances in molecular genetics links RBP triggered stress response and SG formation to a variety of neurodegenerative diseases, such as amyotrophic lateral sclerosis (ALS), Huntington's disease and frontotemporal lobar degeneration³³⁻³⁵. Increased reactivity of G3BP1 and formation of aberrant SGs was found dependent on the disease state in whole brain lysates of JNPL3 mouse amygdala³⁶.

These examples implicate G3BP1 in the development of a promising therapeutically target. Especially in the field of SG-related neurodegenerative disease G3BP1 could contribute to the understanding of disease progression and alike operate as a target.

1.3 Specific aims of this thesis

This thesis aims to better characterize the function of the G3BP1 mRBP family in post-transcriptional gene regulation. To do so, single G3BP1 and G3BP2 knockouts as well as double knockout HEK293 cells should be generated and consequences on cellular viability and cellular stress responses will be monitored using a variety of cell-based assays, such as monitoring the growth of the generated knockout cells, as well as microscopy approaches. Furthermore, to study the role of G3BP1 proteins in target mRNA metabolism, the transcriptome-wide RNA targets of G3BP1 will be identified using Photoactivatable Ribonucleoside Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) and the effects on target mRNA metabolism will be studied using polysome as well as ribosome-profiling approaches. Finally, interaction partners of G3BP1 will be identified by Immunoprecipitation (IP) combined with mass spectrometry (MS) to gain further information on biochemical processes and signaling pathways in which G3BP1 proteins might be involved in.

2 Material

Table 1: Chemicals and materials

Chemical	Commercial name	Manufacturer
BSA	Bovine serum albumin	Sigma Aldrich
$C_{12}H_{22}O_{11}$	Sucrose	Gibco
$C_{24}H_{40}O_4$	Sodium deoxycholate	Sigma Aldrich
C_3H_8O	2-Propanol	Thermo Fisher Scientific
$C_4H_{10}O_2S_2$	Dithiothreitol	Sigma-Aldrich
DMEM	Dulbecco's Modified Eagle's Medium	Thermo Fisher Scientific
EDTA-free protease inhibitor cocktail		Roche
EtOH	Ethanol	Decon Laboratories Inc
FBS	Fetal bovine serum	Sigma Aldrich
HCl	Hydrogen chloride	Thermo Fisher Scientific
HEPES	4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid	Thermo Fisher Scientific
KCl	Potassium chloride	Thermo Fisher Scientific
KOH	Potassium hydroxide	Gibco
MOVIOL	17951	Polysciences Inc.
$MgCl_2$	Magnesium chloride	Thermo Fisher Scientific
NaCl	Sodium chloride	Thermo Fisher Scientific
Nitrocellulose Membran (0,45 μ m)		Amersham
NuPAGE Novex 4-12% Bis-Tris Protein ge		Thermo Fisher Scientific
NuPAGE MOPS SDS Running Buffer		Thermo Fisher Scientific
NuPAGE Novex 10% Bis-Tris Midi Protein Gels		Thermo Fisher Scientific
PFA (16% w/v)	Para formaldehyde	Thermo Fisher Scientific
SDS	sodium dodecyl sulfate	Sigma Aldrich

Chemical	Commercial name	Manufacturer
Tris-HCl		Sigma Aldrich
TRIzol Reagent		Life Technologies
Trypan blue dye (0,4%)		Bio-Rad
Tween20		Sigma Aldrich

Table 2: Buffers and solutions

Chemical	Ingredients	Final concentration
1x Transfer buffer	10x SLAB 4 buffer	10%
	Methanol	20%
	Water	70%
1x TBS-T	Tris-HCl	50.0 mM
	NaCl	150.0 mM
	Tween20	0.1%
10x SLAB 4 buffer	Tris-HCl	250 mM
	Glycine	1.9 M
	SDS	0,5%
Hypotonic buffer	Tris-HCl	5 mM
	MgCl ₂	2.5 mM
	KCl	1.5 mM
	DTT	2 mM
	Triton X-100	0.5% v/w
	Sodium deoxycholate	0.5% v/w
NP40 lysis buffer	HEPES-KOH	50 mM
	KCl	150 mM
	MgCl ₂	2 mM
	NP40	2% v/v
	DTT	0.5 mM

Chemical	Ingredients	Final concentration
RIPA lysis buffer	Tris-HCl	50 mM
	NaCl	150 mM
	NP40	1%
	Sodium deoxycholate	0.5%
	SDS	0.1%
SDS loading dye	SDS	1.7%
	Glycerol	5%
	Bromophenol blue	0.002%
	Tris-HCl	60 mM
	DTT	100 mM
Washing buffer	HEPES-KOH	50 mM
	KCl	300 mM
	MgCl ₂	2 mM
	NP40	0,5% v/v

Table 3: Enzymes and other materials

Chemical	Catalog number	Manufacturer
anti-FLAG-M2 magnetic beads	M8823	Sigma Aldrich
Blasticidin	Ant-bl-5	Sigma Aldrich
Calf intestinalphosphatase (CIP)	M029	New England Biolabs
Cycloheximide	C4859	Sigma Aldrich
FlagPeptide	F3290	Sigma Aldrich
Glycoblue	AM9516	Ambion
Hygromycin	Ant-hg-5	Invivogen
Lipofectamin	M04500	Sigma Aldrich
L-Glutamin	25030081	Thermo Fisher Scientific
pDONR221 plasmid	12536017	Thermo Fisher Scientific

Chemical	Catalog number	Manufacturer
Penicillin-streptomycin	M04300	Sigma Aldrich
pENTR4 plasmid	A10465	Thermo Fisher Scientific
Precision Plus Protein Dual Color Standards	1610374	BioRad
Proteinase K	03115879001	Roche
RNase I	AM2294	Invitrogen
SUPERase In	AM2694	Invitrogen
T4 Polynucleotide Kinase (T4 PNK)	M0201	New England Biolabs
T4 Rnl1	EL0021	Thermo Fisher Scientific
T4 Rnl2(1-249) K227Q		Addgene
Zeocin	Ant-zn-5	Invivogen
γ - ³² P-ATP	NEG002Z001MC	Perkin Elmer

Table 4: Antibodies

Antibodies	Dilution	Catalog number	Manufacturer
Primary antibodies			
Mouse anti- β -tubulin	1:1000	T4026	Sigma Aldrich
Rabbit anti-RPS19	1:1000	Ab57643	Abcam
Mouse anti-HA	1:2000	501030108	Santa Cruz
Mouse anti-G3BP1	1:1000	Ab56574	Abcam
Rabbit anti-G3BP1	1:2000	AV37713	Sigma Aldrich
Rabbit anti-G3BP2	1:1000	Ab86135	Abcam
Rabbit anti-RPL7a	1:1000	Ab70753	Abcam
Mouse anti-EIF4E	1:1000	Ab1126	Abcam
Mouse anti-P-EIF2SI	1:1000	Ab32157	Abcam

Antibodies	Dilution	Catalog number	Manufacturer
Secondary antibodies			
Goat anti-mouse	1:1000	P0047	Dako
Goat anti-rabbit	1:1000	P0447	Dako
Fluorescent labeled secondary antibodies			
Goat anti-mouse-IgG-Alexa Fluor 488	1:1000	A32723	Invitrogen
Goat anti-rabbit-IgG-Alexa Fluor 546	1:1000	A11010	Invitrogen

Table 5: Commercial kits

Kit	Catalog number	Manufacturer
GATEWAY BP recombinase system	11789020	Thermo Fisher Scientific
GATEWAY LR recombinase system	11791020	Thermo Fisher Scientific
Pierce™ BCA Protein Assay Kit	23225	Thermo Fisher Scientific
Pierce™ BCA Protein Assay Kit	23225	Thermo Fisher Scientific
QIAprep Spin Miniprep Kit	27104	Qiagen
QIAquick Gel Purification Kit	28704	Qiagen
SuperScript III Reverse Transcriptase	18080-044	Thermo Fisher Scientific
Zero Blunt PCR Cloning Kit	K270040	Thermo Fisher Scientific

Table 6: Cell line

Cell line	Commercial name	Manufacturer
Flp-In T-REx HEK293	R78007	Thermo Fisher Scientific

Table 7: Equipment

Equipment	Manufacturer
Amersham Imager 600	GE
Analyst GT Multimode Reader	Molecular Devices
Axiovert 200 inverted microscope	Zeiss
BBD6220 Co2 Incubator	Thermo Fisher Scientific
Centrifuge 5804	Eppendorf
Discovery 90SE Centrifuge	Sorvall
NanoDrop 2000	Thermo Fisher Scientific
Precision 280-series water bath	Thermo Fisher Scientific
Semi-dry Blotter	BioRad
Sorval Legend Micro 21 R Centrifuge	Thermo Fisher Scientific
TC20 Automated Cell Counter	BioRad
TH-641 Swinging Bucket Rotor	Sorvall
Thyphoon FLA 7000	GE
Vortex-Genie 2	Scientific Industries
VS110	Olympus

3 Methods

3.1 Cell cultivation

Human kidney Flp-In T-REx HEK293 cells (Thermo Fisher Scientific, R78007) cultivated in high glucose Dulbecco's Modified Eagle's Medium (DMEM, Thermo Fisher Scientific, 11965118) containing 10% v/v FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine, 15 µg/ml blasticidin and 100 µg/ml zeocin were grown at 37°C and 5% CO₂-atmosphere and threatened under sterile conditions. Cell lines cloned to inducible express FlagHA-G3BP1 or FlagHA-G3BP1-ΔRRM truncation mutants were grown using selective media supplemented with 100 µg/ml hygromycin. For protein expression, 1 µg/ml doxycycline was added for 48 h.

3.2 Generation of stable Flp-In T-REx HEK293 cell lines and Plasmids

The generation of a stable cell line expressing FlagHA-G3BP1-NTF2, FlagHA-G3BP1-ΔRRM, FlagHA-G3BP1-RRM in Flp-In T-REx 293 cells, was operated according to a previous published protocol³⁷. Therefore, pENTR4 plasmid (Thermo Fisher Scientific, A10465) of DNA fragments encoding the amino acids sequence from aa 1 until aa 341 of human G3BP1, termed gene of interest (GOI), was amplified by PCR technology. Sequence specific primers lined by attB1 and attB2 recombination sites were used:

Table 8: Primer for BP reaction

Primer	Sequence
G3BP1_NTF2_fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCatgggatggagaagcctagtc
G3BP1_NTF2_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtcatcagtatctgaagatatcattgtgaaca
G3BP1_ΔRRM_fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCatgggatggagaagcctagctccc
G3BP1_ΔRRM_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtcatgggtgactgacgggtgtctcacc
G3BP1_RRM_fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCctctcattggcaacctgcctcatgaagt
G3BP1_RRM_rev	GGGGACCACTTTGTACAAGAAAGCTGGGTtcaactgccgtggcgcaagccc

pDONR221 plasmid (Thermo Fisher Scientific, 12536017) was applied for recombination of the purified PCR product by the GATEWAY BP recombinase system (Thermo Fisher Scientific, 11789020). Using the GATEWAY LR recombinase system (Thermo Fisher Scientific, 11791020), pENTR4-GOI plasmids were recombined with the pFRT/TO/FLAGHA-DEST destination vector.

As a screening for positive clones, cells were grown in selective culture media supplemented with 100 µg/ml hygromycin followed by western blot and sequencing verification. The expression of FlagHA-tagged protein was achieved by the addition of 1 µg/ml doxycycline for 48 h.

3.3 CRISPR/Cas9

To generate single knockout of G3BP2 (G3BP2-KO) in cells and G3BP1/2-DKO in previous generated Flp-In T-REx HEK293 G3BP1-KO, CRISPR/Cas9 technique was used as described in ^{38,39}. 100,000 cells were seeded into one well of a 6-well-plate and incubated over night at 37°C and 5% CO₂. On the following day, cells were transfected with a previous designed pSpCas9(BB)-2A-GFP plasmid containing the guide sequence for G3BP2-KO. After an additionally incubation for 24 h, GFP-positive cells were detected by FACS cell sorting according to GFP-Hi signal. Each transfected cell was placed in one well of a 96-well plate and further cultivated by adding 50 µl fresh media until a colony was detectable. As protein depletion screening, western blot analysis was performed. For clone detection on genomic level, PCR products were cloned according to the Zero Blunt PCR Cloning Kit (Thermo Fisher Scientific, K270040) and 4 to 6 clones send for Sanger sequencing.

3.4 Growth Curves

200,000 parental Flp-In T-REx HEK293 cells and CRISPR/Cas9 mediated mutants were seeded into one well of a 6-well-plate at day 0 and cultivated in the corresponding cell culture media until 90% confluency. Every 2 days, cells were counted and transferred to a new culture dish containing fresh media. If necessary, they were expanded to 10 or 20 cm dishes.

3.5 SDS-PAGE and Western Blot

For protein detection, cells were collected by centrifugation at 500xg for 5 min at 4°C and lysed using 50 µl RIPA lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS and EDTA-free protease inhibitor cocktail (Roche, 04693159001)) for 10 min on ice. Subsequently, the lysate was separated by centrifugation at 13,000xg for 10 min at 4°C before determining the protein concentration using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, 23225).

For sample preparation, 30 µg protein diluted in 15 µl dest.H₂O was adjusted in premixed SDS loading dye (1.7% SDS, 5% glycerol, 0.002% bromophenol blue, 60 mM Tris-HCl pH 6.8, 100 mM DTT), heated to 95°C for 1 min and loaded to a NuPAGE™ Novex™ 10% Bis-Tris Midi Protein Gels (Thermo Fisher Scientific, WG1202). The gel was further blotted on a 0.45 µm pore size nitrocellulose membrane using a semi-dry blotter (Bio-Rad) for 1 h at 300 mA. Blocking was performed for 1 h using 5% BSA solution in 1x TBS-0.1% Tween20 (Sigma-Aldrich, P1379) followed by three wash steps using 1x TBS-0.1% Tween. The specific primary antibody incubation was taken place over night at 4°C. The membrane was then washed three times in 1x TBS-0.1% before the corresponding secondary antibody, conjugated to horseradish peroxidase, was added and incubated for 1 h at RT. The protein detection was performed on an Amersham Imager 600 (GE) using chemiluminescence reagent.

3.6 Total RNA extraction

TRIzol Reagent (Thermo Fisher Scientific, 15596018) was used as described in the manufacturer's instructions to isolate total RNA from Flp-In T-REx 293 cells (Thermo Fisher Scientific, R78007) and CRISPR/Cas9-mediated G3BP1/2-DKO cells with FlagHA-tagged G3BP1 after 0, 1, 3 and 6 days of 1 µg/ml doxycycline treatment to detect specific mRNA targets. Sequencing was performed on an Illumina HiSeq 2500 platform at The Rockefeller University Genomics Resource Center.

3.7 PAR-CLIP

For the identification of RNA targets of G3BP1 protein, 4-thiouridine (4SU) photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) was performed following a previous published protocol⁴⁰. Cells were expanded to 20 20 cm plates and incubated with the photoactivatable nucleoside 4SU for 16 h to incorporate the photoactivatable nucleosides into nascent transcripts. Labeled cells were further UV-crosslinked by 365 nm. Following cell lysis a single RNase A digestion step was performed and protein-RNA complexes immunoprecipitated using pre-conjugated anti-FLAG-M2 magnetic beads (Sigma, M8823). Further 5' phosphorylation and 5'-³²P-radiolabeling enables the detection of RNA using autoradiography. Detected band representing the RNA-Protein complex was cut and treated by a proteinase K digest. Recovered RNA was then size-fractionated on a 15% UREA-Gel and by applying 19nt and 35nt size markers, corresponding RNA-fragments were isolated. cDNA library was generated by 3' and 5' adapter ligation followed by a reverse transcriptase step and PCR amplification.

Deep sequencing was performed on Illumina HiSeq 2500 at The Rockefeller University Genomics Resource Centre. For analysis PAR-CLIP suite⁴⁰ and PARalyzer software⁴¹ were used.

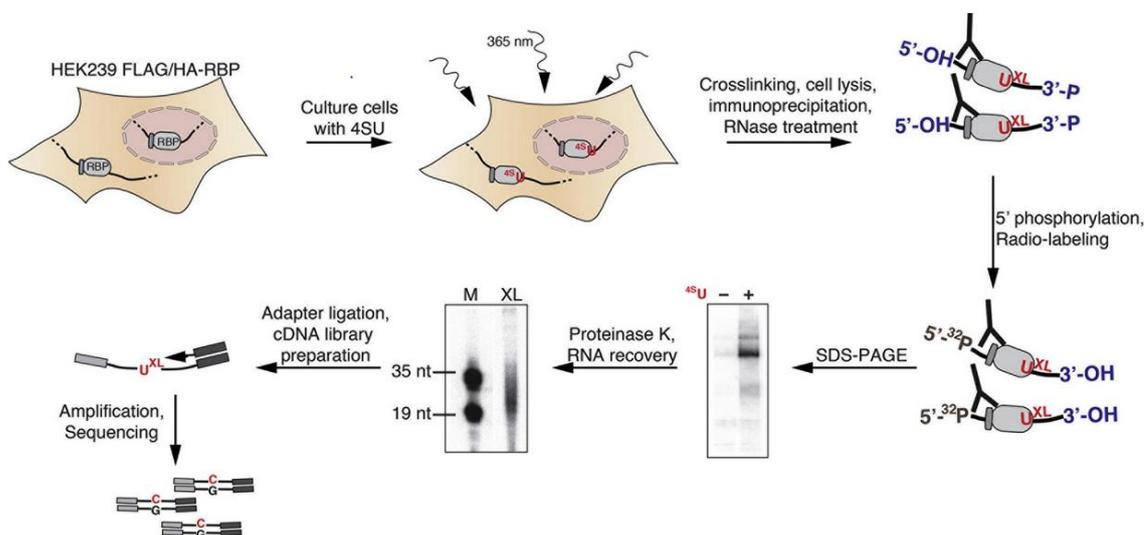


Figure 1: Workflow of a PAR-CLIP experiment⁴⁰.

Cells stable expressing FlagHA-G3BP1 were incubated with 4SU to incorporate the photoactivatable nucleosides into nascent transcripts. UV_{365nm}-Crosslinked and RNase A digested lysates were anti-HA-immunoprecipitated. After 5' phosphorylation and 5'-³²P-radiolabeling, RNA-RBP complexes were separated by SDS-PAGE and visualized using autoradiography. Corresponding band of the phosphoimage was subjected to proteinase K digest and recovered RNA size-fractionated on a 15% UREA-gel. Using 19nt and 35nt size-markers, corresponding RNA-molecules were isolated for cDNA library generation.

3.8 Fluorescence microscopy

RNA-Fluorescence *in situ* hybridization (FISH) and Immunofluorescence (IF) experiments were performed according to a previously described protocol⁴² for the simultaneous detection of cellular RNAs and proteins under various stress conditions. 50,000 Flp-In T-REx 293 cells were seeded on an 8-well chamber slide pre-coated with 0,1% gelatine 24 h before they were threatened with either 400 μ M sodium arsenite for 30 min at 37°C (oxidative stress) or incubated at 45°C for 30 min (heat stress) for stress granule assays. Following, cells were fixed by adding 250 μ l 4% PFA in 1xTBS for 12 min at RT. Chamber slides containing the fixed cells were hybridized over night at 40°C with fluorophore-linked polyT-21d29-Atto647N in 1x Hybridization buffer for the detection of polyadenylated mRNAs. After three wash steps, slides were incubated over night at 4°C with corresponding primary antibody for protein detection followed by an overnight incubation at 4°C with fluorophore-labeled secondary antibody and DAPI for nuclei staining.

3.9 Flag-IP and Mass Spectrometry

FlagHA tagged G3BP1 cells from five 20-cm plates were used for Immunoprecipitation (IP) in three replicates. They were harvested by centrifugation at 500xg for 15 min at 4°C and lysed using NP40 lysis buffer (50 mM HEPES-KOH, pH 7.5, 150 mM KCl, 2 mM MgCl₂, 2% v/v NP40, 0.5 mM DTT, 1x complete EDTA-free protease inhibitor cocktail (Roche, 04693159001)). For IP, anti-FLAG-M2 magnetic beads were added to the cell suspension and incubated for 2 h at 4°C on a rotating wheel. Following three wash steps with washing buffer (50 mM HEPES-KOH, pH 7.5, 300 mM KCl, 2 mM MgCl₂, 0.5% v/v NP40), the immunoprecipitates were eluted using Flag peptide 200x abundantly. The samples were separated on NuPAGE Novex 4-12% Bis-Tris protein gel (Thermo Fisher Scientific, NP0322BOX) for a short run of approximately 10 mm and transferred to the Rockefeller High-Throughput and Spectroscopy Resource Center for nano LC-MS/MS analysis. G3BP1-interactors were identified using MaxQUANT.

3.10 Polysome profiling

In terms of G3BP1-having effects on translational regulation polysome profiling was performed. Cells were grown in a 20 cm dish to 90% confluency, pretreated with 100 µg/ml cycloheximide (Sigma Aldrich, C4859) for 5 min and harvested by 500xg centrifugation for 5 min at 4°C. 500 µl hypotonic buffer (5 mM Tris-HCl, pH 7.5, 2.5 mM MgCl₂, 1.5 mM KCl, complete EDTA-free protease inhibitor cocktail (Roche, 04693159001), 100 µg/ml cycloheximide, 2 mM DTT, 0.5% v/w Triton X-100, and 0.5% v/w sodium deoxycholate) were used to lyse the cell pellet followed by centrifugation at 13,000xg for 5 min at 4 °C. The OD_{260nm} of the cleared lysate was measured using NanoDrop 2000 (Thermo Fisher Scientific) and subsequently diluted in hypotonic buffer to a final OD_{260nm} of 2 in 400 µl. Ribosomal bound mRNA was separated by ultracentrifugation on a 10-60% sucrose gradient by 35,000xg for 2 h at 4°C in a Discovery 90SE Centrifuge (Sorvall) on TH-641 Swinging Bucket Rotor (Sorvall).

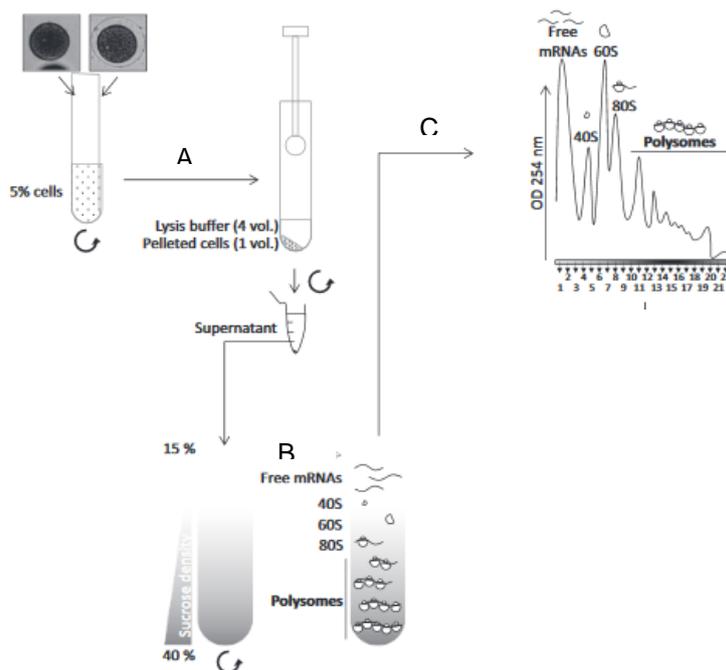


Figure 2: Workflow of a polysome profiling experiment ⁴³.

A) Cells supplemented with 100 µg/ml cycloheximide were harvested and the cleared lysate transferred on a 10-60% sucrose gradient. **B)** RNA fragments were further separated on the 10-60% sucrose gradient by centrifugation at 35,000xg for 2 h. **C)** Monosome and polysome fractionation was detected by OD₂₅₄ absorbance.

3.11 Ribosome profiling

For ribosome profiling, cells were cultivated and lysed as described for polysome profiling in 3.10. After lysate purification using centrifugation at 13,000xg for 5 min at 4 °C, 18 U RNase I (Invitrogen, AM2294) per OD_{260nm} were added and incubated for 30 min on ice to digest mRNA to single fragments observed by only one ribosome. 5 µl of SUPERase In (Invitrogen, AM2694) were used to stop the reaction. 400 µl of the digested RNA were further separated on a 10-60% sucrose gradient by centrifugation at 35,000xg for 2 h and 4°C in a Discovery 90SE Centrifuge (Sorvall) on TH-641 Swinging Bucket Rotor (Sorvall). Following, the monosome fraction was collected and RNA extracted using TRIzol Reagent (Thermo Fisher Scientific, 15596018) according to manufacturer`s protocol. RNA was recovered by ethanol precipitation for 30 min at -20°C followed by centrifugation for 30 min at 15,000xg and 4°C. Samples were separated on a 15% UREA-Gel and stained with ethidium bromide for 5 min. Using 26nt and 34nt size markers, corresponding RNA-fragments were cut and isolated by shaking the gel piece in 400 µl 0.3M NaCl for 2 h at 37°C. cDNA preparation was further performed according to ⁴⁰. Quality analysis was performed at the The Rockefeller University Genomics Resource Centre using High Sensitivity D1000 ScreenTape and TapeStation Analysis Software A.02.01.

4 Results

This thesis addressed the functional characterization of the evolutionary conserved G3BP1-binding protein family comprising two family members: G3BP1 and G3BP2. G3BP1 and G3BP2 proteins show a total amino acid identity of 63% and both members comprise two domains, a N-terminal nuclear transport factor 2 (NTF2) domain with an identity of 81% and a C-terminal RNA recognition motif (RRM) with an identity of 79% (Figure 3A). Based on poly(A) mRNA-seq of Flp-In T-REx HEK293 cells, the mRNA expression of G3BP1 was fivefold higher than G3BP2. RPKM (reads per kilobase exon per million mapped reads) is a method to quantify gene expression by normalizing total read length and the number of RNA sequencing reads⁴⁴. The G3BP1 protein family was suggested to play a pivotal role in cellular stress response and signal transduction pathways⁹. However their function in RNA metabolism remains poorly understood. Using a variety of biochemical and molecular biological methods, such as PAR-CLIP, polysome profiling or RNA-FISH combined with IF; we aimed to get deeper insight into the role of the G3BP1 protein family in target RNA metabolism.

4.1 G3BP1 protein binds mRNA in CDS and 3'UTR

In order to identify the molecular targets of G3BP1 proteins along with binding sites on target mRNAs, we performed 4-thiouridine (4SU) photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP). Therefore, Flp-In T-REx HEK293 cells expressing FlagHA-tagged G3BP1 (FlagHA-G3BP1) under control of a doxycycline (Dox)-inducible promoter were grown in presence of Dox and 4SU, UV-crosslinked and harvested. After cell lysis, FlagHA-G3BP1 was immunoprecipitated using anti-FLAG M2 magnetic beads. Crosslinked immunoprecipitated protein-RNA complexes were radiolabelled and separated by SDS-PAGE. Using phospho-imaging, crosslinked G3BP1-RNA-complexes were detected. The phosphoimage showed a single crosslinking band at around 70 kDa representing crosslinked RNAs in complex with FlagHA-G3BP1 (Figure 3B, upper panel). Western blot analysis of the same samples confirmed that the crosslinking band indeed depicts G3BP1-bound RNA fragments (Figure 3B, lower panel). In a next step, the G3BP1-bound RNA fragments were recovered, converted into a cDNA library and deep sequenced. PAR-CLIP was performed in replicates. Obtained sequencing reads of both replicates were aligned to a laboratory intern annotated human reference transcriptome using PAR-CLIP suite pipeline⁴⁰. From a total of 1.6×10^7 reads (1.3×10^7 for the replicate) the majority mapped to mature mRNAs (Figure 3C, upper panel).

In total 71% of the annotated reads mapped to mRNAs what accredits G3BP1 as mRNA binding protein. This is in agreement with G3BP1 proteins being localized predominantly in the cytoplasm. Sequence reads were assigned to each mRNA-category as perfect matches (d0), with one mismatch caused by a T-to-C conversion (d1 T-to-C), with one mismatch other than a T-to-C conversion (d1 other), or with two mismatches (Figure 3D). 63% of reads mapping to mature mRNAs showed the characteristic thymidine-to-cytidine (T-to-C) conversion indicated by d1 T-to-C. T-to-C conversion events signal specific crosslinking to G3BP1 protein compared to background events, which are detected as d0, d1 other or d2 and have a low percentage of T-to-C converted reads. We used PARalyzer⁴¹ as a further analysis method and found the binding sites were predominantly located in coding sequences (CDS) and 3'untranslated regions (UTRs) (Figure 3C, lower panel). Binding sites in CDS were equally distributed across the whole length other than binding sites in 3'UTRs. Here binding sites are strongly enriched close to the 3' and 5' ends of the 3'UTR as detected by a more detailed analysis which is not based on clustered fragments (Figure 3E).

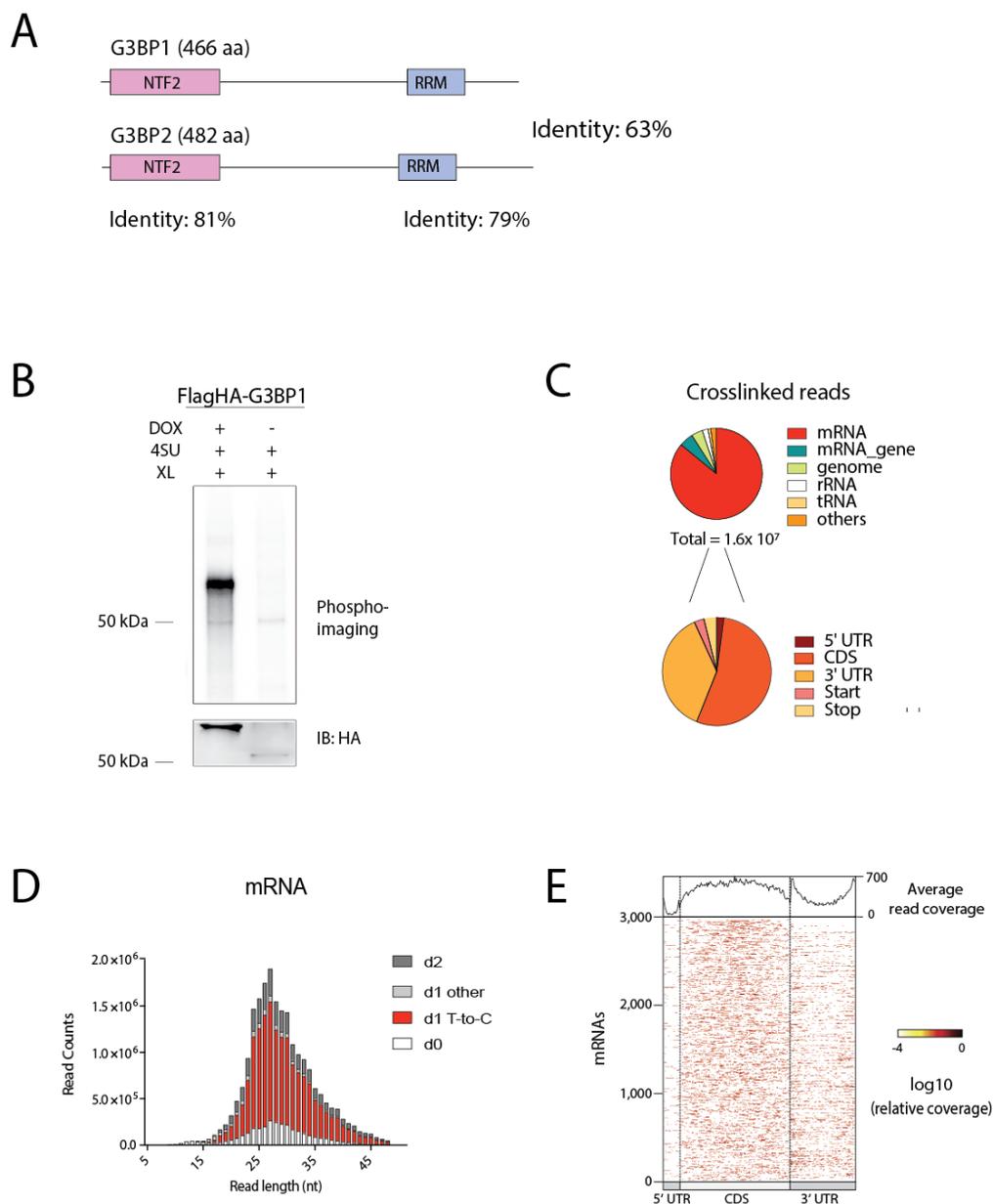


Figure 3: G3BP1 binds target mRNAs in CDS and 3'UTR as determined by PAR-CLIP.

A) Domain structure of human G3BP1 and G3BP2 as described by Pfam. The proteins length is indicated in amino acids (aa) and the total identity as well as the identity of both domains, NTF2 and RRM, is listed. **B)** Autoradiograph of crosslinked and immunoprecipitated RNA-FlagHA-G3BP1 complex from 4SU PAR-CLIP experiment separated by SDS-PAGE. **C)** Relative composition of G3BP1 PAR-CLIP reads crosslinked to various RNA categories with up to two mismatches. mRNA mapped reads are further plotted divided into functional regions. **D)** Reads mapped to mRNA, the most abundant RNA category for G3BP1 PAR-CLIP reads, assigned as error distance 0 (d0, white), error distance 1 with T-to-C conversion (d1 T-to-C, red), other error distance 1 (d1 other, light grey) and as error distance 2 (d2, dark grey). **E)** Upper panel: Average read coverage of the 3,000 most abundant mRNA-mapped reads of G3BP1 PAR-CLIP distributed over functional regions. Lower panel: Meta-gene plot of mRNA-mapped reads representing the relative coverage of one normalized mRNA read in each row.

4.2 CAPRIN1 and USP10 are direct interactors of G3BP1

We next aimed to identify direct protein interactors of G3BP1 for a better insight in biochemical processes, signaling pathways or complexes that are formed with G3BP1. Therefore, FlagHA-G3BP1 expressing HEK293 cells grown in presence of Dox to induce protein expression as well as parental HEK293 cells, serving as a control, were harvested. After cell lysis and RNase I digest, anti-Flag immunoprecipitation was performed using pre-conjugated anti-FLAG M2 magnetic beads. Immunoprecipitated FlagHA-G3BP1 protein complexes were recovered by Flag-peptide elution. The experiment was performed in triplicates including three replicates of anti-Flag immunoprecipitation in FlagHA-G3BP1 expressing HEK293 cells as well as three replicates of anti-Flag immunoprecipitation in parental HEK293 cells. Eluted complexes were separated on a SDS-Gel to remove free Flag-peptide. The gel was subjected for mass spectrometric (MS) analysis to the Proteomics Resource Center at The Rockefeller University. MS data revealed 10 protein candidates that directly interacted with G3BP1 indicated by a LFQ (label-free quantification) enrichment (Figure 4A). Among all interactors, CAPRIN1 (Cytoplasmic activation/proliferation-associated protein-1), USP10 (Ubiquitin specific protease 10) and G3BP2 showed the highest LFQ values. This is in agreement with a recent study published by Kedersha and co-workers where they described G3BP1 in complex with CAPRIN1 and USP10 to mediate (SG) condensation (Figure 4B)²¹.

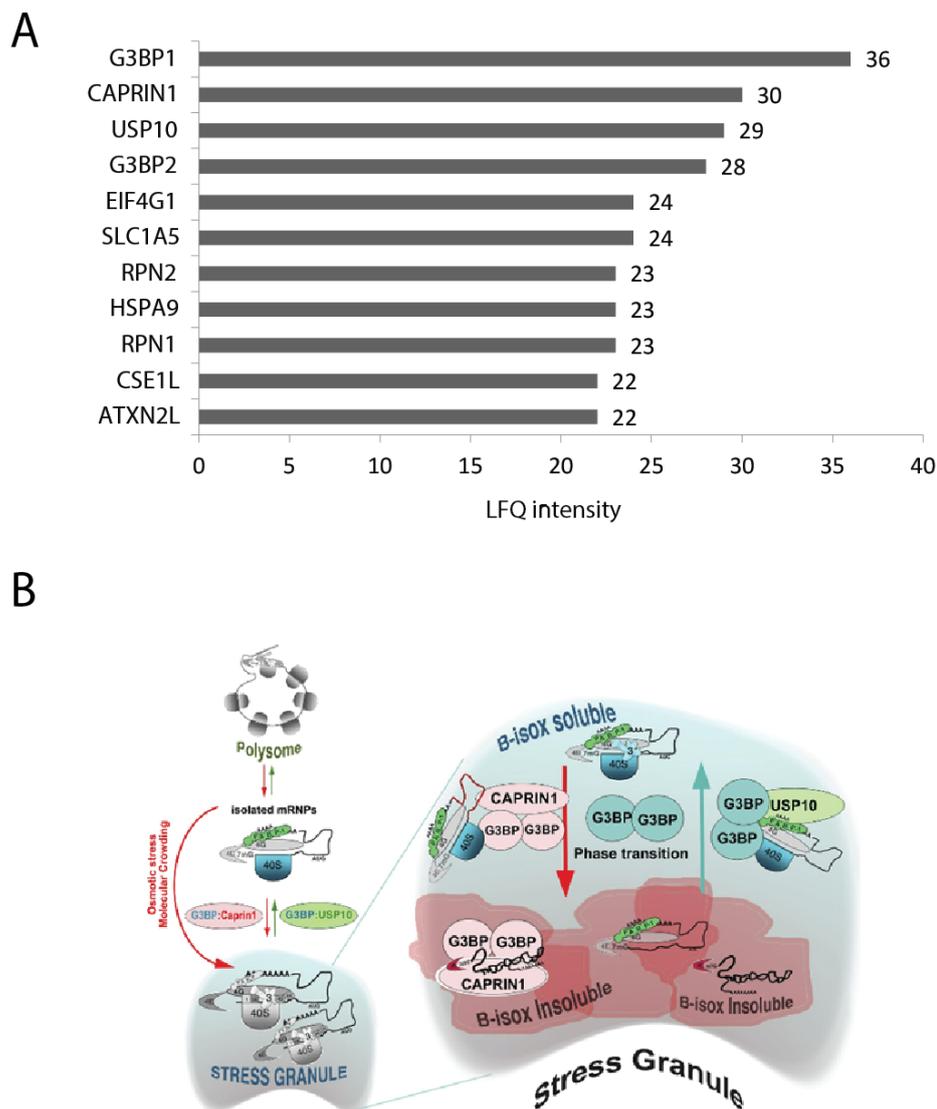


Figure 4: G3BP1 proteins interact with CAPRIN1 and USP10.

A) IP and Mass Spectrometry analysis to detect RNase independent interactors of G3BP1. FlagHA-G3BP1 lysate was immunoprecipitated using anti-Flag-M2 magnetic beads and eluted with Flag peptide. The gene names of the eight most abundant interactors of G3BP1 were listed according to their LFQ enrichment calculated by MaxQUANT. Parental Flp-In cells served as a blank. **B)** Suggested model of G3BP-CAPRIN1-USP10 complex that modulates stress granules formation²¹. CAPRIN1 is assumed to promote SG formation by the interaction with G3BP, whereas USP10 acts as an inhibitor as it causes SG disassembly.

4.3 Double knockout of G3BP1 and G3BP2 causes cell growth defects

To investigate the effect of each family member of the G3BP1 RNA-binding protein on cell growth, CRISPR/Cas9-mediated G3BP2 single knockouts (G3BP2-KO) as well as G3BP1/2 double knockout (G3BP1/2-DKO) were generated in parental Flp-In T-REx HEK293 cells. G3BP1 single knockout cells (G3BP1-KO) used in this project were generated previous to this work according to the same protocol as G3BP2-KO and G3BP1/2-DKO cells were generated. Knockout verification of each cell clone was first assessed by western blotting using human protein-specific antibodies for G3BP1 and G3BP2 (Figure 5A) Next, knockouts were verified on genomic level by sanger sequencing (Figure 5B). After successful established G3BP1 and G3BP2 single knockout as well as double knockout HEK293 cells, we investigated the functional consequences of each knockout on cellular homeostasis. Therefore, we monitored the growth of the corresponding cell lines and seeded at day 0 200,000 cells in a 6-well plate. We counted the cell number over eight days with two-day intervals. G3BP1-KO and G3BP2-KO cells did not show any alterations in growth if compared to parental HEK293 cells. However, loss of G3BP1 and G3BP2 in G3BP1/2-DKO cells caused severe cell growth defects (Figure 5C). Although G3BP1/2-DKO was not lethal, cell growth was dramatically impaired, suggesting a fundamental role of G3BP1 proteins for cellular homeostasis.

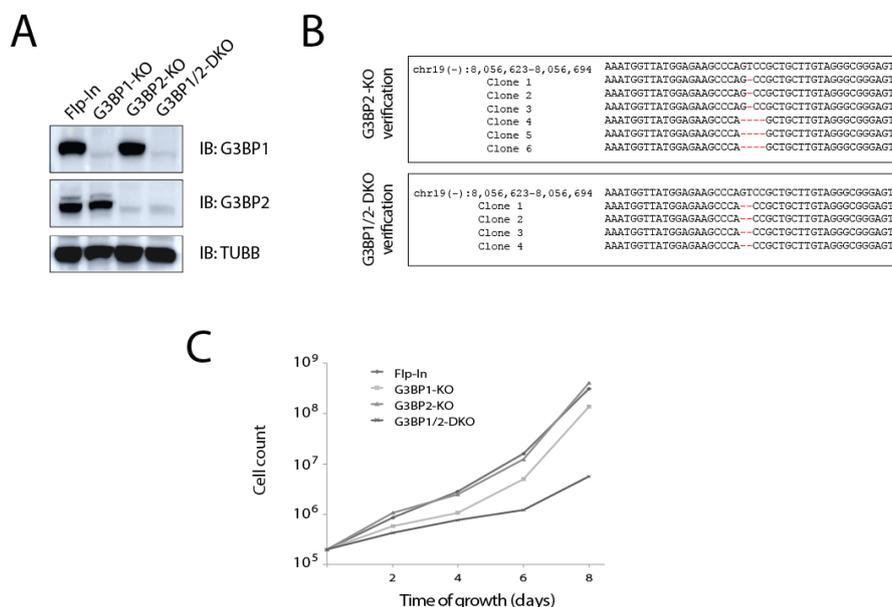


Figure 5: G3BP1 and G3BP2 double knockout causes cell growth defects.

A) Western blot verification of CRISPR/Cas9-mediated G3BP1 and G3BP2 single knockout and G3BP1/2 double knockout using human G3BP1 and G3BP2 specific antibodies. Parental Flp-In cells and anti-TUBB detection served as controls. **B)** Knockout verification on genomic level by sanger sequencing. **C)** Growth curves of parental Flp-In cells, G3BP1 and G3BP2 single knockout and G3BP1/2 double knockout cells over a time period of eight days.

4.4 G3BP1 proteins are essential for stress granule formation

Members of the G3BP1 family of RNA-binding proteins localize to cytoplasmic SGs^{4,21}. Stress granules are described as cytoplasmic RNA-protein complexes containing non-translating mRNAs, translation initiation components and many RNA-binding proteins affecting mRNA function²⁴. Formation of SG can have several reasons including reduced translation initiation rates caused by stress response²³. To investigate if G3BP1 proteins are essential for SG formation, we analyzed the subcellular localization of G3BP1 proteins in absence or presence of environmental stress by multicolor RNA-FISH combined with IF⁴². This method allows the simultaneous detection of proteins and RNAs in fixed cells. Thereby, proteins are visualized by immunostaining with protein-specific antibodies, whereas the detection of polyadenylated mRNAs was enabled by hybridization to fluorophore-labeled oligo(dT) probes. If cells are exposed to harmful conditions, untranslated polyadenylated mRNAs accumulate in SGs, and therefore serve as SG marker.

We first studied the localization of G3BP1 in unstressed cells. As G3BP1 family proteins have a NTF2 domain giving it the possibility to enter the nucleus, a translocation event could proceed. The signal for G3BP1 is equally distributed throughout the cytoplasm but not in the nucleus (Figure 6A). This result confirmed previous published studies that describe G3BP1 protein family as cytoplasmic²¹.

When stress was applied, G3BP1 was recruited to cytoplasmic SGs (Figure 6C). To investigate if G3BP1 protein family proteins are essential for SG formation, CRISPR/Cas9-mediated G3BP1 and G3BP2 single knockout Flp-In HEK293 cells (G3BP1-KO and G3BP2-KO) as well as G3BP1/2-DKO cells were examined to stress conditions, too. Cells were exposed to 400 μ M sodium arsenite causing oxidative stress response (Figure 6C) or incubated at 44°C to activate heat stress response (Figure 6E). Comparable to parental Flp-In HEK293 cells, G3BP2-KO cells revealed SG formation under both stress conditions (Figure 6C/E). However; SG formation in G3BP1-KO as well as G3BP1/2-DKO cells was limited (Figure 6D/F). G3BP1/2-DKO completely prevented assembly of SGs under arsenite stress (Figure 6D) and reduces the assembly to 11% under heat stress (Figure 6F), highlighting that the G3BP1 protein family is essential for SG formation. G3BP1-KO cells do not completely prevent SG formation, but compared to parental heat stressed cells, only 20% of SGs are formed under same stress treatment. Additionally, as shown in Figure 6D, G3BP1 knockout cells almost completely abolished SG formation in response to sodium arsenite stress. Also for G3BP2-KO cells was a reduced SG formation measured.

The reduction in terms of SG formation however was weaker with a remaining SG formation of 71% for heat stressed cells and 62% for sodium arsenite treatment.

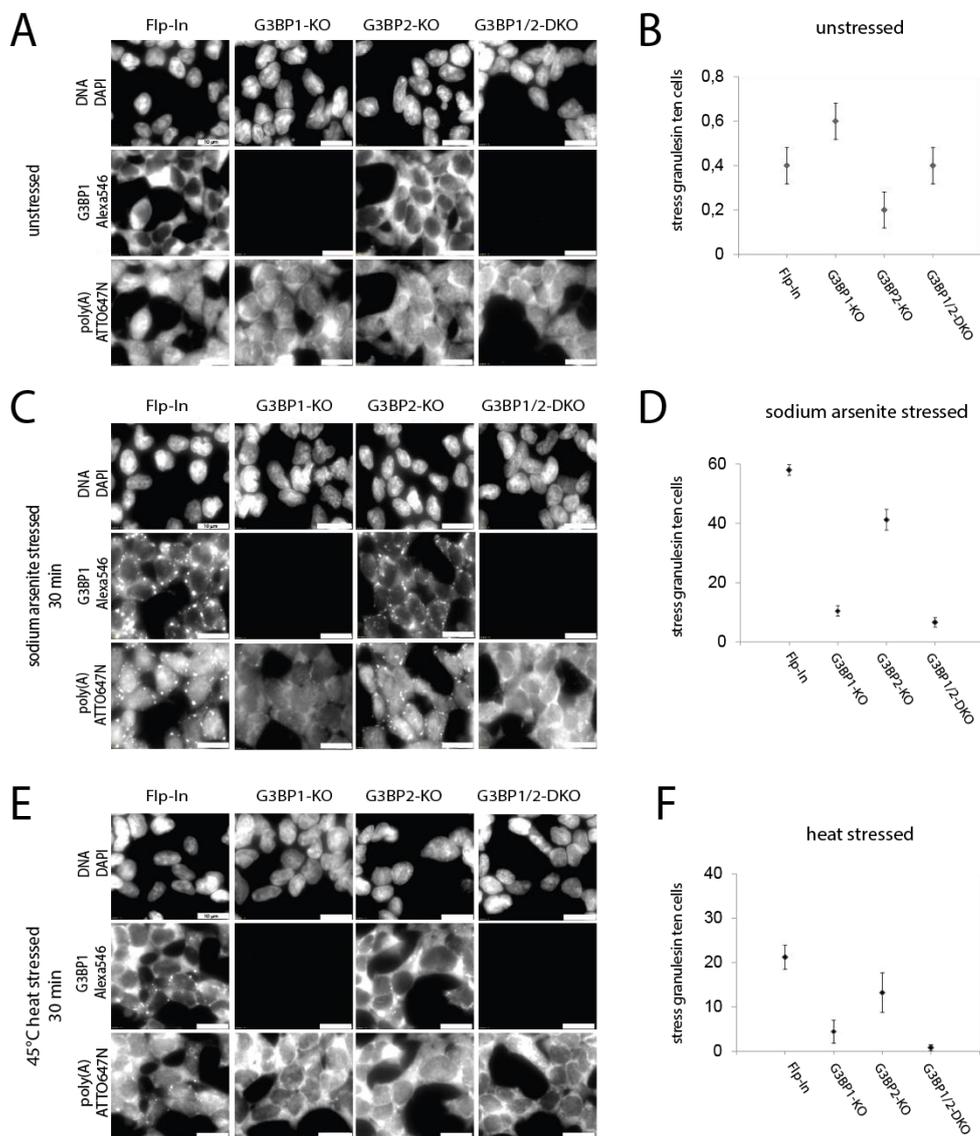


Figure 6: G3BP1 proteins are essential regulators of cytoplasmic stress granules.

A) RNA-FISH and IF experiments for subcellular localization of G3BP1 proteins. Parental Flp-In, G3BP1-KO, G3BP2-KO and G3BP1/2-DKO cells were fixed and stained by IF using human G3BP1 specific antibody and corresponding fluorescently labeled secondary antibody. Poly(A)-mRNAs were hybridized to fluorescently labeled oligo(dT) probes, nuclei were detected using DAPI. **B)** Calculation of SG ratio based on representative micrographs described in (A). The average of 5 times 10 counted cells is listed. **C)** RNA-FISH and IF experiments to detect oxidative SG formation. Parental Flp-In, G3BP1-KO, G3BP2-KO and G3BP1/2-DKO cells were arsenite stressed with a final concentration of 400 μ M, fixed and stained by IF using human G3BP1 specific antibody and corresponding fluorescently labeled secondary antibody. Poly(A)-mRNAs were hybridized to fluorescently labeled oligo(dT) probes, nuclei were detected using DAPI. **D)** Calculation of stress granules ratio based on representative micrographs described in (C). The average of 5 times 10 counted cells is listed. **E)** RNA-FISH and IF experiments to detect heat SG formation. Parental Flp-In, G3BP1-KO, G3BP2-KO and G3BP1/2-DKO cells were heat stressed at 45°C, fixed and stained by IF using human G3BP1 specific antibody and corresponding fluorescently labeled secondary antibody. Poly(A)-mRNAs were hybridized to fluorescently labeled oligo(dT) probes, nuclei were detected using DAPI. **F)** Calculation of stress granules ratio based on representative micrographs described in (E). The average of 5 times 10 counted cells is listed.

4.5 Double knockout of G3BP1 and G3BP2 causes translational repression

The binding of G3BP1 across CDS and 3' UTRs suggested that the G3BP1 protein family might be involved in translational regulation. We therefore performed polysome profiling analyses of parental Flp-In HEK293 cells and compared the profile to G3BP1 and G3BP2 single knockout Flp-In HEK293 cells as well as G3BP1/2-DKO cells. Polysome profiling is a method to infer the translational status of mRNA in an overall manner combined with the specific detection of individual proteins⁴³. Therefore, cells were treated for 5 min with cycloheximide to inhibit protein synthesis by blocking the translocation step during translational elongation. Cell lysates were separated on a 10-60% sucrose gradient by centrifugation at 35,000xg. Subsequent fractionation and UV detection (254 nm) revealed characteristic pattern of a polysome profile where fractions containing free RNA were followed by fractions containing the small 40S ribosomal subunit only, followed by fractions of the large 60S subunit, one assembled monosome as well as polysomal fractions indicative of heavy translated mRNAs. In G3BP1-KO and G3BP2-KO cells we observed a slight signal shift to monosomal fractions if compared to parental Flp-In HEK293 cells. However, in G3BP1/2-DKO cells translation was strongly repressed as indicated by a strong shift from heavy polysomes to monosomes (Figure 7A).

We subsequently performed western blotting analyses of collected monosome and polysome fractions to get further information about the association of G3BP1 proteins with polyribosomes. Using protein specific antibodies for G3BP1 and G3BP2, we detected both proteins in each fraction of the polysome profiling experiments supporting their association with ribosomes (Figure 7B). Western blotting was also conducted for proteins of the small (RPS19) or large ribosomal subunit (RPL7a), respectively (Figure 7B). RPS19 was enriched in the first fractions in accordance with the peak of the 40S subunit in the polysome profile, while almost no signal was detected in the following fractions of the 60S subunit since RPS19 is not part of the 60S ribosomal subunit. RPS19 is again detectable in fractions containing fully assembled mono-ribosomes and polysomes. Similar pattern was detectable for RPL7a. The first few fractions containing the 40S ribosomal subunit did not show a signal, whereas the following fractions, starting with the 60S subunit only, showed an enrichment of RPL7a proteins. As a further control, the eukaryotic translation initiation factor 4E (EIF4E) was detectable in the first fractions only (Figure 7B).

In order to identify specific mRNAs whose translation is limited upon loss of G3BP1 protein function, we aimed to perform ribosome profiling. Ribosome profiling allows the identification of ribosome-protected mRNA fragments thereby giving the opportunity to gain insights into the active translome⁴⁵. First, it was important to find optimal conditions for the RNase I treatment of cell lysates in order to obtain mRNAs fragments protected by one ribosome only. Therefore, we treated cell lysates of parental Flp-In HEK293 cells, derived from five 20 cm plates, with different amounts of RNase I and incubated them either at room temperature or on ice. Digested RNase-treated samples were further separated on 10-60% sucrose gradients and analyzed by UV detection at 254 nm (Figure 7C). Excessive concentration of 150 U/OD RNase I caused the digestion even of assembled monosomes. In case the RNase I concentration was too low (13 U/OD), mRNAs were still occupied by more than one ribosome what results in a faulty sequencing result as a highly mRNA would not be detected as such a one. A final concentration of 18 U RNase I per OD₂₆₀ of the lysate incubating for 30 min on ice was further used as optimal condition to get the footprints of ribosomal protected mRNA fragments. Using these optimized conditions, ribosome profiling was performed for parental Flp-In HEK293 cells, G3BP1-KO and G3BP1/2-DKO in triplicates. Digested lysates were separated on a 10-60% sucrose gradient and centrifuged by 35,000xg for 2 h and fractionated under UV detection. Monosom fraction of each sample was collected and RNA extracted using TRizol reagent. Two size markers with the length of 26nt and 34nt displayed the fragments on the SDS-Gel that were supposed to be occupied by one single ribosome. After a phosphatase treatment, 5'hydroxyl terminus is radiolabeled followed by adapter ligation using the barcode technology and cDNA library generation. Amplified and purified cDNA was submitted to The Rockefeller University Genomics Resource Center for DNA quality control using the TapeStation software. Results showed one fragment at the expected size of about 100nt, what revealed the product of ribosome obsessed mRNA ligated to the 3' and 5'ligands.

4.6 Double knockout of G3BP1 and G3BP2 causes EIF2S1-mediated stress response

Since double knockout of G3BP1 and G3BP2 resulted in translational repression in Flp-In HEK293 cells as showed by polysome profiling, we aimed to elucidate if phosphorylation of the eukaryotic translation initiation factor 2 α (EIF2S1/eIF2 α) caused the general translational repression.

Phosphorylated EIF2S1 is a key regulator for translational repression resulting in SG formation triggered by at least one of four kinases: EIF2AK1 to EIF2AK4. Phosphorylation of EIF2S1 at serine 51 inhibits efficient GDP-GTP exchange, leading to decreased levels of translationally competent ternary complexes eIF2/tRNAⁱMet/GTP and therefore inhibits general translation initiation under stress conditions^{46,47}.

Using an antibody specific for phosphorylated EIF2S1, western blotting analyses of cell lysates of parental HEK293 cells as well as G3BP1-KO, G3BP2-KO and G3BP1/2-DKO HEK293 cells were performed (Figure 7D). While single knockout cells revealed only a slight increase of phosphorylated EIF2S1 compared to parental HEK293, loss of both G3BP1 and G3BP2 proteins caused a strong signal in phosphorylation of EIF2S1 indicating that G3BP1/2-DKO led to the activation of an intracellular EIF2S1-mediated stress response possibly being the reason for stalled translation.

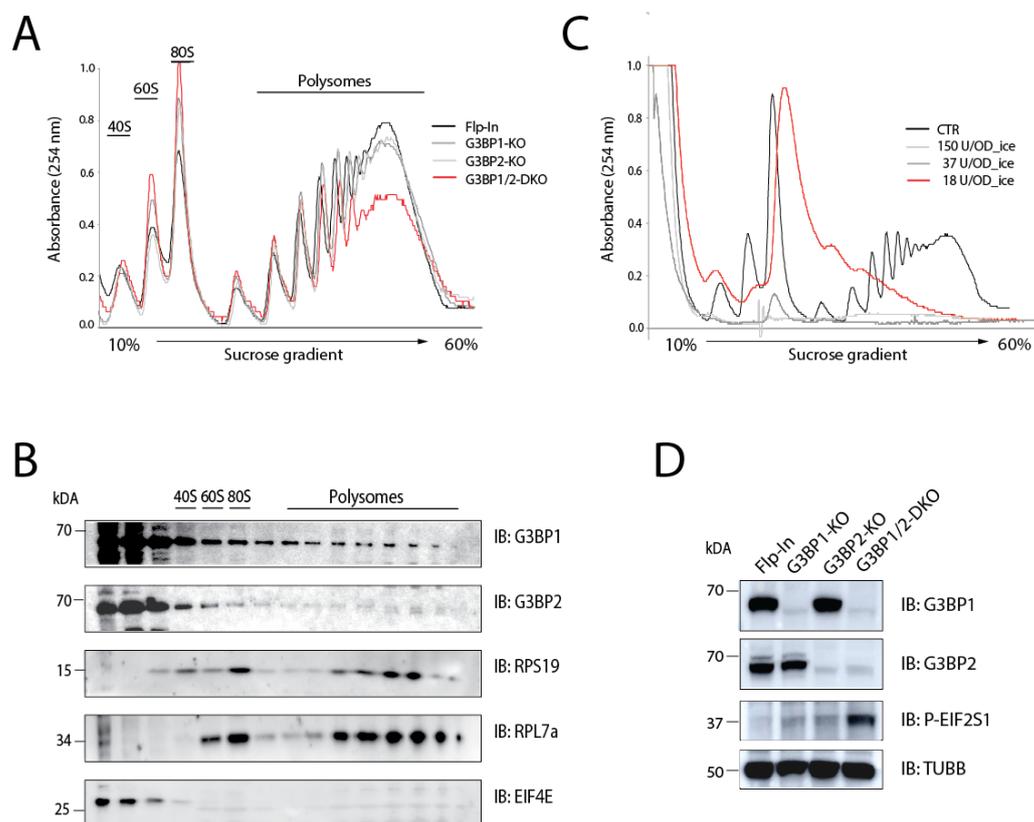


Figure 7: G3BP1 proteins involved in translation regulation.

A) Polysome profiling of parental Flp-In cells, G3BP1 and G3BP2 single knockout and G3BP1/2 double knockout cells. **B)** Western blot analysis of parental Flp-In cell lysates after sucrose gradient separation and fractionation during polysome profiling. Human protein specific antibodies were used as indicated. **C)** Ribosome profiling of parental Flp-In cells using different RNase I concentrations for lysate digestion. **D)** Western blot analysis of parental Flp-In cells, G3BP1 and G3BP2 single knockout and G3BP1/2 double knockout for phosphorylated EIF2S1 antibody to probe for stress response caused by CRISPR/Cas9-mediated knockouts. Anti-TUBB detection served as loading control.

4.7 RRM domain is crucial for RNA-binding and functionality of G3BP1

To investigate the contribution of each domain of G3BP1 on its function, we cloned different G3BP1 truncation mutants in parental Flp-In T-REx HEK293: FlagHA-G3BP1-NTF2, FlagHA-G3BP1- Δ RRM, FlagHA-G3BP1-RRM (Figure 8A). Corresponding coding sequence of each construct was cloned into pDONR221 and subsequently recombined into the destination vector pFRT/TO/FlagHA/dest according to the Gateway strategy offered by Invitrogen. The resulting pFRT/TO/FlagHA/CDS plasmids additionally encoded a hygromycin resistance gene to allow for clonal selection. The mentioned plasmids were used to transfect parental Flp-In T-REx HEK293 cells. After hygromycin selection, cell clones expressing the FlagHA-tagged G3BP1 truncation mutants under control of doxycycline were analyzed by western blotting for expression of FlagHA-tagged proteins (Figure 8B). Western blotting using anti-HA antibody only showed a signal for FlagHA-G3BP1 full length and the truncation mutant FlagHA-G3BP1- Δ RRM under doxycycline control, though not for the short mutants FlagHA-G3BP1-NTF2 and FlagHA-G3BP1-RRM (Figure 8B). The short variations only consisting of either of the domains thus could not be detected or expressed. Further experiments to evaluate the domain crucial for the functionality of G3BP1 protein family were hence only performed for FlagHA-G3BP1- Δ RRM cells.

In order to evaluate the RBD that is crucial for RNA binding in G3BP1, a pilot-scale PAR-CLIP was performed on doxycycline induced and uninduced FlagHA-G3BP1- Δ RRM mutants and doxycycline induced full length FlagHA-G3BP1. The phosphoimage of separated FlagHA-G3BP1-RNA complexes of FlagHA-G3BP1- Δ RRM cells and full length FlagHA-G3BP1 under the control of doxycycline only revealed a crosslinking band for full length FlagHA-G3BP1 but not for the FlagHA-G3BP1- Δ RRM truncation mutant (Figure 8C, upper panel). This result was expected, as RRM domain is a well-established RBD recognizing target RNAs with high affinity and specificity. Thus the FlagHA-G3BP1- Δ RRM truncation mutant failed to crosslink RNA, western blotting analysis confirmed a successful anti-Flag immunoprecipitation of FlagHA-G3BP1- Δ RRM (Figure 8C, lower panel).

Next we asked whether loss of RNA-binding of G3BP1- Δ RRM was accompanied by a failure to co-localize with cytoplasmic stress granules. Therefore, we again conducted RNA-FISH and IF on doxycycline-inducible FlagHA-G3BP1- Δ RRM cells. Cells were stressed with 400 mM sodium arsenite for 30 min at 37°C. Since the FlagHA-G3BP1- Δ RRM cell line was generated in parental HEK293 cells, the obtained cell line also expressed endogenous G3BP1. Therefore cells were capable of SG assembly anyway as shown before (Figure 6).

Nevertheless, the sub-cellular localization and assembly of FlagHA-G3BP1- Δ RRM in SGs under cellular stress conditions was evaluated using an anti-HA antibody. The panel detecting anti-HA staining did not show a signal for SG indicating that the RRM domain of G3BP1 is essential for its localization to SGs (Figure 8D).

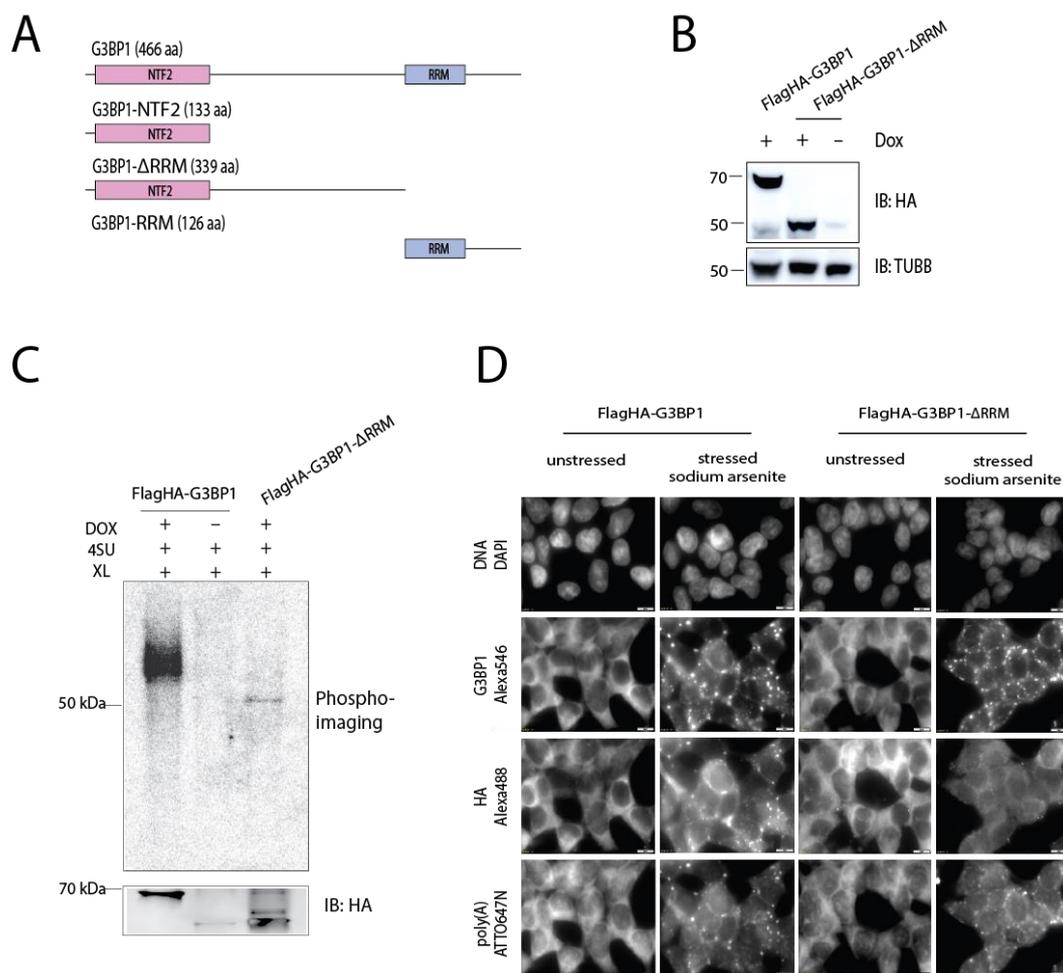


Figure 8: RRM domain crucial for RNA-binding and stress granules formation.

A) Domain structure and amino acid length of full length G3BP1, G3BP1-NTF2, G3BP1- Δ RRM and G3BP1-RRM truncation mutants. **B)** Western blot verification of full length FlagHA-G3BP1 and truncation mutant FlagHA-G3BP1- Δ RRM using anti-HA and anti-TUBB antibody. **C)** Upper panel: Autoradiographic RNA detection of crosslinked and immunoprecipitated RNA-protein complex from pilot scale FlagHA-G3BP1- Δ RRM PAR-CLIP experiment separated on a SDS-Gel. Lower panel: Western blot analysis to verify FlagHA-G3BP1- Δ RRM immunoprecipitation using anti-HA and anti-TUBB antibody. **D)** Fluorescence analysis of RNA-FISH and IF stress experiment. Unstressed and sodium arsenite FlagHA-tagged G3BP1- Δ RRM truncation mutants and full length FlagHA-G3BP1 cells, respectively, were fixed and stained by IF using anti-human G3BP1 and anti-HA primary antibody corresponding to fluorescently labeled secondary antibody. Poly(A)-mRNAs were hybridized to fluorescently labeled oligo(dT) probes, nuclei were detected using DAPI.

5 Discussion

The G3BP1 family of RNA binding proteins is essential for stress granules formation²⁰. Further it is supposed to be involved in a variety of cell biologic processes; however, a comprehensive characterization of the G3BP1 protein family remains elusive. Therefore the aim of this thesis was to systematically investigate the role of the G3BP1 protein family in post-transcriptional gene regulation.

5.1 G3BP1 protein binds mRNA in CDS and 3'UTR

In order to investigate the RNA binding properties and simultaneously identify the RNA targets of G3BP1 in a transcriptome wide manner, 4SU PAR-CLIP was performed. Data analysis after cDNA-sequencing revealed G3BP1 as mRNA binding protein binding CDS and 3'UTRs. This result is consistent with the cytoplasmic localization of G3BP1 proteins⁴ as well as their association with ribosomes⁶¹. On the other hand, our data are in contrast to a recent study by Martin and co-workers who reported on G3BP1 binding preferentially to intron-retaining transcripts in the cerebellum⁶. As G3BP1 binds targeted mRNA with a single RRM, it suggests that G3BP1 proteins need to interact with diverse RNA sites in an indiscriminate fashion. In agreement with this, a single RRM domain is mostly too small to recognize enough RNA nucleotides to bind sequence specific⁴⁸. Nevertheless, electrophoretic mobility shift assays (EMSAs) should be performed to either find a sequence specific RNA recognition element (RRE) of G3BP1 proteins or to prove that G3BP1 binds target mRNAs rather unspecifically.

5.2 CAPRN1 and USP10 are direct interactors of G3BP1

The identification of protein interactors of G3BP1 proteins might be essential to functionally and mechanistically classify the protein, since information about molecular interactors may be helpful to detect, predict or characterize unknown biological process the protein of interest is involved in. As mechanisms of post-transcriptional gene regulation (PTGR) often functionally depend on the interaction of RBP with other RBP⁴⁹, it is often enigmatic, how an interacting protein influences the function of another protein in certain cellular processes such as development, metabolism and cancer progression⁵⁰.

To identify direct interactors of G3BP1, Flag-IP experiments combined with MS analysis were conducted on RNase I-treated cell lysates of HEK293 cells inducible expressing FlagHA-tagged G3BP1. MaxQUANT analysis revealed USP10 and CAPRIN1 as the most abundant interactors of G3BP1 (Figure 4A). Based on various stress response assays G3BP1 has already been shown to form complexes with CAPRIN1 and USP10 to mediate stress granules condensation in a mutually exclusive manner²¹.

CAPRIN1 has been described as a cytoplasmic phosphoprotein with effects on cellular activation and proliferation⁵¹. A study by Solomon and co-workers showed that its overexpression induced phosphorylation of EIF2S1 accompanied with translational repression¹². Previous published PAR-CLIP data for CAPRIN1 described it as a mRNA-binding protein targeting mRNAs equally distributed over CDS and 3'UTR⁵². CAPRIN1 has also been shown to be localized to cytoplasmic SGs, suggesting to play a role in mRNA transport or translational regulation⁵³.

USP10 belongs to a complex protein family known for its ubiquitination properties in terms of production and recycling. Further it is supposed to have functional effects on cell growth control, differentiation and apoptosis^{54,55}.

However, it is not yet known whether those proteins play distinct, overlapping or complementary roles in the cell next to SG formation. Also it is still enigmatic whether they regulate each other in terms of protein expression or function. To evaluate the interaction but also the extent of G3BP1 being directly involved in SG formation or only in a recruiting manner, it would be interesting to perform knockout cells for CAPRIN1 and USP10 followed by SG assays. When performing an RNase I digestion, we ensured a close interaction between G3BP1 and enriched proteins. Certainly to get a better understanding about the structural composition of G3BP1 and the character of the interaction, another experiment using Flag-tagged G3BP1- Δ RRM truncation mutant should be performed. As G3BP1 consists of one RRM and one NTF2 domain, this result could contribute to an overall structural characterization of the binding site of G3BP1. Even so Flag-tagged G3BP1-NTF2 and G3BP1-RRM truncation mutants should be generated and analyzed by IP and MS.

Further interactions between G3BP1 and proteins we could detect are not yet characterized or confirmed in literature. However they give further reference about the biological processes G3BP1 proteins are assumed to be involved in. With ATX2L (Ataxin-2-like protein) another regulator of SG and P-body formation was detected.

Another interactor, EIF4G1 (Eukaryotic translation initiation factor 4 gamma 1) reaffirms the idea of G3BP1 being involved in translation initiation. EIF4G1 is a member of the EIF4F protein complex⁵⁶ that is involved in the recruitment of mRNAs to the ribosome as it recognizes mRNA cap and regulates the ATP-dependent unwinding of 5'-terminal secondary structure⁵⁷. HSPA9, a mitochondrial stress-70 protein that is involved in cell proliferation and cellular stress⁵⁸ is also detected as an interactor.

5.3 G3BP1 proteins are essential for stress granule formation

Stress response assays using RNA-FISH and IF revealed G3BP1 necessary for SG formation (Figure 6) as CRISPR/Cas9 mediated single knockouts of G3BP1 and G3BP2 as well as double knockout of G3BP1/2 prevent accumulation of SGs. This result confirmed previous published data in which cytoplasmic G3BP1 was described as SG marker in HeLa and CoS cells²⁰. Solomon and co-workers further investigated G3BP1 localized in SGs in various cell lines including NIH 3T3 mouse fibroblasts stable expressing human G3BP1 what proved G3BP1 family to assembly in SGs rudimental across various cell types¹². SG prevention in G3BP2 knockout cells was weaker than the effect of the G3BP1 single knockout for sodium arsenite and heat stress conditions (Figure 6D/F). We suggest that this is caused by the less abundance of G3BP2. Based on poly(A) mRNA-seq data, G3BP2 is fivefold less abundant than G3BP1.

5.4 Double knockout of G3BP1 and G3BP2 causes translational repression

mRNA translation is a complex mechanism mainly proceeded in three distinct phases: translation initiation, elongation and termination, whereas translation initiation mostly represents the rate-limiting step⁵⁹. Polysome profiling not only reveals proteins involved in translational regulation but also gives the opportunity to specifically distinguish which step of protein synthesis is regulated. In case of changes in translational initiation, an inhibition would lead to a reduction of polysome peaks accompanied by an increase of 40S, 60S and 80S monosome peak as the formation of the translation initiation complex is blocked. For regulatory effects of a certain protein on translational elongation, an increase of polysome peaks would be detectable in the polysome profile. Changes in terms of translational termination would show up by larger polysome peaks compared to the control profile⁶⁰.

We performed polysome profiling of CRISPR/Cas9-mediated G3BP1 and G3BP2 single or double knockout Flp-In HEK293 cells to evaluate the impact of both proteins on translational regulation of gene expression. As highlighted in (Figure 7A), a reduction in polysome peaks was detectable for G3BP1/2-DKO cells, indicating a function of G3BP1 protein family in translational initiation. Inhibition of translational initiation is often accompanied by blocked EIF2S1 activity thereby preventing the recruitment of eIF2/GTP/Met-tRNAi to the 40S subunit⁴⁶.

Since we could detect a phosphorylation of EIF2S1 in G3BP1/2-DKO cells (Figure 7D), the intracellular stress response triggered upon G3BP1/2 double knockout was potentially the reason for translational inhibition detected by polysome profiling. Further experiments will be necessary to give a valid explanation for stalled translation. In order to confirm or disprove that this translational inhibition is a stress related effect, CRISPR/Cas9 mediated knockout of any of the four EIF2S1 kinases, EIF2AK1 to EIF2AK4, should be generated in G3BP1/2-DKO cells. Subsequent polysome profiling will reveal whether stalled translation can be recovered.

Western blot analysis of polysome fractions identified G3BP1 and G3BP2 as ribosomal associated proteins, as they are detectable in every fraction of the polysome profile (Figure 7C). This is in agreement with previously published data gained from ribosome affinity purification⁶¹. In this study, Simsek and co-workers performed Flag-IPs on Flag-eL36 expressing mouse embryonic stem cells once with and once without RNase A treatment to distinguish between directly to the ribosome bound ones and other, that are associated with ribosomes by interactions mediated with mRNAs. Those experiments systematically identified hundreds of ribosome-associated proteins. Among others, G3BP2 was found to be an RNase-dependent interactor of the ribosome like many poly(A)-binding proteins that are translation-related (LARP1, ZNF622). Moreover, G3BP1 proteins are also described to be associated with the 40S ribosomal subunit by²¹. This finding states an anomaly among other SG markers like TIA1/TIAL1 what qualifies G3BP1 proteins to cause stalled translation suggested in a guidance manner.

5.5 Double knockout of G3BP1 and G3BP2 causes EIF2S1-mediated stress response

Double knockout of G3BP1 and G3BP2 resulted in the phosphorylation of EIF2S1 indicating that the G3BP1 protein family is essential to prevent a stress response in unstressed human cells. As previously mentioned there are four kinases (EIF2AK1-4) that trigger the phosphorylation of EIF2S1. Further experiments using western blotting combined with antibodies against phosphorylated EIF2AK1-4, respectively, or generation of knockout cells should clarify which of the kinases is reasonable for EIF2S1 activation by G3BP1/2-DKO.

To put this result into context with the findings gained by polysome profiling, it is possible, that assumed translational regulation, determined by the shift from heavy polysomes to monosomes could be stress-induced by G3BP1/2-DKO. Performing polysome-profiling analyses on EIF2AK1-4 knockout cells should clarify if the translational repression observed in G3BP1/2-DKO HEK293 cells is caused by an activated stress response.

Our mass spectrometric analyses revealed that G3BP1 interacts with CAPRIN1 (Figure 4A), a protein that has been shown to trigger the phosphorylation of EIF2S1 resulting in stress response when overexpressed in Flp-In HEK293T cells¹². CAPRIN1-mediated induction of EIF2S1 phosphorylation occurred through a mechanism that was depending on its ability to bind mRNAs; however, the deeper insights remain elusive. Since G3BP1 and CAPRIN1 show comparable expression levels in HEK293 cells (based on poly(A)-RNAseq), we suggest, that overexpression of CAPRIN1 or loss of G3BP1 function activate the same stress response pathway due to an imbalance of the CAPRIN1-G3BP1-USP10 complex. Similar to the G3BP1 protein family, also CAPRIN1 has another family member: CAPRIN2, which is tenfold less abundant. Additionally, as both RBP families mediate SG formation in an agonistic manner²¹ further western blot analysis will be necessary to test whether CAPRIN1 is upregulated in G3BP1/2-DKO cells. To evaluate a possible G3BP1/2-DKO regulatory event on CAPRIN1, CRISPR/Cas9-mediated knockout of CAPRIN1 should be performed in G3BP1/2-DKO cells. Further polysome profiling of those G3BP1/2 and CAPRIN1 triple knockout cells could clarify if G3BP1/2-DKO mediated translational regulation is caused by a regulatory event on CAPRIN1.

5.6 RRM domain is crucial for RNA-binding and functionality of G3BP1

Using Flp-In HEK293 cells expressing a FlagHA-tagged G3BP1- Δ RRM truncation mutant for a pilot-scale PAR-CLIP, we could prove that the RRM domain was crucial for binding of G3BP1 to target mRNAs as there was no protein-RNA crosslinking band detectable for FlagHA-G3BP1- Δ RRM cells using autoradiography (Figure 8C). As the RRM domain is the most abundant functional domain in eukaryotic RBPs known for its binding abilities to RNA², this result was expected. G3BP1 proteins combine only two domains, one RRM and one NTF2, why by exclusion of the NTF2 domain RRM is told to be important. We also cloned FlagHA truncation mutants only consisting of the G3BP1 NTF2 or RRM domain, respectively. However we had issues to express FlagHA-tagged G3BP1-NTF2 and FlagHA-tagged G3BP1-RRM.

To investigate which domain of G3BP1 was essential for SG assembly, we performed RNA-FISH and IF method on Flp-In HEK293 cells expressing a FlagHA-tagged G3BP1- Δ RRM truncation mutant (Figure 8D). As the truncation mutant was cloned in parental Flp-In cells, SG formation was not inhibited. However, the G3BP1- Δ RRM truncation mutant did not co-localize to cytoplasmic SGs being evidence that the functionality of G3BP1 proteins in SG formation is dependent on its ability to bind target mRNAs. As mentioned before, truncation mutants with RRM only have to be cloned as the positive control.

5.7 Conclusion

This thesis provides fundamental work on the characterization of the G3BP1 mRNA-binding protein family in terms of post-transcriptional gene regulation. We identified target mRNAs of G3BP1 in a transcriptome-wide manner by PAR-CLIP and showed that G3BP1 binds target mRNAs preferentially in CDS and 3' UTR¹². Continuing, G3BP1 single and G3BP1/2 double knockout cells show defects in stress granule assembly. Using G3BP1- Δ RRM truncation mutants, RRM domain was proven to be essential for RNA binding and localization to SG. CAPRIN1 and USP10 proteins were detected as direct interactors forming a complex with G3BP1. G3BP1/2-DKO cells showed defects in cell growth and caused translational arrest as monitored by polysome profiling. Further G3BP1/2-DKO cells triggered phosphorylation of EIF2S1. Whether phosphorylation of EIF2S1 conditionally affects translational arrest has to be evaluated by further experiments.

6 Outlook

Within this study, we identified target mRNAs of G3BP1 and could show that the RRM domain is crucial for the functionality of the protein. However, further experiments are necessary to understand the role of G3BP1 and G3BP2 in target mRNA metabolism. Although we described the binding sites for G3BP1, PAR-CLIP should be performed on FlagHA-tagged G3BP2 expressing Flp-In HEK293 cells. Although G3BP1 and G3BP2 are from the same protein family and exhibit a structural and functional high identicality, detailed analyses are needed to evaluate the overlap of target mRNAs. Also overlapping targets of G3BP1 and CAPRIN1 should be analyzed as there is already a CAPRIN1 PAR-CLIP data set available⁵². In order to detect a RNA-binding specificity *in vitro*, gel shift analysis for recombinant produced G3BP1 and G3BP2 proteins should be performed. Therefore, both proteins need to be bacterially expressed and purified.

G3BP1-KO and CAPRIN1 overexpression, as mentioned in literature¹² trigger the phosphorylation of EIF2S1. This makes sense if the complex with USP10 is regulated. For this reason it is further interesting which effect a USP10 knockout would have.

Finally the unsolved question whether phosphorylation of EIF2S1 triggers translational arrest should be answered by the generation of CRISPR/Cas9-mediated knockouts of EIF2AK1-4 in existing G3BP1/2-DKO cells.

List of figures

<i>Figure 1: Workflow of a PAR-CLIP experiment⁴⁰</i>	14
<i>Figure 2: Workflow of a polysome profiling experiment⁴³</i>	16
<i>Figure 3: G3BP1 binds target mRNAs in CDS and 3'UTR as determined by PAR-CLIP.</i> 19	
<i>Figure 4: G3BP1 proteins interact with CAPRIN1 and USP10.</i>	21
<i>Figure 5: G3BP1 and G3BP2 double knockout causes cell growth defects.</i>	22
<i>Figure 6: G3BP1 proteins are essential regulators of cytoplasmic stress granules.</i>	24
<i>Figure 7: G3BP1 proteins involved in translation regulation.</i>	27
<i>Figure 8: RRM domain crucial for RNA-binding and stress granules formation.</i>	29

List of tables

<i>Table 1: Chemicals and materials</i>	5
<i>Table 2: Buffers and solutions</i>	6
<i>Table 3: Enzymes and other materials</i>	7
<i>Table 4: Antibodies.</i>	8
<i>Table 5: Commercial kits</i>	9
<i>Table 6: Cell line</i>	10
<i>Table 7: Equipment</i>	10
<i>Table 8: Primer for BP reaction</i>	11

List of references

1. Glisovic, T., Bachorik, J.L., Yong, J. & Dreyfuss, G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters* **582**, 1977–1986 (2008).
2. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins (2014).
3. Burd, C.G. & Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. *Science (New York, N.Y.)* **265**, 615–621 (1994).
4. Parker, F. *et al.* A Ras-GTPase-activating protein SH3-domain-binding protein. *Molecular and cellular biology* **16**, 2561–2569 (1996).
5. Lunde, B.M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nature reviews. Molecular cell biology* **8**, 479–490 (2007).
6. Martin, S. *et al.* Preferential binding of a stable G3BP ribonucleoprotein complex to intron-retaining transcripts in mouse brain and modulation of their expression in the cerebellum. *Journal of neurochemistry* **139**, 349–368 (2016).
7. Prigent, M., Barlat, I., Langen, H. & Dargemont, C. IkappaBalpha and IkappaBalpha/NF-kappa B complexes are retained in the cytoplasm through interaction with a novel partner, RasGAP SH3-binding protein 2. *The Journal of biological chemistry* **275**, 36441–36449 (2000).
8. Tourriere, H. *et al.* RasGAP-associated endoribonuclease G3BP: selective RNA degradation and phosphorylation-dependent localization. *Molecular and cellular biology* **21**, 7747–7760 (2001).
9. Irvine, K., Stirling, R., Hume, D. & Kennedy, D. Rasputin, more promiscuous than ever: a review of G3BP. *The International journal of developmental biology* **48**, 1065–1077 (2004).
10. Annibaldi, A., Dousse, A., Martin, S., Tazi, J. & Widmann, C. Revisiting G3BP1 as a RasGAP binding protein: sensitization of tumor cells to chemotherapy by the RasGAP 317-326 sequence does not involve G3BP1. *PloS one* **6**, e29024 (2011).
11. Costa, M., Ochem, A., Staub, A. & Falaschi, A. Human DNA helicase VIII: a DNA and RNA helicase corresponding to the G3BP protein, an element of the ras transduction pathway. *Nucleic acids research* **27**, 817–821 (1999).
12. Solomon, S. *et al.* Distinct structural features of caprin-1 mediate its interaction with G3BP-1 and its induction of phosphorylation of eukaryotic translation initiation factor 2alpha, entry to cytoplasmic stress granules, and selective interaction with a subset of mRNAs. *Molecular and cellular biology* **27**, 2324–2342 (2007).
13. Soncini, C., Berdo, I. & Draetta, G. Ras-GAP SH3 domain binding protein (G3BP) is a modulator of USP10, a novel human ubiquitin specific protease. *Oncogene* **20**, 3869–3879 (2001).
14. Pickart, C.M. & Eddins, M.J. Ubiquitin: structures, functions, mechanisms. *Biochimica et biophysica acta* **1695**, 55–72 (2004).
15. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annual review of biochemistry* **67**, 425–479 (1998).
16. Sun, L. & Chen, Z.J. The novel functions of ubiquitination in signaling. *Current opinion in cell biology* **16**, 119–126 (2004).
17. Wilkinson, K.D. Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **11**, 1245–1256 (1997).
18. Wilkinson, K.D. *et al.* Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. *Biochemistry* **34**, 14535–14546 (1995).
19. Ortega, A.D., Willers, I.M., Sala, S. & Cuezva, J.M. Human G3BP1 interacts with beta-F1-ATPase mRNA and inhibits its translation. *Journal of cell science* **123**, 2685–2696 (2010).
20. Tourriere, H. *et al.* The RasGAP-associated endoribonuclease G3BP assembles stress granules. *The Journal of cell biology* **160**, 823–831 (2003).

21. Kedersha, N. *et al.* G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *The Journal of cell biology* **212**, 845–860 (2016).
22. Martin, S. & Tazi, J. Visualization of G3BP stress granules dynamics in live primary cells. *Journal of visualized experiments : JoVE* (2014).
23. Kedersha, N.L., Gupta, M., Li, W., Miller, I. & Anderson, P. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *The Journal of cell biology* **147**, 1431–1442 (1999).
24. Buchan, J.R. & Parker, R. Eukaryotic stress granules: the ins and outs of translation. *Molecular cell* **36**, 932–941 (2009).
25. McCormick, C. & Khapersky, D.A. Translation inhibition and stress granules in the antiviral immune response. *Nature reviews. Immunology* (2017).
26. Panas, M.D., Ivanov, P. & Anderson, P. Mechanistic insights into mammalian stress granule dynamics. *The Journal of cell biology* **215**, 313–323 (2016).
27. Guitard, E., Parker, F., Millon, R., Abecassis, J. & Tocque, B. G3BP is overexpressed in human tumors and promotes S phase entry. *Cancer letters* **162**, 213–221 (2001).
28. Barnes, C.J. *et al.* Heregulin induces expression, ATPase activity, and nuclear localization of G3BP, a Ras signaling component, in human breast tumors. *Cancer research* **62**, 1251–1255 (2002).
29. Zhang, H. & Shao, R.-g. G3BP: a promising target for cancer therapy. *Yao xue xue bao = Acta pharmaceutica Sinica* **45**, 945–951 (2010).
30. Winslow, S., Leandersson, K. & Larsson, C. Regulation of PMP22 mRNA by G3BP1 affects cell proliferation in breast cancer cells. *Molecular cancer* **12**, 156 (2013).
31. French, J., Stirling, R., Walsh, M. & Daniel Kennedy, H. The Expression of Ras–GTPase Activating Protein SH3 Domain-binding Proteins, G3BPs, in Human Breast Cancers. *The Histochemical Journal* **34**, 223–231 (2002).
32. Zekri, L. *et al.* Control of fetal growth and neonatal survival by the RasGAP-associated endoribonuclease G3BP. *Molecular and cellular biology* **25**, 8703–8716 (2005).
33. Monahan, Z., Shewmaker, F. & Pandey, U.B. Stress granules at the intersection of autophagy and ALS. *Brain research* **1649**, 189–200 (2016).
34. Liu-Yesucevitz, L. *et al.* Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS one* **5**, e13250 (2010).
35. Wolozin, B. Regulated protein aggregation: stress granules and neurodegeneration. *Molecular neurodegeneration* **7**, 56 (2012).
36. Vanderweyde, T. *et al.* Contrasting pathology of the stress granule proteins TIA-1 and G3BP in tauopathies. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**, 8270–8283 (2012).
37. Spitzer, J., Landthaler, M. & Tuschl, T. Rapid creation of stable mammalian cell lines for regulated expression of proteins using the Gateway(R) recombination cloning technology and Flp-In T-REx(R) lines. *Methods in enzymology* **529**, 99–124 (2013).
38. Le Cong & Zhang, F. Genome engineering using CRISPR-Cas9 system. *Methods in molecular biology (Clifton, N.J.)* **1239**, 197–217 (2015).
39. Ran, F.A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281–2308 (2013).
40. Garzia, A., Meyer, C., Morozov, P., Sajek, M. & Tuschl, T. Optimization of PAR-CLIP for transcriptome-wide identification of binding sites of RNA-binding proteins. *Methods (San Diego, Calif.)* **118-119**, 24–40 (2017).
41. Corcoran, D.L. *et al.* PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. *Genome biology* **12**, R79 (2011).

42. Meyer, C., Garzia, A. & Tuschl, T. Simultaneous detection of the subcellular localization of RNAs and proteins in cultured cells by combined multicolor RNA-FISH and IF. *Methods (San Diego, Calif.)* **118-119**, 101–110 (2017).
43. Chasse, H., Boulben, S., Costache, V., Cormier, P. & Morales, J. Analysis of translation using polysome profiling. *Nucleic acids research* **45**, e15 (2017).
44. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**, 621–628 (2008).
45. Ingolia, N.T. Ribosome Footprint Profiling of Translation throughout the Genome. *Cell* **165**, 22–33 (2016).
46. Sonenberg, N. & Hinnebusch, A.G. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136**, 731–745 (2009).
47. Kedersha, N. & Anderson, P. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochemical Society transactions* **30**, 963–969 (2002).
48. Auweter, S.D., Oberstrass, F.C. & Allain, F.H.-T. Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic acids research* **34**, 4943–4959 (2006).
49. Moore, M.J. From birth to death: the complex lives of eukaryotic mRNAs. *Science (New York, N.Y.)* **309**, 1514–1518 (2005).
50. Halbeisen, R.E., Galgano, A., Scherrer, T. & Gerber, A.P. Post-transcriptional gene regulation: from genome-wide studies to principles. *Cellular and molecular life sciences : CMLS* **65**, 798–813 (2008).
51. Grill, B. *et al.* Activation/division of lymphocytes results in increased levels of cytoplasmic activation/proliferation-associated protein-1: prototype of a new family of proteins. *Journal of immunology (Baltimore, Md. : 1950)* **172**, 2389–2400 (2004).
52. Baltz, A.G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Molecular cell* **46**, 674–690 (2012).
53. Anderson, P. & Kedersha, N. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nature reviews. Molecular cell biology* **10**, 430–436 (2009).
54. Rolen, U. *et al.* Activity profiling of deubiquitinating enzymes in cervical carcinoma biopsies and cell lines. *Molecular carcinogenesis* **45**, 260–269 (2006).
55. Ovaa, H. *et al.* Activity-based ubiquitin-specific protease (USP) profiling of virus-infected and malignant human cells. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 2253–2258 (2004).
56. Gosselin, P. *et al.* Tracking a refined eIF4E-binding motif reveals Angel1 as a new partner of eIF4E. *Nucleic acids research* **41**, 7783–7792 (2013).
57. Akabayov, S.R., Akabayov, B. & Wagner, G. Human translation initiation factor eIF4G1 possesses a low-affinity ATP binding site facing the ATP-binding cleft of eIF4A in the eIF4G/eIF4A complex. *Biochemistry* **53**, 6422–6425 (2014).
58. Chen, T.H.-P. *et al.* Knockdown of Hspa9, a del(5q31.2) gene, results in a decrease in hematopoietic progenitors in mice. *Blood* **117**, 1530–1539 (2011).
59. Jackson, R.J., Hellen, C.U.T. & Pestova, T.V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews. Molecular cell biology* **11**, 113–127 (2010).
60. Coudert, L., Adjibade, P. & Mazroui, R. Analysis of translation initiation during stress conditions by polysome profiling. *Journal of visualized experiments : JoVE* (2014).
61. Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051-1065.e18 (2017).