

Characterization of a phloem-mobile RNA-binding  
protein and the circRNA content of phloem in  
*Brassica napus*

Dissertation zur Erlangung des Doktorgrades

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Biologie  
der Universität Hamburg

Vorgelegt von  
Kim Lara Lühmann

Hamburg, 2022

Datum der Disputation: 01.03.2023

1. Gutachterin: Prof. Dr. Julia Kehr
2. Gutachter: Prof. Dr. Tim Gilberger

## Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den 21.11.2022



.....  
Kim Lara Lümann

## Zusammenfassung

Aufgrund ihres sessilen Lebensstils sind Pflanzen nicht in der Lage vor abiotischen und biotischen Stressfaktoren zu fliehen. Dadurch müssen Pflanzen sich mithilfe einer effektiven Kommunikation und Reaktion diesen Problemen stellen. Ein besseres Verständnis des Langstreckentransport und der Langstreckenkommunikation kann für die Züchtung von Pflanzen genutzt werden, welche resistenter gegenüber abiotischen und biotischen Stressfaktoren sind. Aus diesem Grund wurde in dieser Arbeit der Langstreckentransport und die Langstreckenkommunikation von Pflanzen untersucht. Dies wurde durch die Charakterisierung des bekannten, Phloem-mobilen Proteins GRP7 und die Identifizierung von zirkulären RNAs im Phloem erreicht.

Um ein besseres Verständnis für den Einfluss von GRP7 auf den Langstreckentransport von RNAs zu erhalten wurde untersucht welche RNAs das *Arabidopsis thaliana* und *Brassica napus* GRP7 binden kann. Zudem wurde getestet wie stark die Bindungsaffinität von GRP7 für RNAs verschiedener Längen und mit verschiedenen Modifikationen ist. Auch der Einfluss der Glycin-reichen Region von GRP7 auf die Bindungsaffinität von RNA wurde untersucht. Dies führte zu dem Ergebnis, dass GRP7 kürzere RNAs mit einer höheren Affinität bindet als längere RNAs. Jedoch gibt es auch einige Ausnahmen an längeren RNAs, wie zum Beispiel die prä-mRNA von GRP7, welche ebenfalls mit einer hohen Affinität gebunden werden. GRP7 zeigte jedoch keine Präferenz gegenüber doppelsträngigen oder methylierten RNAs. Selbst kleine Änderungen in der Glycin-reiche Region von GRP7 zeigten eine veränderte Bindeaffinität von GRP7 insbesondere gegenüber langer RNAs. Die Entfernung des kompletten Glycin-reichen Bereichs von GRP7 führte dazu, dass kurze RNAs nur noch schwach und lange RNAs gar nicht mehr gebunden werden konnten. Daraus lässt sich schließen, dass die Glycin-reiche Region von GRP7 eine wichtige Funktion in der RNA-Bindung des Proteins innehat. In weiteren Experimenten wurde eine hohe Struktur-Flexibilität der Glycin-reichen Region und die Fähigkeit von GRP7 Kondensate zu bilden nachgewiesen.

Zudem wurde der Phloemsaft von *B. napus* auf seinen Gehalt an zirkulären RNAs untersucht. Zirkuläre RNAs entstehen durch das *back-splicing* von linearen Transkripten, welches in einer kovalenten Bindung zwischen 3' und 5'-Ende der RNA und damit in einer besonders stabilen RNA resultiert. Um den Gehalt an zirkulären RNAs im Phloemsaft zu untersuchen, wurden lnc-RNA Sequenzierdaten und circRNA-Sequenzierdaten von Blatt- und Phloemproben von *B. napus* verwendet. Es wurden 93 (lncRNA Sequenzierung) bzw. 15 zirkuläre RNAs (circRNA Sequenzierung) im Phloem identifiziert. Neun der circRNAs wurden in beiden Sequenzierungen gefunden. Zur weiteren Überprüfung identifizierter circRNAs wurde eine RT-PCR mit konvergenten (für lineare RNAs) und divergenten (für circRNAs) Primern durchgeführt, welche in Amplifikaten für fünf der sieben getesteten divergenten Primerpaare resultierte. Eine Sanger-Sequenzierung der *back-splicing junction* konnte für vier RNAs ihre Zirkularität bestätigen. Weitere Versuche zur Überprüfung der Zirkularität sowie der Funktion von zirkulären RNAs müssen durchgeführt werden, um die Funktion der circRNAs im Langstrecken-Transport und Kommunikations-System der Pflanzen zu verstehen.

## Abstract

Due to the sessile lifestyle of plants, they cannot escape abiotic and biotic stresses. Therefore, they must employ effective signaling to encounter obstacles by effective signaling and response. A better understanding of the long-distance transport and signaling system within plants would allow these systems for breeding plants more resistant against abiotic and biotic stresses. Therefore, this work investigates the long-distance transport and signaling of RNAs in phloem of plants. Here, the prominent phloem-mobile RNA-binding protein GRP7 and the circular RNA (circRNA) content of the phloem were analyzed.

To better understand the involvement of GRP7 in RNA long-distance transport and signaling, the general RNA-binding ability of both, AtGRP7 (*Arabidopsis thaliana* GRP7) and BnGRP7 (*Brassica napus* GRP7) was studied, for more information about the target RNAs of GRP7. Following this, the affinities of both GRP7s were tested for different RNA length and modifications to identify possible preferences of the protein. Finally, the involvement of the glycine-rich region (RGG-domain) of GRP7 in RNA-binding was further analyzed. GRP7 showed a very broad range of RNA-binding, suggesting it is a none-specific RNA-binding protein. Additionally, GRP7 bound shorter RNAs with a higher affinity than long RNAs. Some exception occurred, like GRP7s own pre-mRNA, which was bound with an equally high binding affinity. No preference of GRP7 could be detected for methylated or double-stranded RNAs. The glycine-rich region of GRP7 proved to be an important element in the RNA-binding ability of GRP7. Small changes within the glycine-rich region showed significant changes in the RNA-binding affinity of GRP7 and the removal of the RGG-domain resulted in a lower affinity towards short RNAs and no binding at all for long RNAs. Thus, the RGG-domain of GRP7 has an important role in the RNA-binding of this protein. In further experiments like SEC-SAXS, the high flexibility of the glycine-rich region and the ability of GRP7 to form condensates was demonstrated.

In addition to GRP7, the content of circular RNAs in the phloem sap of *B. napus* was studied. Circular RNAs are derived from back-splicing of pre-mRNAs, resulting in the covalent link of the 3' and 5' end and thus stable, circRNAs. To investigate the circRNA content of *B. napus* phloem, sequencing of a long non-coding library and a circRNA enriched library of leaf and phloem sap was analyzed. In phloem sap, 93 and 15 circRNAs were identified from the two sequencing libraries respectively. Nine of these RNAs were found in both sequencing libraries. To further prove the predicted circRNAs, RT-PCR was done for circRNAs with divergent (for circular RNAs) and convergent (for linear RNAs) primer pairs, resulting in amplicons for five of seven tested circular RNAs. The Sanger sequencing of the back-splicing junction (BSJ) for four of these RNAs proved their circularity. Further experiments to demonstrate the circularity of RNAs in the phloem as well as to investigate their roles and functions within the long-distance transport and signaling system of plants.

## List of Abbreviations

bp	base pair
BSJ	back-splicing junction
CC	companion cell
CDS	coding DNA sequence
CircRNA	circular RNA
DMSO	Dimethylsulfoxid
DNA	Deoxyribonucleic acid
dNTPs	Desoxynucleosidtriphosphate
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
<i>et. al</i>	lat. et alii, and others
GRP	glycine-rich RNA-binding protein
GRR	glycine-rich region
IDR	intrinsically disordered region
kb	kilo base pairs
LB-medium	lysogeny broth medium
LLPS	Liquid-liquid phase separation
lnc	long non-coding
nt	nucleotide
OD <sub>600</sub>	optical density measured at 600 nm
ori	origin of replication
PAGE	polyacrylamid gel electrophoresis
PCR	polymerase chain reaction
PCD	programmed cell death
PD	plasmodesmata
PP	phloem parenchyma
PPU	pore plasmodesmata unit
RBP	RNA-binding protein
RGG-domain	arginine-glycine-glycine-rich domain
RRM	RNA recognition motif

SE	sieve element
SEL	size exclusion limit
SP	sieve pores
TAE	Tris-Acetat-EDTA
Tris	Tris(hydroxymethyl)-aminomethan
UTR	untranslated region
UV-light	ultra violet light
v/v	<i>volume/volume</i>
w/v	<i>weight/volume</i>
x g	multiples of gravity

# Table of Contents

Zusammenfassung.....	III
Abstract .....	IV
List of Abbreviations.....	V
1 Introduction.....	1
1.1 The vascular tissue in plants.....	1
1.2 Content of phloem and long-distance transport and signaling.....	4
1.2.1 RNA content and long-distance transport and signaling.....	4
1.2.2 Protein content and long-distance transport.....	5
1.3 RNA-binding proteins .....	6
1.3.1 Intrinsically disordered RNA-binding domains .....	7
1.3.2 Glycine-rich RNA-binding protein 7.....	8
1.4 Circular RNAs .....	10
1.4.1 Circular RNAs in plants .....	12
1.5 Objective.....	13
2 Material and methods.....	14
2.1 Materials.....	14
2.1.1 Equipment .....	14
2.1.2 Consumables .....	15
2.1.2 Chemicals.....	15
2.1.3 Oligonucleotids.....	15
2.1.4 Enzymes.....	16
2.1.5 Bacterial cultivation media .....	16
2.1.6 Buffers .....	17
2.1.7 Plasmids.....	19
2.1.7 Bacterial strains .....	19
2.1.8 Plants .....	20
2.1.9 Genes of purified proteins.....	20
2.1.10 Bioinformatic tools .....	21

2.2 Methods .....	22
2.2.1 Bacterial methods.....	22
2.2.1.1 Overnight cultures of <i>E. coli</i> .....	22
2.2.1.2 Glycerol stocks of <i>E. coli</i> .....	22
2.2.1.3 Preparation of chemically competent <i>E. coli</i> .....	22
2.2.1.4 Transformation of <i>E. coli</i> .....	22
2.2.1.5 Expression culture .....	22
2.2.2 Plant methods .....	23
2.2.2.1 Cultivation of <i>B. napus</i> .....	23
2.2.2.2 Phloem sap sampling.....	23
2.2.3 Nucleic acid methods .....	23
2.2.3.1 Genomic DNA isolation .....	23
2.2.3.2 RNA isolation .....	23
2.2.3.3 RNA precipitation .....	24
2.2.3.4 Plasmid isolation.....	24
2.2.3.5 Isolation of PCR amplicons .....	24
2.2.3.6 DNase digestion.....	24
2.2.3.7 Reverse transcription .....	25
2.2.3.8 Polymerase chain reaction (PCR) .....	25
2.2.3.8.1 Colony PCR.....	26
2.2.3.9 Agarose gel electrophoresis .....	26
2.2.3.10 Site directed mutagenesis .....	26
2.2.3.11 Restriction digest of plasmids and PCR products.....	26
2.2.3.12 Ligation .....	27
2.2.3.13 <i>Sma</i> I-cloning.....	27
2.2.3.14 Golden gate cloning.....	27
2.2.3.15 T7 <i>In vitro</i> transcription of RNAs .....	28
2.2.3.16 Verification of RNA integrity .....	29

2.2.3.17 5-mC RNA Methylation ELISA .....	29
2.2.3.18 RNaseR digestion .....	30
2.2.3.19 Sanger sequencing.....	30
2.2.3.20 Illumina RNA-sequencing .....	30
2.2.4 Protein methods.....	31
2.2.4.1 Protein purification of AtGRP7, BnGRP7, AtGRP7mut and AtGRP7 short .....	31
2.2.4.1.1 Lysis of <i>E. coli</i> .....	31
2.2.4.1.2 Nickel-NTA affinity column.....	31
2.2.4.1.3 Dialysis .....	31
2.2.4.1.4 Size exclusion chromatography .....	31
2.2.4.1.5 Concentrating of protein .....	31
2.2.4.2 SDS-PAGE.....	32
2.2.4.3 Coomassie staining .....	32
2.2.4.4 Mass spectrometry .....	32
2.2.4.4.1 Identification of proteins through peptide mass fingerprint .....	32
2.2.4.5 UV-crosslinking of protein and RNA .....	33
2.2.4.6 CnBr-column .....	33
2.2.4.6.1 Immobilisation of GRP7 on CnBr sepharose beads .....	33
2.2.4.6.2 RNA-binding of immobilized GRP7 .....	33
2.2.4.7 ThioflavinT Assay .....	34
2.2.4.8 Liquid-liquid-phase-separation Assay .....	35
2.2.4.9 SEC-SAXS measurement .....	35
2.2.4.10 Microscale thermophoresis (MST) .....	36
2.2.5 Bioinformatic methods.....	37
2.2.5.1 Pairwise sequence alignment.....	37
2.2.5.2 RNA-seq data mapping with STAR.....	37
2.2.5.3 Detection of circular RNAs with circTools.....	38
2.2.5.4 Data analysis with Access .....	39

2.2.5.5 SAXS data analysis .....	39
3 Results .....	40
3.1 Purification of glycine-rich protein 7 .....	40
3.2 RNA-binding of GRP7.....	42
3.2.1 <i>B. napus</i> leaf RNAs bound by AtGRP7 .....	42
3.2.1.1 Illumina sequencing of AtGRP7 enriched leaf RNAs.....	43
3.2.2 Phloem RNAs bound by BnGRP7 .....	46
3.2.2.1 Illumina sequencing of BnGRP7 enriched phloem RNAs.....	49
3.2.3 RNA-binding affinities of GRP7 towards different types of RNA.....	51
3.2.4 RNA methylation of total phloem RNAs and CnBr-column enriched RNAs .....	56
3.3 The glycine-rich region of GRP7 .....	59
3.3.1 Comparison of AtGRP7 and BnGRP7 .....	59
3.3.2 Interaction of the RGG-domain with RNA during RNA-binding .....	62
3.3.3 The glycine-rich region of AtGRP7.....	64
3.3.3.1 Condensation behavior of AtGRP7, AtGRP7short and AtGRP7mut .....	66
3.3.3.2 The overall structure of partly disordered AtGRP7 .....	68
3.3.3.3 RNA-binding of AtGRP7, AtGRP7short and AtGRP7mut .....	70
3.4 Circular RNAs in phloem.....	73
3.4.1 circRNA detection in long-noncoding sequencing library of total phloem RNA and leaf RNA.....	74
3.4.2 circRNA detection in circRNA enriched sequencing library of total phloem and leaf RNA .....	81
4 Discussion .....	84
4.1 Successful purification of GRP7.....	84
4.2 GRP7 possess a degenerative RNA-binding affinity .....	84
4.2.1 GRP7 binds a broad range of phloem-mRNAs.....	85
4.2.2 RNA-binding affinity of AtGRP7 and BnGRP7 towards different RNAs reveals a preference of binding sRNAs.....	87
4.3 The function of the RGG-domain of GRP7 .....	90

4.3.1 Differences between AtGRP7 and BnGRP7 in RNA-binding affinity and RGG-domain	90
4.3.2 Interaction of the RGG-domain with RNA upon RNA-binding	93
4.3.3 The glycine-rich region of AtGRP7 holds an important function in RNA-binding	94
4.3.3.1 Fibrillization and liquid-liquid phase separation of AtGRP7 caused by its glycine-rich C-terminus	96
4.4 The phloem sap of <i>B. napus</i> contains circRNAs	98
5 Conclusion and Outlook	103
6 List of figures	105
7 List of tables	109
8 References	111
9 Supplements	124
9.1 Additional Material and Methods	124
9.1.1 Oligonucleotides	124
9.1.2 Plasmid maps	126
9.2 Additional figures for results	128
9.2.1 Solubility of AtGRP7	128
9.2.2 Degradation of AtGRP7	129
9.2.3 CnBr-Sepharose beads negative control	130
9.2.4 Additional Illumina sequencing results	132
9.2.5 Normalized MST-time traces and binding curves of AtGRP7	142
9.2.6 Normalized MST-time traces and binding curves of AtGRP7short	153
9.2.7 Normalized MST-time traces and binding curves of AtGRP7mut	156
9.2.8 Tables with dissociation constants of AtGRP7mut and AtGRP7short	166
9.2.9 SEC-SAXS measurement of AtGRP7	167
9.2.10 circRNAs identified in lnc and circRNA sequencing	169

### 1 Introduction

In contrast to many other living organisms, most plants are photoautotroph and sessile. Because they are sessile, they are not able to escape any biotic or abiotic factors, like drought, nutrient deprivation and plant pathogens. Additionally, they encounter a wide range of changing circumstances throughout different seasons and during their development. Consequently, plants must tackle every condition they are experiencing as quickly and efficiently as possible to secure their development, growth and survival. For this, plants must communicate efficiently by transmitting signaling molecules.

One way of communication within plants is cell-to-cell transport of signals through plasmodesmata (PD) which are cytoplasmic channels. PD provide intercellular connections, thereby interconnecting the whole plant [1]. By this, plants could conduct long-distance transport of signals, however cell-to-cell transport is slow. To be able to provide each cell with water, nutrients, sugar and signals, multicellular plants evolved a faster long-distance transport system with a high conductivity [1]. This transport system is the vascular system, which spans the complete plant axial from root to shoot and spreads into leaves [2]. Signaling molecules like phytohormones were identified to travel as signals through the vascular system and it was discovered that proteins and mRNAs can act as signals as well [3,4]. To achieve a better understanding for long-distance signaling of plants, RNAs, especially mRNAs, which can act as signals need to be analyzed. More knowledge of the signaling system in plants can provide new options to breed plants resistant to biotic and abiotic stresses. Thus, the loss of harvest due to climate change causing new biotic and abiotic stresses could be reduced.

#### 1.1 The vascular tissue in plants

The vascular system of higher plants consist of two parts, the xylem and the phloem. The xylem is responsible for the transport of water and nutrients taken up by the roots while the phloem transports mainly photoassimilates and signaling molecules [1]. The xylem and phloem are arranged in vascular bundles, which are organized differently depending on the plant species and on the plant part [1]. The vascular bundles of *Arabidopsis thaliana* (*A. thaliana*) are organized in open collateral bundles as in many dicotyledone plants. These bundles are arranged with the phloem located towards the outside of the stem whereas the xylem is located towards the inside. Within leaves, the xylem is arranged on the adaxial and the phloem on the abaxial site [1]. Open collateral bundles contain a procambium. In

## 1 Introduction

case of *A. thaliana*, it is located between the xylem and the phloem. The cambium provides the meristematic cells of the vascular system.

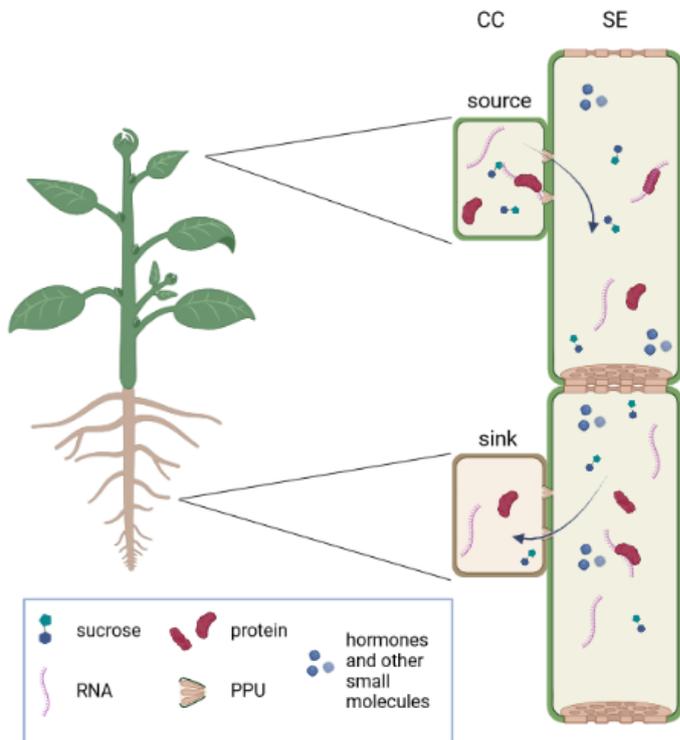
In angiosperms like *A. thaliana* the xylem consists of different cell types. The imperforate tracheary are long and narrow cells. They do not have a perforated plate as the name already suggests but are conducting tracheid cells with a large amount of bordering pits adjacent to perforated tracheary cells. Vessel cells, which are perforated tracheary cells, contain a perforation plate, providing an end-to-end connection of the cells. With this connection, they form a long vessel that is capable of conducting the transport of water and minerals [2]. Both types of tracheary cells contain thickened and lignified secondary cell walls and are undergoing a programmed cell death (PCD) to reach full maturity [1]. The cell content is removed from cells by the breakdown of the vacuolar membrane, releasing the vacuolar content. This leads to the autolytic hydrolysis of the cell content and thereby to cell death [5]. The third cell type are parenchyma cells. They function in synthesis and storage of molecules and they are the only cell type still living at maturity within the xylem [2]. The transport of water and nutrients is driven by transpiration through the roots to stomata openings of the leaves, since the water potential within the soil is higher than the water potential of the air. Thereby, the transport within the xylem is unidirectional [2].

In contrast, the transport within the phloem is bidirectional and the phloem does not consist mostly of dead but of living cells. The four cell types existing in the phloem are the sieve elements (SE), companion cells (CC), phloem parenchyma cells (PP) and fibers [2,6]. The fibers are providing the other phloem cells with stability on the abaxial site [6]. SE contain thickened cell walls without lignification and even though they are living cells surrounded by a plasma membrane they do not contain a nucleus or functional ribosomes nor vacuoles and Golgi apparatus [1,6,7]. Some organelles, like the smooth endoplasmic reticulum (ER), mitochondria and plastids still remain within the mature SEs [1]. SEs are connected to each other by sieve plates (SP) which contain pores for the transport of molecules. The pores emerge from plasmodesmata (PD) which penetrate the sieve plate. Callose deposition occurs at the cell walls surrounding the PDs until the callose replaces the previously deposited cell wall in this area so that a cylinder is formed around the PD. Large, expanded pores are formed by the removal of the ER within the PD as well as the detachment of the surrounding callose and the middle lamella during the development of SPs [7].

CCs are neighboring cells of SEs, securing the survival and function of mature SEs [1]. They are specialized parenchyma cells containing an extraordinarily high number of mitochondria, ribosomes and plastids, thereby having a very dense cytoplasm [1,7]. CCs are connected through specialized plasmodesmata, so called pore plasmodesmata units (PPU), to SEs. Multiple PD-branches of one PPU

## 1 Introduction

exist at the CC, which are fused into one single pore at the SE side [8]. Compared to the size exclusion limit (SEL) of 0.8 kDa of mesophyll and epidermal PDs [9], the SEL of PPU is up to 60 kDa [10] and therefore considerably higher. Due to the close connection of SEs and CCs as well as their shared origin, they are often referred to as the SE-CC complex [11].



**Figure 1: Transport of molecules through the phloem.** Within the phloem, different molecules like RNA, hormones, proteins and sucrose are transported from the source tissue (high content of the molecule) towards the sink tissue (low content of the molecule). This transport is bi-directional. The figure was created with BioRender.

The transport of resources through the phloem underlies the principle of a source to sink movement driven by the high pressure difference in between source and sink which has been already described in 1930 by Münch [1,7] (Figure 1). The loading of the phloem is proposed to work in different ways, apoplastic or symplastic [12] depending on the plant species [13]. During apoplastic loading, sugars are loaded from the apoplast actively into the CC or directly into the SEs by membrane transporters against a high concentration gradient [12,14]. For symplastic loading, passive diffusion of sugars from surrounding cells through PDs into CCs is possible along a concentration gradient from mesophyll cells (high sugar concentration) into the phloem (lower sugar concentration) [15]. Therefore, CCs need a high number of PDs connecting them to adjacent cells [12]. Another mechanism of symplastic loading is the polymer trap. Sucrose diffuses from mesophyll cells into specialized CCs, so called intermediary cells [16]. There, sucrose is converted into larger oligosaccharides like raffinose. Such large sugars cannot diffuse back through PDs into mesophyll cells and are thereby accumulated in the phloem [17].

## 1 Introduction

For other molecules like RNAs and proteins it is proposed that they reach the SE-CC complex by cell-to-cell movement through PDs [18].

### 1.2 Content of phloem and long-distance transport and signaling

Besides sugars, a broad spectrum of molecules is found within the phloem. Everything from inorganic molecules [19], phytohormones like auxin, abscisic acid (ABA), gibberellin and cytokinin [20] to proteins [21,22] and RNAs [23,24] is present. SEs do not contain a nucleus and no functional ribosomes [6] but tRNA halves which inhibit translation [25]. Thus, RNAs and proteins present within the phloem sap have to move from adjacent cells into the SEs. The first intercellular movement of endogenous macromolecules was reported for *KNOTTED1* mRNA and protein [26,27]. How the movement, loading and unloading of RNAs and other macromolecules in the phloem functions is largely unknown. By grafting and *Cuscuta* experiments, the mobility of proteins and RNAs is studied [28–30] and to identify proteins and RNAs within the phloem, phloem sap can be sampled. A variety of different sampling methods exist, like aphid stylet sampling [31,32] or the puncturing of the phloem and spontaneous exudation [21]. By all these approaches, a better understanding on phloem sap content and long-distance movement of molecules can be achieved.

#### 1.2.1 RNA content and long-distance transport and signaling

Around 1000 different mRNAs have been identified in the phloem sap of varied plant species like *Ricinus communis*, cucumber and watermelon [33–35]. For the vast majority of these mRNAs the function is still unknown and around 50 % are not conserved between species [1]. The long-distance signaling function of only a few mRNAs has been shown so far. *BEL5* mRNA for example moves through the phloem and is involved in tuber formation [3], while *GAI1* and *GAI2* mRNA, two homologs from the GRAS family, are involved in long distance trafficking and regulate the leaf development in plants [36].

Besides mRNAs, also small RNAs like siRNAs, miRNAs and tRNAs have been identified within the phloem [23,37]. For example, miR395, miR398 and miR399 in *Brassica napus* (*B. napus*). All of them show a differential expression within the phloem during nutrient deprivation: miR399 during phosphate [38], miR395 during sulfate stress [23], and miR398 during copper deprivation [23]. Due to their change in abundance within the phloem during stress they are likely responding as long-distance signals. Just recently, another class of RNA was found in phloem sap of apple trees, circular RNAs [39]. So far circular RNAs have not been considered as long-distance signals, even though it has been known

## 1 Introduction

for several decades that viroids are circular [40] and that they are able to move through the phloem [41,42].

By the analysis of RNA sequencing data from grafted *A. thaliana*, more than 2000 mobile transcripts were identified [30], which might travel through the phloem across graft junctions. Mobile transcripts were analyzed for similarities in their structure, modifications and sequences to find a common motif which leads to their selection for long-distance transport. A structural motif that was found are tRNA-like structures, which seem to be enriched, but not significantly, within the RNAs identified as mobile [43]. This structure was predicted for over 50 mobile transcripts [43]. In addition to this structural feature, around a quarter (562 RNAs) of graft-mobile RNAs showed 5-methyl-cytosin methylation sites, which might be another feature for the selection of mobile RNAs [44]. The identified possible selection features do not cover even half of mobile mRNAs. Thus, the selection of mRNAs for long-distance transport and signaling is still an open question.

### 1.2.2 Protein content and long-distance transport

Besides RNAs, phloem sap also contains proteins. The number of identified non-redundant proteins varied between species, in pumpkin 1209 proteins were identified by LC-MS/MS analysis whereas 103 different proteins in *B. napus* [21] and 16 in cucumber were detected by 2D-PAGE and MALDI-MS [21,45]. In rice phloem, 107 proteins have been observed [46]. It is assumed that these proteins are transported through the PPU from CCs into SEs, due to the fact that mature SEs do not contain a nucleus and functional ribosomes [6]. That phloem proteins are likely to be transported through PPU from CC to SE was shown by a study observing the change of SEL of mesophyll plasmodesmata up to over 20 kDa by injection of phloem sap protein fractions [47]. For example, phloem protein 2 (PP2) was identified to mediate such SEL changes [47]. Proteins found in phloem samples are involved in a broad range of different biological processes, like stress response, protein degradation, signaling, RNA-binding, and translation [21,22].

Some of the identified phloem proteins are proteins with chaperone function, like cyclophilins and heat shock proteins, e.g. heat shock protein 70 (HSP70) [21,22]. Since chaperons help with protein or RNA folding and cell-to-cell transport, the proteins found might be involved in the transport of proteins into the SEs [48]. Additionally, proteins involved in protein degradation have been found in pumpkin and *B. napus*. Ubiquitin and ubiquitin-like proteins as well as parts of the proteasome were observed within the phloem [21,22,49], opening the possibility of protein degradation within SEs or another function of the ubiquitin proteins.

## 1 Introduction

Furthermore, proteins involved in translation have been identified in phloem. Translation initiation and elongation factors were found in phloem of pumpkin as well as *B. napus*. Even parts of the small and large subunits of ribosomes were detected [21,22], despite the fact that functional ribosomes are not present in mature SEs [6]. Additionally, by the isolation of ribonucleoprotein complexes from phloem sap, parts of the ribosome subunits were found to be associated with ribosomal RNA [49].

Other phloem proteins are known to be involved in stress responses, one example are lectins, which are present in phloem sap of *B. napus* [21]. Over a hundred stress-related proteins were identified in pumpkin phloem sap [22]. Aside from stress proteins, also proteins involved in developmental processes, like the long-distance signal FLOWERING LOCUS T (FT) which is participating in flower induction, were identified in phloem sap of *B. napus* [21,50].

RNAs are known to be present within the phloem sap and it is assumed that RNAs are usually associated with RNA-binding proteins (RBPs) for transport over short and long distances. This was shown for CsPP2 which was able to translocate viroid and viral RNA [42,51]. Therefore, it is not surprising that multiple RNA-binding proteins were identified in phloem sap of different plants [21,22,46,52]. Besides PP2, the phloem protein 1 from *Cucurbit maxima* (*C. maxima*, pumpkin) has RNA-binding capacity and was found in phloem sap of pumpkin [22] as well as CmPP16 [22]. This protein was shown to move together with its own mRNA into SEs and is able to transport other RNAs [53], thereby indicating the importance of RBPs for RNA transport. Another class of RBPs was found in *B. napus* phloem sap: the glycine-rich RNA-binding proteins [21]. The glycine-rich RNA-binding protein 7 (GRP7) was not only found in *B. napus*, but also in pumpkin [22], rice [46] and castor bean [52] phloem sap samples, indicating that the function of this protein in the phloem is conserved.

### 1.3 RNA-binding proteins

As the name suggests, RNA-binding proteins (RBP) are capable of binding RNAs. RBPs are involved in an array of different cellular processes such as splicing [54,55], translation [56], RNA transport [53] and stress response [57–59].

Typical RBPs contain highly conserved RNA-binding domains, like the K homology (KH) domain [60], RNA recognition motif (RRM) [61,62], ZINC-finger domain [63] and DEAD-box helicase domain [64,65].

The RRM is one of the most abundant RNA-binding domains and consist of two conserved sequence motifs, RNP1 and RNP2 [62]. The around 80 amino acids long RRM consists of a four-stranded antiparallel beta-sheet and two alpha-helices [62]. Between the beta-sheets and alpha-helices lay variable, unstructured regions and the C- and N-terminal regions of RRM-containing proteins are often

## 1 Introduction

highly disordered [62]. The RNAs interact with RNP1 and RNP2, thereby laying on top of the beta-sheets [62]. The RNA-binding selectivity of RBPs is not determined by the RRM but by the unstructured parts of the RRM and the N- and C-terminal regions of RRM-containing proteins [62].

The analysis of the RBP-proteome of *A. thaliana* revealed that more than 50% of the identified proteins (461) did not contain a highly conserved RNA-binding domain, 168 RBPs contained classical and 181 proteins non-classical RNA-binding domains [66]. This finding is in agreement with other RBP-proteome studies [67].

### 1.3.1 Intrinsically disordered RNA-binding domains

Besides classical RNA-binding domains, it was shown that many RBPs harbor intrinsically disordered regions (IDR) which enable them to bind RNAs [67]. One example for such RNA-binding IDR is the arginine-glycine-glycine (RGG) domain (or RGG-box) [68]. Another IDR often found within RBPs are [G/S]Y[G/S] motifs [69]. The functions of proteins containing RNA-binding IDR are broad. Among other things, they act in splicing [54] translation [70,71] and formation of granules like stress granules [72].

Proteins containing molecule-interaction domains like the RRM, long IDR (like [G/S]Y[G/S] motif) as well as repetitive motifs tend to undergo liquid-liquid phase separation (LLPS) [73] and aggregation [74]. LLPS was reported for proteins like fused in sarcoma (FUS) and hnRNPs [73,75], both RBPs containing RRM and IDR. FUS and hnRNP induced LLPS was reported with and without RNA present and was disrupted by phosphorylation within the IDR like tyrosine phosphorylation in the [G/S]Y[G/S] motif of FUS [76]. LLPS was also disrupted by arginine methylation of FUS RGG-domain [77].

The RGG-domain is a low complexity region since no general structure can be determined [78]. It is usually located towards either terminus of the protein and consists of glycine-glycine repeats, often interspersed with arginine and aromatic amino acids [68,79]. Besides arginine, no other positively charged amino acids are observed in the RGG-domain [78]. The spans of glycine can vary between different proteins [78] and appear to function as spacers between arginine, thereby positioning the arginine at the right place for RNA-binding [80]. This might define the RNA-binding specificity and affinity of the protein towards certain RNAs [67]. In general, the interaction of RGG-domain with RNAs are dynamic and of low affinity [81]. Through methylation of arginine in the RGG domain, the function of the protein can be regulated by abolishing the RNA-binding capability [78,80]. Another way of regulating proteins with RGG-domain is altering the length of the RGG-domain via alternative splicing, thereby resulting in proteins with shorter RGG-domain missing their original function [78]. RNA-binding of RGG-domain relies much on arginine. Arginine side chains can form hydrogen-bonds with the phosphate backbone of RNA as well as with ribose and bases, preferably guanine [81,82]. In

## 1 Introduction

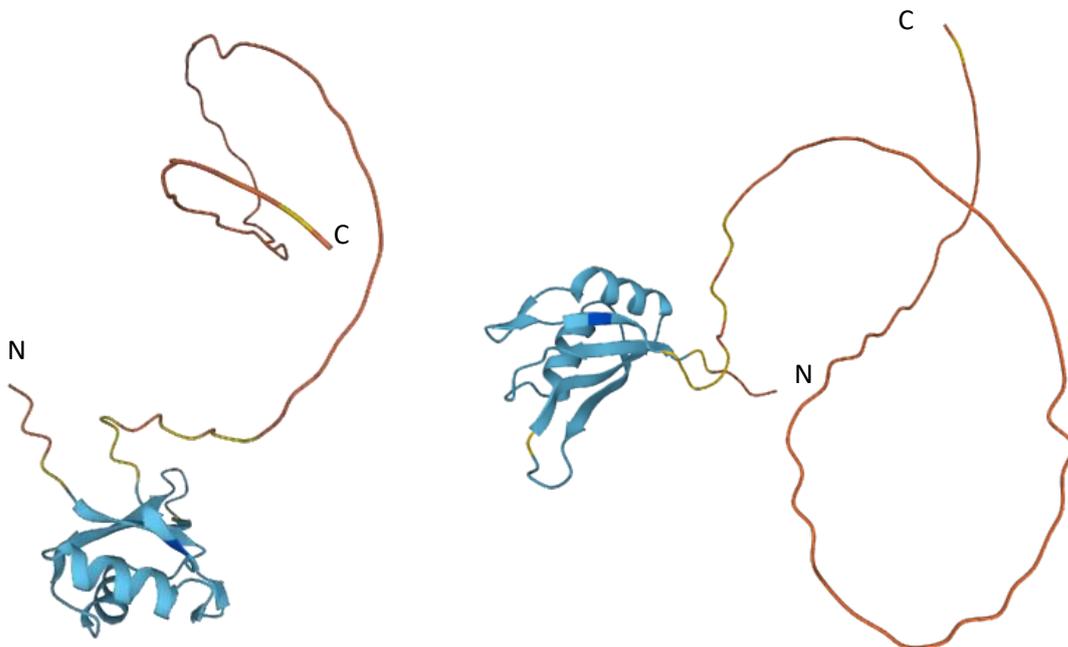
addition, arginine side chains are involved in pi-stacking and cation pi-interaction [81,83]. Likewise, interspersed tyrosines in the RGG are involved in pi-stacking and hydrogen-bonds with nucleotid-bases [81].

The RGG-domain not only contributes to RNA-binding of proteins, but is also involved in LLPS [81]. Arginine can not only interact by pi-pi stacking, electrostatic interactions, pi-cation stacking and hydrogen bonds with RNA but also promotes the formation of LLPS [83]. Glycine can perform pi-pi interaction too, due to its exposed backbone. In addition, glycine can increase pi-pi stacking of neighboring amino acids by providing flexibility and higher accessibility of their side chains [81]. Tyrosine as well as phenylalanine can promote LLPS by pi-pi stacking as well [83].

In phloem sap, the glycine-rich RNA-binding protein 7 is known to contain an RGG-domain.

### 1.3.2 Glycine-rich RNA-binding protein 7

The glycine-rich RNA-binding protein 7 (GRP7) is a roughly 17 kDa big protein, containing a N-terminal RRM-domain and a glycine-rich C-terminal domain [84] (Figure 2).



**Figure 2: Predicted protein structure of AtGRP7.** AtGRP7 contains two major domains, the RRM at the N-terminus and the C-terminal glycine-rich region. The RRM has a highly conserved structure with two alpha helices and four beta sheets[62]. The glycine-rich region is unstructured. The confidence of the model ranges from highly confident (dark blue), over confident (light blue) through low (yellow) to very low (orange). Therefore, the prediction of the RRM is confident, whereas a very low confidence is stated for the prediction of the glycine-rich region. N: N-terminus, C: C-terminus. The protein structure prediction was received through AlphaFold [85,86].

## 1 Introduction

The RRM provides GRP7 the ability to interact with RNAs, a single point mutation within the RRM inhibits this function [87]. The glycine-rich domain of GRP7 also plays an important role in RNA-binding, since the truncation of it results in a reduced RNA-binding capability [88,89]. The glycine-rich domain is intrinsically disordered [90] and contains repetitions of poly-glycine interrupted by small patches of serine, tyrosine and arginine [84].

Generally, GRP7 is located in the cytosol as well as in the nucleus, to which it can be imported by the nuclear import receptor transportin1 (TRN1) [59]. Just recently, AtGRP7 was also reported to be located within the apoplast together with GRP8 [91].

GRP7 is one of the proteins that have been identified in phloem sap of different plant species [21,22]. Therefore, GRP7 might be important for the selection of specific RNAs for long-distance transport and signaling, which makes it an interesting candidate to study. GRP7 was reported to bind different RNAs, like mRNAs [92] and vsiRNAs [88]. So far, GRP7 is known to be involved in a variety of different biological processes like circadian rhythm [87] and alternative splicing [55], as well as stress response [57,93] and pathogen response [88,94].

GRP7 as well as its closely related relative glycine-rich RNA-binding protein 8 (GRP8) are undergoing a circadian oscillation influenced by the circadian clock core proteins [95]. By a negative feedback loop, GRP7 stays in a stable oscillation during the day by influencing its own transcript via alternative splicing (AS) through a cryptic splicing site [96]. This leads to a transcript containing 166 additional nucleotides originating from the intron [95]. Since this additional sequence is harboring a prematurely stop codon, a loss-of-function truncated version of GRP7 is translated [95,96] and the AS transcript undergoes degradation by nonsense-mediated decay (NMD). GRP7 functions as a slave oscillator of the circadian clock and is influencing not only its own transcript levels but also the transcripts of many other genes by AS, leading to degradation of these transcripts through the NMD pathway [55,97].

Besides its involvement in AS, GRP7 is upregulated during cold stress in *A. thaliana* [93,95]. Furthermore, it increases the viability of *E. coli* during cold shock by substituting the function of an endogenous chaperone responsible for cold stress tolerance, leading to the hypothesis that GRP7 is involved in cold response [58]. This function seems to be conserved in glycine-rich proteins throughout monocotyledons and dicotyledons [93]. The glycine-rich domain enables GRP7 to undergo liquid-liquid phase separation (LLPS) which might play a key-role in the cold adaption process. The condensates formed by GRP7 LLPS contain RNAs as well as the RNA chaperons cold shock protein 1 and 3 [90]. Similar proteins like FUS and hnRNP1 were reported to undergo LLPS as well [75,98].

Furthermore, GRP7 participates in pathogen response. During viral infections of zucchini yellow mosaic virus (ZYMV) and tobacco rattle virus (TRV), GRP7 inhibits the spread of the virus during the early

## 1 Introduction

infection phase [88]. This inhibitory effect was abolished after removal of the C-terminal glycine-rich region. Additionally, in GRP7 overexpression plants, the spread of tobacco mosaic virus (TMV) is impaired and necrotic spots are reduced in size and abundance compared to wild type (WT) plants [94]. In *grp7*-knock out plants, the virus spread was increased for all three viruses [88,94]. The same inhibitory effect on pathogen spread of GRP7 overexpression plants was observed during *Pectobacterium carotovorum* infection [94]. In contrast to these results, the spread of *Botrytis cinerea* was increased in GRP7 overexpression plants compared to WT and *grp7*-knock out plants [94]. The GRP7 homologue small RNA-binding protein 1 (CsSRBP1) impairs the spread of pathogen virus by binding the vsRNAs [88]. These are small viral RNAs generated by DICER-like enzymes and GRP7 might deliver the vsRNAs to the RISC-complex which is then able to digest complementary mRNAs. GRP7 could also travel with vsRNAs through PDs, since the glycine-rich region provides GRP7 the ability to move from cell to cell [88]. This would raise the possibility that GRP7 is thereby warning neighboring cells before the virus is able to spread locally [88]. It was also suggested, that the glycine-rich region is important for the inhibitory effect of GRP7 on the spread of ZYMV, since ZYMV degrades the glycine-rich region of GRP7 during the infection and the truncated version of GRP7 was not able to impair the local and systemic spread of ZYMV [88].

During abiotic stress, GRP7 shows differing responses. During oxidative stress, GRP7 as well as GRP2 and GRP8 were upregulated [99] while during salt stress, the expression of GRP7 was reduced in contrast to the increased expression of other GRPs [100]. Not only high salt concentrations repressed the expression of GRP7 but also ABA and mannitol. In addition, *grp7*-knock out plants were more sensitive towards high salt, ABA and mannitol [101]. Two stress and ABA responsive genes, *RD29A* and *RAB18* were upregulated in *grp7*-knock out plants compared to WT plants grown with and without ABA [101]. This suggests an involvement of GRP7 in ABA stress response.

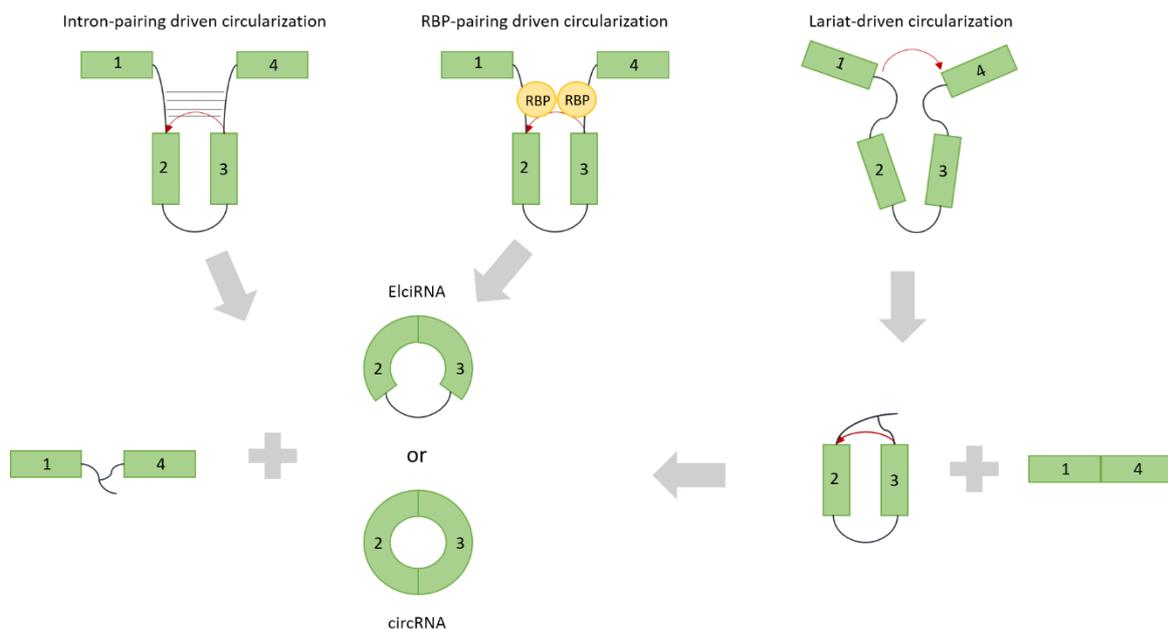
According to a recent publication, AtGRP7 is one of the few RNA-binding proteins identified within the apoplastic fluid of *A. thaliana* together with a high amount of circular RNAs (circRNAs), thus suggesting that AtGRP7 is capable of interacting with and binding circRNAs [91].

### 1.4 Circular RNAs

Circular RNAs (circRNAs) are single-stranded, long-non coding RNAs with their 5' end covalently linked to the 3' end [40]. This link is formed by so called back-splicing, mediated through the spliceosome and canonical splicing sites [102–105]. The back-splicing junctions (BSJ) are often flanked with especially long and insufficiently spliced introns [104,106]. Splicing of circRNAs is regulated by intronic complementary sequences (cis-factors) and splicing factors (trans-factors) [107].

## 1 Introduction

Three different mechanisms drive the circularization of RNAs (Figure 3). The first mechanism is based on reverse complementary sequences, like ALU-repeats [108], in introns flanking the circularized part of the RNA. Through the complementary sequence introns pair with each other, thereby bringing the 5' donor and the 3' acceptor back-splicing sites close together [109]. The second mechanism includes RBPs, binding at binding sites close to the donor and acceptor site and bringing them into close proximity probably by dimerization [110,111]. The third mechanism for the biogenesis of circRNAs is exon skipping during mRNA splicing [103,112]. By this, the GU/AG splicing of the mRNA precursor leads to an exon-intron intermediate which undergoes back-splicing and thereby circularization [113]. This mechanism seems to be widespread within plant circRNA formation [112]. Most circRNAs are derived from exons of coding genes, but circRNAs can also contain introns and microexons [106,114,115] or originate from introns, UTRs and intergenic regions [116]. How circRNAs are degraded is mostly unknown. Two possible ways of circRNA degradation are small-RNA mediated degradation, which has been shown for one of the best studied circRNAs in human brain, CDR1as [117], and the recruitment of endonucleases by N<sup>6</sup>-methylation of adenosine [118].



**Figure 3: Biogenesis of circRNA.** Circular RNAs can be derived in three different ways: Intron-pairing, RBP-pairing and lariat-driven circularization. Either elciRNA (Intron containing circRNAs) or circRNAs (containing only Exons) can be the result. Graphic was adapted from Ebbesen *et. al* 2016 [119].

In 1976 circRNAs have been described for the first time. They were observed in viroids, pathogenic small RNAs, which can infect plants [40]. The circular form of the viroid RNA offers higher stability against degradation since the 3' and 5' ends of RNAs are not accessible for exonucleases. This is especially important because they are not protected by coat proteins [40]. Since their discovery, circRNAs have been reported in many different eukaryotes [120] like drosophila [121], mammals

## 1 Introduction

[106,122], plants [120,123] and archaea [124]. Most research is found on circRNAs in mammals, especially human. At the beginning, circRNAs were thought to be just by-products or miss-spliced products of canonical and alternative splicing without any function [102]. This changed after identification of circRNAs in deep sequencing results with even higher expression than their linear counterpart mRNA [104,106]. Until now, circRNAs were shown to be involved in miRNA sponging [117], protein sponging [121] and some are even translated [121,125]. The expression of circRNAs is depending on the tissue [111] and developmental stage [121,126], indicating that they do have a specific function.

### 1.4.1 Circular RNAs in plants

Looking into plant research, not much is known about circRNAs and their function within plants despite the fact that circRNAs were first described in plant viroids [40]. In plants, circRNAs were predicted to be generated from exons, exons-introns, untranslated regions (UTRs), intergenic regions and they were reported to even span over multiple genes [127]. The last type of circRNAs could not be confirmed by further experiments [127].

Plant-circRNAs were identified by lncRNA sequencing data or sequencing data from circRNA enriched sequencing libraries of different plant species, different tissues and during different developmental stages and biotic and abiotic stresses [39,112,123,127–130]. In *A. thaliana*, 970 circRNAs have been predicted [112], 3528 in kiwifruit [130] and only 62 in barley [128], thereby displaying a broad variety in circRNA expression in plants. Additionally, some circRNAs showed a differential expression under various biotic and abiotic stresses, thereby indicating an involvement in stress response [127,130,131].

In apple trees, circRNA expression was investigated from five different tissues, roots, leaves, flowers, fruits and phloem. In total, over 6000 circRNAs have been predicted in all five tissues [39]. Of these RNAs, only 175 circRNAs were found in all five tissues [39], possibly indicating a tissue specific function for most of the circRNAs. In phloem, 1026 tissue specific circRNA have been identified and the genes from which they derive show a Cluster of Gene (COG) enrichment in different functional classifications, for example signal transduction mechanism [39].

Some circRNAs have been validated by additional experiments besides next generation sequencing (NGS) for example by RT-qPCR, but in most cases only few circRNAs have been tested [39,112,123,127]. The functions of circRNAs in plants are still largely unknown [132], only five circRNAs have been found to be potential miRNA sponges, but the function has been demonstrated only for one of the five RNAs experimentally [133]. Other functions of circRNAs and their biogenesis only have been predicted by bioinformatic tools, therefore circRNA research in plants lacks of experimental validations [132].

## 1 Introduction

### 1.5 Objective

The aim of this work is a better understanding of the mechanisms of the RNA long-distance transport and signaling in the phloem of plants. Plants are sessile organisms; thus, they cannot escape biotic and abiotic stresses they encounter. The vascular system interconnects the whole plant and takes part in the transmission of information from one plant part towards another. Especially important for the long-distance signaling system is the phloem where a broad spectrum of molecules, like different classes of RNAs, are transported over long-distances. Besides RNAs, RNA-binding proteins (RBPs) were found within the phloem and it was suggested that the RNA transport is conducted by certain RBPs [21,22,47,52]. The glycine-rich RNA-binding protein 7 (GRP7) can be found in the phloem of different plants [21,22,47,52]. Consequently, GRP7 is an interesting candidate to study. Thus, AtGRP7 and BnGRP7 should be studied to understand the involvement of GRP7 in RNA long-distance transport.

For this, the RNA-binding ability of both GRP7s should be analyzed. The measurement of the binding affinity of both proteins towards different RNAs with different modifications should reveal patterns of their selectivity and maybe even of RNAs that are phloem-mobile. Additionally, the function of the glycine-rich region of GRP7 regarding the RNA-binding should be investigated. Thus, different GRP7 mutants should be tested and the differences between AtGRP7 and BnGRP7 are also a matter of interest.

Furthermore, the circular RNA content of *B. napus* phloem should be researched. Circular RNAs are a class of long-noncoding RNAs which have not been in the focus of research in plant science. Due to their circular form, these RNAs are extraordinary stable and thereby provide a great foundation for long-distance transport and signaling. CircRNAs were just recently discovered in the phloem [39]. Additional interest lays within circRNAs since viroids, small plant pathogens consisting only of a circRNA as genome [40], were found to be traveling through the plant over long-distances [41,134,135]. Since pathogens as reduced as viroids are using the cellular machinery of their hosts, it is likely that they might have adapted their circular form to achieve their mobility and to mimic other functions of the plant [41,42,134]. This indicates that circRNAs might harbor a crucial role in plant signaling.

A better understanding of the long-distance signaling system would provide crucial benefits. Most importantly, it could be used to increase the resistance of plants against abiotic and biotic stresses, which in turn could provide a vital tool to tackle the new problems emerging due to climate change.

## 2 Material and methods

## 2.1 Materials

## 2.1.1 Equipment

The equipment which was used is listed in Table 1.

**Table 1: Machines/equipment used in this thesis.**

<b>Equipment</b>	<b>company</b>
-80 °C freezer	Binder, Tuttlingen (D)
-20 °C freezer	Liebherr, Biberach (D)
+4 °C fridge	Liebherr, Biberach (D)
ÄKTA™ start	GE Healthcare Bio-Sciences AB, Uppsala (S)
ÄKTAprime™ PLUS	GE Healthcare Bio-Sciences AB, Uppsala (S)
Avegene SLite 140S gel-documentation	Pacific Image Electronics, New Taipei City (TWN)
Balance EG	Kern&Sohn GmbH, Balingen (D)
Branson Sonifier 250	Branson Ultrasonics, Eemnes (NL)
Bioanalyzer 2100	Agilent, Santa Clara (US)
Centrifuge MiniSpin	Eppendorf, Hamburg (D)
Centrifuge Sigma 1-16 k	Sigma, Osterode am Harz (D)
Centrifuge Hettich ROTINA 380R	Hettich, Tuttlingen (D)
Centrifuge Beckman Coulter Avanti JXN-30	Beckman Coulter Diagnostics, Krefeld (D)
Chemidoc Touch Gel	BioRad, Munich (D)
Heating block Thermomixer comfort	Eppendorf, Hamburg (D)
Magnetic stirrer MR 3001	Heidolph Instruments GmbH & CO. KG, Schwabach (D)
Monolith NT.115™	NanoTemper, Munich (D)
NanoDrop one <sup>c</sup> UV/Vis spectrophotometer	NanoDrop products, Wilmington (US)
Olympus MVX10 Macroscope	Olympus, Hamburg (D)
Peristaltic pump 2232 Microperplex S	Pharmacia LKB, Uppsala (S)
pH-meter	Mettler-Toledo, Gießen (D)
Photometer	Eppendorf, Hamburg (D)
PowerPac™ Basic Power Supply	BioRad Munich (D)
Precision balance ABJ	Kern&Sohn GmbH, Balingen (D)
Shaker HT CH-4103	Infors HT AG, Bottmingen (CH)
Stratalinker UV Crosslinker 2400	Stratagene, San Diego (US)
TECAN Spark plate reader	Tecan Group Ltd, Männedorf (CH)
Thermocycler PCR cycler T3000	Biometra, Göttingen (D)
ThermoShaker TS1	Biometra, Göttingen (D)
Ultraflex III MALDI-TOF-TOF	Bruker, Bremen (D)
Vacuum pump Aeromat	KNF, Freiburg (D)
Vortexer VF2	IKA® Labortechnik, Staufen (D)

## 2 Material and methods

### 2.1.2 Consumables

Consumables (except enzymes) used in this thesis are listed in Table 2. Pipet tips and plastic reaction tubes (50 ml, 15 ml, 2 ml and 1.5 ml) were acquired from Sarstedt (Sarstedt, D) and plastic reaction tubes of 1.5 ml, 0.5 ml (low binding) and 0.2 ml were ordered from Eppendorf (Hamburg, D).

**Table 2: Consumables used in this work.**

<b>product</b>	<b>company</b>
Agilent RNA 6000 Nano Kit	Agilent, Santa Clara (US)
Agilent RNA 6000 Pico Kit	Agilent, Santa Clara (US)
Avegene Plant Genomic DNA Mini kit	Avegene, Taipei Hsien (TWN)
GelRed DNA Stain	Biotium, Hayward (USA)
GeneRuler™ 1 kb plus DNA ladder	Thermo Fisher Scientific (Darmstadt, D)
HisTrap™	GE Healthcare, Uppsala (S)
Kimtech Science Precision Wipes Tissue	Kimberly-Clark Professional, Koblenz-Rheinhafen (D)
Monolith NT.115 Standard treated capillaries	NanoTemper, Munich (D)
Ni-NTA Agarose	Qiagen, Hilden (D)
NucleoSpin Gel and PCR Clean up kit	Macherey&Nagel, Düren (D)
NucleoSpin Plasmid easy pure kit	Macherey&Nagel, Düren (D)
Parafilm M	Brand, Wertheim (D)
PD10 empty columns	GE Healthcare, Uppsala (S)
TRIZOL™	Thermo Fisher Scientific, Darmstadt (D)
TRIZOL™ LS	Thermo Fisher Scientific, Darmstadt (D)
Vivaspin centrifugal concentrators	Sartorius, Göttingen (D)
Zymo RNA Clean up and concentrator Kit (-25)	Zymo research (Freiburg, D)

### 2.1.2 Chemicals

All chemicals that have been used were ordered from Roth (Karlsruhe (D)), AppliChem (Darmstadt (D)), Merck (Darmstadt (D)), Sigma Aldrich (Taufkirchen (D)) and Serva (Heidelberg (D)) unless stated otherwise.

### 2.1.3 Oligonucleotids

Oligonucleotides were synthesized by EurofinsGenomics, Ebersberg (D). A list of all used oligonucleotides can be found in supplements in Table 33.

## 2 Material and methods

### 2.1.4 Enzymes

All enzymes used in this work are listed in Table 3.

**Table 3: Enzymes used in this work.** Manufacturers of each enzyme are listed.

<b>enzyme</b>	<b>manufacturer</b>
<i>Bsal</i>	New England Biolabs (Frankfurt am Main, D)
<i>DpnI</i>	New England Biolabs (Frankfurt am Main, D)
<i>HindIII</i>	New England Biolabs (Frankfurt am Main, D)
<i>NdeI</i>	New England Biolabs (Frankfurt am Main, D)
<i>SmaI</i>	New England Biolabs (Frankfurt am Main, D)
<i>XhoI</i>	Thermo Fisher Scientific (Darmstadt, D)
Phusion high fidelity polymerase	Thermo Fisher Scientific (Darmstadt, D)
Dream Taq Polymerase	Thermo Fisher Scientific (Darmstadt, D)
T4 DNA Ligase	Thermo Fisher Scientific (Darmstadt, D)
T4 RNA Ligase II	Thermo Fisher Scientific (Darmstadt, D)
Revert Aid reverse transcriptase	Thermo Fisher Scientific (Darmstadt, D)
Maxima H- reverse transcriptase	Thermo Fisher Scientific (Darmstadt, D)
RiboLock	Thermo Fisher Scientific (Darmstadt, D)
RNaseA	Thermo Fisher Scientific (Darmstadt, D)
RNaseR	Lucigen (Teddington, UK)
T7 polymerase	In house
TaqPolymerase	In house
Trypsin	Promega (Madison, US)
Thrombin	GE Healthcare (Uppsala, S)

### 2.1.5 Bacterial cultivation media

All media used for bacterial cultivation are listed in Table 4 and additives for these media are listed in Table 5. For solid media, 15 g of Agar Agar was added in 1 l.

**Table 4: Different media used for bacterial cultivation.**

<b>Medium</b>	<b>Components (per liter)</b>
<b>LB</b>	10 g NaCl 10 g trypton 5 g yeast extract
<b>ZY</b>	10 g tryptone 5 g yeast extract
<b>ZYP-5052 auto-induction medium</b>	930 ml ZY 1 ml MgSO <sub>4</sub> 20 ml 50x 5052 50 ml 20x NPS
<b>TB</b>	12 g trypton 24 g yeast extract 5 g glycerol
<b>TB medium</b>	900 ml TB 100 ml 10x TB salts

## 2 Material and methods

**Table 5: Different additives for the media used for bacterial cultivation.**

Additives	Components
<b>20x NPS</b>	0.5 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> 1 M KH <sub>2</sub> PO <sub>4</sub> 1 M Na <sub>2</sub> HPO <sub>4</sub>
<b>50x 5052</b>	250 g/l glycerol 25 g/l glucose 100 g/l α-lactose
<b>10x TB salts</b>	0.17 M KH <sub>2</sub> PO <sub>4</sub> 0.72 M K <sub>2</sub> HPO <sub>4</sub>

### 2.1.6 Buffers

General buffers that have been used are listed in Table 6.

**Table 6: General buffers used in during this work.** Buffer components and pH for each buffer is listed in the table.

Buffer	Buffer components	pH
<b>50x TAE-buffer</b>	2 M Tris 1 M acetic acid 0.05 M EDTA	~8.5
<b>6x DNA loading dye</b>	10 mM Tris HCl 0.15 % (w/v) orange G 0.03 % xylene cyanol 60 % (v/v) glycerol 60 mM EDTA	7.6
<b><i>In vitro</i> transcription buffer</b>	500 mM Tris-HCl 150 mM MgCl <sub>2</sub> 50 mM DTT 20 mM Spermidin	7.5
<b>10x SDS running buffer</b>	0.25 M Tris 1.92 M Glycerol 1 % (w/v) SDS	~8.3
<b>6x SDS loading buffer</b>	0.375 M Tris-HCl 60 % (v/v) Glycerin 0.6 M DTT 9 % SDS 0.02 % (w/v) Bromphenoleblue	6.8
<b>Colloidal coomassie staining solution</b>	5 % (w/v) Aluminiumsulfate (14-18) hydrate 10 % (v/v) Ethanol (96 %) 0.02 % (w/v) CBB G-250 2 % (v/v) ortho-Phosphoric acid (85%)	
<b>MST-buffer</b>	25 mM Tris HCl 150 mM NaCl 1 mM DTT	~8.0

## 2 Material and methods

<b>10x crosslinking buffer</b>	10 mM Tris HCl 150 mM NaCl 1 mM Mg <sub>2</sub> Cl	7.5
<b>CnBr-column assay buffer</b>	25 mM HEPES 150 mM sodium acetate	7.0

Buffers for purification of AtGRP7, AtGRP7mut and BnGRP7 proteins are listed in Table 7.

**Table 7: Buffers used for GRP7 protein purification.** The components and pH of lyse-, elution-, dialysis- and SEC-buffer are listed.

<b>Buffer</b>	<b>components</b>	<b>pH</b>
<b>Lyse buffer</b>	50 mM HEPES 200 mM KCl 1 % Glycerol 30 mM Imidazol 1 mM PMSF 1 mM DTT	8.0
<b>Elution buffer</b>	50 mM HEPES 300 mM KCl 1 % Glycerol 1 M Imidazol 1 mM DTT	8.0
<b>Dialysis buffer</b>	25 mM HEPES 150 mM KCl 1 mM DTT	7.0
<b>SEC-buffer</b>	25 mM HEPES 150 mM KCl 3 % Glycerol 1 mM PMSF 1 mM DTT	7.0

## 2 Material and methods

Buffers for AtGRP7 short protein purification are listed in Table 8.

**Table 8: Buffers used for AtGRP7short protein purification.** The components and pH of lyse-, elution-, dialysis- and SEC-buffer are listed.

Buffer	components	pH
<b>Lyse buffer</b>	50 mM HEPES 500 mM NaCl 1 % Glycerol 30 mM Imidazol 1 mM DTT	8.0 (4° C)
<b>Elution buffer</b>	50 mM HEPES 500 mM NaCl 1 % Glycerol 1 M Imidazol 1 mM DTT	8.0 (4° C)
<b>Dialysis buffer</b>	25 mM HEPES 300 mM NaCl 1 mM DTT	7.0 (4° C)
<b>SEC-buffer</b>	25 mM HEPES 300 mM NaCl 1 mM DTT	7.0 (4° C)

### 2.1.7 Plasmids

Plasmids which have been used are listed in Table 9. The Plasmid maps can be found in supplements 9.1.2 (Figure 40 and Figure 41).

**Table 9: Overview of plasmids used in this work.**

Plasmid	Company
<b>pet28a+ expression plasmid</b>	Merck, Millipore, Darmstadt (D)
<b>pUC57 cloning plasmid</b>	Genescript, Rijswijk, (NL)

### 2.1.7 Bacterial strains

All bacterial strains used in this work are listed in Table 10.

**Table 10: Bacterial strains and their genotype and antibiotic resistance**

species	strain	Genotype/chromosomal background	Resistance
<i>E. coli</i>	XL10 Gold	<i>E. coli</i> TetrD(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacIqZDM15 Tn10 (Tet') Amy Cam <sup>r</sup> ].	Tetracyclin, Chloramphenicol
<i>E. coli</i>	BL21-Gold (DE3)	<i>E. coli</i> B F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal λ(DE3) endA Hte	Tetracyclin
<i>E. coli</i>	BL21-Codon RIPL (DE3)	Plus <i>E. coli</i> B F <sup>-</sup> ompT hsdS(r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) dcm <sup>+</sup> Tet <sup>r</sup> gal λ(DE3) endA Hte [argU proL Cam <sup>r</sup> ] [argU ileY leuW Strep/Spec <sup>r</sup> ]	Tetracyclin, Streptomycin, Chloramphenicol

## 2 Material and methods

### 2.1.8 Plants

All plants that were used in this work are listed in Table 11.

**Table 11: Plants used in this work.**

<b>Plant</b>	<b>origin</b>
<i>Arabidopsis thaliana</i> Columbiana 0	MPI (AG Kragler), Golm (D)
<i>Brassica napus</i> cv. Drakkar	CBGP, Madrid (E)

### 2.1.9 Genes of purified proteins

The gene-IDs for the purified proteins are listed in Table 12.

**Table 12: Genes that have been used for protein expression and purification.**

<b>Gene</b>	<b>Usage</b>
At2g21660	Protein expression, AtGRP7
BnaC08g49360D	Protein expression, BnGRP7

## 2 Material and methods

### 2.1.10 Bioinformatic tools

All bioinformatic tools used in this thesis are listed in Table 13.

**Table 13: Software used in this thesis.**

<b>Tool</b>	<b>Publisher</b>
Agilent 2100 Expert Software	Agilent, Santa Clara (US)
ATSAS 3.1.3	Manalastas-Cantos <i>et. al</i> (2021) [136]
Benchling	<a href="http://www.benchling.com/">www.benchling.com/</a>
BioRender	<a href="http://www.biorender.com/">www.biorender.com/</a>
Circtools version 1.2.1	<a href="https://www.github.com/dieterich-lab/circtools">www.github.com/dieterich-lab/circtools</a> [137]
CHROMIX	Panjekovich & Svergun (2018) [138]
CRY SOL	Svergun <i>et. al</i> (1995) [139]
DAMMIF	Franke & Svergun (2009) [140]
Galaxy EU server	<a href="http://www.usegalaxy.eu/">www.usegalaxy.eu/</a>
GNOM	Semenyuk & Svergun (1991) [141]
flexControl Software	Bruker Daltonik, Bremen (D)
iTasser	<a href="http://www.zhanggroup.org/I-TASSER/">www.zhanggroup.org/I-TASSER/</a> [142]
MASCOT server	<a href="http://www.matrixscience.com/">www.matrixscience.com/</a>
mMass	<a href="http://www.mMass.org/">www.mMass.org/</a>
MO.Affinity Analysis Software version	NanoTemper, Munich (D)
OriginPro 2021b	OriginLab, Northampton (US)
Pairwise sequence alignment by EMBOSS Needleman	<a href="http://www.ebi.ac.uk/Tools/psa/emboss_needle/">www.ebi.ac.uk/Tools/psa/emboss_needle/</a>
ProtParam	<a href="http://www.web.expasy.org/protparam/">www.web.expasy.org/protparam/</a>
PRIMUS	Konarev <i>et. al</i> (2003) [143]
PyMOL	Schrödinger, <a href="https://www.github.com/schrodinger/pymol-open-source">www.github.com/schrodinger/pymol-open-source</a>
SASpy	Panjekovich & Svergun (2016) [144]
SnapGene	SnapGene, San Diego (US)
SREFLEX	Panjekovich & Svergun [145]
STAR version 2.5.2b-2	<a href="https://www.github.com/alexdobin/STAR">www.github.com/alexdobin/STAR</a> [146]

## 2 Material and methods

### 2.2 Methods

#### 2.2.1 Bacterial methods

##### 2.2.1.1 Overnight cultures of *E. coli*

*E. coli* was grown overnight at 37 °C and 150 rpm in 5-10 ml of LB-medium and the according antibiotic fitting to the resistance of the bacterium. The overnight cultures were used as pre-cultures for expression cultures or for plasmid isolation.

##### 2.2.1.2 Glycerol stocks of *E. coli*

Glycerol stocks of *E. coli* were prepared by mixing 300 µl of 100 % Glycerol with 700 µl of *E. coli* overnight culture. The glycerol stocks were frozen and stored at -80 °C.

##### 2.2.1.3 Preparation of chemically competent *E. coli*

A 5 ml pre-culture of *E. coli* was grown at 37 °C and 150 rpm in LB medium overnight. This overnight culture was used to start an *E. coli* culture in 200 ml of LB medium. This culture was grown overnight at 24 °C and 150 rpm up to an OD<sub>600</sub> of 0.4-0.6. The culture was chilled on ice for at least 10 min and then harvested by centrifugation at 4 °C and 4800 xg for 10 min in two 50 ml reaction tubes. This step was repeated once so the whole culture was pelleted. The pellets were gently dissolved in 10 ml ice-cold TB buffer each and incubated for 10 min on ice. Centrifugation was carried out at 4 °C and 4800 xg for 10 min. The supernatant was discarded and the pellets were gently resuspended in 1.86 ml of ice-cold TB-buffer and 0.14 ml of 100 % DMSO. The competent cells were incubated for 10 min on ice and aliquoted in 100 µl aliquots, frozen in liquid nitrogen and stored at -80 °C.

##### 2.2.1.4 Transformation of *E. coli*

Transformation of *E. coli* was done by heat shock. Therefore, the competent cells were thawed on ice and 2-10 µl of plasmid DNA or ligation mix was added to 100 µl of cells and shortly mixed by flipping the reaction tube. The cell-plasmid mix was incubated on ice for 20 min. The heat shock was done at 42 °C for 45 s. 200 µl LB was added and the cells were incubated for one hour at 37 °C and 150 rpm shaking.

After the incubation, bacteria were plated on LB-agar plates with an antibiotic in working concentration according to the resistance gene carried on the plasmid. The bacteria were grown overnight at 37 °C

##### 2.2.1.5 Expression culture

The overexpression of proteins was done in either in TB- or ZY-medium and the antibiotic in working concentration fitting to the resistance of the used *E. coli*. 400 ml of culture was grown in one 2.5 l triple baffled flask at 37 °C until an OD<sub>600</sub> of 0.7-0.9. To induce the protein expression in TB-medium 1 mM

## 2 Material and methods

of IPTG was used. ZY-medium used an autoinduction and therefore did not need any additional IPTG. For overnight expression the cultures were incubated at 24 °C and 150 rpm overnight after induction. For 3 h expression the culture was further grown at 37 °C for 3 h.

The bacteria were harvested by centrifugation for 30 min at 20 °C and 7500 xg. The pellets were either directly used for protein purification or stored at -20 °C.

### 2.2.2 Plant methods

#### 2.2.2.1 Cultivation of *B. napus*

*B. napus* (cv. Drakkar) plants were grown in a green house at 15 °C, 70 % humidity and long-day conditions. They were grown pots with a diameter of 18 cm and a soil-sand mixture of 3:1. Watering was done by an automatic watering system.

#### 2.2.2.2 Phloem sap sampling

Phloem sap sampling of *B. napus* was usually done around noon. Around 8-week-old plants which did not start to flower by that time were watered before sampling and punctured at the inflorescence stem multiple times. The first drop of all sites was removed with a tissue. All following drops were collected in a pre-cooled 1.5 ml safe seal reaction tube. The collection was carried out for around an hour, collected phloem sap was frozen in liquid nitrogen and stored at -80 °C.

### 2.2.3 Nucleic acid methods

#### 2.2.3.1 Genomic DNA isolation

Genomic DNA was isolated from *B. napus* and *A. thaliana* leaves. The leaves were frozen in liquid nitrogen and ground in a mortar cooled with liquid nitrogen. Approximately 100 mg plant material was filled into 1.5 ml reaction tubes. The genomic DNA isolation was carried out with the Plant Genomic DNA Mini from Avegene kit. The instructions of the provided protocol were followed. The genomic DNA was eluted by adding 50 µl of H<sub>2</sub>O.

Genomic DNA was frozen in liquid nitrogen and stored at -80 °C until further use.

#### 2.2.3.2 RNA isolation

RNA was isolated from different plant parts of *B. napus* as well as from leaves from *A. thaliana*.

For the RNA isolation of solid plant material like leaves and flowers, the plant material was frozen in liquid nitrogen and ground in a mortar cooled with liquid nitrogen. Around 100 mg of ground plant material was distributed in 1.5 ml reaction tubes. Following, TRIzol was used according to the protocol provided by the company. The optional centrifugation step number three after adding TRIzol for fatty samples was carried out to remove excess of plant material after 5 min of incubation with TRIzol. The washed and dried RNA pellets were dissolved in 50 µl of H<sub>2</sub>O and heated at 70 °C for 10 min.

## 2 Material and methods

For the RNA isolation of liquid plant material like phloem sap, LS TRIzol was used. The purification was carried out according to the manual provided by the company until step nine. Afterwards, the RNA was further purified through the ZYMO RNA and concentrator Kit -25 by following the instructions of the provided protocol from the kit. The elution of the RNA was done by adding 50 µl of RNase and DNase free water.

RNA was frozen in liquid nitrogen and stored -80 °C until further use.

### 2.2.3.3 RNA precipitation

For RNA precipitation 2.5 to 3 volumes of 100 % ethanol or 1-2 volumes of 100 % isopropanol were added to the RNA containing sample. Precipitation was carried out over night at -20 °C. Following, samples were centrifuged for 30 min at 4 °C and 20000 xg. The supernatant was removed and the pellet was washed with 1 ml of 75 % ethanol. The pellet was centrifuged down at 4 °C and 7500 xg for 5 min. Supernatant was removed and the pellet was dried for 10 min at RT under a hood. The pellet was dissolved in 50 µl of DEPEC-treated RNase and DNase free H<sub>2</sub>O and heated at 70 °C for 10 min. RNA was frozen in liquid nitrogen and stored at -80 °C until further use.

### 2.2.3.4 Plasmid isolation

For the isolation of plasmids, 2 ml of bacterial overnight culture was centrifuged in 2 ml reaction tubes at 7500 xg for 5 min at room temperature, the supernatant was removed and the centrifugation of 2 ml of culture was repeated one more time. The plasmid was further purified using the NucleoSpin plasmid easy pure kit from Macherey and Nagel according to the manufacturer's instructions. The plasmid DNA was eluted with 50 µl H<sub>2</sub>O.

### 2.2.3.5 Isolation of PCR amplicons

The isolation of PCR amplicons was performed either after an agarose gel run or directly after a PCR. For both, the PCR and gel clean up kit from Macherey and Nagel was used. The instructions of the kit were followed and the amplicon was eluted with 25 µl of H<sub>2</sub>O.

### 2.2.3.6 DNase digestion

DNase digestion of purified RNA was carried out with 1-10 U of DNase for 30 min at 37 °C. The reaction was stopped with 25 mM EDTA. As alternative, DNase digestion was carried out on the RNA Clean & Concentrator-25 RNA-Kit (Zymo Research) columns according to the manufactures protocol.

## 2 Material and methods

### 2.2.3.7 Reverse transcription

The synthesis of cDNA was done with the reverse transcriptase RevertAid from ThermoFisher according to the manufactures protocol. Between 150-1000 ng of RNA was used for the reaction. A typical reaction mix is shown in Table 14.

**Table 14: Reverse transcription reaction mix.** All components and the volume or amount of each is listed.

<b>component</b>	<b>Volume/amount</b>
5x reaction buffer	4 $\mu$ l
RNA	0.15-1 $\mu$ g
10 pmol Primer	1 $\mu$ l
10 mM dNTPs	2 $\mu$ l
20 U RiboLock	0.5 $\mu$ l
Reverse transcriptase	1 $\mu$ l
H <sub>2</sub> O	Up to 20 $\mu$ l total volume

For RNAs with high GC-content or many secondary structures it is recommended to mix primer, water and RNA, incubate them at 65 °C for 5 min prior adding the rest of the components.

The reaction was incubated at 42-45 °C for 1 h and stopped by 10 min incubation at 70 °C.

### 2.2.3.8 Polymerase chain reaction (PCR)

PCRs that were carried out usually had a total reaction volume of 20-100  $\mu$ l. A typical PCR reaction is shown in Table 15. For genes with a high GC-content 5-10 % DMSO was added to the reaction mix. When Phusion Polymerase (ThermoFisher Scientific) was used, either 5x HF-or 5x GC-buffer was selected depending on the GC-content of the template.

**Table 15: PCR reaction mix.** The table includes components and volume of each component.

<b>component</b>	<b>volume</b>
Phusion polymerase (2 U/ $\mu$ l)	1 $\mu$ l
5x buffer	5-10 $\mu$ l
10 mM Forward Primer	1 $\mu$ l
10 mM Reverse Primer	1 $\mu$ l
10 mM dNTPs	2 $\mu$ l
Template	1-2 $\mu$ l
H <sub>2</sub> O	Up to a total volume of 50 $\mu$ l

## 2 Material and methods

### 2.2.3.8.1 Colony PCR

The colonies were tested for successful cloning and transformation by a colony PCR after a transformation. Therefore, each tested colony was picked by a pipet tip and transferred into 20 µl of H<sub>2</sub>O. 10 µl of this water-colony mix was added to 10 µl of PCR mix which included forward and reverse Primer, dNTPs, Taq Polymerase, a 10x reaction buffer and 2 µl of 50 mM MgCl<sub>2</sub>.

For pet28a+ constructs pLic forward and reverse primer and for pUC57 constructs pUC-M13 forward and reverse primer were used.

### 2.2.3.9 Agarose gel electrophoresis

Agarose gel electrophoresis was performed within 1-3 % agarose gels with a volume of 50 ml or 100 ml 1xTAE buffer (Table 6) and with 2.5 µl or 5 µl of GelRed. 10-100 µl of sample containing 1x sample buffer were loaded on the gel. Usually gels run for 45 min at 120 V and were documented through Avegene SLite 140S gel-documentation.

### 2.2.3.10 Site directed mutagenesis

Two complementary primer have been designed for site directed mutagenesis. Both primers contain the desired nucleotide change within their middle and on both sites of the change, 15 nucleotides fitting to the template were added. A PCR with these primers was performed followed by a restriction digest with *DpnI* for the removal of the template plasmid containing the gene that was mutated. After the removal, 5 µl of this mix was transformed into *E. coli* XL10-Gold.

### 2.2.3.11 Restriction digest of plasmids and PCR products

Restriction endonuclease enzymes were bought from NEB or ThermoFisher and the restriction digestion was performed for 1 h at 37 °C and stopped by incubation at 80 °C for 10 min. The typical reaction mix is shown in Table 16. The instructions of the manufacturer were followed if not stated otherwise.

**Table 16: Restriction digest reaction mix.** The table includes components and the volume or amount of each component.

component	volume/amount
10 x CutSmart buffer	2 µl
Restriction enzyme 1	0,5-1 µl
Restriction enzyme 2	0,5-1 µl
Template DNA	1-2 µg for plasmids
H <sub>2</sub> O	Up to a total volume of 20 µl

## 2 Material and methods

### 2.2.3.12 Ligation

For the ligation of plasmids and inserts, T4 Ligase from ThermoFisher was used. A typical ligation reaction mix is shown in Table 17. The ligation was performed either for 2 h at 22 °C or for 16 h at 16 °C and stopped with an incubation at 70 °C for 10 min.

**Table 17: Ligation reaction mix.** The table includes components and the volume or amount of each component.

component	volume/amount
10 x T4 DNA Ligase buffer	2 µl
T4 DNA Ligase (5 U/µl)	1 µl
10 mM ATP	1 µl
Plasmid	100-200 ng
Insert	3-5x molar excess to plasmid
H <sub>2</sub> O	Up to a total volume of 20 µl

### 2.2.3.13 *Sma*I-cloning

Cloning of PCR-products into pUC57 was done by *Sma*I cloning in case the PCR products didn't contain a *Sma*I recognition and cutting sequence. For this cloning method, the restriction digestion and the ligation were combined into one step. A typical reaction mix is shown in Table 18. The reaction was incubated for 16 h at 16 °C and 5 µl of this were used further for the transformation into *E. coli*.

**Table 18: *Sma*I-cloning reaction mix.** The table includes all components and the volume or amount of each component.

component	volume/amount
10 x T4 DNA Ligase buffer	2 µl
T4 DNA Ligase (5 U/µl)	1 µl
<i>Sma</i> I (20 U/µl)	1 µl
10 mM ATP	1 µl
Plasmid	100-200 ng
Insert	3-5x molar excess to plasmid
H <sub>2</sub> O	Up to a total volume of 20 µl

### 2.2.3.14 Golden gate cloning

Golden Gate cloning was performed with modified pet28a+ (Figure 41). These plasmids contain two *Bsa*I-recognition sites with different and specific cutting sites, which enables directed cloning of the insert into the plasmid with only one restriction enzyme.

## 2 Material and methods

A typical golden gate reaction contained *Bsa*I and T4 Ligase at the same time, like described in *Sma*I cloning (Table 18). Therefore, restriction digestion and ligation were combined in one reaction. The reaction was incubated at 37 °C and 16 °C and inactivated at 80 °C (Table 19).

Table 19: Thermocycler-program for Golden Gate cloning.

temperature	time
37 °C	2 min
16 °C	5 min
37 °C	60 min
50 °C	10 min
80 °C	10 min

} x 50

### 2.2.3.15 T7 *In vitro* transcription of RNAs

The *in vitro* transcription of RNA was done for RNA sizes of 21 nucleotides to up to over 1000 nucleotides. Depending on the size, the double stranded DNA template with T7 promotor was generated in different ways.

For small RNAs up to 100 nucleotides two complementary DNA oligonucleotides with the T7 promotor sequences at the 5' ends were ordered. The complementary oligonucleotides were undiluted combined one to one and cooked at 95 °C for 5 min. Afterwards they were cooled down at RT for 10-20 min.

For larger RNAs starting from 100 nucleotides, a PCR with a forward primer containing the T7 promotor was performed to generate a double stranded DNA template. Since the *in vitro* transcription of most RNAs were carried out multiple times, the genes were cloned into the pUC57 plasmid for easy amplification without the need of prior RNA isolation and cDNA synthesis.

The *in vitro* transcription was carried out in a total volume of 100 µl. The general compounds of the reaction mix are shown in Table 20. For longer RNAs 10 % of DMSO was added in the reaction mix. For cy5-labeled RNAs, 0.05 - 0.25 mM of cy5 UTP was added and the amount of UTP was reduced to 1.8-1.9 mM UTP per reaction. For methylated RNAs, usually 1 mM of NTP and 1 mM of methylated NTP like m<sup>5</sup>CTP per reaction were used.

*In vitro* transcription for small RNAs was performed over night at 37° C. For long RNAs *in vitro* transcription was conducted at 30° C for 2-3 h. Following the transcription, the DNA Template was removed by DNase digestion as described in 2.2.3.6 and further purified through RNA Clean & Concentrator-25 RNA- Kit (Zymo Research, Freiburg, D).

## 2 Material and methods

**Table 20: *In vitro* transcription reaction mix.** The table includes the components and volume or amount of each component used.

<b>component</b>	<b>volume/amount</b>
10x reaction buffer	10 $\mu$ l
Pyrophosphatase (5 U/ $\mu$ l)	5 $\mu$ l
100 mM ATP	2 $\mu$ l
100 mM GTP	2 $\mu$ l
100 mM CTP	2 $\mu$ l
100 mM UTP	2 $\mu$ l
RiboLock (20 U/ $\mu$ l)	1 $\mu$ l
T7-Polymerase (1000 U/ $\mu$ l)	5 $\mu$ l
template DNA	5-10 pmol
H <sub>2</sub> O	Up to a total volume of 100 $\mu$ l

### 2.2.3.16 Verification of RNA integrity

The bioanalyzer was used to verify the integrity of RNAs. Nanochips were used and loaded according to the manufacturer's instructions and total eukaryotic RNA assay was selected.

### 2.2.3.17 5-mC RNA Methylation ELISA

The percentage of 5-mC methylated RNA within total phloem RNA and in the third elution of BnGRP7 bound RNAs was assessed through MethylFlash™ 5-mC RNA Methylation ELISA Easy Kit (Fluorometric) from Epigentek (Farmingdale, US). As suggested by the manual of the manufacturer, 200 ng of RNA was used and the assay was carried out as described in the manual.

## 2 Material and methods

### 2.2.3.18 RNaseR digestion

To test whether or not circRNAs are present in phloem and leaf of plants, RNaseR digestion of total phloem and leaf RNA was performed. Therefore, different concentrations of RNaseR and RNaseA were used to confirm a complete digestion. An overview of the reaction mix is shown in Table 21.

The RNase digestion was carried out for 30 min at 37 °C. Afterwards, the leftover RNA was purified with ZYMO Clean up and concentrator-25 kit from Zymo Research (Freiburg, D).

**Table 21: RNaseR digestion reaction mix.**

<b>component</b>	<b>volume/amount</b>
RNaseA or RNaseR	1 U
total RNA	1 µg
10x Reaction buffer	2 µl
H <sub>2</sub> O	up to 20 µl

### 2.2.3.19 Sanger sequencing

Sanger sequencing of plasmids was carried out by Eurofins genomics (Ebersberg, D) and Microsynth Seqlab (Göttingen, D).

### 2.2.3.20 Illumina RNA-sequencing

RNA for RNA sequencing was checked for quality through the bioanalyzer as well as the nanodrop. Phloem RNA was additionally checked for its purity by performing a PCR targeting the small RuBisCO subunit and Thioredoxin H. RNA that was qualified for RNA sequencing was send to Novogene (Cambridge, UK). RNA sequencing library preparation and Illumina sequencing was performed by Novogene (Cambridge, UK).

## 2 Material and methods

### 2.2.4 Protein methods

#### 2.2.4.1 Protein purification of AtGRP7, BnGRP7, AtGRP7mut and AtGRP7 short

The purification of all proteins was performed quite similar in the following five steps. All steps have been performed at RT, except for AtGRP7 short. For this protein, the purification was done at 4 °C. The ÄKTA™start is at RT, the ÄKTAprime™ PLUS at 4 °C.

##### 2.2.4.1.1 Lysis of *E. coli*

For the protein purification, *E. coli* cells containing the expressed protein needed to be lysed. Therefore, the pelleted cells were resuspended in 50 ml of lysis buffer (Table 7 and Table 8) for each protein), containing additional 1 mM PMSF, 1 mM AEBSF and 100 µg/ml Lysozyme. The lysate was incubated for 20 min at RT stirring. Next the cells were lysed by sonification on ice. The sample was sonificated for 30 s at 50 % duty cycle and energy output 5 and then incubated for 30 s on ice. This step was repeated eight times. To remove the cell debris, centrifugation was carried out at 15 °C (4 °C for AtGRP7 short) and 40000 xg for 30 min. The supernatant was filtered with a 0.4 µm filter.

##### 2.2.4.1.2 Nickel-NTA affinity column

After the lysis of *E. coli*, the supernatant was used to perform a Nickel-NTA affinity chromatography. The supernatant loaded onto a 5 ml Nickel-NTA column with a flowrate of 1 ml/min. The loading of the column was followed by washing with 30 ml lysis buffer containing 1 M KCl (for AtGRP7 short, normal lysis buffer was used). The column was connected to the ÄKTA™ start (for AtGRP7 short ÄKTAprime™ PLUS) and a gradient of imidazol up to 1 M was used to elute the protein.

##### 2.2.4.1.3 Dialysis

Fractions (3 ml per fraction) from the nickel-NTA affinity column containing the protein were collected, combined and 20 U of Thrombin was added. Everything was transferred into a dialysis tube with a MWCO of 3.0 kDa. The dialysis was performed for 2 h at RT (for AtGRP7 short the dialysis was performed overnight at 4 °C) against 1 l of dialysis buffer.

##### 2.2.4.1.4 Size exclusion chromatography

The dialyzed protein was concentrated up to 2-5 ml using a Vivaspin50 centrifugal filter with a MWCO of 5000 Da (for AtGRP7 short a Vivaspin50 centrifugal filter with MWCO of 3500 Da was used) at 4800 xg and RT. Then, 2-5 ml of protein was loaded into a sampling tube at the ÄKTA™ start (for AtGRP7 short, ÄKTAprime™ PLUS was used). The size exclusion chromatography (SEC) was performed with a HiLoad™ 16/600 Superdex™ 75 µg SEC-column with a volume of 120 ml.

##### 2.2.4.1.5 Concentrating of protein

After collection of protein containing fractions (3 ml per fraction), the fractions were checked on an SDS-PAGE for purity of the protein and combined. These fractions were concentrated using a

## 2 Material and methods

Vivaspin50 filter with a MWCO of 5000 Da (for AtGRP7 short a Vivaspin50 filter with MWCO of 3500 Da was used) at 4800 xg and RT. The concentration was carried out until a protein concentration of at least 100 µM was reached.

### 2.2.4.2 SDS-PAGE

#### 2.2.4.3 Coomassie staining

SDS-PAGE gels were stained with a colloidal Coomassie brilliant blue staining solution (Table 6). Therefore, the SDS-PAGE was washed three times with H<sub>2</sub>O for 5 min. Afterwards, the gel was incubated in the staining solution for at least 30-60 min or overnight.

#### 2.2.4.4 Mass spectrometry

##### 2.2.4.4.1 Identification of proteins through peptide mass fingerprint

To identify proteins after performing an SDS-PAGE, peptide mass fingerprint was performed.

Therefore, the protein band that needed identification was cut out of the SDS-PAGE and transferred into a 0.5 ml low binding tube. To wash the gel, 100 µl of NH<sub>4</sub>CO<sub>3</sub> was added, mixed and incubated for 5 min. After washing, the gel was de-stained by adding 50 µl NH<sub>4</sub>CO<sub>3</sub> and 50 µl of 100 % acetonitril. The gel piece was incubated at 37 °C and occasionally mixed during de-staining. Following the de-staining, the supernatant was removed and 100 µl 100 % acetonitril was added to the gel pieces. After 5 min of incubation, acetonitril was removed and the gel pieces were dried with open lid at RT. For the trypsin digestion, 20 µl trypsin (0,001µg/µl) was added to the gel pieces and an incubation at 37 °C for 2-16 h was conducted.

Following the trypsin digestion, the Anchor chip or stainless-steel target was loaded with the sample. Therefore, 1 µl of TA30 (30 % Acetonitril, 70 % H<sub>2</sub>O mit 0.1 % TFA) and 1 µl of sample were pipetted onto a free space of the anchor chip target and dried. The dried spot was covered with 0.5 µl of HCCA matrix (alpha-Cyano-4-hydroxycinnamic acid in TA30 (saturated)) and again dried. For the stainless-steel target first 0.5 µl of HCCA was used, dried and 1 µl of sample was added. This was dried again and covered with another 0.5 µl of HCCA. Besides the samples, the peptide standard (Peptide Calibration Standard (#206195, Bruker Daltronik, Bremen (D))) was loaded on the target and covered with 0.5 µl of HCCA matrix.

The peptide standard was measured first to adjust the MS. Following, the samples were measured with 15-20 % laser power on anchor chip or stainless-steel targets.

The program mMass was used to analyze the MS-data. Peaks were selected by hand and selected when they matched peaks observed for *in silico* trypsin digestion of the protein and had an intensity that was at least 3-5 times higher than the background. For analysis of peptide mass fingerprint, MASCOT server

## 2 Material and methods

with either Swissprot or NCBI database was used. Search parameters were default settings, Oxidation (M) was selected as a variable modification and the search was conducted either specifically for *A. thaliana* or *Viridiplantae*.

### 2.2.4.5 UV-crosslinking of protein and RNA

To crosslink BnGRP7 and RNA, the 10x crosslinking buffer (Table 6), 20  $\mu\text{M}$  of BnGRP7 and 50  $\mu\text{M}$  of *bna-miR164* were combined in a finale volume of 20  $\mu\text{l}$ . The mix was incubated for 15 min at RT. Following the incubation, the mix was irradiated 3 min for four times with UV-light (15 W per UV-bulb, 4000  $\mu\text{watts/cm}^2$ ) with Stratalinker UV-crosslinker.

### 2.2.4.6 CnBr-column

#### 2.2.4.6.1 Immobilisation of GRP7 on CnBr sepharose beads

For the immobilization of AtGRP7, AtGRP7short and BnGRP7, 500  $\mu\text{g}$  of protein and 40-50 mg of CnBr-sepharose 4G beads were used.

Sepharose beads were washed 3 times with 300  $\mu\text{l}$  of 1 mM HCl. After the first and second washing step, the beads were centrifuged for 30 s at 3000 xg and the supernatant was removed. Following the third washing step beads were transferred to a microcentrifugation spin column and supernatant was removed by centrifugation at 700 xg for 1 min. 500  $\mu\text{g}$  of protein and 1 mg of dextran sulfate with 250  $\mu\text{l}$  of coupling buffer were added to the beads and incubated over night at 4 °C rolling.

On the next day supernatant was removed by centrifugation at 4 °C and 700 xg for 1 min. The beads were washed twice with coupling buffer to remove any remaining unbound protein.

To block any open binding space on the beads, 400  $\mu\text{l}$  of 0.1 M Tris-HCl pH 8.0 was added to the beads and incubated over night at 4 °C rolling. Subsequent the beads were centrifuged at 4 °C and 700 xg for 1 min. The beads were stored in 500  $\mu\text{l}$  SEC-buffer of the immobilized protein containing 20 % glycerol at 80 °C after freezing in liquid nitrogen.

For negative control, beads were treated in the same way as described above without adding 500  $\mu\text{g}$  of protein (dextranulfat was added) for the first overnight incubation.

#### 2.2.4.6.2 RNA-binding of immobilized GRP7

To test the RNA-binding capability of immobilized GRP7, 70  $\mu\text{g}$  total leaf RNA and 25  $\mu\text{g}$  total phloem RNA was used. After storage at -80 °C, the beads were thawed on ice and washed three times with assay buffer (25 mM HEPES pH 7.5, 150 mM sodium acetate) in a microcentrifuge spin column. Following 200  $\mu\text{l}$  assay buffer and 70 or 25  $\mu\text{g}$  of RNA were added to the beads and incubated for 10 min at RT rolling. Flowthrough was collected by centrifugation at 4 °C and 700 xg for 30 s in a clean

## 2 Material and methods

1.5 ml reaction tube. The beads were washed five times with 300  $\mu$ l assay buffer, each washing step was collected in 1.5 ml tubes by centrifugation at 4 °C and 700 xg for 30 s.

The elution was done in one step for leaf RNA to test the protocol and after optimization in three steps for phloem RNA. The one step elution was carried out by incubating the beads with assay buffer containing 2 M of sodium acetate for 5 min rolling at RT. The elution was collected by centrifugation at 4 °C and 700 xg for 30 s in a clean 1.5 ml reaction tube.

For the three-step elution, the beads were first incubated with assay buffer containing 250 mM sodium acetate for 5 min rolling at RT. The elution was collected as described before. The second elution was carried out the same way as the first elution with assay buffer containing 500 mM sodium acetate. After the second elution was collected, the third elution was done in the same manner with assay buffer containing 2 M sodium acetate and collected by centrifugation.

Assay buffer with 2 M sodium acetate was added to the flowthrough, the washing steps and the first elution to increase the salt concentration to 300 mM. Following 2.5 to 3 volumes of 100 % ethanol was added to flowthrough, washing steps and elutions for RNA precipitation.

The one-step elution as well as the third elution of the three-step elution protocol were sent for RNA sequencing together with input leaf- and phloem RNA.

The same protocol with three elution steps was carried out with negative-control beads without immobilized protein.

### 2.2.4.7 ThioflavinT Assay

To measure protein aggregation caused by amyloid fiber-like structures, the fluorescence dye ThioflavinT and a slightly modified version of the protocol from Wördehofft *et. al* [147] was used for the measurements. The aggregation of AtGRP7, AtGRP7mut and AtGRP7short was monitored over 80 h in a Tecan Spark plate reader with and without RNA. Therefore, black 96-well plates were used and a total assay volume of 100  $\mu$ l. In all measurements, 20  $\mu$ M of ThioflavinT was included, for protein measurements 45  $\mu$ M of protein was added and for measurements including RNA, 100  $\mu$ M of *bnamiRNA164* and 1 mM of BnPARCL mRNA was used. Negative controls included everything except of protein.

The measurement was carried out at 27 °C  $\pm$  1 °C in a cycle of first mixing the sample orbital for 5 s with a frequency 510 rpm, then waiting for 30 s before measuring the fluorescence with an excitation 448 nm and the emission at 483 nm. The next cycle started after a waiting time of 15 min.

## 2 Material and methods

### 2.2.4.8 Liquid-liquid-phase-separation Assay

To test if AtGRP7 is able to induce liquid-liquid-phase-separation (LLPS) in presence of RNA, different protein concentrations were tested with a constant RNA concentration of 1  $\mu$ M. The RNA used was R21- labeled with cy3. RNA, 10x phase separation buffer and protein (1  $\mu$ M, 25  $\mu$ M and 50  $\mu$ M) were combined in a total volume of 20  $\mu$ l. After 10 min incubation at RT, PEG3350 was added with a finale concentration of 10 % (w/v). The samples were observed after 15 min of incubation time at RT or after 5 min incubation time on ice with an Olympus MVX10 Macroscope (Objective: MV Plapo 1x & 2x) using an RFP-filter.

### 2.2.4.9 SEC-SAXS measurement

AtGRP7 was purified as described above and handed over to EMBL-Hamburg (D). The measurement was conducted by Cy M. Jeffries (EMBL, Hamburg (D)) as well as the analysis. The sample details were listed in Table 22. Detailed information about the data collection can be found in Table 23.

Table 22: Sample overview for SEC-SAXS measurement.

Sample details	
Organism	<i>A. Thaliana</i>
UniProt sequence ID	Q03250
Extinction coefficient $\epsilon$ ( $A_{280}$ , 0.1%(w/v))	1.533
MW (Da)	16889.79
Energy (eV)	12400.4
Solvent	25 mM HEPES, 100 mM KCl, 1 mM DTT pH 7.0
Sample concentration	GRP7 2.1 mg/ml BSA batch 1.6 mg/ml

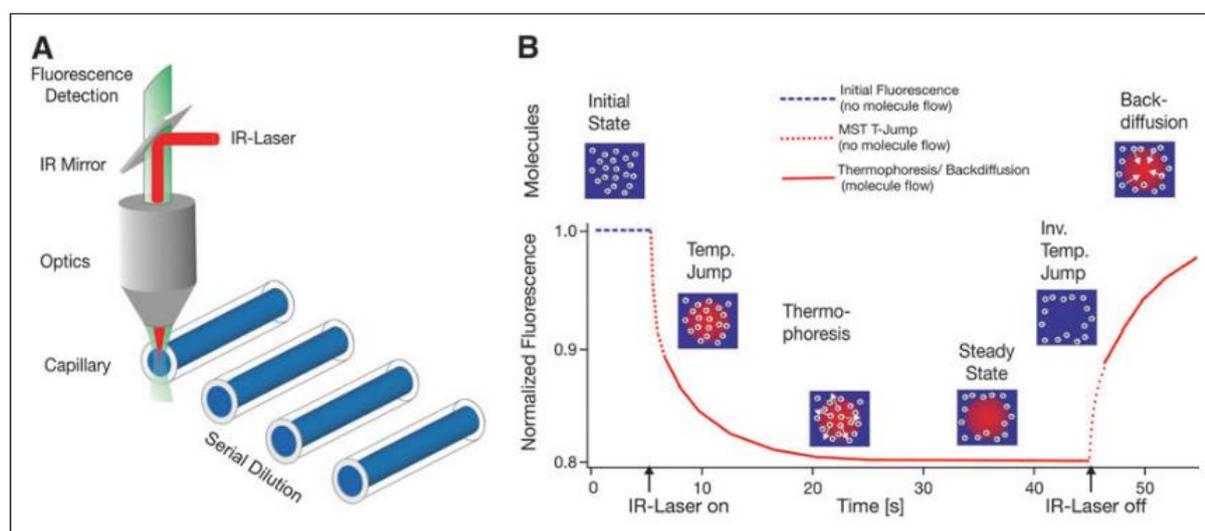
Table 23: SEC-SAXS data collection details.

SAXS data-collection	
Instrument/data processing	PETRA III Beamline P12 BioSAXS at DESY, Hamburg. PILATUS 6M detector
Wavelength (nm)	0.099984
Sample-to-detector distance (m)	3
Absolute scaling method	Relative scattering of pure water
Monitoring for radiation damage	Frame comparison
Exposure time (s/frame)	0.245000
Sample configuration	Monomer GRP7
Sample T ( $^{\circ}$ C)	23

## 2 Material and methods

### 2.2.4.10 Microscale thermophoresis (MST)

Microscale thermophoresis (MST) is a technique to track the movement of fluorescence labeled molecules through a temperature gradient on a microscopic scale in a volume of only a few  $\mu\text{l}$  [148]. The method is based on the thermophoresis of molecules, which describes the movement of molecules along a temperature gradient. This movement is influenced by the size, charge and interaction of the molecule with the solvent. All these factors can be influenced by an interaction between two molecules. Thereby, this effect is used in MST to investigate binding events between molecules. To achieve this, capillaries filled with a dilution row of the ligand and a constant target concentration. These capillaries are heated by an infrared laser in a small, defined spot. The movement of the molecules away from the heated area is tracked through the emission of the fluorescence labeled target within the heated area [148].



**Figure 4: Microscale thermophoresis (MST).** A) Visualization of the experimental set up with capillaries holding the serial dilution of the ligand. The infrared laser heats each capillary at a specific area and the fluorescence of the labeled target is measured. B) Graph showing a typical thermophoresis curve with different states during MST. The graphic used was published by Jerabek-Willemsen *et. al* [148].

MST-curves are generated for each dilution. From these curves normalized fluorescence can be calculated for each concentration used and applied on a diagram with concentration on a logarithmic x-axis and the normalized fluorescence on the y-axis. A binding curve can be fitted through the measured points, from which the dissociation constant can be obtained.

MST was conducted to determine the binding affinity of different proteins towards different RNAs. To achieve this, all RNAs were labeled with cy5-UTP as described in RNA *in vitro* transcription. A 16-step dilution row of protein was prepared with an input concentration of usually around 100  $\mu\text{M}$ , therefore the highest measured concentration was 50  $\mu\text{M}$ . the same volume as the diluted protein of 40 nM

## 2 Material and methods

(20 nM finale concentration) of labeled RNA was added to the dilution row and everything was measured in standard capillaries at RT. For each replicate a separate dilution row was prepared.

For the analysis of MST results, all long RNAs and all small RNAs were analyzed at the same MST-on time. The decision, which MST-on time was chosen for each protein, depended on the response amplitude. For a valid RNA-binding, a minimal response amplitude of 5 was expected and the MST-on time with most RNAs displaying a response amplitude of  $\geq 5$  was chosen. Therefore, the MST-on time of 6.93-7.93 s was used for the analysis of the  $K_d$  of AtGRP7 towards long RNAs as well as for all  $K_d$ s of AtGRP7mut and AtGRP7 short.  $K_d$ s for AtGRP7 towards short RNAs were determined at 6.09-7.09 s. For BnGRP7, an MST-on time of 0.11-1.11 s was chosen for the  $K_d$  analysis of long RNAs and 6.09-7.09 s for the  $K_d$  analysis of short RNAs.

### 2.2.5 Bioinformatic methods

#### 2.2.5.1 Pairwise sequence alignment

Pairwise sequence alignments were performed with the global pairwise sequence alignment tool EMBOSS-Needle from EMBL-EBI. The default settings were used.

#### 2.2.5.2 RNA-seq data mapping with STAR

RNA sequencing data from lncRNA sequencing as well as circRNA sequencing were mapped to the *B. napus* genome assembly AST\_PRJEB5043\_v1 [149] with STAR version 2.5.2b-2 [146], since this is the last STAR-version giving out chimeric junction reads on its own. STAR ran on the galaxy EU-server. The STAR settings suggested for circTools (DCC) [137,150] were used and are displayed in Figure 5 as the command line command.

## 2 Material and methods

A

```
mkdir -p tempstargenomedir && STAR --runMode genomeGenerate --genomeDir 'tempstargenomedir' --genomeFastaFiles
'/data/dnb07/galaxy_db/files/b/4/3/dataset_b430e57b-9f9f-4ed1-8781-01d816338464.dat' --sjdbGTFfile '/data/dnb06
/galaxy_db/files/3/2/6/dataset_326f750f-0762-4f76-8e16-1a6c10980128.dat' --sjdbOverhang '149' --runThreadN
${GALAXY_SLOTS:-4} && STAR --runThreadN ${GALAXY_SLOTS:-4} --genomeLoad NoSharedMemory --genomeDir
'tempstargenomedir' --sjdbGTFfile '/data/dnb06/galaxy_db/files/3/2/6/dataset_326f750f-
0762-4f76-8e16-1a6c10980128.dat' --sjdbOverhang '149' --readFilesIn '/data/dnb06/galaxy_db/files/1/e/7
/dataset_1e747c23-f226-4594-9fe5-d1c69a1d8bb3.dat' --readFilesCommand zcat --outSAMtype BAM
SortedByCoordinate --quantMode GeneCounts --outSAMattributes All --outSAMstrandField None
--outFilterIntronMotifs None --outSAMunmapped None --outSAMprimaryFlag OneBestScore --outSAMmapqUnique "255"
--outFilterType Normal --outFilterMultimapScoreRange "1" --outFilterMultimapNmax "20" --outFilterMismatchNmax
"999" --outFilterMismatchNoverLmax "0.5" --outFilterMismatchNoverReadLmax "1.0" --outFilterScoreMin "1"
--outFilterScoreMinOverLread "0.66" --outFilterMatchNmin "0" --outFilterMatchNminOverLread "0.7"
--seedSearchStartLmax "50" --seedSearchStartLmaxOverLread "1.0" --seedSearchLmax "0" --seedMultimapNmax "10000"
--seedPerReadNmax "1000" --seedPerWindowNmax "50" --seedNoneLociperWindow "10" --alignIntronMin "20"
--alignIntronMax "1000000" --alignMatesGapMax "1000000" --alignSJoverhangMin "15" --alignSJDBoverhangMin "10"
--alignSplicedMateMapLmin "0" --alignSplicedMateMapLminOverLmate "0.66" --alignWindowsPerReadNmax "10000"
--alignTranscriptsPerWindowNmax "100" --alignTranscriptsPerReadNmax "10000" --alignEndsType Local --twopassMode
"Basic" --twopass1readsN "-1" --chimSegmentMin "15" --chimScoreMin "15" --chimScoreDropMax "20"
--chimScoreSeparation "10" --chimScoreJunctionNonGTAG "-1" --chimJunctionOverhangMin "15" --limitBAMsortRAM "0"
--limitOutSJoneRead "1000" --limitOutSJcollapsed "1000000" --limitSjdbInsertNsJ "1000000" && samtools view -@
${GALAXY_SLOTS:-4} -Shb Chimeric.out.sam | samtools sort -@ ${GALAXY_SLOTS:-4} - ChimericSorted
```

B

```
mkdir -p tempstargenomedir && STAR --runMode genomeGenerate --genomeDir 'tempstargenomedir' --genomeFastaFiles
'/data/dnb07/galaxy_db/files/b/4/3/dataset_b430e57b-9f9f-4ed1-8781-01d816338464.dat' --sjdbGTFfile '/data/dnb06
/galaxy_db/files/3/2/6/dataset_326f750f-0762-4f76-8e16-1a6c10980128.dat' --sjdbOverhang '149' --runThreadN
${GALAXY_SLOTS:-4} && STAR --runThreadN ${GALAXY_SLOTS:-4} --genomeLoad NoSharedMemory --genomeDir
'tempstargenomedir' --sjdbGTFfile '/data/dnb06/galaxy_db/files/3/2/6/dataset_326f750f-
0762-4f76-8e16-1a6c10980128.dat' --sjdbOverhang '149' --readFilesIn '/data/dnb07/galaxy_db/files/7/e/7
/dataset_7e7c5dc9-d1c6-486e-a204-5f637c3e0a4c.dat' '/data/dnb06/galaxy_db/files/4/f/4
/dataset_4f4f97f2-1e86-49c3-9638-67c5d567a04b.dat' --readFilesCommand zcat --outSAMtype BAM
SortedByCoordinate --quantMode GeneCounts --outSAMattributes All --outSAMstrandField None
--outFilterIntronMotifs None --outSAMunmapped None --outSAMprimaryFlag OneBestScore --outSAMmapqUnique "255"
--outFilterType Normal --outFilterMultimapScoreRange "1" --outFilterMultimapNmax "20" --outFilterMismatchNmax
"999" --outFilterMismatchNoverLmax "0.5" --outFilterMismatchNoverReadLmax "1.0" --outFilterScoreMin "1"
--outFilterScoreMinOverLread "0.66" --outFilterMatchNmin "0" --outFilterMatchNminOverLread "0.7"
--seedSearchStartLmax "50" --seedSearchStartLmaxOverLread "1.0" --seedSearchLmax "0" --seedMultimapNmax "10000"
--seedPerReadNmax "1000" --seedPerWindowNmax "50" --seedNoneLociperWindow "10" --alignIntronMin "20"
--alignIntronMax "1000000" --alignMatesGapMax "1000000" --alignSJoverhangMin "15" --alignSJDBoverhangMin "10"
--alignSplicedMateMapLmin "0" --alignSplicedMateMapLminOverLmate "0.66" --alignWindowsPerReadNmax "10000"
--alignTranscriptsPerWindowNmax "100" --alignTranscriptsPerReadNmax "10000" --alignEndsType Local --twopassMode
"Basic" --twopass1readsN "-1" --chimSegmentMin "15" --chimScoreMin "15" --chimScoreDropMax "20"
--chimScoreSeparation "10" --chimScoreJunctionNonGTAG "-1" --chimJunctionOverhangMin "15" --limitBAMsortRAM "0"
--limitOutSJoneRead "1000" --limitOutSJcollapsed "1000000" --limitSjdbInsertNsJ "1000000" && samtools view -@
${GALAXY_SLOTS:-4} -Shb Chimeric.out.sam | samtools sort -@ ${GALAXY_SLOTS:-4} - ChimericSorted
```

**Figure 5: Command line input for STAR 2.5.2b-2 used for the mapping. A)** Command line for paired end reads. **B)** Command line for single end reads.

The mapping was done for paired end reads as well as for each read individually.

### 2.2.5.3 Detection of circular RNAs with circTools.

Multiple tools are available to detect circular RNA. In this thesis, the module DCC run through circTools was used to detect circRNAs, while Novogene used CIRCexplorer2 and the circRNAs detected by Zsófia Fekete (Agricultural Biotechnology Center, Hungary) were received through her own script. To run circTools, the mapping of sequencing reads with STAR was done as described and recommended for DCC. All files were named accordingly to their type and if they originate from paired or single end read (paired end reads were only named by their replicate number (first, second, third) and single end reads

## 2 Material and methods

were additionally named mate1 or mate2). For more detailed information about the use of DCC and circTools see [www.docs.circ.tools/en/latest/](http://www.docs.circ.tools/en/latest/) [137,150].

The input for the command line to run the detect module (DCC) in circTools is:

```
circTools detect @samplesheet -mt1 @mate1 -mt2 @mate2 -B @bam_files.txt -D -an  
Brassica_napus_Ensemble42.exon.gtf -Pi -F -M -Nr 5 6 -fg -G -A  
Brassica_napus_Ensemble42.dna.fa_uncompressed.fasta
```

### 2.2.5.4 Data analysis with Access

To compare how many identical circRNAs could be found in lncRNA libraries compared to circRNA enriched libraries, Microsoft Access was used. There, two Excel tables were compared and new Excel tables were generated with circRNAs found in both tables.

### 2.2.5.5 SAXS data analysis

The SAXS data processing and analysis was done by Cy M. Jeffries (EMBL, Hamburg (D)). The software used for data analysis can be found in Table 13. With iTasser, the structure of AtGRP7 was predicted and one model was used for further refinement for a better fit of the predicted structure model towards the *ab initio* modeled AtGRP7 by DAMMIF. The refined structure was then superimposed to the DAMMIF model of AtGRP7 with SASpy.

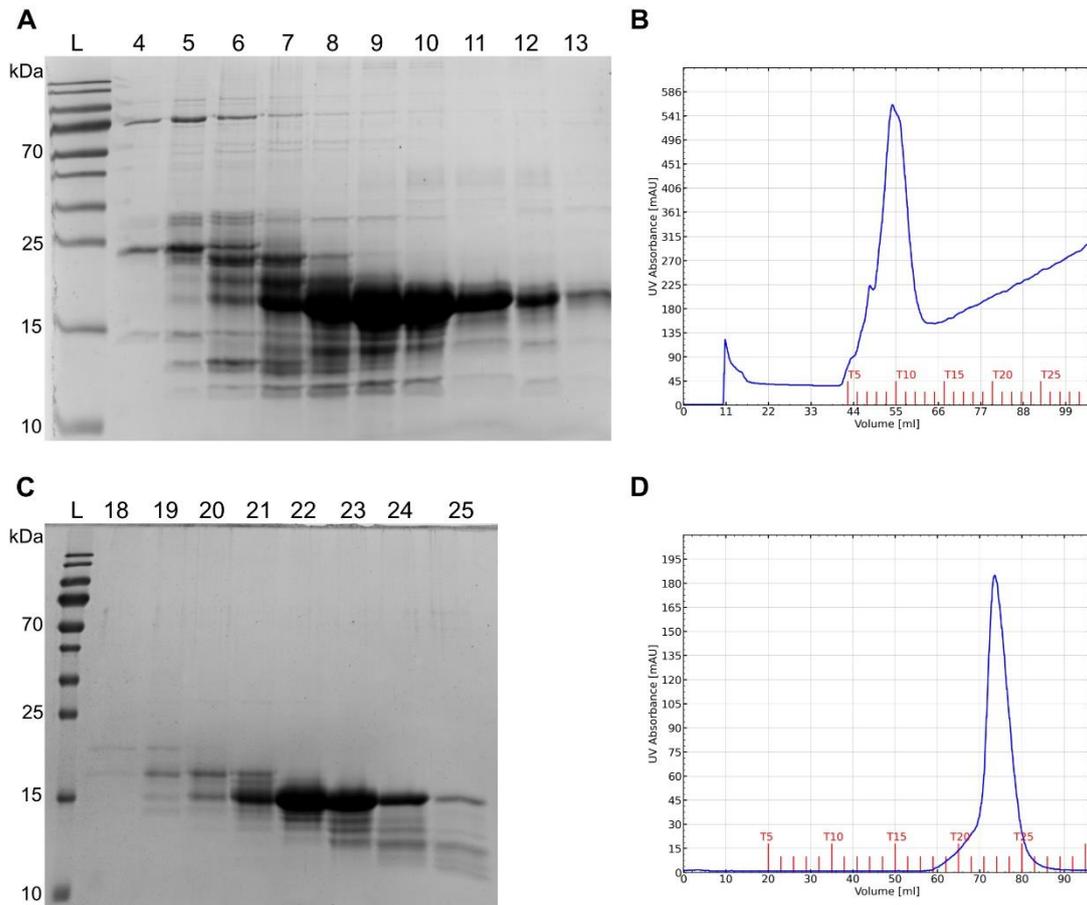
## 3 Results

In this thesis, the phloem-mobile RNA-binding protein GRP7 was investigated. Besides other RNA-binding proteins (RBPs), GRP7 was found in phloem samples of various plants [21,22,47,52]. Since the loading, transport and unloading of RNAs into the sieve elements is yet not understood, RBPs might play a crucial role. To better understand the RNA-binding behavior of GRP7, the protein was purified and analyzed. The results are presented in the following chapters.

### 3.1 Purification of glycine-rich protein 7

To analyze the phloem-mobile RNA-binding protein AtGRP7 *in vitro*, the protein needed to be purified. To accomplish this, GRP7 was first heterologously expressed in *E. coli*. The best expression was achieved by overnight-expression at 22 °C. AtGRP7 was already purified by other researchers. They used a GST-tag to tag the protein and purified it accordingly to the GST-column purification without any special adjustments of the protocol [151,152]. Since AtGRP7 was cloned into the expression vector pet28a+ in frame with a N-terminal His-tag without any large solubility tag, a purification through a Ni-NTA column with a simple buffer composed of TrisHCl pH 8.0 and 200 mM NaCl and further purification steps were planned. During the purification AtGRP7 showed visible, reversible precipitation or condensation at 4° C which was especially apparent during overnight dialysis at 4° C. To prevent these precipitates, all purification steps were carried out at RT and the purification buffers were adapted to the purification protocol for *Nicotiana tabacum* GRP1 (NtGRP1) [153]. Besides precipitation, a large amount of protein was lost during the lysis and stayed within the insoluble fraction in the pellet. Addition of 1 M of hexandiol 1.6 helped with the solubility of AtGRP7 and therefore increased the amount of purified protein (Figure 44). Furthermore, the purification was carried out as described. After purification of the protein by affinity chromatography with a Ni-NTA column as well as after the following size exclusion chromatography (SEC)-column, the purity of the protein was controlled by SDS-PAGE (Figure 6 A and C). The protein containing fractions for SDS-PAGE were identified by the UV-absorption of different fractions (Figure 6 B and D).

### 3 Results



**Figure 6: SDS-PAGE with purification steps of GRP7 and UV-absorbance in mAU of different fractions during the Nickel NTA and SEC run. A)** Äkta fractions 4-13 from Nickel NTA column on SDS-PAGE on 15 % SDS-PAGE, L: PageRuler Prest prestained protein ladder from ThermoFisher Scientific. **B)** UV-chromatogram for Ni-NTA purification run. **C)** Äkta fractions 18-25 from SEC-Column on 15 % SDS-PAGE, L: PageRuler Prest prestained protein ladder from ThermoFisher Scientific. **D)** UV-chromatogram of SEC purification run.

Since the protein started to form visible precipitation during dialysis at 4° C, the whole purification was performed at RT, even though there was always a slight degradation detected. As seen in Figure 6, below the 16.8 kDa AtGRP7 are multiple lighter bands. To make sure that these protein bands are not a contamination of other proteins, they were analyzed by MS. These bands contained degraded GRP7, each missing more of the glycine-rich part than full-length GRP7 (Figure 45). The problem of degradation could not be overcome. Hence, all following assays had to be conducted with a mix of full-length AtGRP7 and slightly degraded versions of this protein.

### 3.2 RNA-binding of GRP7

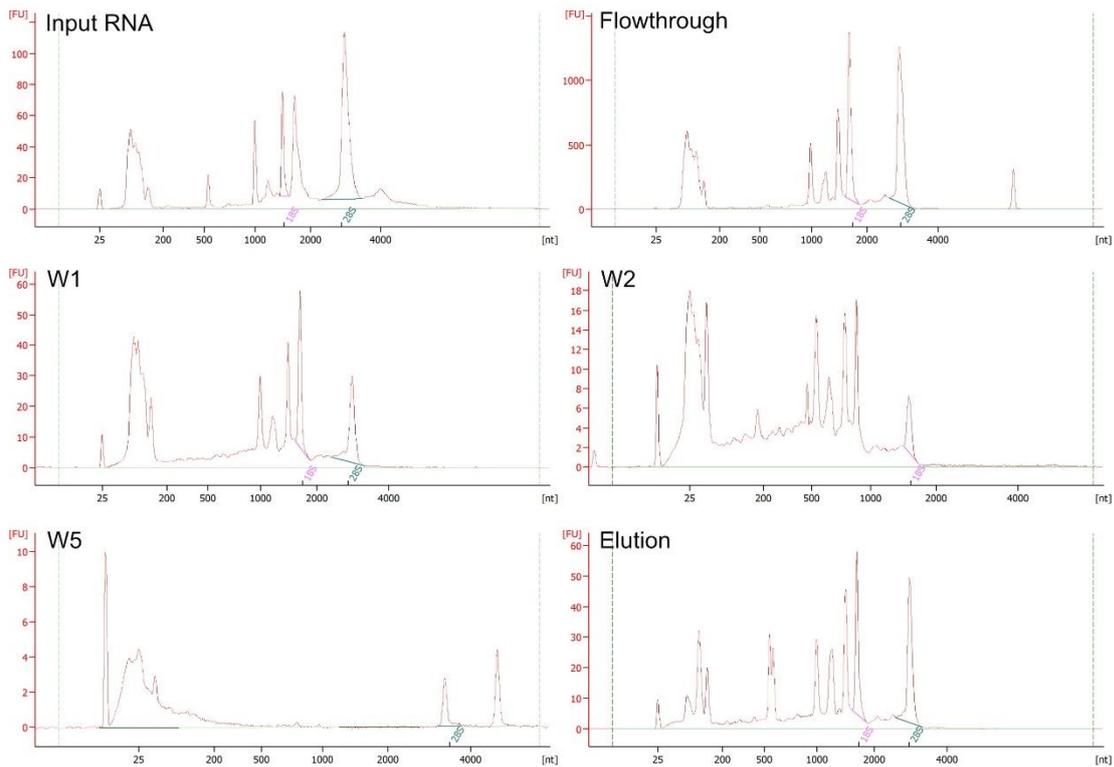
Since GRP7 contains an RRM at the N-terminal region, it is supposed to bind RNAs. This was shown by several publications [88,154]. One publication mentioned that GRP7 is mostly binding short RNAs in contrast to other papers stating the ability of GRP7 to bind larger mRNAs. Therefore, the aim was to better understand which types of the RNA are bound by GRP7.

#### 3.2.1 *B. napus* leaf RNAs bound by AtGRP7

To further investigate which length of RNAs are bound by GRP7, a pilot-experiment was conducted with AtGRP7 and *B. napus* RNA, since both was available at this timepoint. AtGRP7 was immobilized on CnBr-sepharose beads and incubated with *B. napus* leaf RNA. To prevent blocking of the RNA-binding domain of GRP7 by the beads, dextran sulfate was added during the immobilization step and removed afterwards. Furthermore, sepharose beads were prepared as a negative control without protein to test if the beads themselves could bind RNA after blocking open groups with Tris (Figure 46). No RNA was eluted from blocked beads without protein. Additionally, AtGRP7short was immobilized on sepharose beads as well and incubated with *B. napus* leaf RNA. No RNAs were bound by AtGRP7short (Figure 47).

The incubation of beads with and without AtGRP7 with *B. napus* leaf RNA was followed by the elution of bound RNA with a high salt concentration. To check the sizes of bound RNA, input RNA, flow through, washing steps and the elution were measured on a Bioanalyzer NanoChip (Figure 7).

### 3 Results



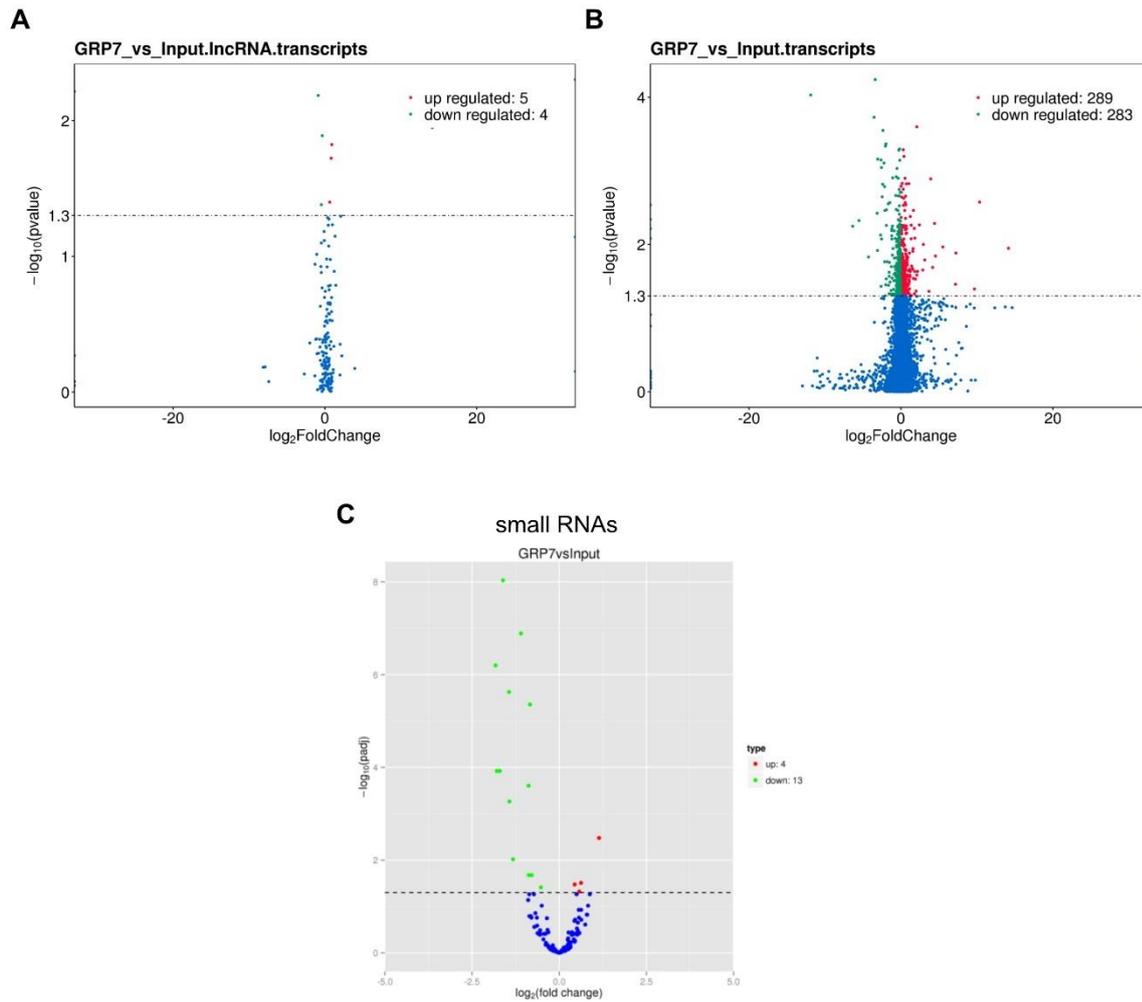
**Figure 7: Electropherograms of leaf RNA and RNA bound by immobilized AtGRP7.** The input leaf RNA was total *B. napus* leaf RNA, the flowthrough contained unbound RNAs, washing step 1 to 5 showed a decreasing amount of weakly bound RNA washed of and the elution contains all RNAs bound strong enough by AtGRP7 to withstand all washing steps. The x-axis shows the size of RNA in nucleotides (nt), the y-axis shows the fluorescence units (FU). Due to the fact that the RNA concentrations differed between samples, the FU differed a lot, therefore it was not possible to use a global y-axis over all electropherograms. The different y-axis needs to be considered when comparing the peak heights of the electropherograms.

RNA eluted from immobilized AtGRP7 ranged from a length of <100 nt up to nearly 4000 nt. In proportion to the other RNA peaks, the RNA peak at 500 nt increased in the elution compared to the input. In addition, also the small RNA peak in proportion to the other RNA peaks increased within the elution. Thus, the RNAs in the elution showed a broad range of sizes and thereby a broad range of bound RNAs by GRP7.

#### 3.2.1.1 Illumina sequencing of AtGRP7 enriched leaf RNAs

To get a better understanding which RNAs are bound by AtGRP7, total leaf RNAs and the elution of the AtGRP7-column were sequenced with Illumina sequencing. Two libraries, one for long-noncoding RNAs and one for small RNAs, were prepared and sequenced by Novogene (Cambridge, UK). The expression, or in this case enrichment, of RNAs between total leaf RNA and eluted RNAs was compared. The sequencing showed that up to 294 long RNAs were significantly enriched within the elution compared to total leaf RNA while 4 small RNAs were significantly enriched (Figure 8).

### 3 Results



**Figure 8: Volcano plots of differentially expressed (enriched) RNAs between total leaf RNA and eluted RNAs from AtGRP7 column.** The elution of AtGRP7 column (GRP7) compared to total leaf RNA (input) shows enriched RNAs. **A)** Five significantly enriched lncRNA transcripts were found in the elution. **B)** 289 significantly enriched mRNAs were found in the elution of AtGRP7 column. **C)** In total, four sRNAs show a significant enrichment in the elution while 13 sRNAs seem to be less present within the elution than within total leaf RNA. The volcano plots and analysis were done by Novogene.

In the analysis of Novogene, only 10 of the 289 enriched transcript RNAs had a normalized read count (RC) for the elution of at least 10 and a Log2FoldChange of at least 1 (Table 24).

### 3 Results

**Table 24: AtGRP7-enriched transcript RNAs from *B. napus* leaf RNA analyzed by Novogene.** Gene name, elution normalized read count (RC), Input normalized RC, Log2FoldChange and p-value are shown.

Gene name	Elution normalized RC	Input normalized RC	Log2Foldchange	p-value
BnaC01g26930D	18.375228	0	Inf	0.000773778
BnaA07g16530D	22.0209177	0	Inf	0.005497456
BnaCnng12930D	49.3951307	0	Inf	0.025677013
BnaC05g32140D	185.76289	0	Inf	0.031149701
BnaA01g34180D	24.837755	1.06429	4.54457156	0.014334895
BnaC09g08190D	15.027747	0.699036	4.426118167	0.005179095
BnaCnng12690D	18.9416287	3.52914267	2.424170726	0.005416894
BnaA09g32640D	12.0709207	4.072743	1.567463032	0.012390577
BnaC09g25620D	17.1072653	6.699113	1.352567165	0.008367994
BnaC09g16500D	27.1040147	12.5386713	1.11212208	0.031953443

In the analysis of the enriched transcripts of Novogene revealed mostly uncharacterized transcripts which had been enriched. Additionally, the read counts are low and the Log2foldChange is also quite low with a maximum of 4.5 for *BnaA01g34180D* transcript. The results haven't been further investigated, especially since the CnBr column with AtGRP7 and *B. napus* leaf RNA was in general a test run for the following CnBr column with total phloem sap RNA.

In parallel to the analysis of Novogene for enriched RNAs in the elution of the AtGRP7 column, another analysis was done by Huge Wolfeenden (John-Innes-Center, Norwich, UK) as well. He provided the following list of significantly enriched RNAs (Table 25).

### 3 Results

**Table 25: AtGRP7-enriched transcript RNAs from *B. napus* leaf RNA analyzed by Huge Wolfeenden (John-Innes-Center, Norwich, UK).**

Gene name	Log2foldChange	p-value
BNAA01G36730D	10.6942796	1.270000e-18
BNACNNG53670D	9.5829894	1.517202e-14
BNAC03G69500D	2.1782769	1.628679e-14
BNAC04G41400D	8.7047817	1.171340e-10
BNAC01G27050D	0.9451713	1.879925e-09
BNAC08G46660D	1.4601753	1.675438e-08
BNAC06G07050D	1.2643981	4.400723e-08
BNAC09G16500D	1.2226283	3.099328e-07
BNAC08G28020D	7.8378630	3.957287e-07
BNAC07G31170D	0.8633866	4.906636e-06
BNAANNG40960D	0.8707820	6.958311e-06
BNAA03G12990D	2.0551310	8.909360e-06
BNAA02G36030D	0.9068988	4.327004e-06
BNAC03G26040D	0.9664196	1.524328e-05

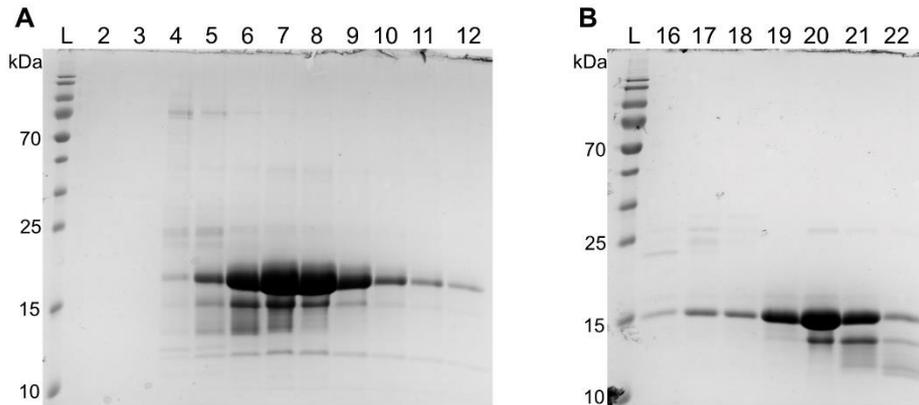
From this list, two RNAs, *BnaCNNG53670D* and *BnaC03g26040D*, were selected for *in vitro* transcription and MST-experiments.

#### 3.2.2 Phloem RNAs bound by BnGRP7

Since GRP7 is present within the phloem sap and might bind preferable some of the RNAs within the phloem sap to possibly provide their mobility, the immobilization of GRP7 towards sepharose beads was repeated and incubated with total phloem sap RNA. The phloem sap was sampled as described from *B. napus* plants.

To prevent abnormalities in RNA-binding through the use of the GRP7 protein of one family member of the Brassicaceae and the total phloem RNA of another, GRP7 protein of *B. napus* (BnGRP7) was used. The purification of BnGRP7 was carried out at RT with two purification steps, the first one done with a nickel column, the second one by SEC. Protein containing fraction were tested for protein purity on SDS-PAGE (Figure 9).

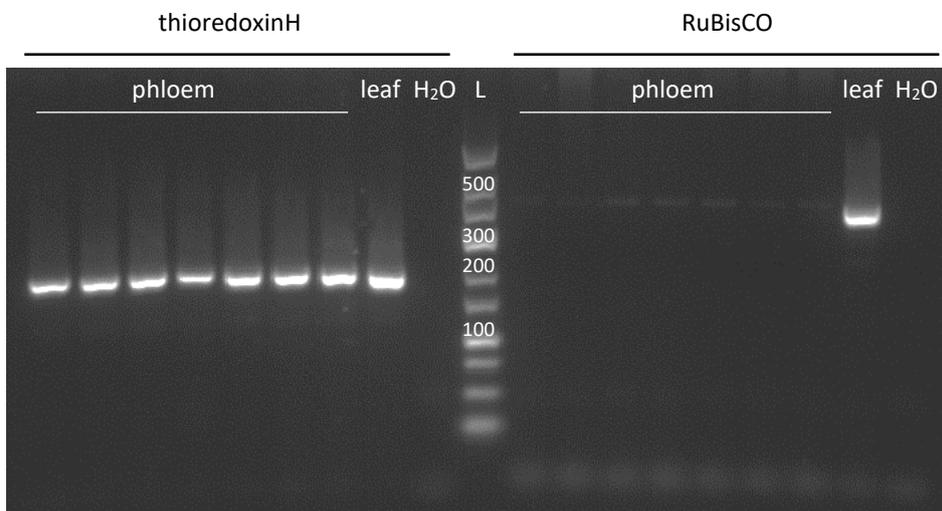
### 3 Results



**Figure 9: SDS-PAGE of BnGRP7 purification steps.** A) Nickel column fractions 2-12 containing the 16.5 kDa BnGRP7 protein. B) SEC column fractions 16-22 containing the BnGRP7 protein. A 15 % SDS-PAGE was used and run for 10 min at 150 V and at 200 V for 45 min. The protein ladder, PageRuler Prest Prestained from ThermoFisher was used, 10, 15, 25 and 70 kDa are highlighted in the ladder.

Clean, protein-containing SEC-fractions were chosen and concentrated. Purified BnGRP7 was then immobilized on sepharose beads.

For the incubation of immobilized BnGRP7 with RNA, only pure total phloem RNA was used. The purity of phloem RNA was revised by reverse transcription-PCR (RT-PCR) to detect RuBisCO subunit RNA as well as thioredoxin h RNA. As a positive control total leaf RNA was used which contains both RNAs. All phloem RNAs extracted were tested for purity by this method (Figure 10).



**Figure 10: PCR-test for phloem purity.** The purity of phloem sap RNA was tested by cDNA synthesis followed by the amplification of a 180 nt long fragment of thioredoxin h as well as a 450 nt long fragment of RuBisCo subunit. The PCR was load onto a 3 % agarose gel and run for 1 h at 120 V. The right site contained the thioredoxinH amplicons and the left site contained the RuBisCO subunit amplicons from phloem and leaf RNA. L: low range DNA ladder from ThermoScientific.

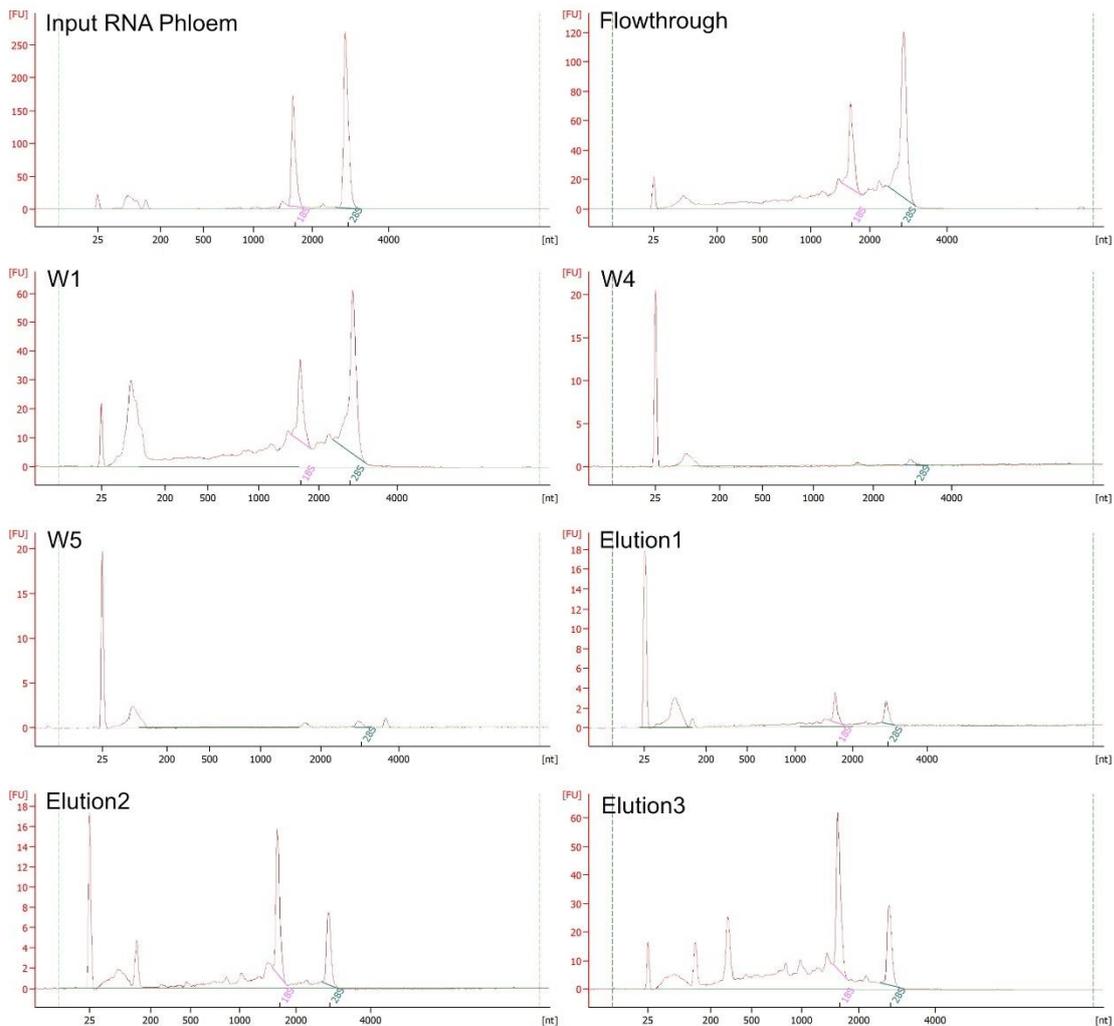
For pure phloem sap, the amount of RuBisCO amplicon is expected to be visibly noticeable less than from leaf since the phloem sap was reported to contain thioredoxin h mRNA but not mRNA of the small

### 3 Results

subunit of RuBisCO [21]. The thioredoxin h amplicon was supposed to be strongly visible in phloem as well as leaf samples [21]. The RT-PCR performed on phloem sap RNA as well as leaf RNA showed strong amplicons for thioredoxin h in both samples and strong amplicons for the small subunit of RuBisCO for leaf samples but not for phloem samples. Thus, the phloem samples are of high quality and do contain only minimal contamination of surrounding, damaged cells.

Phloem RNA analyzed as above was used to be incubated with BnGRP7 immobilized on CnBr-sepharose beads. Input RNA, flowthrough, washing steps and elution steps one to three were checked on a Bioanalyzer NanoChip to get an overview of the length of bound RNA (Figure 11). To increase the amount of RNAs bound with high affinity, three elution steps were performed with increasing amounts of salt (250 mM sodium acetate, 500 mM sodium acetate and 2 M sodium acetate). As with the leaf RNA, RNAs of different sizes were bound, smaller than 100 nt and up to over 1000 nt. The washing steps showed a decreasing amount of unbound RNA washed off while the elutions contain BnGRP7 bound RNAs. Elution three was selected and further investigated by RNA sequencing to gain information which phloem RNAs were bound with high affinity by BnGRP7 withstanding not only the five washing steps but also the first two elutions.

### 3 Results

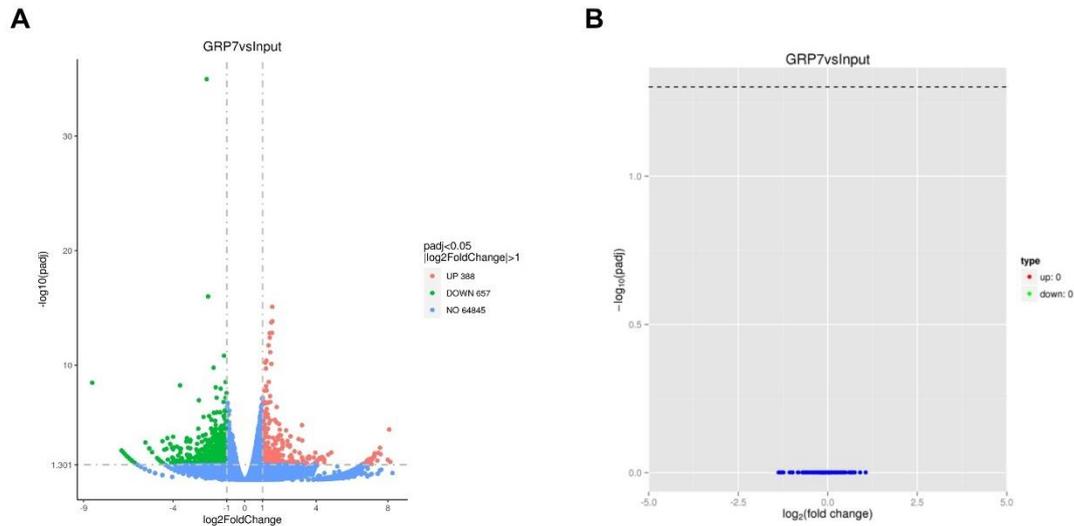


**Figure 11: Electropherograms of phloem RNA and RNA bound by immobilized AtGRP7.** The input phloem RNA is total *B. napus* phloem RNA, the flowthrough contains unbound RNAs, washing step 1 to 5 (W1, W4 and W5) show a decreasing amount of unbound RNA washed of and the elution1-3 contain all RNAs bound strong enough by BnGRP7 to withstand all washing steps. The x-axis shows the size of RNA in nucleotides (nt), the y-axis shows the fluorescence units (FU). Due to the fact that the RNA concentrations differ, the FU differ a lot, therefore it was not possible to use a global y-axis over all electropherograms. The different y-axis needs to be considered when comparing peak heights of the electropherograms.

#### 3.2.2.1 Illumina sequencing of BnGRP7 enriched phloem RNAs

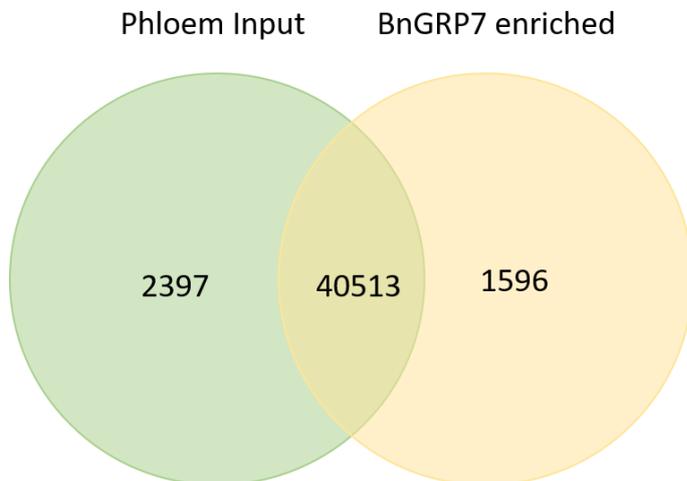
For the investigation of which particular phloem RNAs have been enriched by BnGRP7, total phloem sap RNA as well as the elution three were send for library preparation and Illumina sequencing to Novogene. A small RNA library and a polyA-library were prepared and sequenced. The expression of input phloem RNA and elution, or, in case of the BnGRP7 bound RNAs, the enrichment of RNAs was compared by Novogene. The sequencing showed that up to 388 mRNAs were significantly enriched in the elution while no enriched small RNAs were detected (Figure 12).

### 3 Results



**Figure 12: Volcano Plots of differentially expressed RNAs between total phloem sap RNA and eluted RNAs from BnGRP7 column.** A) Volcano plot of significantly enriched and decreased mRNAs. 388 mRNAs were significantly enriched by BnGRP7. B) Volcano plot of significantly enriched and decreased small RNAs. No small RNAs have been enriched by BnGRP7. Volcano plots and the analysis was done by Novogene.

To compare the overall match of mRNAs with total phloem RNA and eluted BnGRP7 RNAs, a venn diagram was prepared. It compares how many of the RNAs are found in both groups and how many RNAs were only found in one of the groups (Figure 13).



**Figure 13: Venn diagram of Phloem RNA Input and enriched BnGRP7.**

The venn diagram shows that most of the RNAs (40,513) found in phloem input were also found in the elution fraction of the BnGRP7-CnBr column. Only 2,397 RNAs were found exclusively in total phloem RNA and 1,596 only in BnGRP7 elution.

From a few RNAs most enriched in BnGRP7 elution a small overview is prepared in Table 26.

### 3 Results

**Table 26: Overview of some BnGRP7 enriched phloem sap RNAs.**

Gene name	Protein name	Elution FPKM	Input FPKM	Log2foldchange	p-value
BnaA07g32150D	defensin-like protein 3	270.7691025	114.2231344	1.275702914	2.00E-09
BnaC05g32740D	Uncharacterized protein	16.8493523	0	7.413363312	0.00069732
BnaA09g50830D	TPT-domain containing protein (triose phosphate transporter)	3.018649445	0	8.136800065	0.001086443
BnaA03g14400D	Notchless protein homologe	7.338898332	0.35220296	4.404259056	0.000364731
BnaC09g46650D	Remorin C domain containing protein	4.639808436	0	8.076737966	1.13E-07

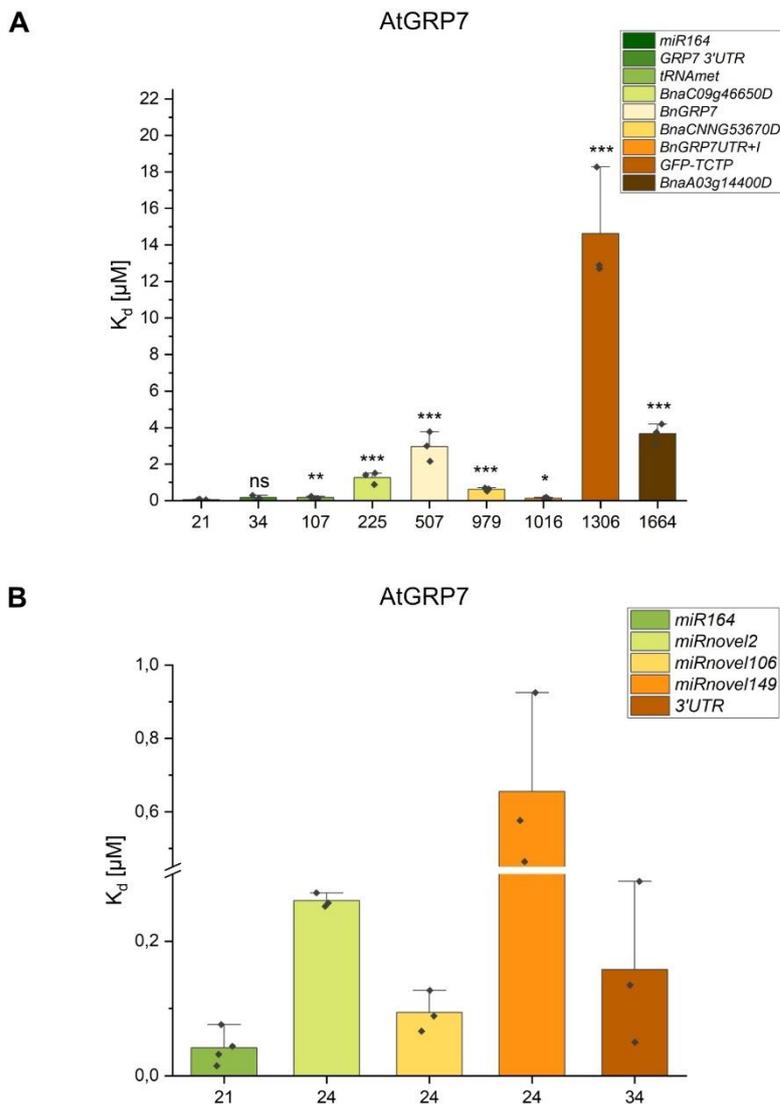
The RNAs listed in Table 26 were also later on *in vitro* transcribed. The RNAs were selected, because the FPKM was quite high, the number of reads between the three replicates of BnGRP7 elution were similar (see supplementary Table 35) and the Log2FoldChange was above one. Additionally, some of the gene functions sound interesting. For example, *BnaA07g32150D* translates a defensin-like protein 3. Since GRP7 is known to be involved in biotic stress response, it might interact with RNAs that translate for defense proteins.

#### 3.2.3 RNA-binding affinities of GRP7 towards different types of RNA

Since GRP7 binds a broad spectrum of RNAs, the question if GRP7 possesses different binding affinities towards different types of RNAs opened up. To tackle this question, all kinds of RNAs were produced by *in vitro* transcription. Some of these RNAs were enriched in the sequencing results and some RNAs showed no enrichment in sequencing. Other RNAs were tested because they are known to be phloem-mobile like *GFP-TCTP* RNA or *BnPARCL* RNA and not mobile like *GFP* RNA. The binding affinity of GRP7 towards the RNAs were measured by Microscale Thermophoresis (MST). To this end, the RNAs were partially labeled with cy5-UTP while all GRP7 proteins remained unlabeled to prevent possible inhibition of the RNA-binding capability of the protein.

A general overview of the RNA-binding affinity of AtGRP7 towards RNAs of different length is provided in Figure 14.

### 3 Results



**Figure 14: RNA-binding affinities of AtGRP7 towards RNAs of different length. A)** The binding affinity of AtGRP7 towards eight RNAs with different length ranging from 21 nt up to 1664 nt which is displayed on the x-axis. The legend displays the names of the RNAs and the Y-axis shows the  $K_d$  in  $\mu$ M, between 0 and 22.5  $\mu$ M. **B)** The binding affinity of AtGRP7 towards small RNAs are shown, with an adjusted y-axis from 0 to 1  $\mu$ M with a cut between 0.3 and 0.45  $\mu$ M. The x-axis displays the length of the RNA and the legend contains the names of the RNAs. The affinity measurements were done with Monolith NT.115 from NanoTemper. Significant differences between *miR164* and longer RNAs were tested by One-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*), and 0.001 (\*\*\*) were tested, no significant difference (ns) is claimed when none of the tested p-values showed a significant difference.

AtGRP7 showed binding of all tested RNAs. The RNA-binding affinity of AtGRP7 ranged from a low nanomolar dissociation constant ( $K_d$ ) to a high micromolar  $K_d$ . The smaller the  $K_d$  the higher the affinity. Therefore, smaller RNAs like *miR164a* showed a high binding affinity while longer RNAs like *BnaA03g14400D* showed a lower binding affinity. The binding affinity of the small RNA *bnamiR164* was significant higher compared to RNAs longer than 100 nt. In Table 27 an overview of all measured RNAs and the mean binding affinity of AtGRP7 for these RNAs is provided.

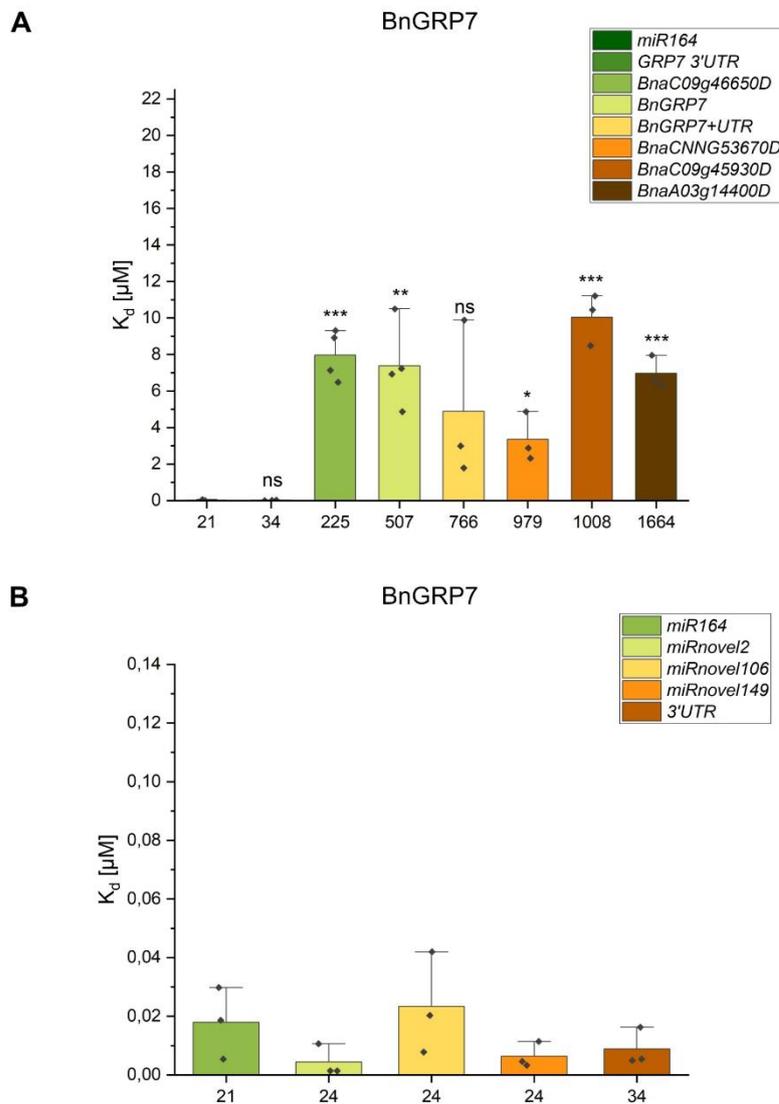
### 3 Results

**Table 27: AtGRP7 dissociation constant ( $K_d$ ) towards different RNAs.** RNA, mean  $K_d$ ,  $K_d$  standard deviation (SD) and the number of replications (N) are displayed. Each replication was done with an independent prepared dilution row of AtGRP7 and addition of the RNA. RNAs are sorted from highest to lowest binding affinity.

RNA	Protein name	Kd mean [ $\mu$ M]	Kd SD [ $\mu$ M]	N
<b>BnPARCL RNA m6A</b>	BnPARCL	0.02373	0.00536	3
<i>bna-miR164</i>	/	0.04173	0.02592	4
<b>BnPARCL RNA GA-motif</b>	BnPARCL	0.06767	0.03465	3
<i>bna-miRnovel106</i>	/	0.0941	0.03063	3
<b>BnPARCL RNA</b>	BnPARCL	0.0997	0.07866	4
<b>BnGRP7+UTR</b>	BnGRP7+UTR	0.12307	0.05458	3
<b>BnGRP7+UTR+Intron</b>	BnGRP7+UTR+Intron	0.1275	0.04898	3
<i>bna-miR164m5C</i>	/	0.14233	0.01704	3
<i>bna-miR164m6A</i>	/	0.14273	0.0656	3
<b>GRP7 3'UTR</b>	GRP7	0.158	0.12115	3
<i>tRNAm<sup>et</sup></i>	/	0.166	0.0592	4
<b>BnPARCL RNA m5C</b>	BnPARCL	0.2079	0.04629	3
<b>BnaCNNG53670D m6A</b>	Putative thioredoxin H10	0.24667	0.08279	3
<b>BnaC03g26040D m5C</b>	Calvin cycle protein CP12-1	0.25567	0.0229	3
<b>AtGRP7</b>	AtGRP7	0.25895	0.14	2
<i>bna-miRnovel2</i>	/	0.26033	0.01041	3
<b>21R' RNA</b>	/	0.26713	0.31348	3
<b>AtGRP7+UTR+Intron</b>	AtGRP7+UTR+Intron	0.36433	0.08551	3
<b>dsBnPARCL RNA</b>	BnPARCL	0.41867	0.26263	3
<b>BnaC03g26040D m6A</b>	Calvin cycle protein CP12-1	0.429	0.09379	3
<i>dsmiR164</i>	/	0.55325	0.27056	4
<b>BnaCNNG53670D</b>	Putative thioredoxin H10	0.62675	0.07531	4
<i>bna-miRnovel149</i>	/	0.655	0.24044	3
<b>BnaC03g26040D</b>	Calvin cycle protein CP12-1	1.20233	0.66485	3
<b>BnaC09g46650D</b>	Remorin C domain containing protein	1.26	0.33287	3
<b>BnaCNNG53670D m5C</b>	Putative thioredoxin H10	1.62667	0.21572	3
<b>BnaCNNG53670D w/o GA</b>	Putative thioredoxin H10	1.91333	0.16442	3
<b>BnaA07g32150D</b>	defensin-like protein 3	2.72	1.25575	3
<b>BnGRP7</b>	BnGRP7	2.97	0.81019	3
<b>BnaA03g14400D</b>	Notchless protein homologue	3.67	0.5583	3
<b>GFP-RNA</b>	GFP	9.16667	3.49401	3
<b>GFP-TCTP-RNA+GRP7 3'UTR</b>	TCTP and GFP	11.54	4.63862	2
<b>GFP-TCTP-RNA</b>	TCTP and GFP	14.62333	3.15939	3
<b>GFP-RNA+3'UTR</b>	GFP	25.65	5.57818	3
<b>GFP-TCTP-RNA m5C</b>	TCTP and GFP	35.50667	28.37558	3
<b>GFP-RNA m5C</b>	GFP	42.34667	17.65422	3
<b>BnaCNNG53670D 120nt UTR</b>	Putative thioredoxin H10	65.88	26.47408	2

### 3 Results

Since BnGRP7 was also immobilized on the CnBr-beads and the bound and eluted phloem-RNAs were sequenced, the binding affinity of BnGRP7 towards different types of RNAs was analyzed by MST as well. A small overview of the binding affinity of BnGRP7 towards different RNA length is given in Figure 15.



**Figure 15: RNA-binding affinity of BnGRP7 towards RNAs with different length.** **A)** The binding affinity of BnGRP7 towards seven different RNAs with different length ranging from 21 nt up to 1664 nt which is displayed on the x-axis. The legend displays the names of the RNAs and the Y-axis shows the  $K_d$  in  $\mu$ M between 0 and 22.5  $\mu$ M **B)** The binding affinity of AtGRP7 towards small RNAs are shown, with an adjusted y-axis from 0 to 0.2  $\mu$ M. The x-axis displays the length of the RNA and the legend contains the names of the RNAs. The affinity measurements were done with a Monolith NT.115 from NanoTemper. Significantly differences between *miR164* and longer RNAs were tested by One-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*), and 0.001 (\*\*\*) were tested, no significant difference (ns) is claimed when none of the tested p-values showed a significant difference.

As already observed for AtGRP7, BnGRP7 showed lower  $K_d$ s for smaller RNAs than for long RNAs, and therefore has a higher binding affinity for smaller RNAs than for longer RNAs. A total overview of the binding affinity of BnGRP7 towards all kinds of RNAs can be found in Table 28.

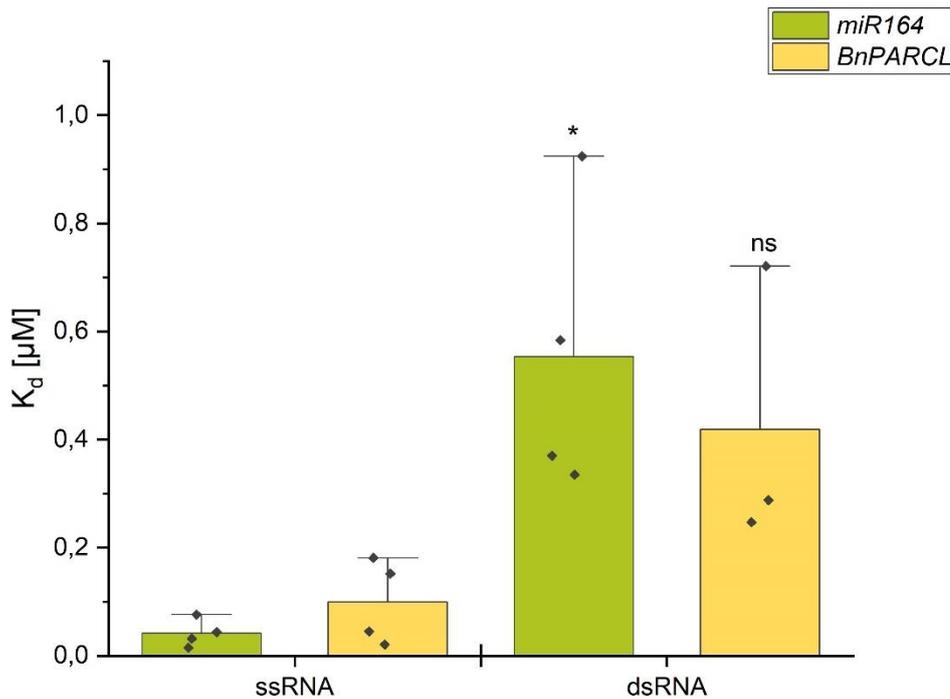
### 3 Results

**Table 28: BnGRP7 dissociation constant ( $K_d$ ) towards different RNAs.** RNA, mean  $K_d$ ,  $K_d$  standard deviation (SD) and the number of replications (N) are displayed. Each replication was done with an independent prepared dilution row of AtGRP7 and addition of the RNA. RNAs are sorted from highest to lowest binding affinity.

RNA	Protein name	Kd mean [ $\mu$ M]	Kd SD [ $\mu$ M]	N
<i>bna-miRnovel149</i>	/	0.00397	0.00331	3
<i>bna-miRnovel2</i>	/	0.00445	0.00533	3
<i>GRP7 3'UTR</i>	AtGRP7	0.01679	0.01274	3
<i>miR164</i>	/	0.021	0.029	3
<i>bna-miRnovel106</i>	/	0.04173	0.03546	3
<i>BnaA09g50830D</i>	TPT-domain containing protein (triose phosphate transporter)	2.75	1.25905	3
<i>BnaA07g32150D</i>	defensin-like protein 3	2.78333	1.28461	3
<i>BnaCNNG53670D</i>	Putative thioredoxin H10	3.35667	1.34016	3
<i>AtGRP7</i>	AtGRP7	3.72	1.63	3
<i>BnPARCL mRNA</i>	BnPARCL	4.07067	0.80806	3
<i>BnGRP7+UTR</i>	BnGRP7	4.89	4.37	3
<i>BnaC03g26040D</i>	Calvin cycle protein CP12-1	6.07667	2.47399	3
<i>BnaC05g32740D</i>	Uncharacterized protein	6.69333	2.00271	3
<i>BnaA03g14400D</i>	Notchless protein homologue	6.97	0.8731	3
<i>BnGRP7</i>	BnGRP7	7.3825	2.32805	4
<i>BnPARCL mRNA m6A</i>	BnPARCL	7.93	2.26	3
<i>BnaC09g46650D</i>	Remorin C domain containing protein	7.9625	1.3671	4
<i>BnaCNNG53670D m6A</i>	Putative thioredoxin H10	9.95333	7.17614	3
<i>BnaC09g45930D</i>	E3 ubiquitin-protein ligase MBR2-like	10.04667	1.40898	3
<i>BnGRP7+UTR+Intron</i>	BnGRP7	10.11667	3.43378	3
<i>BnaC03g26040D m5C</i>	Calvin cycle protein CP12-1	10.38	2.94318	3
<i>BnaCNNG53670D m5C</i>	Putative thioredoxin H10	12.35	1.58528	3
<i>BnaC03g26040D m6A</i>	Calvin cycle protein CP12-1	13.87767	5.4834	3
<i>BnPARCL mRNA m5C</i>	BnPARCL	15.7	1.47309	3
<i>bna-miR164 m5C</i>	/	15.85	6.21	3
<i>AtGRP7+UTR+Intron</i>	AtGRP7	16.23	7.66	3
<i>bna-miR164 m6A</i>	/	16.85	6.50	3

RNAs are not only single stranded, but can also occur double-stranded and RNAs with secondary structures can contain large double-stranded areas as well. The RRM is known to be an RNA-binding motif for single-stranded RNA or DNA, but due to its long RGG-domain, GRP7 might be capable to bind dsRNA as well. Therefore, the binding affinity of AtGRP7 towards single and double stranded RNA was tested. One short (*miR164*) and one longer RNA (*PARCL* mRNA) were used as ssRNA and dsRNA (Figure 16).

### 3 Results

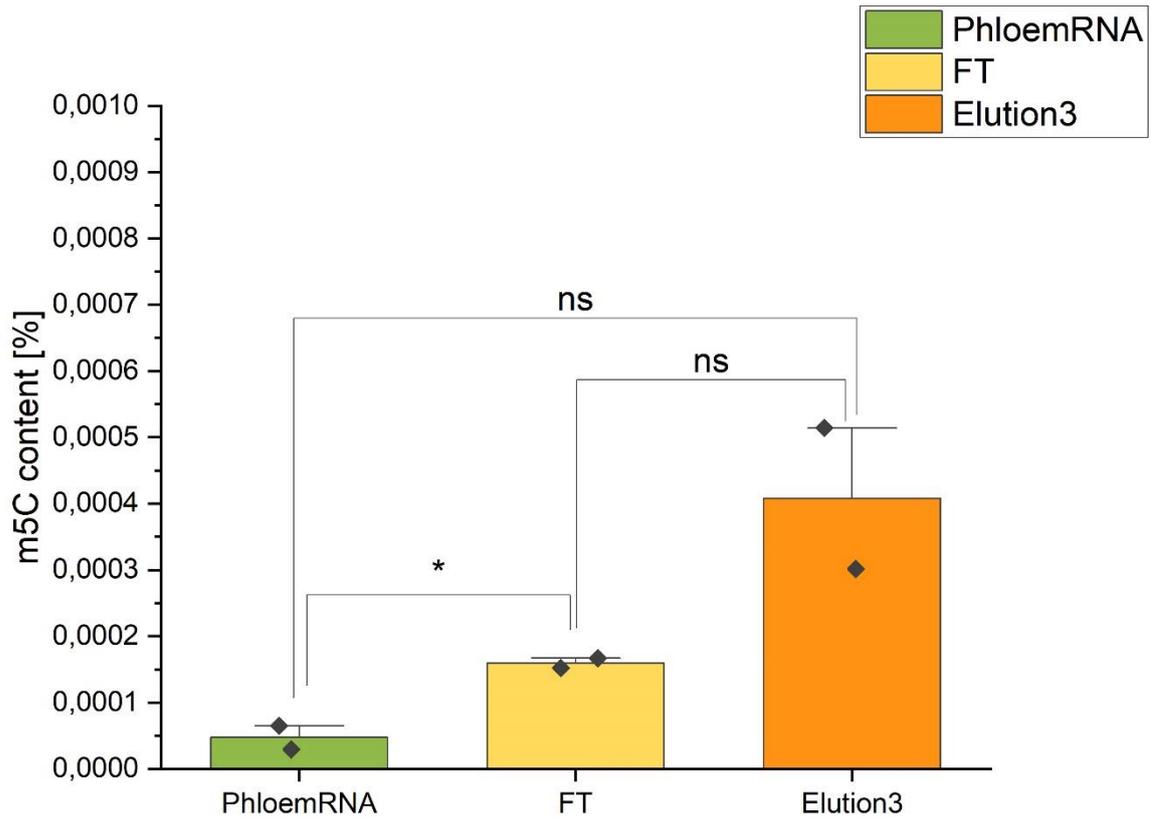


**Figure 16: Affinity of AtGRP7 for single- and double-stranded RNAs.** On the y-axis, the  $K_d$  is displayed from 0 to 1.1  $\mu\text{M}$ . The x-axis harbors the information if the RNA is single- (ss) or double-stranded (ds). *miR164* is displayed in green, *BnPARCL* mRNA in yellow. If the mean  $K_d$  of ss or dsRNA was significantly different was tested by One-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*), and 0.001 (\*\*\*) were tested, no significant difference (ns) is claimed when none of the tested p-values showed a significant difference.

The comparison between ssRNA and dsRNA-binding of AtGRP7 showed only a slight statistical significance in a lower binding affinity towards *dsmiR164* but there was no significant lower binding affinity towards *dsBnPARCL* mRNA. Both dsRNA measurements show large error bars, suggesting that the RNA-binding of AtGRP7 towards dsRNA might differ from that towards ssRNA.

#### 3.2.4 RNA methylation of total phloem RNAs and CnBr-column enriched RNAs

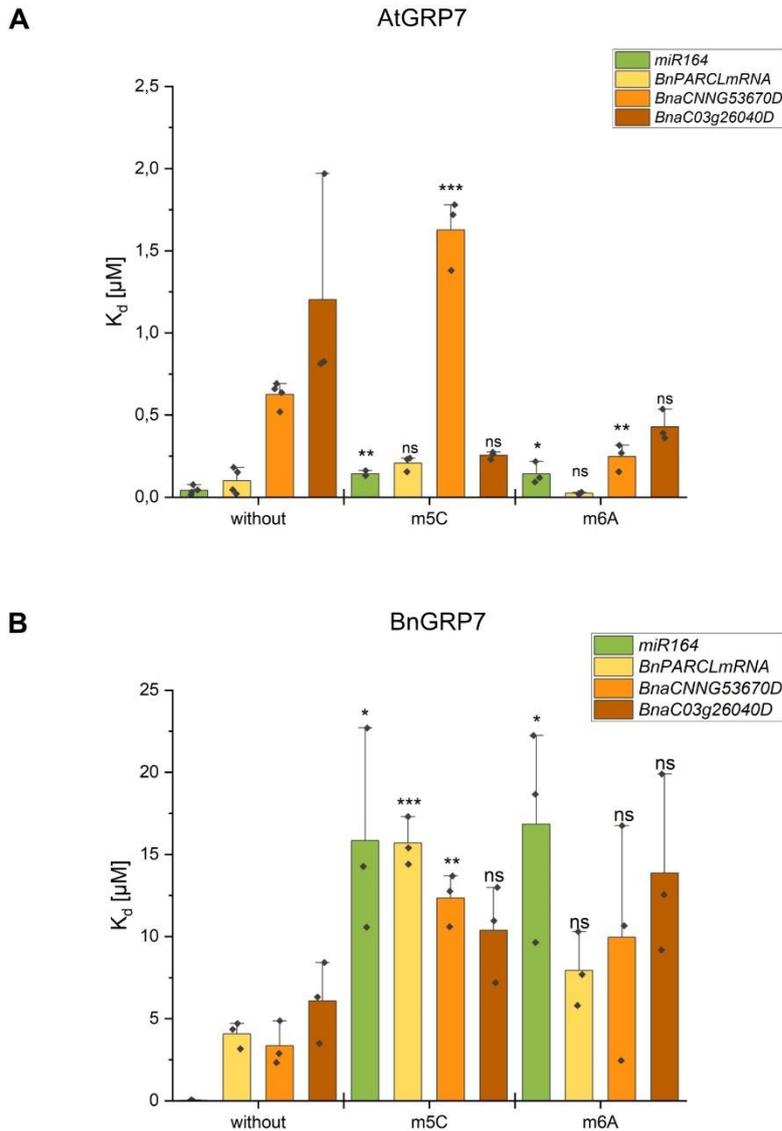
The cytosine methylation ( $\text{m}^5\text{C}$ ) of RNAs might be one of the elements to distinguish phloem-mobile RNAs from non-mobile RNAs, since the number of  $\text{m}^5\text{C}$  methylated graft-mobile RNAs is high [44]. Therefore, it would be interesting to investigate if GRP7 has a selectivity for methylated RNAs. For a first overview, total phloem RNA and the eluted RNAs from the BnGRP7 CnBr-column as well as the flowthrough of the column were analyzed for their  $\text{m}^5\text{C}$  content by ELISA (Figure 17).



**Figure 17:** The content of cytosine methylation in total phloem RNA, flowthrough and elution three of the CnBr-Immobilized BnGRP7. The m<sup>5</sup>C-content is displayed in % on the y-axis, the three different samples are shown on the x-axis. Statistically significant difference was calculated with a One-way-ANOVA. The test was done for n=2 and a p-value of 0.05, \*: significant difference, ns: no significant difference.

The investigation of the m<sup>5</sup>C-content of total phloem RNA, the flowthrough and the elution of the CnBr-column with immobilized BnGRP7 showed no significant difference in the percentage of cytosine methylated RNAs. Even though no significant difference between m<sup>5</sup>C methylated RNAs between phloem RNA and elution three was found, the percentage of methylated RNA tends to be higher in the elution. The lack of a significant difference is due to the low sample size of only two replicates. For this reason, the binding affinity of AtGRP7 as well as BnGRP7 towards methylated RNAs was tested and compared with the binding affinity towards non-methylated RNAs (Figure 18). Not only m<sup>5</sup>C is a common RNA methylation, also the methylation of adenosine is common. Hence, also m<sup>6</sup>A methylated RNAs were investigated.

### 3 Results



**Figure 18: RNA-binding affinity of AtGRP7 and BnGRP7 for RNAs with and without methylation. A)** Binding affinity of AtGRP7 towards four different non-methylated, m<sup>5</sup>C or m<sup>6</sup>A methylated RNAs. On the x-axis, the methylation status of the RNA is shown, the different RNAs are defined in the legend, *miR164* (green), *BnPARCL* mRNA (yellow), *BnaCNNG53670D* mRNA (orange) and *BnaC03g26040D* (brown). The y-axis shows the dissociation constant ( $K_d$ ) in  $\mu\text{M}$  in a range of 0-2.5  $\mu\text{M}$ . **B)** Binding affinity of BnGRP7 towards three different non-methylated, m<sup>5</sup>C or m<sup>6</sup>A methylated RNAs. On the x-axis, the methylation status of the RNA is shown, the different RNAs are defined in the legend, *miR164* (green), *BnPARCL* mRNA (yellow), *BnaCNNG53670D* mRNA (orange) and *BnaC03g26040D* (brown). The y-axis shows the dissociation constant ( $K_d$ ) in  $\mu\text{M}$  in a range of 0-22.5  $\mu\text{M}$ . Significantly different  $K_d$ 's of the two proteins towards non-methylated and methylated RNAs were detected by one-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*) and 0.001 (\*\*\*) were tested, no significant difference (ns) was claimed when none of the tested p-values showed a significant difference. Methylated RNAs were only compared to non-methylated RNAs.

The binding of AtGRP7 towards methylated RNAs showed some significant differences in comparison to non-methylated RNA. Methylated RNAs were bound with a lower binding affinity than non-methylated RNAs. These significant differences in binding were observed in m<sup>5</sup>C and m<sup>6</sup>A methylated *BnaCNNG53670D* mRNA, m<sup>5</sup>C and m<sup>6</sup>A methylated *miR164*. No significant difference was observed for

### 3 Results

the other methylated RNAs. For BnGRP7, m<sup>5</sup>C methylated *BnPARCL* mRNA as well as m<sup>5</sup>C *BnaCNNG53670D* mRNA and m<sup>5</sup>C *miR164* showed a significant higher K<sub>d</sub> and therefore a lower binding affinity compared to the non-methylated RNA. For m<sup>6</sup>A methylation, only *miR164* showed a significantly lower binding affinity. Nevertheless, the binding affinity of the other m<sup>6</sup>A methylated RNAs showed a trend of having a lower binding affinity than not methylated RNAs.

This result is consistent with the m<sup>5</sup>C content of enriched CnBr-BnGRP7-column RNAs, since the m<sup>5</sup>C methylation did not lead in any case to a higher binding affinity for the methylated RNAs. Also, m<sup>6</sup>A methylation did not increase RNA-binding affinity.

#### 3.3 The glycine-rich region of GRP7

The glycine-rich region (RGG-domain) of GRP7 makes up half of the protein. Therefore, it could be expected that a huge part of the protein's function is derived from this flexible region. This was already shown by multiple publications as mentioned in the introduction. To achieve a deeper understanding of the RGG-domain and which part conveys which protein function, the RGG-domain of GRP7 was further investigated.

##### 3.3.1 Comparison of AtGRP7 and BnGRP7

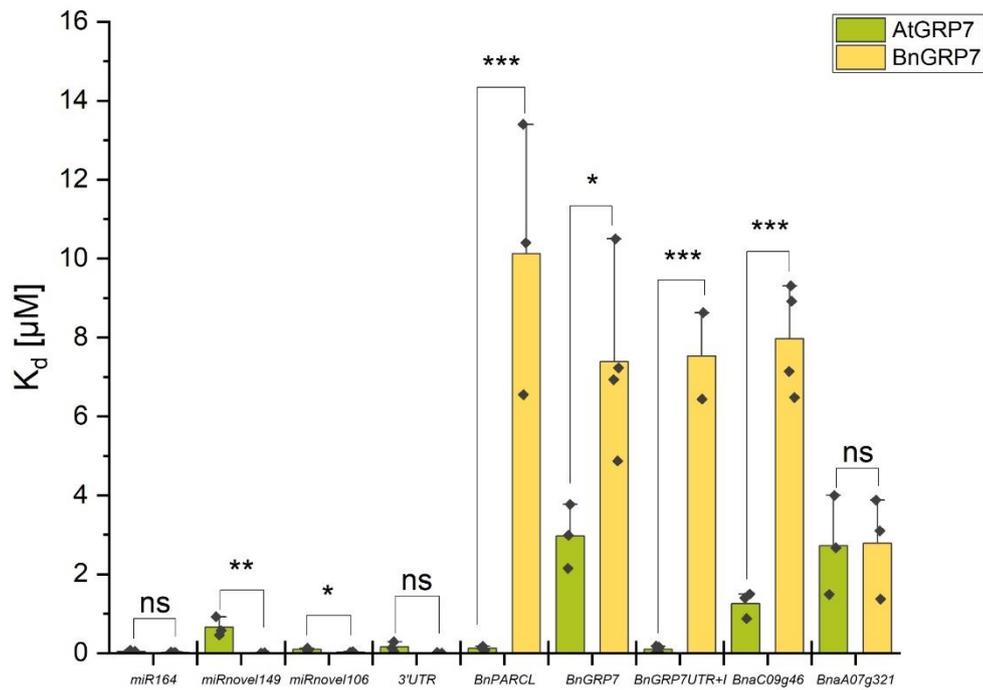
The size of AtGRP7 and BnGRP7 is different. BnGRP7 is only 168 amino acids long while AtGRP7 contains 176 amino acids. To see where the differences in the amino acid sequences of AtGRP7 and BnGRP7 occur, a sequence alignment between the proteins was done.

In the alignment between AtGRP7 and BnGRP7, a part of the AtGRP7 RGG-domain is missing in BnGRP7 (Figure 19).

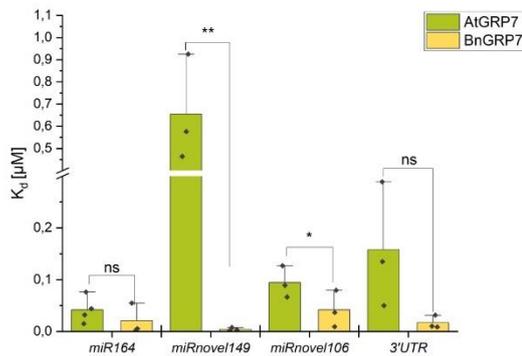


### 3 Results

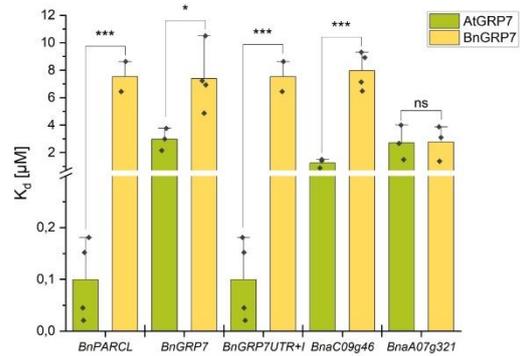
**A**



**B**



**C**

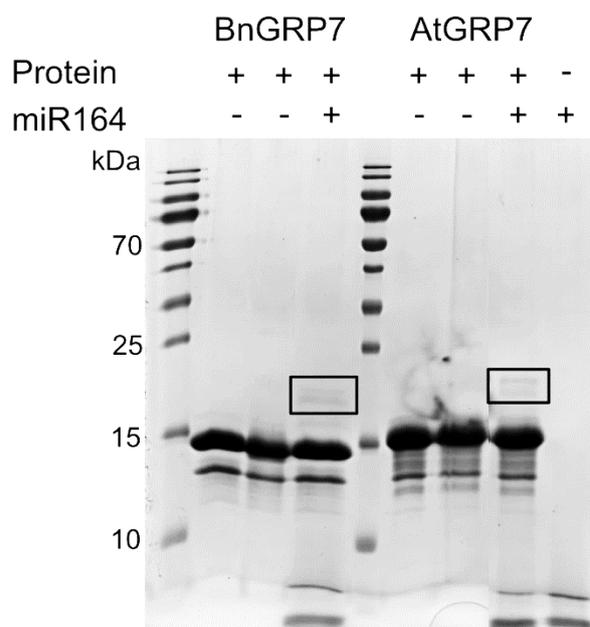


**Figure 20: Binding affinity of AtGRP7 and BnGRP7 towards different RNAs. A)** Dissociation constant in  $\mu$ M as binding affinity of AtGRP7 and BnGRP7 towards small RNAs and long RNAs. **B)** Binding affinities of AtGRP7 and BnGRP7 against small RNAs with a cut in the y-axis between 0.3 and 0.4  $\mu$ M for a better resolution of the low dissociation constants. **C)** Binding affinities of AtGRP7 and BnGRP7 against long RNAs with a cut in the y-axis between 0.3 and 0.7 for a better resolution of the low dissociation constants. The y-axis displays the binding affinity as dissociation constant ( $K_d$ ) in  $\mu$ M. On the x-axis, the RNAs tested are named and the proteins which were used are displayed in two different colors, AtGRP7 in green and BnGRP7 in yellow. Significantly different  $K_d$ s of the two proteins towards each RNAs were detected by one-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*) and 0.001 (\*\*\*) were tested, no significant difference (ns) is claimed when none of the tested p-values showed a significant difference.

### 3 Results

#### 3.3.2 Interaction of the RGG-domain with RNA during RNA-binding

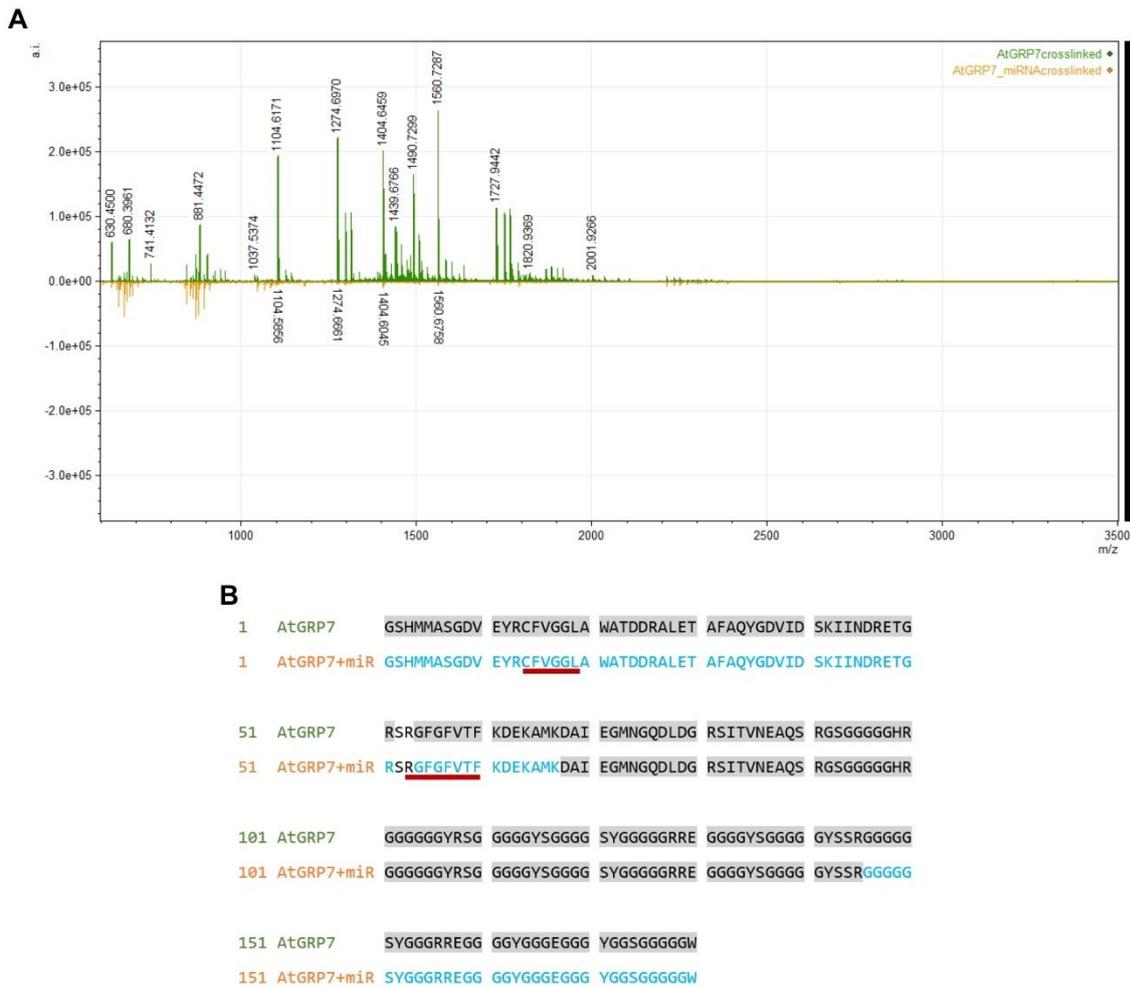
To investigate the interaction of the RGG-domain with RNA during RNA-binding, AtGRP7 and BnGRP7 were crosslinked with *miRNA164a*. For a higher crosslinking success, the ATP in *miRNA164* was subsidized with 8-azido-ATP. After UV-crosslinking, the crosslinked samples were checked on an SDS-PAGE (Figure 21).



**Figure 21: SDS-PAGE of crosslinked BnGRP7 and AtGRP7 with *miRNA164*.** 1) 20  $\mu$ M BnGRP7, 2) 20  $\mu$ M BnGRP7 UV-crosslinked, 3) 20  $\mu$ M BnGRP7+50  $\mu$ M *miR164* UV-crosslinked, 4) 20  $\mu$ M AtGRP7, 5) 20  $\mu$ M AtGRP7 UV-crosslinked, 6) 20  $\mu$ M AtGRP7+50  $\mu$ M *miR164* UV-crosslinked, 7) 20  $\mu$ M AtGRP7+50  $\mu$ M *miR164* UV-crosslinked, 8) 50  $\mu$ M *miR164* UV-crosslinked. 15 % SDS-PAGE, 1.5  $\mu$ l ladder Prestained PageRuler Prest from ThermoFisher Scientific was used, 10, 15, 25 and 70 kDa are highlighted. Two red rectangles highlight the bands of crosslinked protein+*miR164*.

Crosslinked protein+*miR164* showed a slight band between 15 and 25 kDa (lane 3 and 7), which was expected to be the protein crosslinked with *miR164*. The crosslinked protein only did not show an additional band (lane 2 and 5) just like the protein without treatment (lane 1 and 4). For the investigation which parts of AtGRP7 and BnGRP7 are interacting with *miR164*, the highlighted bands were cut, treated with trypsin and analyzed by MALDI-MS. Additionally, the BnGRP7 band of lane one and two as well as the band from AtGRP7 of lane four and five was cut, treated with trypsin and analyzed by MALDI-MS as control (Figure 22).

### 3 Results

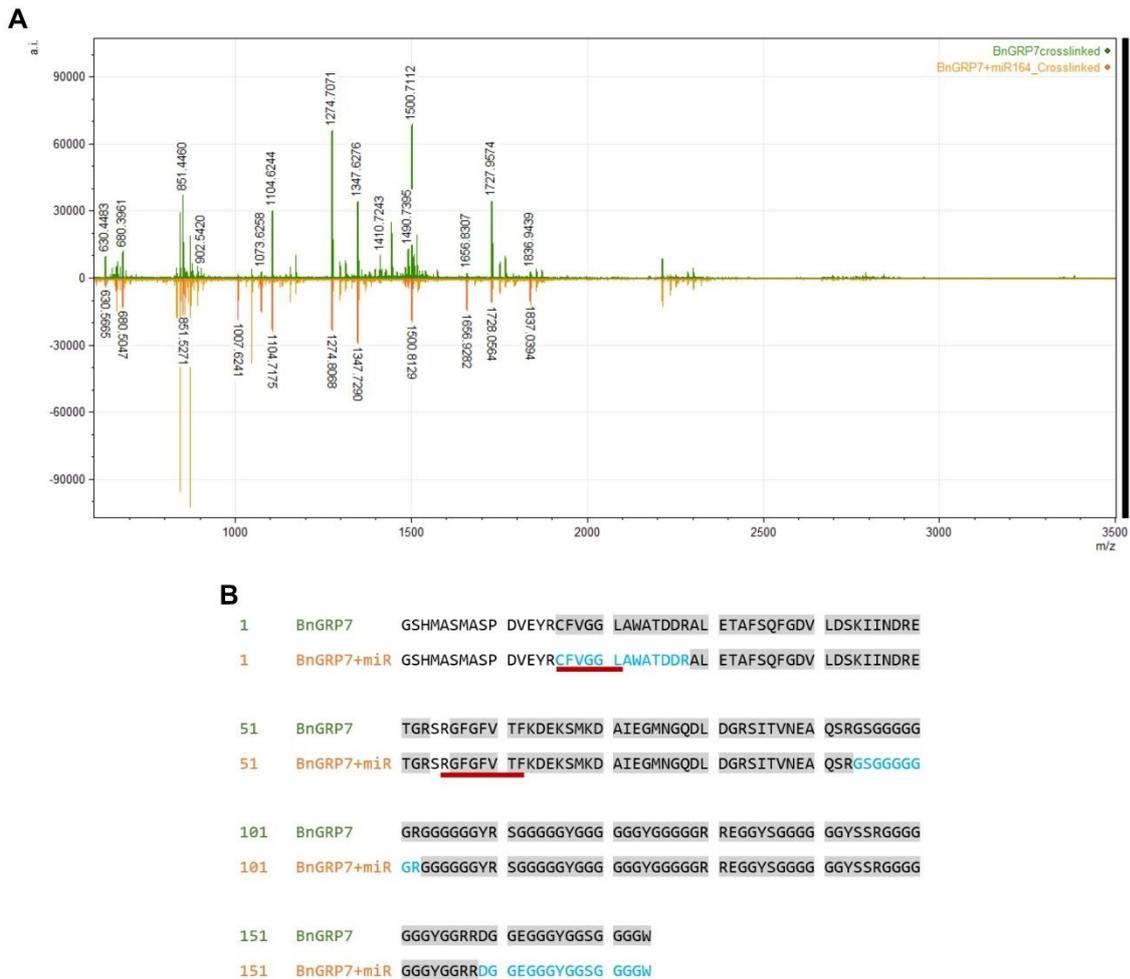


**Figure 22: UV-crosslinking of AtGRP7 with and without *miR164*.** **A)** MALDI-TOF-MS-spectrum of AtGRP7 crosslinked (green) and AtGRP7+*miR164* crosslinked (orange), the x-axis displays the mass to charge ratio of the peptides (m/z) and the y-axis displays the intensity. All peaks matching the peptides expected after a trypsin digestion are selected when they reached at least an intensity four times higher than the background. **B)** Sequence alignment of AtGRP7, highlighting in grey all peptides found in the sample of AtGRP7 crosslinked (green). The peptides for the sequence highlighted in blue in the sequence of AtGRP7 crosslinked with *miR164* (orange) was not found in this sample. RNP1 and RNP2 are underlined in red.

The MS-sample of crosslinked AtGRP7 showed all peptides of AtGRP7 except of the SR-peptide, which was typically not found. In contrast, AtGRP7 crosslinked with *miR164* was missing ten peptides, only four peptides were found for this sample. Especially peptides within the RRM and the end of the RGG-domain were missing (peptide [146-156] and [157-180]), suggesting *miRNA164* interacted with these peptides.

Unlike AtGRP7, crosslinked BnGRP7 harbored nearly all peptides. BnGRP7 crosslinked with *miR164* showed three missing peptides compared to only BnGRP7 crosslinked (Figure 23). One of the missing peptides [158-174] was the last peptide of RGG-domain, indicating an interaction of the RGG-domain with *miR164*. The peptide [16-28] missing was located in the RRM at RNP2.

### 3 Results



**Figure 23: UV-crosslinking of BnGRP7 with and without *miR164*.** **A)** MALDI-TOF-MS-spectrum of BnGRP7 crosslinked (green) and BnGRP7+*miR164* crosslinked (orange), the x-axis displays the mass to charge ratio of the peptides (m/z) and the y-axis displays the intensity. All peaks matching the peptides expected after a trypsin digestion are selected when they reached at least an intensity four times higher than the background. **B)** Sequence alignment of BnGRP7, highlighting in grey all peptides found in the sample of BnGRP7 crosslinked (green). The peptides for the sequence highlighted in blue in the sequence of BnGRP7 crosslinked with *miR164* (orange) was not found in this sample. RNP1 and RNP2 are underlined in red.

To conclude, both AtGRP7 as well as BnGRP7 were missing peptides in the last part of the RGG-domain, indicating that these peptides were probably interacting with *miR164*. Additionally, peptides containing RNP1 and RNP2 were missing in crosslinked AtGRP7 with *miR164*, while only a peptide containing RNP2 was missing in BnGRP7.

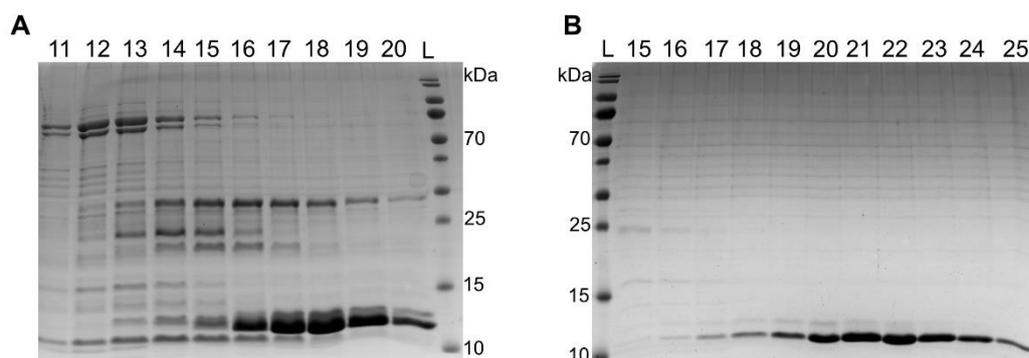
#### 3.3.3 The glycine-rich region of AtGRP7

To get a better understanding how the glycine-rich region of GRP7 is involved in RNA-binding and other functions of GRP7, two different mutants of GRP7 were created. The first GRP7 mutant lacked its glycine-rich tail completely and consisted of the N-terminal RRM only, this mutant was called GRP7short. The other GRP7 mutant contained amino acid changes within its glycine-rich region (RGG-

### 3 Results

domain). All of the eight tyrosine (Y) within the RGG-domain were exchanged against glutamic acid (E). Since FES-kinase was shown to phosphorylate these tyrosines [90], the negative charge of this phosphate was modulated by adding negative charges in form von E instead of the neutral Y. This mutant was called GRP7mut.

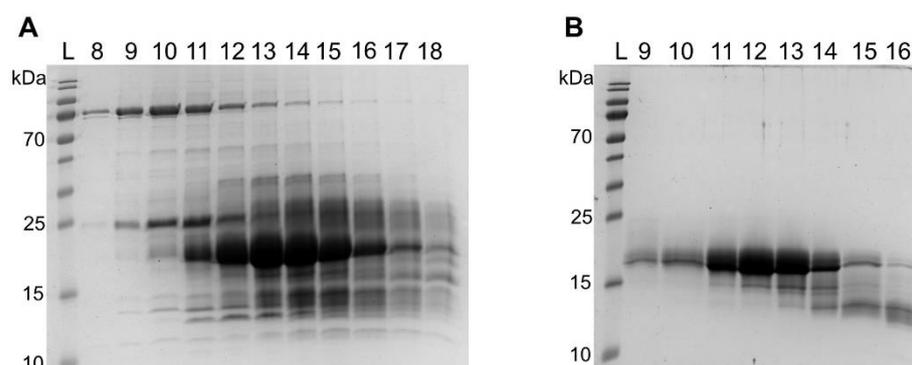
Both of the GRP7 mutants were expressed in *E. coli* and purified as described. The purification of GRP7short was carried out at 4° C with a nickel column and a SEC column, each of the purification steps checked on an SDS-PAGE (Figure 24).



**Figure 24: SDS-PAGE with AtGRP7short purification steps.** A) Nickel column fractions 11-20 containing the 10 kDa AtGRP7short protein. B) SEC column fractions 15-25 containing the AtGRP7short protein. A 15 % SDS-PAGE was used and run for 10 min at 150 V and at 200 V for 45 min. The protein ladder, PageRuler Prest Prestained from ThermoFisher was used, 10, 15, 25 and 70 kDa are highlighted in the ladder.

After SEC, only clean protein fractions were chosen and concentrated.

AtGRP7mut was purified like AtGRP7 at RT with a nickel column and a SEC column, each step observed on an SDS-PAGE (Figure 25).



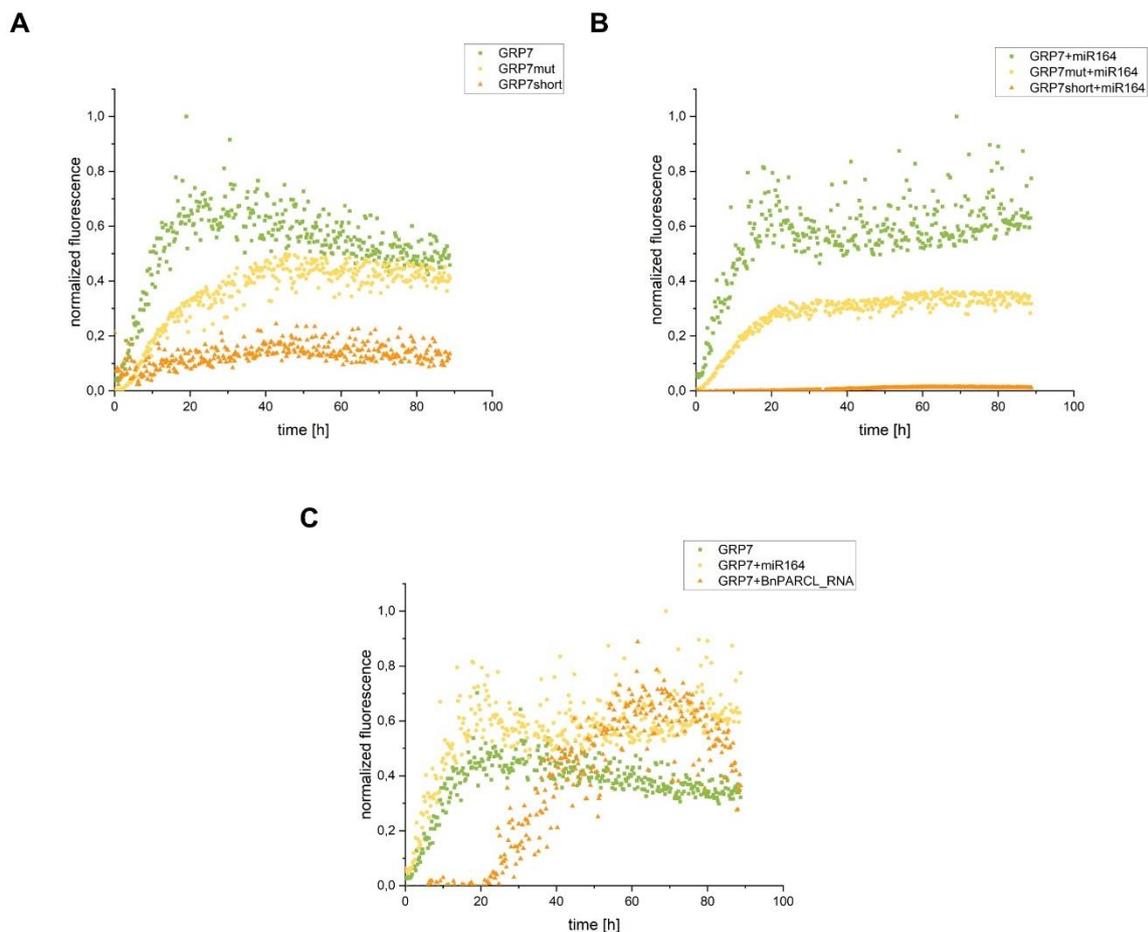
**Figure 25: SDS-PAGE with AtGRP7mut purification steps.** A) Nickel column fractions 8-18 containing the 10 kDa AtGRP7mut protein. B) SEC column fractions 9-16 containing the AtGRP7mut protein. A 15 % SDS-PAGE was used and run for 10 min at 150 V and at 200 V for 45 min. The protein ladder, PageRuler Prest Prestained from ThermoFisher was used, 10, 15, 25 and 70 kDa are highlighted in the ladder.

The cleanest protein-containing fractions after SEC were pooled and concentrated.

### 3 Results

#### 3.3.3.1 Condensation behavior of AtGRP7, AtGRP7short and AtGRP7mut

Whether or not the RGG-domain of AtGRP7 is involved in the formation of filamentous condensates, AtGRP7, AtGRP7short and AtGRP7mut were tested for their condensation potential. All three proteins were observed over 80 h with ThioflavinT, a dye known to interact with filamentous condensates and thereby increasing its fluorescence [147,155]. The fluorescence intensity was measured every 15 minutes. After normalization of the fluorescence intensity against the highest measured fluorescence, the normalized fluorescence intensity of AtGRP7, AtGRP7mut and AtGRP7short with and without *miR164* were compared over time (Figure 26).



**Figure 26: Aggregation of AtGRP7, AtGRP7mut and AtGRP7short visualized by ThioflavinT fluorescence. A)** Normalized fluorescence intensity (y-axis) of AtGRP7 (GRP7, green), AtGRP7mut (GRP7mut, yellow) and AtGRP7short (GRP7short, orange) displayed over 80 h (x-axis, time [h]). **B)** Normalized fluorescence intensity of AtGRP7 + *miR164* (green), AtGRP7mut + *miR164* (yellow) and AtGRP7short (orange) over a time period of 80 h. **C)** Normalized fluorescence intensity of AtGRP7 (green), AtGRP7 + *miR164* (yellow) and AtGRP7 + *BnPARCL* mRNA (orange) over a time period of 80 h.

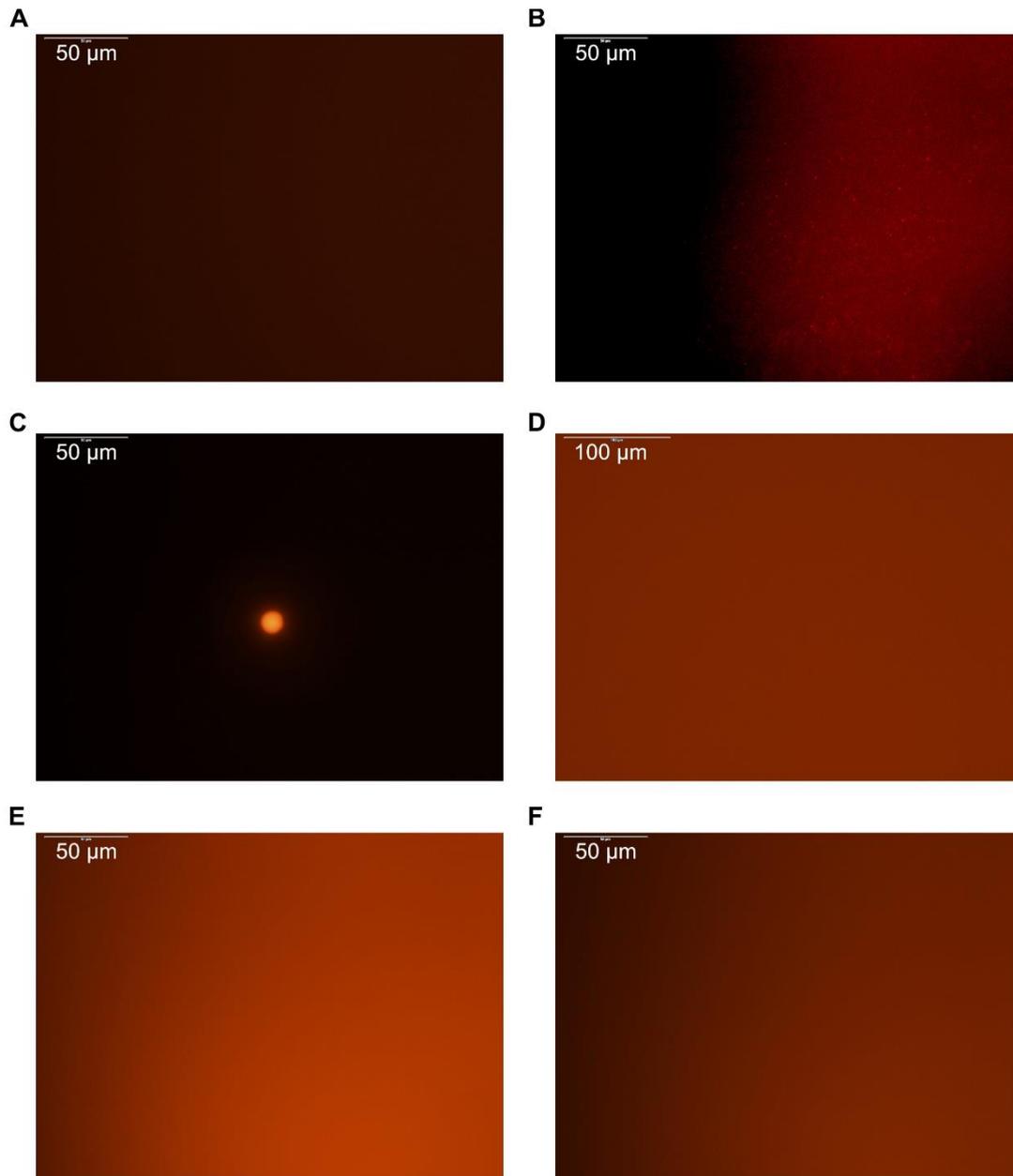
The fluorescence intensity of AtGRP7 raised quickly after starting the measurement and started to decline slightly after 24 h until it hit the same fluorescence intensity as AtGRP7mut (Figure 26, A). The fluorescence intensity of AtGRP7mut raised for 40 h, but not as fast as the fluorescence intensity of AtGRP7 and the fluorescence measured was not as high (Figure 26, A). The fluorescence intensity of

### 3 Results

AtGRP7short stayed quite stable most of the time at a low fluorescence intensity (Figure 26, A). A similar pattern was visible for the fluorescence measurement of the different AtGRP7 versions together with *miR164* (Figure 26, B), but the measuring points showed a more stable fluorescence. Lastly, comparing the fluorescence intensity of AtGRP7, AtGRP7 with *miR164* and AtGRP7 with *BnPARCL* RNA showed in all three cases a rapid increase of fluorescence and overall a high fluorescence. The AtGRP7 samples with RNA did not show a slight decrease of fluorescence like the sample containing only AtGRP7 did (Figure 26, C). This observation showed, that the condensation and thereby the fluorescence was influenced by the glycine-rich region, since AtGRP7short showed only low fluorescence and AtGRP7mut reduced fluorescence compared to AtGRP7. This observation did not change by adding RNA to the reaction.

Besides fibrilization, the capability of AtGRP7 to perform liquid-liquid-phase-separation (LLPS) together with RNA was tested. To detect LLPS, the RNA was labeled with cy3. Therefore, LLPS containing the RNA could be followed with a microscope under UV-light with a GFP filter (Figure 27). No LLPS was visible in AtGRP7 samples with RNA at RT, only after an incubation on ice, the solution became cloudy. This cloudiness appeared to be LLPS droplets formed by AtGRP7, since no LLPS was detected in samples containing only RNA (Figure 27, D). Through centrifugation of the liquid-liquid-phase separated AtGRP7, the small LLPS droplets were merged to larger droplets easily detectable with the microscope (Figure 27, C). No LLPS was visible in AtGRP7mut samples at RT or after incubation on ice (Figure 27, E and F).

### 3 Results

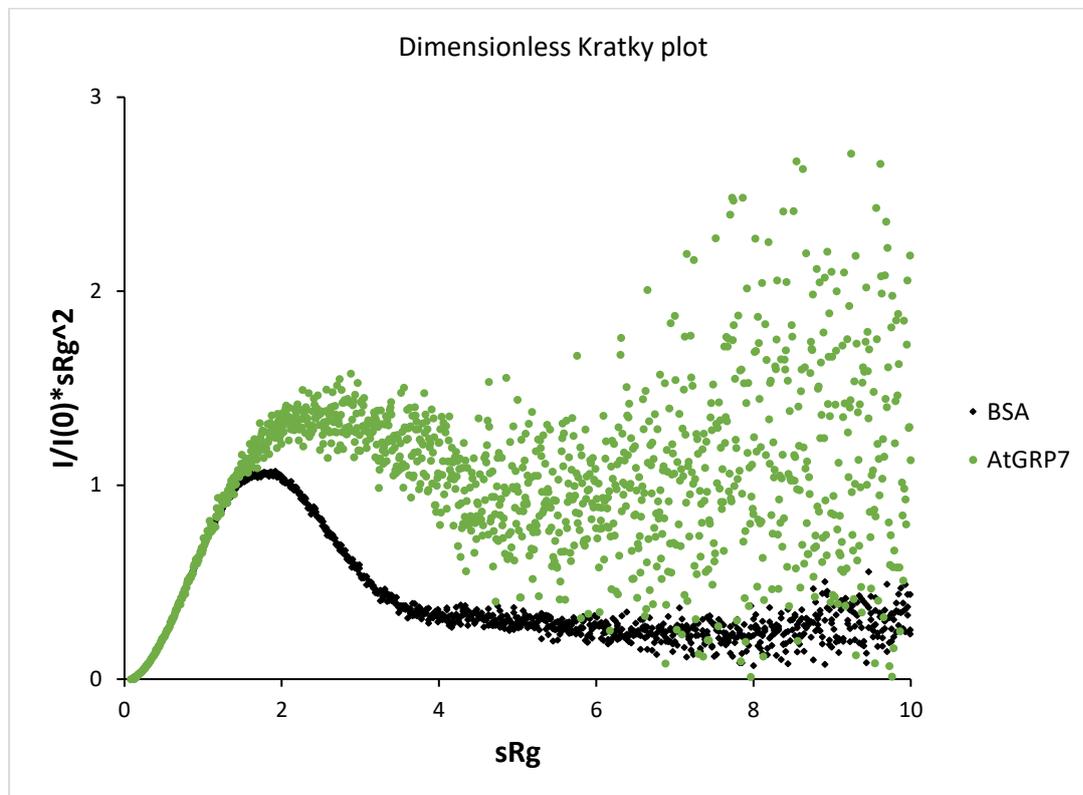


**Figure 27: Detection of liquid-liquid phase separation through a macroscope. A)** 50  $\mu\text{M}$  AtGRP7 with 1  $\mu\text{M}$  21-R RNA at RT. **B)** 50  $\mu\text{M}$  AtGRP7 with 1  $\mu\text{M}$  21-R RNA for incubated on ice for 5 min. Small clusters of fluorescence are visible. **C)** 50  $\mu\text{M}$  AtGRP7 with 1  $\mu\text{M}$  21-R RNA incubated on ice for 5 min and centrifuged for 5 min at 4  $^{\circ}\text{C}$ . The supernatant was removed and the pellet resuspended in 4  $\mu\text{l}$  of LLPS buffer with PEG. **D)** Only 1  $\mu\text{M}$  21-R RNA as negative control. **E)** 50  $\mu\text{M}$  AtGRP7mut with 21-R RNA at RT. **F)** 50  $\mu\text{M}$  AtGRP7mut with 1  $\mu\text{M}$  21-R RNA incubated on ice for 5 min.

#### 3.3.3.2 The overall structure of partly disordered AtGRP7

To determine the overall structure of AtGRP7 especially of the glycine-rich region, small-angle X-ray scattering (SAXS) measurements were performed. By SAXS, low resolution structures of proteins can be elucidated [156]. Such measurements allow the protein to stay in solution [156]. Most interesting for AtGRP7 was the glycine-rich region, since it is assumed to be highly disordered like other RGG-

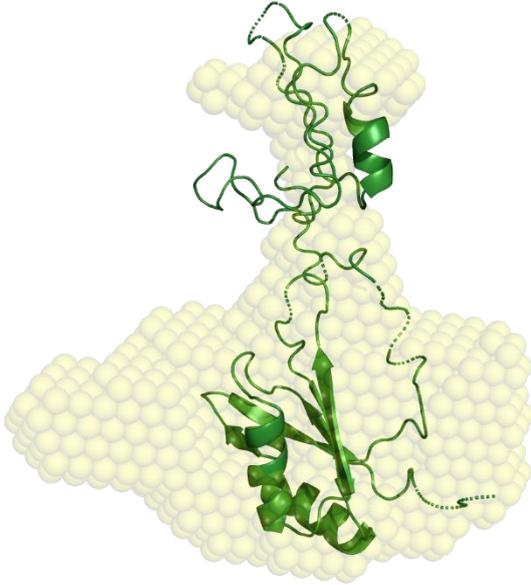
domains. The RRM of AtGRP7 is expected to be highly structured and globular as it is for other RRMs [62]. The dimensionless Kratky plot was used to gain an overview of the structured status of AtGRP7 (Figure 28), as proteins that contain unstructured regions did not show a bell-shaped plot like globular proteins, in this case BSA [157,158].



**Figure 28: Dimensionless Kratky plot of AtGRP7 (green) and BSA (black).** The Kratky plot shows the overall shape of the protein, if it is more globular or elongated. BSA is a globular shaped protein and the dimensionless Kratky plot allows the comparison of different proteins in one plot to estimate the overall shape of proteins. Y-axis:  $I/I(0)*sRg^2$ , x-axis:  $sRg$ .

The Kratky plot of AtGRP7 contains a bell-shaped start and opens up at the end. Compared to the globular protein BSA a difference is visible, this indicates that AtGRP7 is a partly disordered protein.

Besides the Kratky plot, an *ab initio* model of AtGRP7 was modeled with DAMMIF to visualize the overall shape of AtGRP7. The predicted structure of AtGRP7 was fitted to this model by SREFLEX to achieve a  $\chi^2$  of 1.43 from the original  $\chi^2$  of 5. The bead model was superimposed with the SREFLEX adjusted structure by SASpy to display the predicted location of RRM- and RGG-domain (Figure 29). The superimposition yielded a nsd of 4.

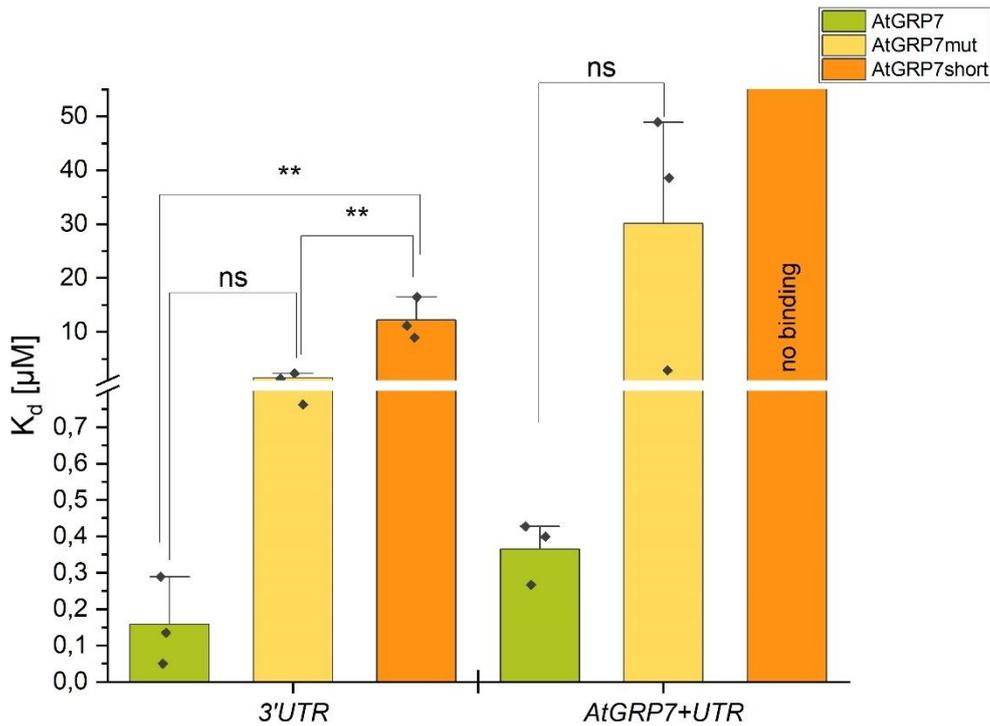


**Figure 29: Superimposed model of AtGRP7.** The bead model of AtGRP7 was constructed with DAMMIF from data of the SEC-SAXS measurement of AtGRP7. The predicted structure of AtGRP7 was generated by iTasser [159] and fitted to the bead model by SREFLEX.  $\chi^2$  of the predicted structure fitted to the bead model was 1.43. The nsd value for the superimposition is 4.

#### 3.3.3.3 RNA-binding of AtGRP7, AtGRP7short and AtGRP7mut

The RNA-binding affinity of AtGRP7 and BnGRP7 towards certain RNAs differed significantly. Due to their difference in the RGG-domain, the two AtGRP7 proteins with mutations in their RGG-domain were investigated for their RNA-binding ability. The RNA-binding affinity of AtGRP7, AtGRP7mut and AtGRP7short for short and long RNAs was compared (Figure 30). The first interesting candidate RNAs for this were the pre-mRNA of *AtGRP7* and the 3'UTR region of *AtGRP7* mRNA which is known to be a binding site of AtGRP7.

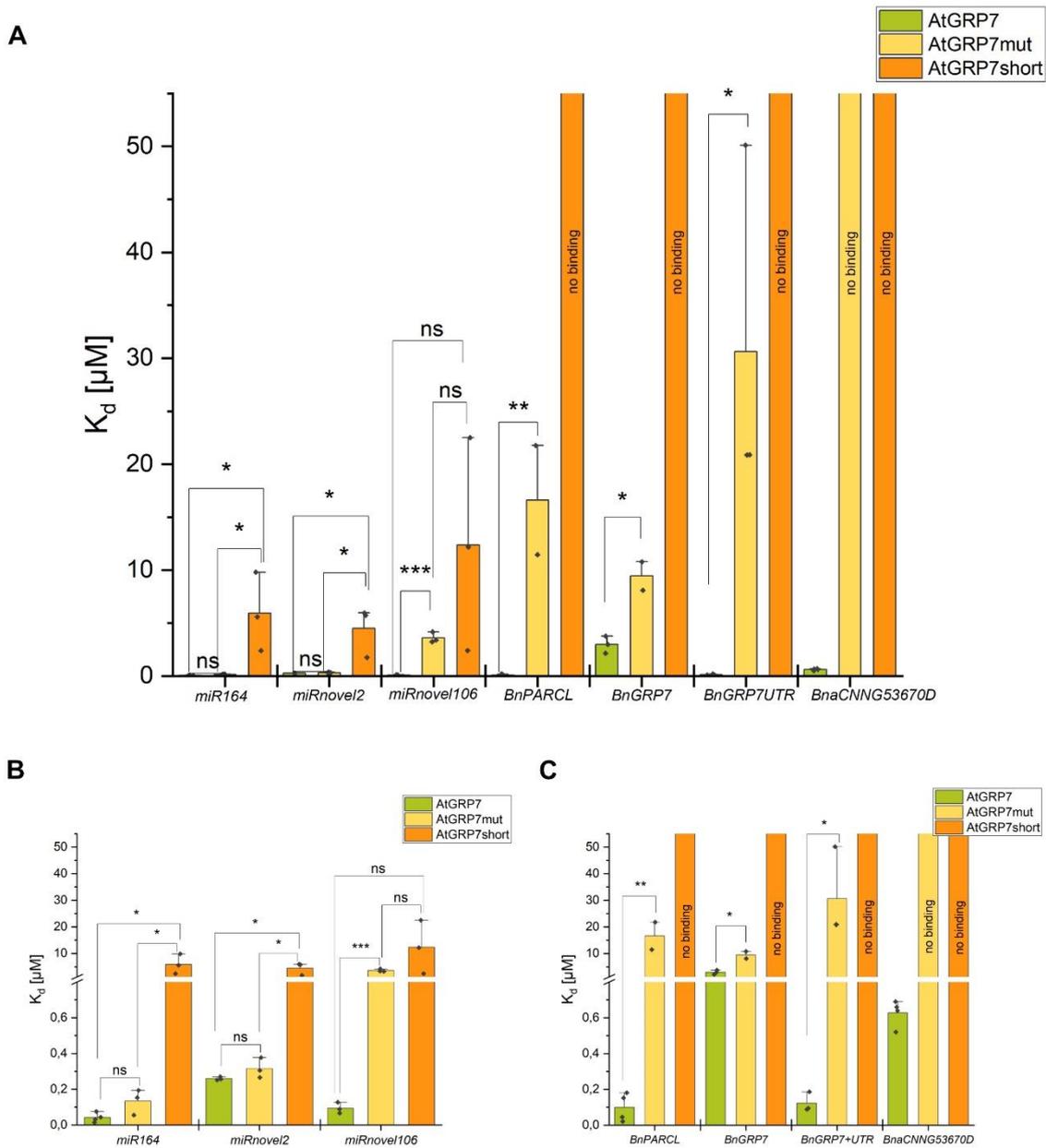
### 3 Results



**Figure 30: Binding affinity of AtGRP7, AtGRP7mut and AtGRP7short against the 3'UTR of AtGRP7 and AtGRP7 pre-mRNA.** The y-axis displays the binding affinity as dissociation constant ( $K_d$ ) in  $\mu\text{M}$  and was cut between 0.8 and 1  $\mu\text{M}$  for a better resolution of the binding affinities below 1  $\mu\text{M}$ . On the x-axis, the RNAs tested are named and the proteins which were used are displayed in three different colors, AtGRP7 in green, AtGRP7mut in yellow and AtGRP7short in orange. Significantly different binding affinities of the three proteins towards each RNAs were detected by one-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*), 0.001 (\*\*\*) were tested, no significant difference (ns) is claimed when none of the tested p-values showed a significant difference between the binding affinities of two proteins towards one RNA. No one-way-ANOVA was performed when no binding was detected.

AtGRP7 showed a statistically significant higher binding affinity towards its own 3'UTR region than AtGRP7short. AtGRP7mut did not show a statistically significant different binding affinity towards the 3'UTR as well as towards *AtGRP7* mRNA with UTRs and Intron. AtGRP7 short did not show binding of *AtGRP7* mRNA with UTR and Intron at all. This indicates, that the RGG-domain is required for RNA-binding. Even though no statistically significant  $K_d$  was observed between AtGRP7 and AtGRP7mut, the mean  $K_d$  of AtGRP7mut was nearly 100-fold higher than the  $K_d$  of AtGRP7 towards *AtGRP7* mRNA with UTRs and Introns. Therefore, the binding of AtGRP7 towards more RNAs was compared with the binding affinity of AtGRP7mut and AtGRP7short (Figure 31).

### 3 Results



**Figure 31: Binding affinity of AtGRP7, AtGRP7mut and AtGRP7short towards different RNAs. A)** Dissociation constant in  $\mu\text{M}$  as binding affinity of AtGRP7, AtGRP7mut and AtGRP7short towards small RNAs and long RNAs. **B)** Binding affinities of different AtGRP7 versions only against small RNAs with a cut in the y-axis between 0.8 and 1  $\mu\text{M}$  for a better resolution of the low dissociation constants. **C)** Binding affinities of different AtGRP7 versions only against long RNAs with a cut in the y-axis for a better resolution of the low dissociation constants. The y-axis displays the binding affinity as dissociation constant ( $K_d$ ) in  $\mu\text{M}$ . On the x-axis, the RNAs tested are named and the proteins which were used are displayed in three different colors, AtGRP7 in green, AtGRP7mut in yellow and AtGRP7short in orange. Significantly different binding affinities of the three proteins towards each RNAs were detected by one-way-ANOVA. P-values of 0.05 (marked with \* if a significant difference was found), 0.01 (\*\*) and 0.001 (\*\*\*) were tested, no significant difference (ns) is claimed when none of the tested p-values showed a significant difference. No one-way-ANOVA was performed when no binding was detected.

The comparison of  $K_d$ s between AtGRP7, AtGRP7mut and AtGRP7short for more RNAs showed a consistent trend. Small RNAs like *miR164* were mostly bound with around the same  $K_d$  by AtGRP7 and AtGRP7mut and therefore the statistical analysis resulted in no significant difference in binding affinity towards most small RNAs. AtGRP7short showed a higher  $K_d$  towards small RNAs and thereby showed

### 3 Results

a statistically significant different binding affinity for two of the three small RNAs compared to the binding affinity of AtGRP7 and AtGRP7mut.

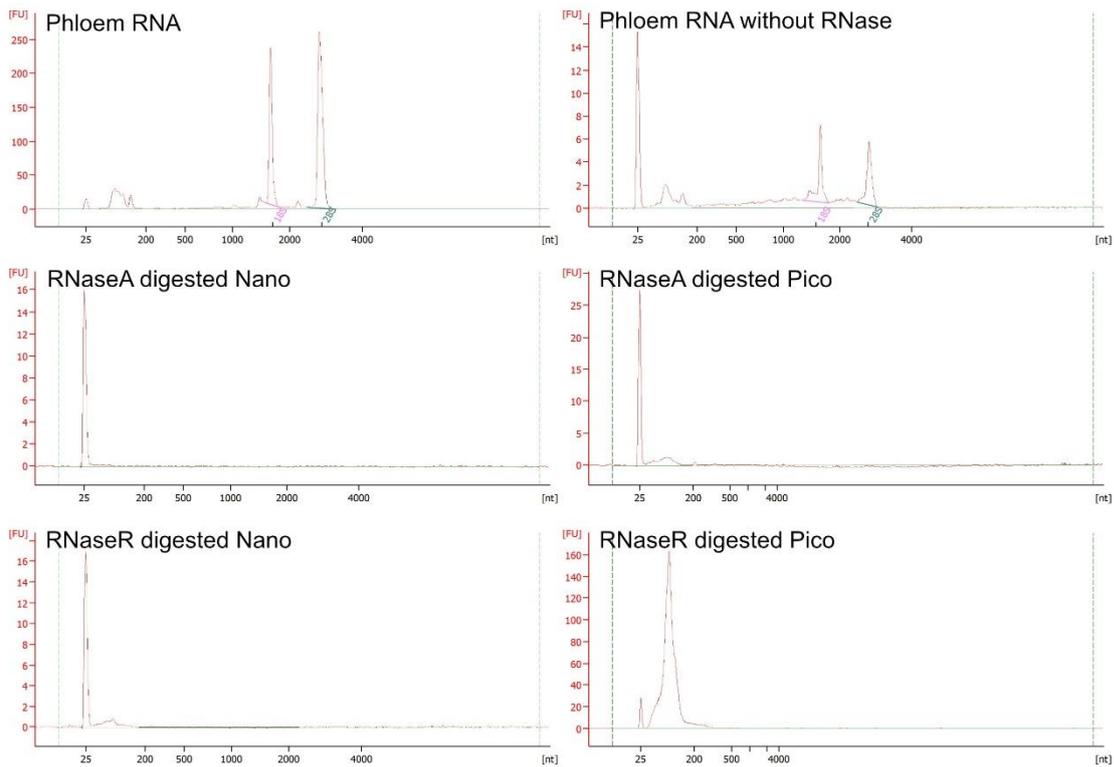
For longer RNAs like *BnPARCL* RNA or *BnGRP7* RNA, it was not possible to detect binding of AtGRP7short. Therefore, no statistical tests for significant differences were performed. AtGRP7mut showed binding towards longer RNAs, only for *BnaCNGG53670D* RNA no binding was detected. For *BnPARCL* and *BnGRP7+UTR* RNA, AtGRP7mut showed a statistically significant higher  $K_d$  and therefore a lower binding affinity compared to AtGRP7.

To conclude, the binding affinity of AtGRP7 for certain RNAs was influenced by the RGG-domain and therefore probably through the phosphorylation status of tyrosine within the glycine-rich region.

#### 3.4 Circular RNAs in phloem

Plants are known to contain circular RNA (circRNA), which originate from back-splicing of linear RNA. Due to the lack of 3'- and 5'-end, circRNAs are very stable. Therefore, they would be perfect candidates for a long-distance transport through the plant. To test whether or not circular RNAs are present within phloem sap, total phloem sap RNA was digested with RNaseR and RNaseA. RNaseR is an exonuclease and since the 3'- and 5'-end of circRNAs are covalently linked, no RNA ends are available leaving RNaseR without a substrate. Thus, an RNaseR treatment of total RNAs should lead to the digestion of linear RNAs but not of circular RNAs. As a control, RNaseA digestion was performed in parallel, which leads to the degradation of linear and circular RNAs. After the RNA digestion, leftover RNA was purified and analyzed on Nano- and PicoChips from Bioanalyzer (Figure 32).

### 3 Results



**Figure 32: Electropherograms of total Phloem RNA and RNase treated Phloem RNA.** Total phloem RNA shows its distinct electropherogram with two large peaks of ribosomal RNA and a small peak for small RNA. The same pattern can be found for Phloem RNA incubated at 37 °C without the addition of RNase. In RNaseA treated phloem RNA, only the marker is visible, on a NanoChip as well as on the PicoChip. For RNaseR treated phloem RNA, only a very small peak is visible on the NanoChip, but it is better visible on the PicoChip. The y-axis displays the fluorescence in fluorescence units (FU) and the x-axis displays the length in nucleotides. Marker peak at 25 nt.

Phloem RNA incubated without RNase showed the same peak-pattern as non-digested phloem RNA. The fluorescence units detected were less due to the high dilution of phloem RNA by the mock-treatment and following RNA purification. Phloem RNA treated with RNaseA was completely digested and no RNA peaks were visible on Nano- or PicoChip. Phloem RNA treated with RNaseR showed a very small peak directly following the marker peak. This peak was strongly visible on the PicoChip.

It can be concluded that the RNAs not digested by RNaseR might be circular. This raised the question which circular RNAs are in the phloem sap.

#### 3.4.1 circRNA detection in long-noncoding sequencing library of total phloem RNA and leaf RNA

Within the sequencing results of the long-noncoding sequencing library of total phloem sap RNA and total leaf RNA, circular RNAs should be detectable since these sequencing libraries have not been enriched for Poly-A containing mRNAs. Considering that the RNaseR digestion of phloem sap RNAs showed residual RNA and it is known that plants contain circRNAs, the sequencing results of total phloem sap RNA and total leaf RNA were further analyzed. These sequencing libraries were prepared

### 3 Results

from phloem RNA that was not sampled from the same plants as the leaves, since both libraries were just re-used and not specifically prepared for circRNA detection and comparison. By a STAR-alignment of the sequencing reads against the same genome assembly of *B. napus* (AST\_PRJEB5043\_v1) as used by Novogene, chimeric reads were identified. Chimeric reads are reads of which one part can be mapped to one gene and the other part towards another area of the same gene or another part in the genome which is not adjacent to the first location. These chimeric reads of the alignment were further analyzed with the DCC tool included in circTools to identify circRNAs, since these reads might contain a back-splicing junction (BSJ) of a circular RNA. Circular RNAs were detected in phloem sap as well as in leaves, but the read counts were quite low. The in phloem detected circRNAs with the highest read counts of at least 50 total read counts over all replicates can be found in table Table 29 and the whole table of predicted phloem circRNAs with an overall read count of at least 5 can be found in the supplements in Table 40.

**Table 29: List of circRNAs from phloem sap with the highest read counts (RC).** The location of circRNA is the 5' (acceptor) and 3' (donor) end of the back-splicing junction of each circular RNA. The read counts for each sample as well as the mean read count is shown.

Location	Gene	strand	RC Phloem1	RC Phloem2	RC Phloem3	mean RC
LK034926:12893-12926	BnaCnng48510D	-	459	342	376	392.3
LK034680:3566-3646	antisense strand of BnaCnng45920D	-	106	96	101	101
LK032891:85002-85037	intergenic	+	86	91	67	81.3
LK034680:3566-3646	BnaCnng45920D	+	68	56	64	62.7
LK032045:141379-143249	BnaC02g25110D-1	-	50	45	41	45.3
LK032045:141379-143249	antisense strand of BnaC02g25110D-1	+	47	46	39	44
LK032144:174755-197745	antisense strand of BnaC06g24630D- 70D	-	53	36	30	39.7
LK032104:439794-439832	intergenic	-	37	16	29	27.3
LK031850:1530258-1530301	Rev strand of BnaA0732420D	-	23	20	13	18.7
LK033217:59750-61637	BnaC02g48340D	-	22	22	12	18.7
LK032105:210538-210579	BnaC04g36770D	-	12	23	20	18.3
LK033217:59750-61637	antisense strand of BnaC02g48340D	+	21	22	9	17.3
LK031850:1530258-1530322	antisense strand of BnaA0732420D	-	15	21	15	17

Some circRNAs seemed to span multiple genes like the circRNA in location LK032144:174755-197745. Other circRNAs like LK034926:12893-12926 (BnaCnng48510D) are a fragment of an exon of one gene.

### 3 Results

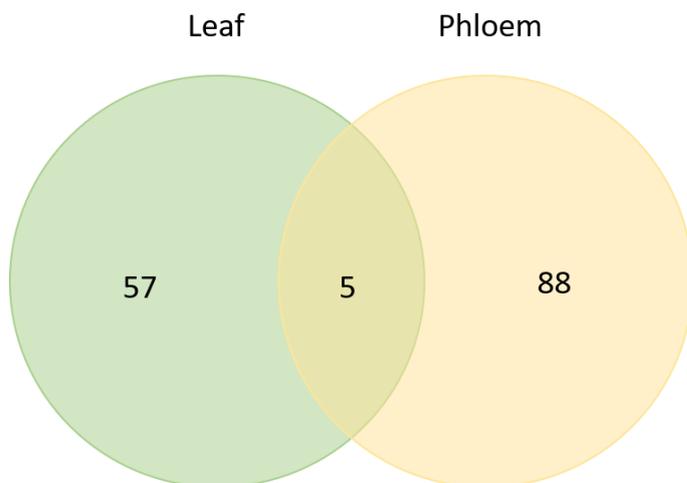
In total, 93 circRNAs were detected in *B. napus* phloem sap, with a cut-off of at least two reads per replicate.

In addition, circRNAs were also predicted in *B. napus* sequencing samples from leaves. Circular RNAs with the highest read counts can be found in Table 30 and a list of all circRNAs detected in leaves can be found in the supplements in Table 39.

**Table 30: List of circRNAs from leaves with the highest read counts (RC).** The location of circRNA is the 5' and 3' end of the back-splicing junction of each circular RNA. The read counts for each sample as well as the mean read count is shown.

Location	Gene	Strand	RC Leaf1	RC Leaf2	RC Leaf3	mean RC
LK034926:12893-3099	antisense strand of BnaCnng48500D and BnaCnng48510D	-	134	232	191	185.7
LK032891:85002-20613	BnaC08g14860D, BnaC08g14830D, BnaC08g14820D, BnaC08g14790D	+	104	102	139	115
LK032223:268479-168146	Intergenic, BnaA06g19440D, BnaA06g19450D, BnaC06g19460D	-	20	43	29	30.7
LK032105:210538-469641	Multiple genes	-	12	21	36	23
LK032116:501707-197745	Multiple genes	-	14	20	18	17.3
LK032334:301970-302466	intergenic	+	17	21	12	16.7

To compare if the same circRNAs were present in leaves and phloem sap of *B. napus*, a venn diagram was prepared (Figure 33).



**Figure 33: Venn diagram of leaf circRNAs and phloem circRNAs.**

### 3 Results

In both, leaf and phloem, 5 circRNAs are present, while 88 were exclusively found in phloem sap and 57 in leaves.

Since the prediction of circular RNAs through sequencing data results often in false positive circRNA candidates, the data analysis was repeated by Zsófia Fekete (Agricultural Biotechnology Center, Hungary) who analyzed the long non-coding phloem RNA sequencing results with her in house circRNA detection script. Thereby, she identified eight promising circRNA candidates (Table 31). Some of the candidates were also detected by circ-tools.

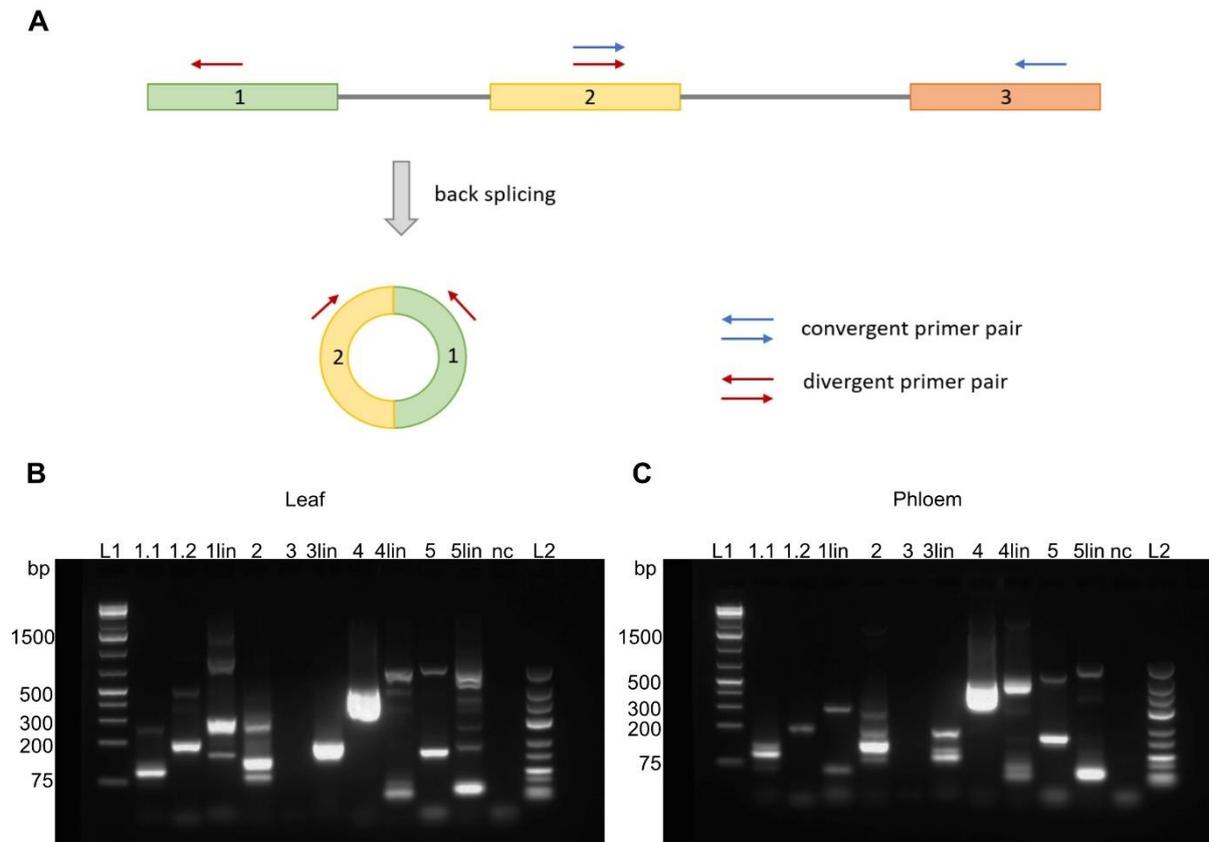
**Table 31: Overview of highly promising circular RNAs in phloem.** The analysis to gain the candidates was done by Zsófia Fekete (Agricultural Biotechnology Center, Hungary) .

circRNA number	circRNA location	Gene	mean RC	mean RC (circ-tools)	Possible length in bp
1	LK031899:250520-251632	BnaC08g16790D exon 1-7	5	Not found	1112
2	LK031823:1895973-1903835	BnaC03g03910D, BnaC03g03930D	5.7	4.3	7862
3	LK032206:21055-21587	intergenic	32.7	Not found	532
4	LK032045:141379-143249	Antisense strand of BnaC02g25110D	46	44	1870
5	LK032103:361866-362269	Antisense strand of BnaA01g28640D	9.7	Not found	403
6	LK033217:59750-61637	BnaC02g48340D	17.3	17.3	1887
7	LK031791:2161670-2161754	Antisense strand of BnaC03g08960D	7	Not found	84
8	LK032960:39014-39157	ENSRNA049475537	8.3	Not found	143

From these eight promising circRNAs, seven RNAs were further tested (circRNA1-5 and circRNA7-8).

One quick method to verify circRNAs is to perform a PCR with divergent primers. Divergent primers face with their 3'-end towards each other on a linear template. On a circular template, the primers also face with their 5'-end towards each other whereby a PCR-product can be obtained. Therefore, divergent as well as convergent primer were ordered to validate circRNAs from Zsófia Feketes (Agricultural Biotechnology Center, Hungary) list of promising circRNA candidates (Figure 34).

### 3 Results

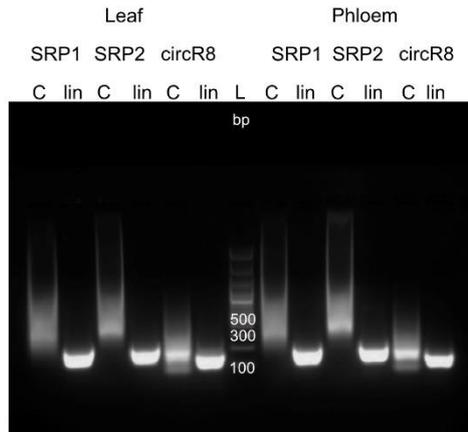


**Figure 34: Agarose gel with PCR amplicons originating from divergent and convergent primer pairs. A)** Visualization of the locations of divergent and convergent primer pairs. **B)** cDNA based on total leaf RNA was reverse transcribed with random hexamer primer. Following PCR contained either divergent or convergent primer pairs. **C)** cDNA based on total phloem sap RNA was reverse transcribed with random hexamer primer. Following PCR contained either divergent or convergent primer pairs. Amplicons originating from divergent primer pairs are numbered according to the circRNA number in Table 31. Amplicons originating from convergent primer pairs are numbered according to Table 31 and contain “lin” in their name. 2 % agarose gel run for 1 h at 120 V. L1: 1 kb plus ladder from ThermoFisher Scientific, L2: LowRange ladder from ThermoFisher Scientific.

The divergent-primer PCR of circRNAs showed amplicons spanning the back-splicing junction (BSJ) for circRNA1, circRNA2, circRNA4 and circRNA5. Thus, these circRNAs were likely to exist. For circRNA3, repetitions of the PCR gave inconsistent results and no amplicon was obtained from divergent primers for circRNA7 in all replications (not shown).

Lastly, circRNA8 was also tested with divergent and convergent primers. CircRNA8 originates from an SRP-RNA, these were found to be present in very high numbers in the phloem sap. SRP-RNAs have a highly ordered secondary structure with their 3' and 5' ends in close proximity [160,161]. Therefore, besides circRNA8, two other SRP-RNAs were tested with divergent and convergent primer pairs to make sure that the detection of circRNA8 was not caused by errors during cDNA synthesis (Figure 35).

### 3 Results



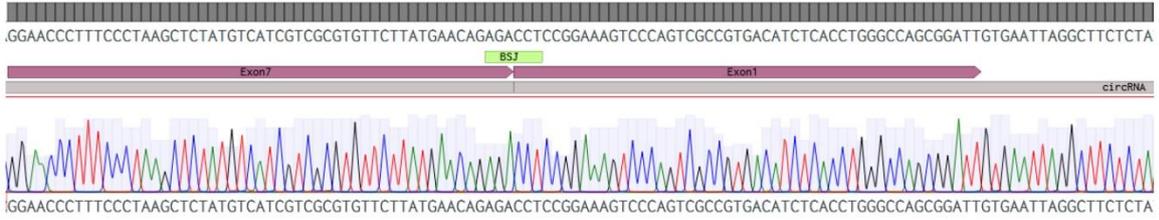
**Figure 35: Agarose gel with divergent and convergent PCR-amplicons of different SRP-RNAs.** SRP-RNA1 (ENSRNA049470465), SRP-RNA2 (ENSRNA049475535) and circRNA8 (ENSRNA049475537) were amplified from cDNA with divergent (C) and convergent (lin) primer pairs. The linear amplicons were expected to have a size of approximately 150 bp, the divergent amplicons were expected to have a size of 300 bp (for SRP-RNA1 and 2) and 150 bp (for circRNA8). 2 % Agarose gel, run for 1 h at 120 V. L: 100 bp plus ladder from ThermoFisher Scientific.

The amplicons from convergent primer pairs showed the expected size of 150 bp for cDNA from leaf and from phloem. The divergent primer amplicons show a smear between 200 and 500 bp for SRP-RNA1 and SRP-RNA2 and a smear with a visible band for circRNA8 for cDNA from leaf and phloem.

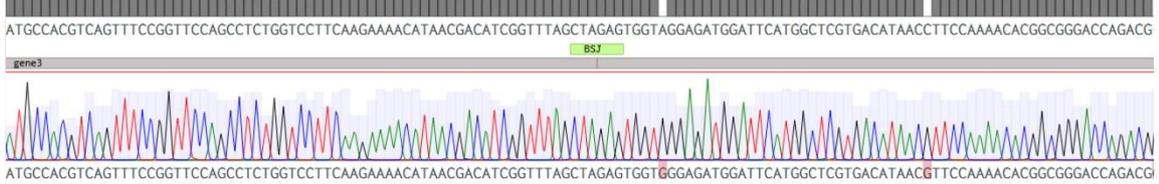
It can be concluded, that six out of eight circRNA candidates were verified by PCR. To be sure the amplicons really span the BSJ of each circRNA, the amplicons were purified and send for Sanger sequencing. The alignment of the sequencing results towards circRNA1, 2, 4 and 5 is shown in Figure 36.

### 3 Results

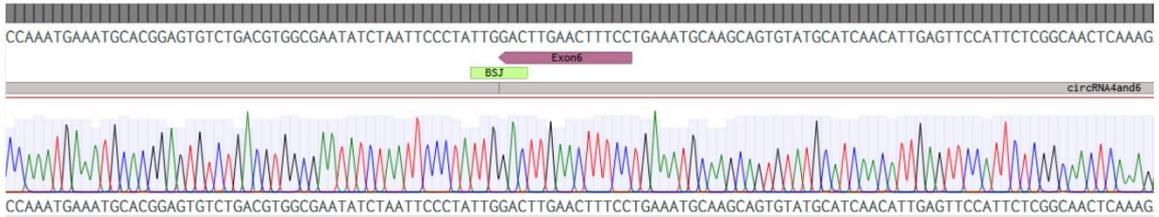
#### circRNA1\_BSJ



#### circRNA2\_BSJ



#### circRNA4\_BSJ



#### circRNA5\_BSJ

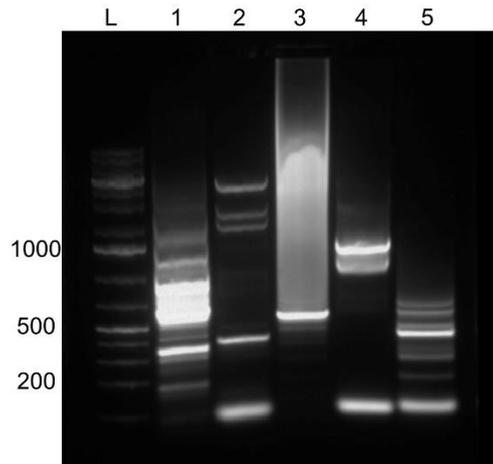


**Figure 36: Pairwise sequence alignment of BSJ amplicons with potential circRNA sequence.** The amplicons of BSJ were cloned into pUC57 and sequenced by Sanger sequencing (Microsynth). The sequencing results of circRNA1, circRNA2, circRNA4 and circRNA5 BSJ were aligned with the designated circRNA with Benchling by pairwise sequencing. For the alignment, standard settings were used. The predicted BSJ are highlighted with a green feature. Exons are highlighted with a lilac feature. For circRNA5, also the donor and acceptor sites were highlighted in blue since the predicted BSJ was not the BSJ detected by Sanger sequencing.

The BSJ of circRNA3 and circRNA8 could not be sequenced. For circRNA1, circRNA2 and circRNA4, the predicted BSJ was covered by Sanger sequencing, while the BSJ of circRNA5 was covered by Sanger sequencing but showed a deletion, thereby indicating that the BSJ of circRNA5 might be different than predicted.

Since circRNAs can span over multiple exons and do not always contain all introns in between these exons, the size and sequence of a circular RNA cannot be determined solely through the distance of the back-splicing donor and acceptor site. To gain knowledge on the full-length of the RNAs, circRNA candidates were fully amplified by using two primers back-to-back to each other (Figure 37).

### 3 Results



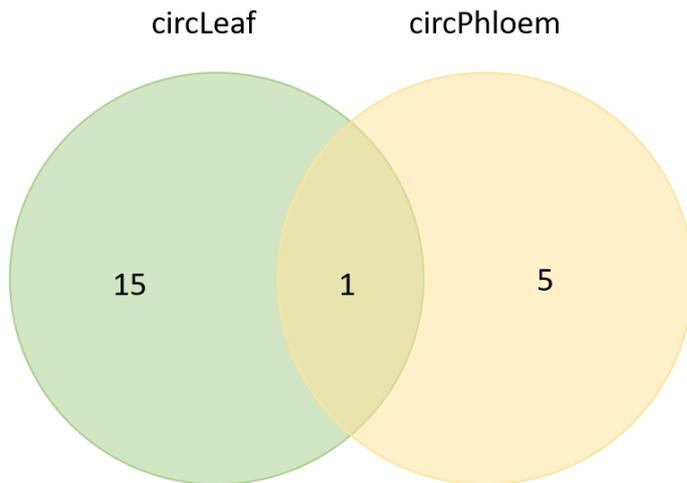
**Figure 37: Agarose gel with total circRNA amplicons.** CircRNA1, 2, 3, 4 and 5 were amplified from cDNA with back-to-back primer. The 1 % agarose gel run for 55 min at 120 V

The amplification of total circRNA resulted for many circRNAs in multiple amplicons. Only circRNA3 showed one distinct band and a lot of smear above. Only two very close amplicons were observed for circRNA4. The attempted to purification of the amplicons and the following sequencing was not successful for all circRNAs.

#### 3.4.2 circRNA detection in circRNA enriched sequencing library of total phloem and leaf RNA

An additional method to verify that circRNAs, identified from chimeric reads of a lnc-RNA sequencing library are really circRNAs is the preparation and sequencing of a circRNA enriched library. Usually, this enrichment of circRNAs can be achieved by RNaseR digestion and the removal of rRNAs. Hence, phloem sap and leaf material were harvested from separate plants to generate biological duplicates. Leaf material and phloem sap from each biological replicate originated from the same plants. After the isolation of total RNA, the samples were send to Novogene for circRNA library preparation, sequencing and bioinformatic analysis.

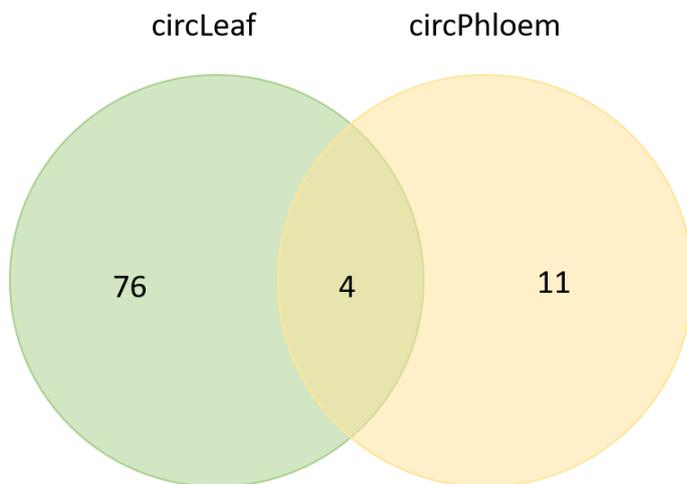
The analysis of Novogene showed that six circRNAs were identified in phloem sap and 16 circRNAs in leaves and one of these RNAs in both (Figure 38).



**Figure 38: Venn diagram of circRNAs in leaf and phloem sap.** Overlap of leaf circRNAs and phloem circRNAs identified by Novogene from circRNA enriched sequencing library.

The amount of identified circRNAs from Novogene in the circRNA enriched sequencing data was lower than the identified circRNAs from the lnc-library by circTools and by the in-house script from Zsófia Fekete (Agricultural Biotechnology Center, Hungary).

To make sure that this difference in numbers comes from the RNaseR digestion and not from sequence analysis, the sequence analysis done for lnc-library was repeated with mapping of circRNA enriched reads by STAR and the discovery of circRNAs through circTools. By this, 15 circRNAs were identified in phloem samples and 80 were identified in leaf samples (Figure 39).



**Figure 39: Venn diagram of circRNAs in leaf and phloem in RNaseR treated sequencing library identified with the detect module (DCC) from circTools.**

The analysis showed, that the amount of leaf circRNA was higher in circRNA enriched sequencing than in the lncRNA sequencing, where only 62 were found. In phloem, 68 circRNAs less were found in circRNA enriched sequencing compared to the circRNAs found in the lncRNA sequencing.

### 3 Results

Upon comparison of the six circRNAs identified by Novogene in the sequencing data, none was found to be matching the 15 circRNAs identified with circTools. Likewise, the circRNAs detected in the lncRNA sequencing of phloem did not show any match to the circRNAs found by Novogene. In contrast, nine of the 15 circRNAs identified with the circTools detect module in RNaseR treated sequencing samples were also identified in lncRNA sequencing. The six circRNAs found in both sequencing samples are interesting targets for further research. They are listed in Table 32.

**Table 32: Phloem circRNAs detected in lncRNA and circRNA sequencing.**

location	first_lnc	second_lnc	third_lnc	mean RC_lnc	first_circ	second_circ	mean RC_circ	strand
LK032334:301970-302005, intergenic	5	4	5	<b>4.7</b>	34	44	<b>26</b>	+
LK032891:85002-85037, intergenic	86	91	67	<b>81.3</b>	479	752	<b>410.3</b>	+
LK032960:6840-6928, BnaA06g02960D, probable carboxylesterase 3	5	5	2	<b>3</b>	152	108	<b>86.6</b>	-
LK034680:3566-3646, Bnacnng45920D, nucleolin isoform2	68	56	64	<b>62.7</b>	3	5	<b>2.7</b>	+
LK034680:3566-3646, antisense strand of BnaCnng45920D	106	96	101	<b>101</b>	3	18	<b>7</b>	-
LK034926:12893-12926, antisense strand of BnaCnng48510D	459	342	376	<b>192.3</b>	22	43	<b>21.7</b>	-

All further identified circRNAs can be found in Table 43 (leaf circRNAs) and Table 44 (phloem circRNAs) in the supplements. Most of the identified circRNAs in Table 27 are very short (<100 nt) and only three are derived from known, protein coding genes.

## 4 Discussion

In this chapter, the results presented in this thesis will be discussed. The RNA-binding capability of GRP7 and the involvement of the glycine-rich region as well as the circRNA content of the phloem will be part of this discussion.

### 4.1 Successful purification of GRP7

During the purification of GRP7, different problems occurred. The first one was the visible precipitation of GRP7 at low temperatures. This precipitation appeared to be liquid-liquid-phase separation (LLPS) (see Figure 27). In later experiments, LLPS caused by GRP7 was observed and will be discussed further (see 4.3.3.1). The protein solution got as cloudy with LLPS during the purification at 4°C. Hence, protein purification was adapted to room temperature to reduce protein loss and hexandiol 1.6 was added, since it disrupts droplet formation [162,163]. The second problem occurring during purification could not be solved. Besides full-length GRP7, multiple degraded versions of GRP7 were purified (Figure 6 and Figure 45). The reduction of protein expression time to 3 h as well as the use of multiple different protease inhibitors like AEBSF, PMSF and a proteinase inhibitor cocktail from Roche did not stop degradation. In a study about the degradation and synthesis rate of proteins in *A. thaliana*, GRP7 was reported to be fast degrading [164]. Additionally, GRP7 seems to be degraded by TUMV during its infection of *A. thaliana*, thereby abolishing the inhibitory effect of GRP7 on the systemic spread of the virus [88]. This indicates the fragility of the C-terminus of GRP7. Since the protease sites responsible for the degradation of the protein could not be identified, GRP7 was purified with degradation as little as possible, but the problem could not be solved completely.

### 4.2 GRP7 possess a degenerative RNA-binding affinity

To investigate the RNA-binding ability of GRP7, the first question to be answered was what kind of RNAs GRP7 can bind. AtGRP7 was earlier shown to bind pre-mRNAs especially within the 3'UTR to facilitate alternative splicing (AS) of these RNAs [55,92]. GRP7 is not only a nuclear protein, but it is also found in the cytosol where it also binds RNA, like the vsiRNA from TUMV [88]. Thus, GRP7 was reported to bind short and long, unprocessed mRNAs. To get a better understanding what types of RNAs are bound by GRP7, the protein was immobilized on CnBr-sepharose beads and incubated with leaf RNA. The immobilized AtGRP7 was still able to bind RNA, since RNA was found to be in the elution while no RNA was left after the fifth washing step (Figure 7). This RNA was bound by GRP7 and not by open binding spots on CnBr-sepharose beads, in consideration of the negative control with CnBr beads

only. The negative control did not show any RNAs in the elution after five washing steps (Figure 46). Additionally, the elution of AtGRP7short immobilized on CnBr-sepharose beads did not show any bound RNAs (Figure 47). In contrast, the elution of AtGRP7 showed peaks for all kind of RNA sizes compared to input RNA on the electropherograms from bioanalyzer NanoChip, indicating a broad range of RNA-binding ability. Therefore, the elution was sequenced and compared with the total leaf RNA input by Novogene. The analysis of enriched RNAs in the elution compared to the input leaf RNA revealed, that the enrichment of RNAs by AtGRP7 was minimal. Only ten of the enriched RNAs had a Log2FoldChange over 1 (Table 24). Four of these RNAs had an infinite Log2FoldChange since no read counts in input samples were found. Also, the read counts for the enriched RNAs were quite low, only for *BnaC05g32140D* the read count was above 100. Additionally, the comparison between input leaf RNA and the elution was done by Huge Wolfeenden (John-Innes-Center, Norwich, UK). The list of enriched candidates had only few overlaps with the enriched RNAs found by the analysis of Novogene. Both analyses showed an enrichment of *BnaC01g27050D*, *BnaC08g46660D*, *BnaC09g16500D* and *BnaA02g36030D*. In contrast, the analysis by Novogene showed no significant enrichment of *BnaC07g31170D*, *Bnaanng40960D* and *BnaC03g26040D*. The other RNAs from the list of Huge Wolfeenden (John-Innes-Center, Norwich, UK) have not been found at all in the sequence analysis of Novogene or were only found in input leaf RNA samples. The variations in enriched RNAs might be due to the different pipelines used for the analysis of the data, since each pipeline works in another way and can result in large variations in the outcome of analysis [165].

Sequencing and data analysis of small RNAs by Novogene revealed only four small RNAs to be significantly enriched by AtGRP7. All of these RNAs are miRNAs, three of them novel miRNAs (*bna-miRnovel2*, *bna-miRnovel106* and *bna-miRnovel149*) and one known miRNA, *bna-miR164a*.

To summarize the enrichment analysis of AtGRP7 CnBr-column, AtGRP7 did not show a high selectivity towards certain RNAs. This might not only be due to a general non-selective binding of AtGRP7. In this assay, immobilized AtGRP7 was incubated with total leaf RNA from *Brassica napus*. Even though *A. thaliana* is closely related to *B. napus*, they are still different plants. This might have caused lower binding selectivity of AtGRP7 towards *B. napus* RNAs and therefore prevented an enrichment of small RNAs.

#### 4.2.1 GRP7 binds a broad range of phloem-mRNAs

Since GRP7 is found in phloem sap, it is likely that GRP7 interacts with RNAs within the phloem. The first observation of phloem sap RNAs bound by GRP7 was done by immobilized GRP7. This time, BnGRP7 (*Brassica napus* GRP7) was immobilized on CnBr-sepharose beads and then incubated with

## 4 Discussion

total phloem sap RNA. Unbound RNAs were removed over the course of five washing steps and the RNAs bound by BnGRP7 were eluted. The elution was changed from one elution with 2 M sodium acetate to three elutions with 0.25, 0.5 and 2 M sodium acetate respectively. The last elution was expected to hold RNAs strongly bound by BnGRP7 because weakly bound RNAs were expected to be washed off by lower salt concentration. With a NanoChip, peaks for different sizes of RNAs were detected in the electropherogram on the bioanalyzer (Figure 11). Elution three, as well as the total input phloem RNA, were sequenced and compared by Novogene. More than 40000 RNAs were found in both total phloem sap RNA and the elution. 2397 RNAs were detected in total phloem sap RNA but not in elution three (Figure 13). 388 RNAs were significantly enriched in the elution compared to the total phloem sap RNA with a Log2FoldChange above 1. From these 388, 202 had a normalized read count above 100 (Table 35). Thus, the enrichment of RNAs within elution three of BnGRP7 CnBr-column showed a higher enrichment than by AtGRP7 CnBr-column. But this time, a poly-A library instead of a Inc-library was used for sequencing, thereby maybe influencing the outcome. This change was done to have a better comparability with other phloem RNA sequencing results. Additionally, the Log2FoldChange of RNAs with normalized read counts above 100 within the elution was only for 30 RNAs higher than 1.5, indicating that the amount of the eluted RNAs, even though it is statistically significant, is low. The sequencing and comparison of small RNAs within phloem RNA input and elution did not show any significantly enriched RNAs. Hence, RNAs bound by AtGRP7 as well as BnGRP7 were identified by sequencing but no selectivity could be inferred from these results. However, it seems like AtGRP7 as well as BnGRP7 bind a large set of different RNAs, especially considering that in the sequencing of BnGRP7, over 40000 RNAs were detected in input RNA and elution and only 2397 RNAs that were detected in input RNA have not been detected in the elution. This leads to the assumption that GRP7 does not have selectivity for RNA length, RNA sequence- or structural motifs. This result is in contrast to the 452 high-confidence RNAs bound by AtGRP7 found in iCLiP and RIPseq data of Meyer *et al.* [92] which share a specific RNA sequence motif. This U/C-rich motif is bound preferably by AtGRP7. Additionally, the authors were able to identify the 3'UTR with a total of 559 binding events to be the preferred binding site of AtGRP7. The 3'UTR is followed by exons and 5'UTR as preferred binding site [92]. In this thesis it might be possible that certain regions of the protein were not able to interact with RNA due to a steric hinderance by the immobilization of GRP7 towards the CnBr-sepharose beads, even though the RNA-binding regions of GRP7 should have been blocked by dextran sulfat during the immobilization. Nevertheless, some RNAs were enriched like *BnaA07g32150D* showing a high normalized RC as well as a Log2FoldChange of at least 1.5, thereby showing some ability to bind RNAs despite the immobilization.

## 4 Discussion

### 4.2.2 RNA-binding affinity of AtGRP7 and BnGRP7 towards different RNAs reveals a preference of binding sRNAs

To further investigate the RNA-binding selectivity, the binding affinity of GRP7 towards different RNAs was tested. Measurements were performed with microscale thermophoresis. First, RNAs with different lengths were tested. Only four small RNAs were significantly enriched in the column elution of AtGRP7 with leaf RNA and no small RNAs were enriched in BnGRP7 column elution with phloem sap RNA. In contrast to this low enrichment of small RNAs was the binding affinity of AtGRP7 and BnGRP7 towards small RNAs (Figure 14 and Figure 15). The binding affinity of AtGRP7 towards the four enriched small RNAs, *bnamiR164a*, *bnamiRnovel2*, *bnamiRnovel106* and *bnamiRnovel149* were in the nanomolar range with  $K_d$ s ranging from 0.0417  $\mu\text{M}$  to 0.655  $\mu\text{M}$  (Table 27). For longer RNAs that have been enriched by AtGRP7 and BnGRP7, AtGRP7 showed  $K_d$ s of 0.627  $\mu\text{M}$  (*BnaCNGG53670D*) and 3.67  $\mu\text{M}$  (*BnaA03g14400D*). Compared to the  $K_d$  of *bnamiR164a* (0.0417  $\mu\text{M}$ ), one of the strongest binding affinities measured for AtGRP7 in this work, the binding affinity towards longer RNAs is significantly lower (Figure 14). This indicates a selectivity of AtGRP7 for small RNAs, since all smaller RNAs measured, including 21-R RNA and tRNA-met show  $K_d$ s in a nanomolar range while the  $K_d$ s for long RNAs span from 0.1  $\mu\text{M}$  up to 65.88  $\mu\text{M}$  (Figure 14 and Table 27). The same binding affinity difference between small RNAs and longer RNAs was observed in binding studies with BnGRP7. There, the  $K_d$  values for small RNAs lay between 0.0039  $\mu\text{M}$  and 0.06  $\mu\text{M}$ , while enriched long RNAs showed  $K_d$  values of 2.75  $\mu\text{M}$  and 15.7  $\mu\text{M}$  (Figure 15 and Table 28).

Nevertheless, a higher number of long RNAs has been measured than small RNA. In addition, only enriched small RNAs were measured but no small RNAs which were not enriched. Hence, it might be that other small RNAs are not bound with such a low  $K_d$ . In contrast, also enriched, long RNAs show higher  $K_d$  values than most of the small RNAs measured and long RNAs which were not enriched did not show a different binding affinity range compared to the enriched long RNAs. Even though many RNA affinity measurements were conducted, more measurements could be conducted to state a finale conclusion about a higher selectivity of AtGRP7 towards small RNAs. This is especially true since the well-studied function of AtGRP7 lays within the interaction with pre-mRNAs [92]. On the other hand, the general interaction of CsSRBP1, a close homologue of AtGRP7 in cucumber, with ss sRNA and vsiRNA has been studied and it was also shown to interact with siRNA and mediate their transport [37,88], which also indicates a preference for short RNA. For the FUS protein, RNA-binding in a length dependent manner was already observed [166], thus this might be true for GRP7 as well.

The binding affinity of AtGRP7 towards single and double stranded RNA has been studied too. A significantly different  $K_d$  was observed for AtGRP7 binding towards *bnamiR164* ssRNA and dsRNA but no significant different affinity was measured for AtGRP7 towards *BnPARCL* ssRNA and dsRNA (Figure

16). Binding towards dsRNA showed large variation between the repetitions, leading to large error bars. Aside from the large error, the trend indicates a lower binding affinity towards dsRNA. This trend could be explained by the preference of the RNA recognition motif (RRM) for ssRNAs [62]. It is assumed that probably not all of the RNAs were double stranded and by that leading to large differences between the repetitions. This might have happened because the annealing of the complementary strands did not work out for each of the strands. Additionally, the concentration for both strands was measured and from this the volume needed to achieve equimolar amounts of both strands was calculated. Nevertheless, the measurement of the concentration is not without error, as well as the pipetting thereafter, which might have led to an unequal amount of both strands. Therefore, it is expected to never achieve 100 % dsRNA and thereby leading to the measurement of single stranded and double stranded RNA at the same time. To minimize this error, hairpin structured RNA with a large complementary sequence could be used as double stranded RNA as it was utilized by Ozdilek *et. al* [167]. Due to the large error for both dsRNAs it would be necessary to repeat the measurements with hairpin RNAs to achieve a better result and understanding about the dsRNA-binding ability of GRP7.

Besides the length of RNAs, specific modifications of RNAs could lead to specific binding and a higher binding affinity. Hence, the percentage of m<sup>5</sup>C methylation of phloem input and the eluted RNAs from BnGRP7 CnBr-column was analyzed through ELISA. This analysis resulted in no significant difference in the content of m<sup>5</sup>C methylated RNAs between input RNA and elution but a trend towards a higher m<sup>5</sup>C content in the elution was observed (Figure 17). Due to the low number of biological replicates (n=2), the error is large. By repeating this experiment, a lower error could be achieved. For now, the results point to no clear selectivity of BnGRP7 for m<sup>5</sup>C methylated RNAs.

The K<sub>d</sub> values of AtGRP7 towards two of four m<sup>5</sup>C methylated RNAs (*bna-miR164a* and *BnaCNNG53670D*) did show a significant difference compared to the same RNA without methylation (Figure 18). In both cases, the K<sub>d</sub> was significantly higher for m<sup>5</sup>C methylated RNA and therefore, the binding affinity lower. Besides m<sup>5</sup>C, also m<sup>6</sup>A methylation was tested, since m<sup>6</sup>A methylation can be found in 3'UTRs of plant RNAs [168], a preferred binding spot of AtGRP7 [92]. Again, m<sup>6</sup>A methylated *bna-miR164* and *BnaCNNG53670D* RNA showed significant higher K<sub>d</sub> values while the other two RNAs showed no significant difference in their K<sub>d</sub> when methylated or not. For BnGRP7, only m<sup>5</sup>C methylated *BnPARCL* RNA had a significantly higher K<sub>d</sub> than the not methylated *BnPARCL* RNA and thereby a significantly lower binding affinity. This indicates that GRP7 does not show selectivity for methylated RNAs. According to other researchers, a high number of graft-mobile RNAs contain a m<sup>5</sup>C methylation [44]. Therefore, it would have been interesting if GRP7 would harbor a higher binding affinity towards m<sup>5</sup>C methylated RNA, but GRP7 does not seem to be involved in the selection of m<sup>5</sup>C methylated RNAs for the transport of these RNAs. To confirm this hypothesis, further affinity measurements should be

performed. For now, only random RNAs were selected and partly methylated at random spots during their synthesis. Thereby, different methylation patterns for each RNA occurred with a different number of methylated sites. Thus, each RNA was methylated at a different site, some at the binding sites of GRP7, some not. By this, huge differences between the repeated measurements for just one RNA were created which can be observed in larger error bars (Figure 18). To get more naturally methylated RNAs, RNAs with known m<sup>5</sup>C and m<sup>6</sup>A methylation sites could be selected like *TCTP1* mRNA or *CAT3* mRNA for m<sup>5</sup>C methylation [44]. Then, these RNAs could be *in vitro* methylated by a methyltransferase like DNMT2 and NSUN2B, since double mutants lead to missing m<sup>5</sup>C methylation of *TCTP1* and thereby to a lack in its mobility [44]. Especially m<sup>6</sup>A methylation within 3'UTRs which are known to be bound by GRP7 should be tested to investigate if m<sup>6</sup>A methylation within the UTR changes the binding affinity of AtGRP7 and might be a method to regulate the binding of GRP7.

Taken together, it was not possible to determine a specific RNA motif or RNA modifications leading to the binding specificity of AtGRP7 and BnGRP7, even though GRP7 does seem to have a higher selectivity towards certain RNAs like its own pre-mRNA and *miRNA164*. This might be due to the RGG-domain of GRP7. In other RBPs containing one or multiple RGG-domain, a similar behavior was observed. The FUS protein binds a huge number of cellular RNAs [169,170] and was reported to bind many different sequence motifs [169,171]. However, the difference in binding affinity of FUS was not huge between these different motifs as well as the binding affinity towards negative control [166]. From this it was concluded that FUS binds a large number of RNAs with comparable binding affinities without a defined RNA sequence motif or structure [166]. Besides FUS, other RGG-domain containing proteins like hRNPs show a broad binding of RNAs [172–174]. Therefore, selectivity of the proteins for certain RNAs could not be defined, even though FUS, FSMR and hnRNPs seem to have a higher selectivity for structured RNAs like G-quadruplex RNAs [80,167,175]. Additionally, FUS has a high binding affinity towards *DNMT* RNA which does not change drastically when mutating the RGG-domain while these alterations result in a greater change in binding affinity towards a huge number of other cellular RNAs [167]. Another indication of a broad RNA-binding spectrum of GRP7 is the huge number of intermolecular interactions of the RGG-domain. It can interact with the phosphate backbone, nucleotide bases and ribose of RNA via pi-pi stacking, hydrogen bonds, electrostatic interactions and pi-cation stacking [81]. Thus, the variety of RNA sequences bound by this domain can be huge, as indicated by scientist who investigated other RGG-domain containing proteins [172–174]. How the RNA-binding selectivity of the RGG-domain functions is still unknown, but it was proposed that the frequency and order of glycine and arginine within the RGG determine the affinity and selectivity for specific RNAs [67]. This is achieved by the arrangement of arginines at specific locations within the RGG-domain by glycines [80].

It is likely that GRP7 shows a similar RNA-binding behavior like other RGG-domain containing proteins and has a degenerative RNA-binding selectivity as FUS and hnRNP since they all harbor an RGG-domain and an RRM. In Addition, they are involved in similar processes like alternative splicing [92,170,176]. To further investigate this possibility, the binding affinity of GRP7 towards highly structured RNAs, like G-Quadruplex RNAs, should be tested as well as other RNAs tested in the publication of Ozdilek *et. al* [167]. Considering that the undefined binding behavior of GRP7 most likely originates from its glycine-rich region, the function of the RGG-domain was further investigated in this thesis as well.

### 4.3 The function of the RGG-domain of GRP7

#### 4.3.1 Differences between AtGRP7 and BnGRP7 in RNA-binding affinity and RGG-domain

Considering AtGRP7 and BnGRP7 originate from two closely related plants, there is not much of a difference expected when investigating their biological function. Other publications showed, that the function of the glycine-rich proteins is conserved between species, like the chaperone function of GRP4 and GRP7 in *A. thaliana* and different GRPs from *Oryza sativa* [93]. However, there are differences in the amino acid sequence of the two GRP7 proteins. The RRM of AtGRP7 and BnGRP7 is highly conserved and only shows minimal differences but within the glycine-rich region (RGG-domain) are more differences (Figure 19). Some of the serines in the RGG-domain of AtGRP7 are exchanged by glycines in BnGRP7. Additionally, BnGRP7 has a shorter glycine-rich region than AtGRP7.

Due to the differences in the RGG-domain, the binding affinities of AtGRP7 and BnGRP7 were compared to investigate if those led to any significant variations in binding affinities. This comparison resulted in two major outcomes. First, the binding affinity of BnGRP7 and AtGRP7 is similar towards most of the shorter RNAs like *bna-miR164a* (Figure 20). For *miRnovel149*, BnGRP7 shows a significantly higher binding affinity compared to the binding affinity of AtGRP7. Second, the binding affinity of BnGRP7 towards longer RNAs is significantly lower compared to AtGRP7 in most cases (Figure 20). Only the  $K_d$  for one RNA (*BnaA07g32150D*) has not been significantly different.

The binding affinity differences between the two proteins is especially high for *BnPARCL* RNA (AtGRP7:  $0.0997 \mu\text{M} \pm 0.079 \mu\text{M}$  vs. BnGRP7:  $4.07067 \mu\text{M} \pm 0.80806 \mu\text{M}$ ) and *BnGRP7 UTR+Intron* RNA (AtGRP7:  $0.1275 \mu\text{M} \pm 0.04898 \mu\text{M}$  vs. BnGRP7:  $10.11667 \mu\text{M} \pm 3.43378 \mu\text{M}$ ). This difference probably does not originate from the amino acid changes within the RRM, since there are only a few and the exchanged amino acids have similar characteristics. Furthermore, the differences in the RGG-domain are more in number and the amino acids exchanged are not as similar to each other. In previous publications, the RGG-domain was reported to be involved in binding of RNA [89] as well as other functions of GRP7 like its chaperone function [151] and cell-to-cell movement [88]. This highlights the importance of the RGG-

## 4 Discussion

domain of GRP7. Moreover, the RRM does confer RNA-binding but it does not establish RNA-binding specificity [62] and thereby probably also not the binding affinity of the RNA-binding protein harboring the RRM. In fact, the RGG-domain was reported to have a high affinity towards RNA while the isolated RRM only showed a low affinity [167]. Even though the RGG-domain of hRNPs, FUS and FMRP contain the same number of RGG-repeats, the binding affinity of FUS towards the RNAs tested was lower compared to the affinity of the other two proteins, indicating that the length is not the only determining factor for binding affinity [167]. This is in line with the results of another publication. Exchanging the RGG-domain of Np13 with the RGG-domain of other proteins did not restore the full function of Np13, thus indicating the importance for the specific RGG-domain for the protein [177]. The length of the RGG-domain was not the main reason for the reduced function of chimeric Np13 but the sequence within the RGG-domain [177]. Therefore, AtGRP7 and BnGRP7 might show the difference in binding for the same reason. In BnGRP7, multiple glycines are missing compared to AtGRP7 RGG-domain. Glycines participate in intermolecular interaction by conferring conformational flexibility [81]. Additionally, pi-pi stack formation through their exposed backbone as well as the formation of hydrogen-bonds towards the nucleotide-bases and ribose contribute to RNA-binding [81,178]. Therefore, less glycine in the RGG-domain could lead to a decrease in binding affinity. Besides glycines, BnGRP7 is missing one tyrosine in its RGG-domain. Tyrosine is known to interact with RNA [81,82], and missing it could be another reason for the reduced binding affinity of BnGRP7 towards long RNAs. Also, as mentioned before, the selectivity and affinity of the RGG-domain could be due to the frequency and order of glycine and arginine within the domain [67,80]. Thus, it is likely that the slightly different sequence in the RGG-domain of AtGRP7 and BnGRP7 is responsible for the different binding affinities towards RNAs.

Although significant differences in the binding affinity of AtGRP7 and BnGRP7 were detected, it does not mean that their biological functions in RNA-binding are significantly different. Except for *BnPARCL* RNA and *BnGRP7 UTR+Intron* RNA, the binding affinities of both proteins are still quite similar when thinking about their biological function and not about statistics. For *BnGRP7* RNA, the  $K_d$  of AtGRP7 is  $2.97 \mu\text{M} \pm 0.81 \mu\text{M}$ , for *BnaC09g46650D* RNA  $1.26 \mu\text{M} \pm 0.33 \mu\text{M}$  and *BnaA07g32150D*  $2.72 \mu\text{M} \pm 1.26 \mu\text{M}$  (Figure 20). The  $K_d$  of BnGRP7 towards *BnGRP7* RNA is  $7.38 \mu\text{M} \pm 2.33 \mu\text{M}$ , for *BnaC09g46650D*  $7.96 \mu\text{M} \pm 1.37 \mu\text{M}$  and for *BnaA07g32150D*  $2.78 \pm 1.28 \mu\text{M}$  (Figure 20). All these  $K_d$  values are within a range of less than  $10 \mu\text{M}$ , thus even though the differences are significant, the binding affinities of AtGRP7 and BnGRP7 are still quite similar.

Another point to consider are the timepoints at which the MST-traces were analyzed. For BnGRP7, the analysis was done at 0.11 to 1.11 s after starting the heating of the capillaries (MST-on time) while the analysis for AtGRP7 was done at 6.93 to 7.93 s MST-on time. For some RNAs bound by AtGRP7 as well

## 4 Discussion

as BnGRP7 a biphasic binding behavior has been observed. Separate binding affinities of these two interactions can only be determined at later MST-on times. There are two phases, the T-jump and the thermophoresis, during an MST experiment which can result in different  $K_d$  values. It is likely that two different species of protein-RNA interaction occur, meaning that there might be two different binding stoichiometries like 1:1 and 2:1. At the T-jump, only a general interaction and not the number of different interactions can be recognized since the observed fluorescence change at this timepoint only depends on the surrounding of the fluorophore [148,179]. The surrounding of the fluorophore is changed by the binding of the interaction partners and therefore can be used for binding affinity calculation but cannot distinguish between the types of interaction [148]. During thermophoresis however, two different species can be observed since the fluorescence intensity change originates from different migration of these species [179]. It is most likely that BnGRP7 binds with a higher affinity towards some RNAs than observed at the MST-on time of 0.11-1.11, since this time is still in the T-jump region. The protein could interact in a second unspecific manner with the RNA. One example could be the attraction of the negative charges of the RNA by the positively charged protein. Thereby, the normal binding affinity is mixed with the binding affinity of the unspecific interaction, leading to a lower observed binding affinity than the real binding affinity. This phenomenon was observed for EcoSSB [179]. Hence, this effect shows how important the analysis of the  $K_d$  at different timepoints could be to determine the true  $K_d$ . Besides the different timepoints of analysis, MST measurements can show large errors of up to 25 % in between the same measurement just on different devices of NanoTemper and executed by different people even though the analysis is done in the same manner for all measurements [180]. Thus, the analysis timepoint and interpretation of MST-data should be considered when using these to determine binding affinities.

To conclude, the difference in the binding affinity of AtGRP7 and BnGRP7 towards longer RNAs exist *in vitro* but if this difference has an impact on their biological function is not known. At this point the importance of the environment of the proteins is also worth noting. In plants, the proteins are surrounded by cell content that could interact with GRP7 and thereby changing its interaction with certain RNAs. Their natural environment could level out some of the differences observed in the *in vitro* binding affinity of AtGRP7 and BnGRP7. Thus, the measured dissociation constants give a good estimation about the overall binding affinities and indicates a difference between the two proteins.

## 4 Discussion

### 4.3.2 Interaction of the RGG-domain with RNA upon RNA-binding

For a better understanding of the function of the RGG-domain of AtGRP7 and BnGRP7, both proteins were UV-crosslinked to *bnamiR164a*, since both proteins bind this RNA with a high affinity. After crosslinking, the samples were analyzed by SDS-PAGE. Crosslinked protein only and not crosslinked protein showed only one protein band and the typical degradation. An additional faint band above the protein band was visible in protein plus *bnamiR164* sample (Figure 21). This band and the crosslinked protein were trypsin digested and analyzed by MALDI-TOF MS. The peak analysis for AtGRP7 crosslinked and AtGRP7 plus *bnamiR164* crosslinked showed that ten peptides were not found in AtGRP7 plus *bnamiR164* crosslinked sample compared to AtGRP7 crosslinked. Peptides covering the RRM and the C-terminal region from the RGG-domain of AtGRP7 were missing (Figure 22). Since the RRM is a known RNA recognition motif that shows interactions with RNA [62], the peptides missing from the RRM region were probably interacting with *bnamiR164* and therefore could not be detected. Additionally, multiple peptides from the C-terminal region of the RGG-domain were not detected. This indicates that the RGG-domain is interacting with *bnamiR164* as well. Since the RGG-domain peptides were not missing in the AtGRP7 crosslinked only sample, it is not likely that these peptides are missing due to their interaction with other parts of the protein. Moreover, it was previously reported that the RGG-domain interacts with RNA [68] and the RNA-chaperone function reported for AtGRP7 is conveyed by its RGG-domain [181]. This also indicates the importance of the interaction between RGG-domain and RNA during RNA-binding.

The peak analysis of BnGRP7 crosslinked and BnGRP7 plus *bnamiR164* crosslinked revealed that only two peptides were missing in the BnGRP7 plus *bnamiR164* sample compared to BnGRP7 crosslinked (Figure 23). Peptide [16-28] is missing in RNP2 of RRM. This is, as described before for AtGRP7, probably due to the interaction of the RRM with RNAs [62]. The second peptide is the last peptide of the C-terminal RGG-domain. As for AtGRP7, this indicates an interaction of the glycine-rich C-terminus with *bnamiR164*. The interaction of RGG-domain with RNA is likely since it was shown for another closely related *B. napus* GRP (BnGRP1) which is involved in cold stress response and acts as RNA-chaperone during cold stress [182]. This indicates a similar mechanism for the chaperone activity observed for AtGRP7 relying on the RGG-domain, which thereby indicates the interaction of RGG-domain and RNA for BnGRPs as well.

The difference between AtGRP7 and BnGRP7 is, that BnGRP7 is only missing the last peptide of the RGG-domain while AtGRP7 is missing two peptides at the end of RGG-domain. This might be an indication why the binding affinity of AtGRP7 is different compared to the binding affinity of BnGRP7, since the further missing peptides of AtGRP7 contain the part of the RGG-domain which contains the largest deletion in BnGRP7.

## 4 Discussion

### 4.3.3 The glycine-rich region of AtGRP7 holds an important function in RNA-binding

To further investigate the function of the RGG-domain of AtGRP7, two mutants were prepared. AtGRP7mut, which contained seven amino acid exchanges from tyrosine (Y) to glutamic acid (E) within its RGG-domain, was used. Such a mutation is used to mimic phosphorylation as glutamic acid introduces a negative charge like the phosphorylation [76,183]. To investigate the impact of introducing negative charges into the RGG-domain, AtGRP7mut was used. For other LC-domain proteins, the introduction of negative charge had an impact on the formation of LLPS [183,184]. The second mutant, AtGRP7short, contains a preliminary stop codon, whereby only the RRM of AtGRP7 is translated and the glycine-rich region is missing. Both proteins were purified and further used in different assays.

As a first experiment, the RNA-binding affinity of the two mutant proteins was measured with MST. RNAs selected were the 3'UTR of GRP7 as a small RNA and *AtGRP7* RNA with UTR and Intron. AtGRP7mut was able to bind both of these RNAs but AtGRP7short was only able to bind the 3'UTR of AtGRP7 (Figure 30). The differences in binding affinity of AtGRP7 and AtGRP7mut were not significant for both RNAs, even though AtGRP7mut bound *AtGRP7* RNA with UTR and Intron with a  $K_d$  of  $30.11 \mu\text{M} \pm 24.5 \mu\text{M}$  while AtGRP7 shows a  $K_d$  of  $0.36 \mu\text{M} \pm 0.09 \mu\text{M}$  towards this RNA. This lack of significance, even though the binding affinities are far from each other is probably due to the large error for AtGRP7mut and could be avoided through more measurements. Still, the binding affinity of AtGRP7 and AtGRP7mut towards *AtGRP7 UTR+Intron* RNA indicates a trend to lower binding affinity of AtGRP7mut, just by the exchange of the tyrosines. Comparing the binding affinity of AtGRP7short towards the binding affinity of AtGRP7 shows a significant difference for the 3'UTR RNA and no binding of AtGRP7short towards *AtGRP7 UTR+Intron* RNA at all (Figure 30). The binding affinity of AtGRP7 towards a part of its own 3'UTR was measured in a previous publication by EMSA (electrophoretic mobility shift assay), revealing a  $K_d$  of  $0.409 \mu\text{M} \pm 0.096 \mu\text{M}$  [87], which is comparable to the  $K_d$  observed for AtGRP7 by MST with  $0.160 \mu\text{M} \pm 0.12 \mu\text{M}$ .

Due to the differences in binding affinity observed for only two RNAs, further MST measurements were conducted to investigate the influence of the RGG-domain. Therefore, small RNAs as well as larger RNAs were tested. AtGRP7short did not show any binding of larger RNAs, but was able to bind small RNAs (Figure 31). The binding affinity of AtGRP7short towards two (*bna-miR164a* and *bna-miRnovel149*) of the three tested short RNA was significantly lower than the binding affinity of AtGRP7 or AtGRP7mut. AtGRP7 and AtGRP7mut did not show significant differences in their binding affinity of two small RNAs (*bna-miR164a* and *bna-miRnovel149*) (Figure 31). In contrast, AtGRP7mut showed a

## 4 Discussion

significant higher  $K_d$  towards *BnGRP7+UTR* RNA and *BnPARCL* RNA. *BnaCNG53670D* RNA was not bound at all by AtGRP7mut.

These results indicate further importance of the glycine-rich region of GRP7 for RNA-binding. The removal of the RGG-domain of AtGRP7 led to a decrease in RNA-binding in a previous publication [89]. For AtGRP7mut, the amino acid exchange from Y to E is most likely the reason for the binding differences observed. One reason might be the negative charge introduced by glutamic acid, which could repel the negatively charged RNA. In addition, the pi-pi stacking ability of tyrosine and its hydrogen-bonds with nucleotide-bases takes an important part in intramolecular interactions [81]. The missing tyrosines probably reduced the ability of AtGRP7mut to interact with RNA and thereby the RNA-binding affinity. Another reason might be the difference in the amino acid sequences within the RGG-domain or also called RGG-domain. For Np13, a RGG-domain containing protein like AtGRP7, the exchange of its RGG-domain with the RGG-domain of other proteins lead to a lack in its original function [177]. This observation is considered to be caused by differences in the AA-sequence and not only a result of the difference in RGG-domain length. For the same reason, the AA-sequence difference, the binding affinity of AtGRP7mut might be lower than the binding affinity of AtGRP7 for long RNAs. Another hint at the RNA-binding of RGG-domain is shown by Ozdilek *et al.*, they observed a high affinity of RGG-domain towards RNA without being associated with a typical RNA-binding domain like the RRM [167]. They also mentioned the low affinity of the isolated RRM [167], supporting the low RNA-binding affinity of AtGRP7short towards small RNAs and the missing ability to bind longer RNAs with an affinity that could be measured. This leads to the conclusion that even small changes in the composition of the RGG-domain might be followed by a change in its functionality. Another reason for the lack of RNA-binding detected by AtGRP7short could be a specific function of the RGG-domain to unwind RNA. The glycine-rich domain of nucleolin is able to unwind RNA and thereby revealing the RNA sequence that can be bound by the rest of the protein [185]. The RGG-domain of AtGRP7 might harbor a similar function, especially since it is, like nucleolin [185], a protein localized in the nucleolus with its RGG-domain at the terminus of the protein [59,84].

In conclusion, the glycine-rich region of AtGRP7 does not only fulfill a supportive role in RNA-binding as suggested by Leder *et. al* [89], but an essential role in RNA-binding since no RNA-binding of AtGRP7short was observed for long RNAs and the RNA-binding affinity of AtGRP7mut was significantly lower towards long RNAs as well.

## 4 Discussion

4.3.3.1 Fibrillization and liquid-liquid phase separation of AtGRP7 caused by its glycine-rich C-terminus

Low-complexity (LC) domains like RGG-domains as well as the G/S Y G/S domains were reported to show liquid-liquid phase separation (LLPS) [73,81]. The formation of LLPS is driven by the multivalence of interaction domains [81,186]. Arginine and tyrosine are responsible for a high number of intra- and intermolecular interactions by pi-pi and pi-cation interactions as well as hydrogen bonding and electrostatic interaction [81,83]. The RGG-domain of AtGRP7 contains a mixture of two known LC domains containing tyrosine and arginine, therefore it is likely that AtGRP7 could form LLPS. Over a period of time, LLPS of other LC domain containing proteins can lead to liquid-to-solid transition due to high protein concentrations. Such transition was reported to result in the formation of aggregates and fibrils [98,187–189]. To investigate the fibrilization of AtGRP7, the protein was observed over time. The fluorescence dye ThioflavinT is known to accumulate in fibril-like aggregates and thereby increasing its fluorescence [147]. The change in fluorescence intensity of ThioflavinT was observed for AGRP7, AtGRP7mut and AtGRP7short with and without RNA over a period of 80 h. A higher fluorescence intensity of AtGRP7 than AtGRP7mut and AtGRP7short was detected, while AtGRP7mut showed a higher fluorescence intensity than AtGRP7short (Figure 26). This indicates that the RGG-domain of AtGRP7 is responsible for the formation of fibril-like aggregates of the protein and thereby increasing the ThioflavinT fluorescence. The fluorescence intensity of AtGRP7 with *bnamiRNA164* and *BnPARCL* RNA was slightly higher than of AtGRP7 alone. Through seeding, the RNA could increase the LLPS and fibrilization of AtGRP7 and thereby the fluorescence intensity. Seeding of protein fibrils through RNA was described before for FUS [190] and might be also a mechanism for AtGRP7. Another finding for FUS and hnRNP2 was the disruption of LLPS and fibrilization through phosphorylation of tyrosine within the RGG-domain [76,183]. AtGRP7mut is a mutant where tyrosine was exchanged to glutamic acid thereby introducing negative charges, which were shown to reduce LLPS formation in hnRNP2 when serine phosphorylation was mimicked by the exchange of serine with glutamic acid [183]. Thus, the tyrosine to glutamic acid mutation probably interrupts or slows down the formation of fibril-like aggregates like the serine mutation or tyrosine phosphorylation.

Since the fibrilization of LC-domain containing proteins starts with LLPS [98], LLPS of AtGRP7 and AtGRP7mut with RNA was studied as well. By the incubation of both proteins with labeled RNA and addition of PEG, LLPS could be visible through a microscope. LLPS was detected in samples with 1  $\mu$ M labeled RNA and 50  $\mu$ M AtGRP7 cooled on ice (Figure 27). This observation also explains the cloudy protein samples, when AtGRP7 was first purified at 4 °C. All other samples did not show any LLPS, including the RNA only sample which suggests that the LLPS observed in the cooled 50  $\mu$ M AtGRP7 sample was not induced by the RNA alone. The occurrence of LLPS only in cooled AtGRP7 samples could be reasonable due to its involvement in cold stress response [58,93]. It might be possible that by

## 4 Discussion

LLPS of AtGRP7 at colder temperatures, specific chaperones and RNAs can be brought into close proximity and thereby being able to preserve function of the RNAs during colder temperatures. A recent study showed that not only AtGRP7 but also two RNA chaperones were included in AtGRP7 LLPS [90], underlining this hypothesis. In addition, LLPS was shown to be induced by colder temperatures for other proteins as well like lysozyme from hen egg white [191]. In general, LLPS is suggested to be sensitive to environmental stresses [192]. Thus, AtGRP7 reacting on cold stress by LLPS is likely.

The results of AtGRP7 showing LLPS at 50  $\mu$ M together with 1  $\mu$ M RNA and low temperatures are consistent with the results of a recent study [90]. This study observed LLPS with 50  $\mu$ M AtGRP7 at colder temperatures but no LLPS when the tyrosine residues in the RGG-domain were phosphorylated [90]. This indicates, LLPS of AtGRP7 is most likely disrupted by phosphorylation of tyrosine or by tyrosine mutation in AtGRP7 as it is the case for FUS [76]. Pi-stacking of the aromatic residue of tyrosine [83] could not be conducted in mutant proteins nor when tyrosines within the low complexity domain of FUS were phosphorylated [76]. Besides the phosphorylation of tyrosine, methylation of arginine within the RGG-domain is common [81,193]. This posttranslational modification not only interferes with RNA-binding but also with LLPS [81]. Many proteins with RGG-domain capable of LLPS are known targets of this methylation [81,193]. The arginine methylation might be another mechanism to control LLPS besides the phosphorylation of tyrosine and it is an interesting modification to investigate further. FUS was reported to conduct less LLPS with methylated arginine within its RGG-domain, clearly hinting at a regulatory function of the methylation concerning LLPS [77]. The same effect occurs when hnRNP arginine within the RGG-domain was methylated [194]. Not only the LLPS of hnRNP is controlled by methylation of arginine within RGG-domain but also its cellular localization. The methylation of arginine within the RGG-domain was required for the translocation of hnRNP into the nucleus [195]. Since AtGRP7 is located in the nucleus as well as in the cytosol [59], the methylation of the RGG-domain of AtGRP7 might have similar effects not only on the LLPS but also on its localization and maybe even cell-to-cell movement. Thereby, it can be assumed that arginine methylation could serve as an important function and regulation for AtGRP7 and should be investigated, since nothing is known about such methylation of GRP7.

To emphasize the disordered structure of the RGG-domain from AtGRP7, SEC-SAXS measurements were conducted to analyze the overall structure of AtGRP7. The RRM-domain of AtGRP7 is known to be a highly structured and conserved domain [62] while the C-terminal RGG-domain is most likely disordered [68]. The dimensionless Kratky plot of BSA, a globular protein, showed a typical bell-shaped curve with a  $q \cdot R_g$  close to 1.75 and a  $I/I(0) \cdot sR_g^2$  of around 1.1 (Figure 28), which is common for globular proteins [156,196]. In contrast, AtGRP7 showed a partially bell-shaped curve with a shift of maximum towards the upper right side and an increased upwards slope of the curve on the right side

(Figure 28). This is typical for elongated, flexible proteins [196]. The larger the shift of maximum and the upwards slope is the higher is the degree of elongation and unfolded regions of the protein [157,158]. Thus, the dimensionless Krakty plot of AtGRP7 indicates a partially unfolded protein, which matches the two domains of AtGRP7: the structured RRM and the unstructured glycine-rich region.

As a final conclusion, the binding affinity of GRP7 was investigated for different RNAs and RNA modifications, resulting in the observation of GRP7 lacking a binding selectivity towards certain RNAs. A higher affinity of GRP7 towards short RNAs was observed, but more measurements should be considered to support this finding. The glycine-rich region of GRP7 plays a key-part in its RNA-binding capability and already small changes within the RGG-domain lead to a difference in its RNA-binding. Additionally, the formation of LLPS of AtGRP7 was shown and a link to the importance of tyrosine in this process was observed. This finding is consistent with results of a recent study [90] and studies of similar proteins like FUS and hnRNP2 [76,183,193]. In Addition, SEC-SAXS measurements confirmed that AtGRP7 is a partially unstructured protein with a disordered C-terminal domain.

### 4.4 The phloem sap of *B. napus* contains circRNAs

Besides RNA-binding proteins, many RNAs are found within the phloem sap from all different kind of RNA types like tRNAs, mRNAs, miRNAs and siRNAs [23,35,37]. One class of RNAs, circRNAs, has not been further investigated within the phloem so far and was only recently detected in phloem sap of apple trees [39]. To investigate if circRNAs are present in the phloem sap of *B. napus*, RNaseR digestion was performed with total phloem sap RNA. The negative control, which contained everything except for RNaseR, still showed the typical peaks of total phloem sap RNA while RNaseA digested sample did not contain any RNA. In RNaseR digested sample remained a peak for small RNAs, indicating that some RNAs have not been digested by RNaseR (Figure 32). These RNAs might be circular, since RNaseR is not able to digest circular RNA [108]. On the other hand, it should be mentioned that other RNAs like highly structured tRNAs or RNAs with G-quadruplex structures are also poorly digested by RNaseR [108]. Consequently, the remaining RNAs could be circular but also remaining linear RNAs. A possibility to decrease the amount of linear, highly structured RNA by RNaseR digestion is the use of lithium chloride instead of potassium chloride in the reaction buffer [197]. Lithium chloride disrupts the structure of G-quadruplex RNA structures [108] and thereby increasing their availability for RNaseR as a substrate.

A second approach to elucidate the circRNA content of *B. napus* phloem was the analysis of Illumina sequencing results originating from a lnc-sequence library of phloem sap RNA and RNA from leaves as a comparison. Therefore, the paired end reads as well as each read by itself were mapped to the Ensemble *B. napus* genome [149] with STAR [146]. Chimeric junction reads were then used by circTools

detect module (DCC) [137,150] to identify back-splicing junction reads of circular RNAs [106,198]. Within the phloem, 93 circRNAs were identified by DCC (Table 29 and Table 40) and 62 circRNAs were detected in leaves (Table 30 and Table 39). The number of identified circRNAs within total leaf RNA is quite low compared to the amount of circRNAs identified in other publications. For example, in *A. thaliana* 6012 circRNAs were identified [116] and in apple tree 912 circRNAs were identified in leaf and 1064 in phloem [39]. It is likely, that the sequencing depth of the lnc-sequencing of around 35 million reads was too low to detect more circRNAs, since circRNAs often have a low read count and thereby are difficult to detect [198]. Another reason might be the annotation of the *B. napus* genome. For *A. thaliana* and other organisms like human and mice, the annotation of the genome is very good. For organisms that are not as common, the genome annotation is often less complete. Since DCC uses annotated splicing sites, splicing sites that are not annotated would not be considered and novel circRNAs derived from unknown splicing sites are thus not identified [199]. Therefore, it is advised to use at least two or more different pipelines for the identification of circRNAs to find high confidence circRNAs which can be identified by more than one program [132,198,200]. Due to the fact that the identification of circular RNAs additionally results often in a huge amount of false positive circRNAs, Zsófia Fekete (Agricultural Biotechnology Center, Hungary) identified circRNAs in phloem lnc-sequencing data with her own pipeline. By this, over 22000 circRNAs (without threshold) were identified. From these, 21 circRNAs matched circRNAs identified with DCC. Besides mapping the sequencing data against the *B. napus* Ensemble genome, the reads were also mapped against a newer *B. napus* genome assembly (found on NCBI WGS project JAGKQM01, BioProject PRJNA627442). After comparing the circRNAs found in both mappings, eight circRNAs present in both were identified (Table 31). These circRNAs were thought of as important candidates, even though only some were identified by DCC.

To verify the eight circRNAs, a PCR with divergent primer were conducted. An amplicon could only be achieved if the RNAs are circular. circRNA1, circRNA2, circRNA4, circRNA5 and circRNA8 resulted in an amplicon (Figure 34). Through whole RNA amplification with divergent primer, also circRNA3 could be amplified (Figure 37). Hence, only circRNA7 could not be detected by RT-PCR and circRNA6 was not additionally detected since it appeared to be a circular RNA from the same gene with the same BSJ as circRNA4. To confirm that the amplicons matches the BSJ detected, the BSJ amplicons were sequenced by Sanger sequencing. The BSJ of circRNA1, circRNA2 and circRNA4 matched the results of Sanger sequencing while Sanger sequencing of circRNA5 showed a slightly different BSJ than predicted (Figure 36). It might be that multiple circRNAs are derived from the gene *BnaA01g28460D* with different BSJ and instead of amplifying the circRNA with the predicted BSJ, another BSJ was identified. It is common that multiple circRNA isoforms can be found for one gene [115]. It might also be that, since *B. napus* is

tetraploid [201], the BSJ from a circRNA of a homologue gene was amplified and therefore not matching the predicted BSJ.

Even though circRNAs were identified within the sequencing data and also verified by RT-PCR and Sanger sequencing, there might be still the chance that the amplicons of divergent primers originate from template switching of the reverse transcriptase. Template switching is a problem during reverse transcription and might not only lead to false BSJ within RT-PCR but also RNA-sequencing [202]. Thus, an RNaseR digestion prior to cDNA synthesis could help prevent this problem [199]. Another way to directly identify circRNAs without the use of reverse transcription is Northern blotting [199]. Both methods should be used to further verify the investigated circRNAs. Many circRNAs identified for plants were only identified by the analysis of Inc-sequencing or circRNA enriched sequencing data [129,203]. Hence, only a small number of plant circRNAs has been verified [204]. The function of an even smaller number of plant circRNAs is known. One example is a circRNA derived from *A. thaliana* *SEPALLATA3* exon six, which is involved in alternative splicing of *SEPALLATA3* mRNA [205]. Even though more than 70000 circRNAs were predicted for plants in total, the majority was never confirmed by further methods and full-length sequences of circRNAs are missing [204].

To gain further information about the high confidence circRNAs, the RNAs were amplified with divergent primer laying back-to-back to each other to amplify the whole circRNA. By this, multiple amplicons were identified for circRNA1, circRNA2, circRNA4 and circRNA5 but not for circRNA3 (Figure 37). Since circRNAs do not always contain introns as well as all exons and some circRNAs span over multiple exons, the size of circRNAs can be different from the expected size judged by the distance of the BSJ [206]. Therefore, it is important to identify the whole sequence of each circRNA that is further analyzed. This can be achieved by amplifying whole circRNAs and sequencing the amplicon by Sanger sequencing or by Nanopore sequencing [115]. In plants, circRNAs have been only analyzed by Sanger or Illumina sequencing. This might be the reason why some reviews claim that circRNAs in plants could have a size of up to 100 kb [132]. Most likely, circRNAs with BSJ over 100 kb apart from each other are either false positive circRNAs and thereby not existing or a large part is spliced out. For the tested circRNAs no full sequence was derived from Sanger sequencing due to the high number of amplicons. The high number of amplicons could not only originate from isoforms of circRNAs, but also from rolling circle amplification [206]. Since the RNA is circular, the reverse transcriptase could amplify the RNA not only once but multiple times. Hence, primers can bind at multiple sites of this cDNA strand, leading to different lengths of amplicons. For a higher chance of successful whole circRNA sequencing, Nanopore sequencing should be considered as it is a well-working method to identify full-length sequences [114,115].

## 4 Discussion

To further verify the predicted circRNAs from *B. napus* lnc-sequencing, total phloem RNA and total leaf RNA was sent to Novogene in biological duplicates for circRNA sequencing library preparation, Illumina sequencing and bioinformatic analysis. The enrichment of circRNAs by Novogene was achieved by RNaseR digestion to remove linear RNA. The analysis of Novogene showed six circRNAs in phloem and 16 in leaf samples (see Figure 38, Table 41 and Table 42). Both numbers were very small compared to the number of circRNAs within the lnc-sequencing library. By RNaseR digest, false-positive circRNAs are removed through the digestion of linear transcripts, thereby reducing the false-positive rate. Therefore, a lower amount of circRNAs detected was expected. Nevertheless, all circRNAs identified by Novogene were detected with CIRCexplorer2 De novo assembly. This tool does not rely on annotated splicing sites and thereby could be able to detect circRNAs from poorly annotated genomes. Unfortunately, CIRCexplorer2 was also the only program that was not able to detect the well-studied circRNA ciRS-7 when searching for de novo circRNAs, while other programs were able to do so [200]. This might be one reason for the low number of identified circRNAs in *B. napus*. To be sure the large difference of detected circRNAs between lnc-sequencing and circRNA sequencing was due to the program used for circRNA detection, the analysis was repeated by STAR alignment and DCC. This resulted in 80 leaf circRNAs and 15 circRNAs detected within the phloem (see Figure 39, Table 43 and Table 44). More circRNAs were found in leaf with circRNA sequencing compared to lnc-sequencing, which might have been caused by the enrichment of circRNAs through RNaseR digestion. Within phloem, less RNAs were found in circRNA sequencing, leading to the assumption that circRNAs identified within the lnc-sequencing but not in circRNA sequencing might be false positive circRNAs [200]. Nine of the 15 phloem circRNAs were also identified in lnc-sequencing, thereby making these nine RNAs highly promising circRNA candidates (Table 32). For further validation, a RT-PCR and northern blotting should be performed. Additionally, a combination of different circRNA detection pipelines would produce an even better distinction of highly promising circRNAs [200]. This is especially important for other plants where more than thousand circRNAs were identified like *O. sativa* and *A. thaliana* [116]. Compared to the small number of circRNAs identified in *B. napus* the question might arise how such a huge difference occur. It could be that the threshold used for DCC (at least two reads of each circRNA in each replicate) was higher than used by other publications, thereby leading to a smaller number of predicted circRNAs. This seems to be a reasonable explanation since they only found 21 circRNAs in rice being present in all replicates and 25 *A. thaliana* circRNAs in all replicates [116]. Additionally, 69 % of over 4000 predicted circRNAs in rice were found in only one replicate as well as 37 % of circRNAs in *A. thaliana* [116]. Thus, the high number of identified circRNAs, at least in this publication, are owed to the low threshold. Another reason for the low number of identified circRNAs especially in total leaf RNA might be their low read count [198]. Nevertheless, the identified number of circRNAs in other species is much lower for example in Barley [128]. Here, only 62 circRNAs specific

#### 4 Discussion

for transfer cells were identified which matches the identification of 90 circRNAs in phloem sap from Inc-sequencing. In contrast, over 1000 circRNAs were identified in apple phloem sap and in all analyzed tissues over 6000 [39]. Since only a very small fraction of these circRNAs were further analyzed, it might be that a large number of the identified circRNAs are false positives and again, a low threshold was selected [39].

In conclusion, identification of circRNAs from RNA sequencing data is error prone and should be done carefully. For an improved identification, multiple tools can be used and an additional circRNA enriched sequencing library could help to identify true circRNAs [198,199,207]. The verification of circRNAs by RT-PCR, northern blot or other methods is important to be sure that circRNAs really exist. Therefore, the identified phloem circRNAs from *B. napus* will need further verification.

## 5 Conclusion and Outlook

Two topics were investigated in this thesis. The first one aimed to understand how the RNA-binding of GRP7 functions and how it might select RNAs for long-distance transport. The second topic focused on circular RNAs in the phloem sap of *B. napus*.

In the first part of this thesis, the RNA-binding ability of two phloem-mobile glycine-rich RNA-binding proteins (AtGRP7 and BnGRP7) was analyzed. A wide range of different RNAs are bound by AtGRP7 and BnGRP7, short as well as long RNAs while no specific RNA motif could be identified. Additionally, the RNA-binding affinity of AtGRP7 and BnGRP7 was significantly higher towards small RNAs (sRNA) compared to long RNAs. This binding behavior was also observed for FUS [166]. However, only a small number of sRNAs was tested, hence testing of additional sRNAs would be beneficial to confirm this binding behavior. In further MST experiments, methylation of RNAs showed only small differences in binding affinity of AtGRP7 and BnGRP7 compared to unmodified RNAs. Since the methylation was not introduced in naturally occurring methylation sites of RNAs, the results might be influenced by the unnatural and random methylation. To further investigate this matter, RNAs with known methylation sites could be methylated with methyltransferases *in vitro*. Wrapping up these findings, it is likely that AtGRP7 and BnGRP7 harbor a degenerated RNA-binding selectivity as it is suggested for FUS and hnRNPs [167]. Both are RGG containing proteins like AtGRP7 and BnGRP7, with a largely unspecific binding behavior. To further proof this hypothesis, RNAs used to investigate the degenerated binding of other RGG-containing proteins could be used, especially RNAs with G-Quadruplex structures, since other RGG-containing proteins favor to bind these RNA structures [175,178].

The comparison of AtGRP7 and BnGRP7 binding affinity as well as further experiments investigating the function of the RGG-domain led to the conclusion that the RGG-domain is important for RNA-binding. Even small changes within the RGG-domain reduced the RNA-binding affinity of GRP7 towards longer RNAs significantly. This was found to be true for BnGRP7 and AtGRP7mut, both showing a higher  $K_d$  for long RNAs. The removal of the whole RGG-domain resulted in a highly reduced binding affinity for sRNAs and no binding of long RNAs could be measured. Additionally, the possible formation of LLPS in presents of RNA at cold temperatures was detected for AtGRP7 but not for its mutated version AtGRP7mut. This indicates that the tyrosines within the RGG-domain are important for LLPS of AtGRP7, which is consistent with the work of other researchers [90]. In addition, the SEC-SAXS measurement of AtGRP7 further underlined the flexibility and disordered nature of AtGRP7s RGG-domain.

To conclude this part, in the conducted measurements of the RNA-binding ability, GRP7 showed a preference for sRNAs, here further investigation would be useful to underline these findings. Further, its ability to bind highly structured RNAs could be subject for future measurements. Additional insights

of the function of GRP7s GRR in RNA-binding and LLPS could be provided by further modifications of this domain. Finally, fluorescently labeled AtGRP7 without RNA could also be tested for its LLPS behavior.

In the second part of this thesis the circular RNA content of phloem sap from *B. napus* was analyzed. Since circRNAs are more stable than linear RNAs [40], they would be perfect candidates in the long-distance transport. To investigate if circRNAs are common in phloem sap, sequencing results from lnc-seq as well as sequencing results from circRNA enriched sequencing were analyzed for circRNAs, resulting in some interesting candidates. The amplification over the back-splicing junction (BSJ) of five circular RNAs resulted in amplicons, suggesting that these five RNAs are in fact circular. Sanger sequencing of four of the amplicons further proofed their circularity. For this part it can be concluded that circRNAs are present within the phloem sap of plants. In further experiments the function and the mobility of these circular RNAs should be tested.

## 6 List of figures

<b>Figure 1: Transport of molecules through the phloem.....</b>	<b>3</b>
<b>Figure 2: Predicted protein structure of AtGRP7. ....</b>	<b>8</b>
<b>Figure 3: Biogenesis of circRNA.....</b>	<b>11</b>
<b>Figure 4: Microscale thermophoresis (MST).....</b>	<b>36</b>
<b>Figure 5: Command line input for STAR 2.5.2b-2 used for the mapping. ....</b>	<b>38</b>
<b>Figure 6: SDS-PAGE with purification steps of GRP7 and UV-absorbance in mAU of different fractions during the Nickel NTA and SEC run.....</b>	<b>41</b>
<b>Figure 7: Electropherograms of leaf RNA and RNA bound by immobilized AtGRP7. ....</b>	<b>43</b>
<b>Figure 8: Volcano plots of differentially expressed (enriched) RNAs between total leaf RNA and eluted RNAs from AtGRP7 column.....</b>	<b>44</b>
<b>Figure 9: SDS-PAGE of BnGRP7 purification steps.....</b>	<b>47</b>
<b>Figure 10: PCR-test for phloem purity.....</b>	<b>47</b>
<b>Figure 11: Electropherograms of phloem RNA and RNA bound by immobilized AtGRP7.....</b>	<b>49</b>
<b>Figure 12: Volcano Plots of differentially expressed RNAs between total phloem sap RNA and eluted RNAs from BnGRP7 column. ....</b>	<b>50</b>
<b>Figure 13: Venn diagram of Phloem RNA Input and enriched BnGRP7. ....</b>	<b>50</b>
<b>Figure 14: RNA-binding affinities of AtGRP7 towards RNAs of different length.....</b>	<b>52</b>
<b>Figure 15: RNA-binding affinity of BnGRP7 towards RNAs with different length.....</b>	<b>54</b>
<b>Figure 16: Affinity of AtGRP7 for single- and double-stranded RNAs.....</b>	<b>56</b>
<b>Figure 17: The content of cytosine methylation in total phloem RNA, flowthrough and elution three of the CnBr-Immobilized BnGRP7. ....</b>	<b>57</b>
<b>Figure 18: RNA-binding affinity of AtGRP7 and BnGRP7 for RNAs with and without methylation.....</b>	<b>58</b>
<b>Figure 19: Protein sequence alignment of AtGRP7 and BnGRP7. ....</b>	<b>60</b>
<b>Figure 20: Binding affinity of AtGRP7 and BnGRP7 towards different RNAs. ....</b>	<b>61</b>
<b>Figure 21: SDS-PAGE of crosslinked BnGRP7 and AtGRP7 with miRNA164.....</b>	<b>62</b>
<b>Figure 22: UV-crosslinking of AtGRP7 with and without miR164. ....</b>	<b>63</b>
<b>Figure 23: UV-crosslinking of BnGRP7 with and without miR164.....</b>	<b>64</b>
<b>Figure 24: SDS-PAGE with AtGRP7short purification steps.....</b>	<b>65</b>
<b>Figure 25: SDS-PAGE with AtGRP7mut purification steps.....</b>	<b>65</b>
<b>Figure 26: Aggregation of AtGRP7, AtGRP7mut and AtGRP7short visualized by ThioflavinT fluorescence. ....</b>	<b>66</b>
<b>Figure 27: Detection of liquid-liquid phase separation through a microscope. ....</b>	<b>68</b>
<b>Figure 28: Dimensionless Krakty plot of AtGRP7 (green) and BSA (black). ....</b>	<b>69</b>
<b>Figure 29: Superimposed model of AtGRP7.....</b>	<b>70</b>
<b>Figure 30: Binding affinity of AtGRP7, AtGRP7mut and AtGRP7short against the 3'UTR of AtGRP7 and AtGRP7 pre-mRNA. ....</b>	<b>71</b>
<b>Figure 31: Binding affinity of AtGRP7, AtGRP7mut and AtGRP7short towards different RNAs. ....</b>	<b>72</b>

6 List of figures

<b>Figure 32: Electropherograms of total Phloem RNA and RNase treated Phloem RNA.</b> .....	74
<b>Figure 33: Venn diagram of leaf circRNAs and phloem circRNAs.</b> .....	76
<b>Figure 34: Agarose gel with PCR amplicons originating from divergent and convergent primer pairs.</b> .....	78
<b>Figure 35: Agarose gel with divergent and convergent PCR-amplicons of different SRP-RNAs.</b> .....	79
<b>Figure 36: Pairwise sequence alignment of BSJ amplicons with potential circRNA sequence.</b> .....	80
<b>Figure 37: Agarose gel with total circRNA amplicons.</b> .....	81
<b>Figure 38: Venn diagram of circRNAs in leaf and phloem sap.</b> .....	82
<b>Figure 39: Venn diagram of circRNAs in leaf and phloem in RNaseR treated sequencing library identified with the detect module (DCC) from circTools.</b> .....	82
<b>Figure 40: pet28a+ and pUC57 cloning and expression plasmids.</b> .....	126
<b>Figure 41: Modified pet28a+ plasmids suitable for Golden Gate cloning via BsaI.</b> .....	127
<b>Figure 42: Expression plasmids (pet28a+ and pet28a+GGThrombin) containing AtGRP7 and AtGRP7mut. .</b> ..	127
<b>Figure 43: Expression plasmid pet28a+ GGTEV containing BnGRP7.</b> .....	128
<b>Figure 44: Solubility of AtGRP7 with and without 1.6 hexandiol during lysis.</b> .....	128
<b>Figure 45: Sequence coverage of AtGRP7 and its degradation products.</b> .....	129
<b>Figure 46: Electropherograms of CnBr-sepharose beads negative control.</b> .....	130
<b>Figure 47: Electropherograms of CnBr-sepharose beads with AtGRP7short.</b> .....	131
<b>Figure 48: AtGRP7 - miRNA164 MST-traces and binding curve.</b> .....	142
<b>Figure 49: AtGRP7 - dsmiRNA164 MST-traces and binding curve.</b> .....	142
<b>Figure 50: AtGRP7 - 21R- MST-traces and binding curve.</b> .....	143
<b>Figure 51: AtGRP7 - tRNAm<sup>et</sup> MST-traces and binding curve.</b> .....	143
<b>Figure 52: AtGRP7 - GRP7 3'UTR MST-traces and binding curve.</b> .....	143
<b>Figure 53: AtGRP7 - BnmiRnovel2 MST-traces and binding curve.</b> .....	144
<b>Figure 54: AtGRP7 - BnmiRnovel106 MST-traces and binding curve.</b> .....	144
<b>Figure 55: AtGRP7 - BnmiRnovel149 MST-traces and binding curve.</b> .....	144
<b>Figure 56: AtGRP7 - miR164 m6A MST-traces and binding curve.</b> .....	144
<b>Figure 57: AtGRP7 - miR164 m5C MST-traces and binding curve.</b> .....	145
<b>Figure 58: AtGRP7 - GFP-RNA MST-traces and binding curve.</b> .....	145
<b>Figure 59: AtGRP7 - GFP+GRP7UTR-RNA MST-traces and binding curve.</b> .....	145
<b>Figure 60: AtGRP7 - GFP-TCTP-RNA MST-traces and binding curve.</b> .....	146
<b>Figure 61: AtGRP7 – GFP-TCTP-GRP7-UTR-RNA MST-traces and binding curve.</b> .....	146
<b>Figure 62: AtGRP7 – BnaCNG53670D_w/o GA MST-traces and binding curve.</b> .....	146
<b>Figure 63: AtGRP7 – BnPARCL MS- traces and binding curve.</b> .....	146
<b>Figure 64: AtGRP7 – BnaCNG53670D MST-traces and binding curve.</b> .....	147
<b>Figure 65: AtGRP7 – BnPARCL_withGAmotif MST-traces and binding curve.</b> .....	147
<b>Figure 66: AtGRP7 – BnaC03g26040D MST-traces and binding curve.</b> .....	147
<b>Figure 67: AtGRP7 - BnaCNG53670D_120ntUTR MST-traces and binding curve.</b> .....	147
<b>Figure 68: AtGRP7 - BnaA03g14400D MST-traces and binding curve.</b> .....	148

6 List of figures

<i>Figure 69: AtGRP7 - BnaC09g46650D MST-traces and binding curve.</i>	148
<i>Figure 70: AtGRP7 - BnaA07g32150D MST-traces and binding curve.</i>	148
<i>Figure 71: AtGRP7 - BnGRP7 MST-traces and binding curve.</i>	148
<i>Figure 72: AtGRP7 - BnGRP7+UTR RNA MST-traces and binding curve.</i>	149
<i>Figure 73: AtGRP7 – BnGRP7+UTR+Intron RNA MST-traces and binding curve.</i>	149
<i>Figure 74: AtGRP7 - dsBnPARCL MST-traces and binding curve.</i>	149
<i>Figure 75: AtGRP7 - AtGRP7-RNA MST-traces and binding curve.</i>	149
<i>Figure 76: AtGRP7 – AtGRP7_ UTR+Intron MST-traces and binding curve.</i>	150
<i>Figure 77: AtGRP7 - GFP-TCTP-RNA m5C MST-traces and binding curve</i>	150
<i>Figure 78: AtGRP7 - miR164 m5C MST-traces and binding curve.</i>	150
<i>Figure 79: AtGRP7 - miR164 m6A MST-traces and binding curve.</i>	151
<i>Figure 80: AtGRP7 - BnaCNNG53670D m6A MST-traces and binding curve.</i>	151
<i>Figure 81: AtGRP7 -BnaCNNG53670D m6A without GA-motif MST-traces and binding curve.</i>	151
<i>Figure 82: AtGRP7 - BnaCNNG53670D m5C MST-traces and binding curve.</i>	152
<i>Figure 83: AtGRP7 - BnPARCL m6A MST-traces and binding curve.</i>	152
<i>Figure 84: AtGRP7 - BnPARCL m5C MST-traces and binding curve.</i>	152
<i>Figure 85: AtGRP7 - BnaC03g26040D m5C MST-traces and binding curve.</i>	152
<i>Figure 86: AtGRP7 - BnaC03g26040D m6A MST-traces and binding curve.</i>	153
<i>Figure 87: AtGRP7short - BnPARCL RNA MST-traces and binding curve.</i>	153
<i>Figure 88: AtGRP7 short – tRNAm<sup>et</sup> MST-traces and binding curve.</i>	153
<i>Figure 89: AtGRP7short - miR164 MST-traces and binding curve.</i>	154
<i>Figure 90: AtGRP7short - BnaCNNG53670D MST-traces and binding curve.</i>	154
<i>Figure 91: AtGRP7short - BnmiRnovel2 MST-traces and binding curve.</i>	154
<i>Figure 92: AtGRP7short - BnGRP7 RNA MST-traces and binding curve.</i>	154
<i>Figure 93: AtGRP7short - BnGRP7+UTR RNA MST-traces and binding curve.</i>	155
<i>Figure 94: AtGRP7short - BnmiRnovel106 MST-traces and binding curve.</i>	155
<i>Figure 95: AtGRP7short - GRP7 3'UTR MST-traces and binding curve.</i>	155
<i>Figure 96: AtGRP7short - AtGRP7 RNA MST-traces and binding curve.</i>	155
<i>Figure 97: AtGRP7short – AtGRP7+UTR+Intron MST-traces and binding curve.</i>	156
<i>Figure 98: AtGRP7mut – BnPARCL MST-traces and binding curve.</i>	156
<i>Figure 99: AtGRP7mut - miR164 MST-traces and binding curve.</i>	156
<i>Figure 100: AtGRP7mut - BnaCNNG53670D MST-traces and binding curve.</i>	157
<i>Figure 101: AtGRP7mut - BnmiRnovel2 MST-traces and binding curve.</i>	157
<i>Figure 102: AtGRP7mut - BnGRP7 RNA MST-traces and binding curve.</i>	157
<i>Figure 103: AtGRP7mut - BnGRP7+UTR RNA MST-traces and binding curve.</i>	157
<i>Figure 104: AtGRP7mut - BnmiRnovel106 MST-traces and binding curve.</i>	158
<i>Figure 105: AtGRP7mut - GRP7 3'UTR MST-traces and binding curve.</i>	158
<i>Figure 106: AtGRP7mut - AtGRP7 RNA MST-traces and binding curve.</i>	158

6 List of figures

<b>Figure 107: AtGRP7mut - AtGRP7+UTR+Intron RNA MST-traces and binding curve.</b> .....	158
<b>Figure 108: BnGRP7 - BnaC05g32740D MST-traces and binding curve.</b> .....	159
<b>Figure 109: BnGRP7 - BnaC09g46650D MST-traces and binding curve.</b> .....	159
<b>Figure 110: BnGRP7 - BnaA07g32150D MST-traces and binding curve.</b> .....	159
<b>Figure 111: BnGRP7 - BnaA03g14400D MST-traces and binding curve.</b> .....	159
<b>Figure 112: BnGRP7 - BnaA09g50830D MST-traces and binding curve.</b> .....	160
<b>Figure 113: BnGRP7 - BnaC09g45930D MST-traces and binding curve.</b> .....	160
<b>Figure 114: BnGRP7 - BnaCNNG53670D MST-traces and binding curve.</b> .....	160
<b>Figure 115: BnGRP7 - BnaC03g26040D MST-traces and binding curve.</b> .....	160
<b>Figure 116: BnGRP7 - BnaC03g26040D m6A MST-traces and binding curve.</b> .....	161
<b>Figure 117: BnGRP7 - BnaC03g26040D m5C MST-traces and binding curve.</b> .....	161
<b>Figure 118: BnGRP7 – BnPARCL MST-traces and binding curve.</b> .....	161
<b>Figure 119: BnGRP7 - BnaCNNG53670D m6A MST-traces and binding curve.</b> .....	161
<b>Figure 120: BnGRP7 - BnaCNNG53670D m5C MST-traces and binding curve.</b> .....	162
<b>Figure 121: BnGRP7 – dsBnPARCL MST-traces and binding curve.</b> .....	162
<b>Figure 122: BnGRP7 - miR164 MST-traces and binding curve.</b> .....	162
<b>Figure 123: BnGRP7 - GRP7 3'UTR MST-traces and binding curve.</b> .....	162
<b>Figure 124: BnGRP7 - BnmiRnovel106 MST-traces and binding curve.</b> .....	163
<b>Figure 125: BnGRP7 - BnmiRnovel149 MST-traces and binding curve.</b> .....	163
<b>Figure 126: BnGRP7 - BnmiRnovel2 MST-traces and binding curve.</b> .....	163
<b>Figure 127: BnGRP7 - miR164 m5C MST-traces and binding curve.</b> .....	163
<b>Figure 128: BnGRP7 - miR164 m6A MST-traces and binding curve.</b> .....	164
<b>Figure 129: BnGRP7 - BnPARCL m6A MST-traces and binding curve.</b> .....	164
<b>Figure 130: BnGRP7 - BnPARCL m5C MST-traces and binding curve.</b> .....	164
<b>Figure 131: BnGRP7 - BnGRP7 RNA MST-traces and binding curve.</b> .....	164
<b>Figure 132: BnGRP7 - BnGRP7+UTR MST-traces and binding curve.</b> .....	165
<b>Figure 133: BnGRP7 - BnGRP7+UTR+Intron RNA MST-traces and binding curve.</b> .....	165
<b>Figure 134: BnGRP7 - AtGRP7+UTR+Intron RNA MST-traces and binding curve.</b> .....	165
<b>Figure 135: BnGRP7 + AtGRP7 RNA MST-traces and binding curve.</b> .....	165
<b>Figure 136: Plots of SEC-SAXS measurements.</b> .....	168

## 7 List of tables

<b>Table 1: Machines/equipment used in this thesis.</b> .....	14
<b>Table 2: Consumables used in this work.</b> .....	15
<b>Table 3: Enzymes used in this work.</b> .....	16
<b>Table 4: Different media used for bacterial cultivation.</b> .....	16
<b>Table 5: Different additives for the media used for bacterial cultivation.</b> .....	17
<b>Table 6: General buffers used in during this work.</b> .....	17
<b>Table 7: Buffers used for GRP7 protein purification.</b> .....	18
<b>Table 8: Buffers used for AtGRP7short protein purification.</b> .....	19
<b>Table 9: Overview of plasmids used in this work.</b> .....	19
<b>Table 10: Bacterial strains and their genotype and antibiotic resistance</b> .....	19
<b>Table 11: Plants used in this work.</b> .....	20
<b>Table 12: Genes that have been used for protein expression and purification.</b> .....	20
<b>Table 13: Software used in this thesis.</b> .....	21
<b>Table 14: Reverse transcription reaction mix.</b> .....	25
<b>Table 15: PCR reaction mix.</b> .....	25
<b>Table 16: Restriction digest reaction mix.</b> .....	26
<b>Table 17: Ligation reaction mix.</b> .....	27
<b>Table 18: Smal-cloning reaction mix.</b> .....	27
<b>Table 19: Thermocycler-program for Golden Gate cloning.</b> .....	28
<b>Table 20: In vitro transcription reaction mix.</b> .....	29
<b>Table 21: RNaseR digestion reaction mix.</b> .....	30
<b>Table 22: Sample overview for SEC-SAXS measurement.</b> .....	35
<b>Table 23: SEC-SAXS data collection details.</b> .....	35
<b>Table 24: AtGRP7-enriched transcript RNAs from <i>B. napus</i> leaf RNA analyzed by Novogene.</b> .....	45
<b>Table 25: AtGRP7-enriched transcript RNAs from <i>B. napus</i> leaf RNA analyzed by Huge Wolfeenden (John-Innes-Center, Norwich, UK).</b> .....	46
<b>Table 26: Overview of some BnGRP7 enriched phloem sap RNAs.</b> .....	51
<b>Table 27: AtGRP7 dissociation constant (<math>K_d</math>) towards different RNAs.</b> .....	53
<b>Table 28: BnGRP7 dissociation constant (<math>K_d</math>) towards different RNAs.</b> .....	55
<b>Table 29: List of circRNAs from phloem sap with the highest read counts (RC).</b> .....	75
<b>Table 30: List of circRNAs from leaves with the highest read counts (RC).</b> .....	76
<b>Table 31: Overview of highly promising circular RNAs in phloem.</b> .....	77
<b>Table 32: Phloem circRNAs detected in lncRNA and circRNA sequencing.</b> .....	83
<b>Table 33: Oligonucleotides used in this work.</b> .....	124
<b>Table 34: List with all significantly enriched RNAs by AtGRP7 CnBr column from <i>B. napus</i> leaf RNA compared to total leaf RNA.</b> .....	132

<b>Table 35: List of all through BnGRP7 enriched phloem mRNAs of <i>B. napus</i> compared to total phloem sap RNA of <i>B. napus</i>.</b> .....	133
<b>Table 36: Overview of dissociation constants (<math>K_d</math>) measured for AtGRP7mut for different RNAs.</b> .....	166
<b>Table 37: Overview of dissociation constants (<math>K_d</math>) measured for AtGRP7short for different RNAs.</b> .....	166
<b>Table 38: Structural parameters of SEC-SAXS measurements of AtGRP7.</b> .....	167
<b>Table 39: Predicted circRNAs in Inc-sequencing of <i>B. napus</i> leaf RNA.</b> .....	169
<b>Table 40: Predicted circRNAs in Inc-sequencing of <i>B. napus</i> phloem RNA.</b> .....	170
<b>Table 41: Novogene-predicted circRNAs in total leaf RNA from circRNA enriched sequencing.</b> .....	173
<b>Table 42: Novogene-predicted circRNAs in total phloem sap RNA from enriched circRNA sequencing.</b> .....	173
<b>Table 43: DCC-predicted circRNAs from leaf circRNA enriched sequencing</b> .....	174
<b>Table 44: DCC-predicted circRNAs from phloem circRNA enriched sequencing.</b> .....	176

## 8 References

1. Lucas, W. J., Groover, A., Lichtenberger, R., Furuta, K., Yadav, S. R., Helariutta, Y., He, X. Q., Fukuda, H., Kang, J., Brady, S. M., Patrick, J. W., Sperry, J., Yoshida, A., López-Millán, A. F., Grusak, M. A. & Kachroo, P. The Plant Vascular System: Evolution, Development and Functions. *Journal of Integrative Plant Biology* vol. 55 294–388 (2013).
2. Hall, D. W. & Stern, W. *Plant Anatomy. Forensic Botany: A Practical Guide* (2012). doi:10.1002/9781119945734.ch7.
3. Banerjee, A. K., Chatterjee, M., Yu, Y., Suh, S.-G., Miller, W. A. & Hannapel, D. J. Dynamics of a Mobile RNA of Potato Involved in a Long-Distance Signaling Pathway. *Plant Cell* **18**, 3443–3457 (2007).
4. Xu, H., Iwashiro, R., Li, T. & Harada, T. Long-distance transport of Gibberellic Acid Insensitive mRNA in *Nicotiana benthamiana*. *BMC Plant Biol.* **13**, (2013).
5. Bollhöner, B., Prestele, J. & Tuominen, H. Xylem cell death: Emerging understanding of regulation and function. *J. Exp. Bot.* **63**, 1081–1094 (2012).
6. Esau, K. Development and structure of the phloem tissue. *Bot. Rev.* **5**, 373–432 (1939).
7. Cronshaw, J. Phloem structure and function. *Annu. Rev. Plant Physiol.* **32**, 465–484 (1981).
8. Esau, K. & Thorsch, J. Sieve Plate Pores and Plasmodesmata, the Communication Channels of the Symplast: Ultrastructural Aspects and Developmental Relations. *Am. J. Bot.* **72**, 1641–1653 (1985).
9. Goodwin, P. B. Molecular size limit for movement in the symplast of the *Elodea* leaf. *Planta* **157**, 124–130 (1983).
10. Stadler, R., Wright, K. M., Lauterbach, C., Amon, G., Gahrtz, M., Feuerstein, A., Oparka, K. J. & Sauer, N. Expression of GFP-fusions in *Arabidopsis* companion cells reveals non-specific protein trafficking into sieve elements and identifies a novel post-phloem domain in roots. *Plant J.* **41**, 319–331 (2005).
11. Oparka, K. J. & Turgeon, R. Sieve elements and companion cells - Traffic control centers of the phloem. *Plant Cell* **11**, 739–750 (1999).
12. Turgeon, R. Phloem loading and plasmodesmata. *Trends Plant Sci.* **1**, 418–423 (1996).
13. Rennie, E. A. & Turgeon, R. A comprehensive picture of phloem loading strategies. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14162–14167 (2009).
14. Kühn, C., Hajirezaei, M. R., Fernie, A. R., Roessner-Tunali, U., Czechowski, T., Hirner, B. & Frommer, W. B. The sucrose transporter StSUT1 localizes to sieve elements in potato tuber phloem and influences tuber physiology and development. *Plant Physiol.* **131**, 102–113 (2003).
15. Turgeon, R. & Medville, R. The absence of phloem loading in willow leaves. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12055–12060 (1998).
16. Turgeon, R. & Hepler, P. K. Symplastic continuity between mesophyll and companion cells in minor veins of mature *Cucurbita pepo* L. leaves. *Planta* **179**, 24–31 (1989).
17. Haritatos, E., Keller, F. & Turgeon, R. Raffinose oligosaccharide concentrations measured in individual cell and tissue types in *Cucumis melo* L. leaves: Implications for phloem loading. *Planta* **198**, 614–622 (1996).
18. Kehr, J., Morris, R. J. & Kragler, F. Long-Distance Transported RNAs: From Identity to Function. *Annu. Rev. Plant Biol.* **73**, (2022).

## 8 References

19. Hall, S. M. & Baker, D. A. The chemical composition of Ricinus phloem exudate. *Planta* **106**, 131–140 (1972).
20. Weiler, E. W. & Ziegler, H. Determination of phytohormones in phloem exudate from tree species by radioimmunoassay. *Planta* 168–170 (1981).
21. Giavalisco, P., Kapitza, K., Kolasa, A., Buhtz, A. & Kehr, J. Towards the proteome of Brassica napus phloem sap. *Proteomics* **6**, 896–909 (2006).
22. Lin, M. K., Lee, Y. J., Lough, T. J., Phimmey, B. S. & Lucas, W. J. Analysis of the pumpkin phloem proteome provides insights into angiosperm sieve tube function. *Mol. Cell. Proteomics* **8**, 343–356 (2009).
23. Buhtz, A., Springer, F., Chappell, L., Baulcombe, D. C. & Kehr, J. Identification and characterization of small RNAs from the phloem of Brassica napus. *Plant J.* **53**, 739–749 (2008).
24. Hannapel, D. J., Sharma, P. & Lin, T. Phloem-mobile messengerRNAs and root development. *Front. Plant Sci.* **4**, 1–12 (2013).
25. Zhang, S., Sun, L. & Kragler, F. The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant Physiol.* **150**, 378–387 (2009).
26. Kim, J. Y., Yuan, Z., Cilia, M., Khalfan-Jagani, Z. & Jackson, D. Intercellular trafficking of a KNOTTED1 green fluorescent protein fusion in the leaf and shoot meristem of Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 4103–4108 (2002).
27. Sinha, N. R., Williams, R. E. & Hake, S. Overexpression of the maize homeo box gene, KNOTTED-1, causes a switch from determinate to indeterminate cell fates. *Genes Dev.* **7**, 787–795 (1993).
28. David-Schwartz, R., Runo, S., Townsley, B., Machuka, J. & Sinha, N. Long-distance transport of mRNA via parenchyma cells and phloem across the host–parasite junction in *Cuscuta*. *New Phytol.* **179**, 1133–1141 (2008).
29. Haupt, S., Oparka, K. J., Sauer, N. & Neumann, S. Macromolecular trafficking between Nicotiana tabacum and the holoparasite Cuscuta reflexa. *J. Exp. Bot.* **52**, 173–177 (2001).
30. Thieme, C. J., Rojas-Triana, M., Stecyk, E., Schudoma, C., Zhang, W., Yang, L., Minãmbres, M., Walther, D., Schulze, W. X., Paz-Ares, J., Scheible, W. R. & Kragler, F. Endogenous Arabidopsis messenger RNAs transported to distant tissues. *Nat. Plants* **1**, 1–9 (2015).
31. Kennedy, J. S. & Mittler, T. E. A Method of obtaining Phloem Sap via the Mouth-parts of Aphids. *Nature* **171**, 528 (1953).
32. Fisher, D. B., Wu, Y. & Ku, M. S. B. Turnover of soluble proteins in the wheat sieve tube. *Plant Physiol.* **100**, 1433–1441 (1992).
33. Newbury, H. J., Couldridge, C. E., Bale, J. S. & Pritchard, J. A phloem-enriched cDNA library from Ricinus : insights into phloem function. **57**, 3183–3193 (2006).
34. Omid, A., Keilin, T., Glass, A., Leshkowitz, D. & Wolf, S. Characterization of phloem-sap transcription profile in melon plants. *J. Exp. Bot.* **58**, 3645–3656 (2007).
35. Sasaki, T., Chino, M., Hayashi, H. & Fujiwara, T. Detection of several mRNA species in rice phloem sap. *Plant Cell Physiol.* **39**, 895–897 (1998).
36. Haywood, V., Yu, T. S., Huang, N. C. & Lucas, W. J. Phloem long-distance trafficking of GIBBERELLIC ACID-INSENSITIVE RNA regulates leaf development. *Plant J.* **42**, 49–68 (2005).

## 8 References

37. Yoo, B., Kragler, F., Varkonyi-gasic, E., Haywood, V., Archer-evans, S., Lee, Y. M., Lough, T. J. & Lucas, W. J. A Systemic Small RNA Signaling System in Plants. *Plant Cell* **16**, 1979–2000 (2004).
38. Pant, B. D., Buhtz, A., Kehr, J. & Scheible, W. R. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *Plant J.* **53**, 731–738 (2008).
39. Wang, D., Gao, Y., Sun, S., Li, L. & Wang, K. Expression characteristics in roots, phloem, leaves, flowers and Fruits of Apple circRNA. *Genes (Basel)*. **13**, 1–12 (2022).
40. Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J. & Kleinschmidt, A. K. Viroids are single stranded covalently closed circular RNA molecules existing as highly base paired rod like structures. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 3852–3856 (1976).
41. Qi, Y., Pelissier, T., Itaya, A., Hunt, E., Wassenegger, M. & Ding, B. Direct role of a viroid RNA motif in mediating directional RNA trafficking accross a specific cellular boundary. *Plant Cell* **16**, 1741–1752 (2004).
42. Gomez, G. & Pallas, V. A Long-Distance Translocatable Phloem Protein from Cucumber Forms a Ribonucleoprotein Complex In Vivo with Hop Stunt Viroid RNA . *J. Virol.* **78**, 10104–10110 (2004).
43. Zhang, W., Thieme, C. J., Kollwig, G., Apelt, F., Yang, L., Winter, N., Andresen, N., Walther, D. & Kragler, F. TRNA-related sequences trigger systemic mRNA transport in plants. *Plant Cell* **28**, 1237–1249 (2016).
44. Yang, L., Perrera, V., Saplaoura, E., Apelt, F., Bahin, M., Kramdi, A., Olas, J., Mueller-Roeber, B., Sokolowska, E., Zhang, W., Li, R., Pitzalis, N., Heinlein, M., Zhang, S., Genovesio, A., Colot, V. & Kragler, F. m5C Methylation Guides Systemic Transport of Messenger RNA over Graft Junctions in Plants. *Curr. Biol.* **29**, 2465-2476.e5 (2019).
45. Haebel, S. & Kehr, J. Matrix-assisted laser desorption/ionization time of flight mass spectrometry peptide mass fingerprints and post source decay: A tool for the identification and analysis of phloem proteins from Cucurbita maxima Duch. Separated by two-dimensional polyacryla. *Planta* **213**, 586–593 (2001).
46. Aki, T., Shigyo, M., Nakano, R., Yoneyama, T. & Yanagisawa, S. Nano Scale Proteomics Revealed the Presence of Regulatory Proteins Including Three FT-Like proteins in Phloem and Xylem Saps from Rice. *Plant Cell Physiol.* **49**, 767–790 (2008).
47. Balachandran, S., Xiang, Y., Schobert, C., Thompson, G. A. & Lucas, W. J. Phloem sap proteins from Cucurbita maxima and Ricinus communis have the capacity to traffic cell to cell through plasmodesmata. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 14150–14155 (1997).
48. Aoki, K., Kragler, F., Xoconostle-Cazares, B. & Lucas, W. J. A subclass of plant heat shock cognate 70 chaperones carries a motif that facilitates trafficking through plasmodesmata. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16342–16347 (2002).
49. Ostendorp, A., Pahlow, S., Krußel, L., Hanhart, P., Garbe, M. Y., Deke, J., Giavalisco, P. & Kehr, J. Functional analysis of Brassica napus phloem protein and ribonucleoprotein complexes. *New Phytol.* **214**, 1188–1197 (2017).
50. Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C. & Coupland, G. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science (80- )*. **316**, 1030–1033 (2007).
51. Gomez, G., Torres, H. & Pallas, V. Identification of translocatable RNA-binding phloem proteins from melon, potential components of the long-distance RNA transport system. *Plant J.* **41**, 107–116 (2005).
52. Barnes, A., Bale, J., Constantinidou, C., Ashton, P., Jones, A. & Pritchard, J. Determining protein identity from sieve element sap in Ricinus communis L. by quadrupole time of flight (Q-TOF) mass spectrometry.

## 8 References

- J. Exp. Bot.* **55**, 1473–1481 (2004).
53. Xoconostle-Cázares, B., Xiang, Y., Ruiz-Medrano, R., Wang, H. L., Monzer, J., Yoo, B. C., McFarland, K. C., Franceschi, V. R. & Lucas, W. J. Plant paralog to viral movement protein that potentiates transport of mRNA into the phloem. *Science (80-. )*. **283**, 94–98 (1999).
  54. Dreyfuss, G., Matunis, M. J., Piñol-Roma, S. & Burd, C. G. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* **62**, 289–321 (1993).
  55. Streitner, C., Köster, T., Simpson, C. G., Shaw, P., Danisman, S., Brown, J. W. S. & Staiger, D. An hnRNP-like RNA-binding protein affects alternative splicing by in vivo interaction with transcripts in *Arabidopsis thaliana*. *Nucleic Acids Res.* **40**, 11240–11255 (2012).
  56. Carroll, A. J., Heazlewood, J. L., Ito, J. & Millar, A. H. Analysis of the *Arabidopsis* cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Mol. Cell. Proteomics* **7**, 347–369 (2008).
  57. Kim, J. S., Jung, H. J., Lee, H. J., Kim, K. A., Goh, C. H., Woo, Y., Oh, S. H., Han, Y. S. & Kang, H. Glycine-rich RNA-binding protein7 affects abiotic stress responses by regulating stomata opening and closing in *Arabidopsis thaliana*. *Plant J.* **55**, 455–466 (2008).
  58. Kim, J. S., Park, S. J., Kwak, K. J., Kim, Y. O., Kim, J. Y., Song, J., Jang, B., Jung, C. H. & Kang, H. Cold shock domain proteins and glycine-rich RNA-binding proteins from *Arabidopsis thaliana* can promote the cold adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **35**, 506–516 (2007).
  59. Ziemienowicz, A., Haasen, D., Staiger, D. & Merkle, T. *Arabidopsis* transportin1 is the nuclear import receptor for the circadian clock-regulated RNA-binding protein AfGRP7. *Plant Mol. Biol.* **53**, 201–212 (2003).
  60. Siomi, H., Choi, M., Siomi, M. C., Nussbaum, R. L. & Dreyfuss, G. Essential role for KH domains in RNA binding: Impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. *Cell* **77**, 33–39 (1994).
  61. Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K. & Dreyfuss, G. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol. Cell. Biol.* **6**, 2932–2943 (1986).
  62. Burd, C. G. & Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. *Science (80-. )*. **265**, 615–621 (1994).
  63. Berg, J. M. Zinc finger domains: Hypotheses and current knowledge. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 405–421 (1990).
  64. Linder, P. & Jankowsky, E. From unwinding to clamping- the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.* **12**, 505–516 (2011).
  65. Linder, P., Lasko, P., Aschburner, M., Leroy, P., Nielsen, P., Nishi, K., Schnier, J. & Slonimski, P. Birth of the DEAD box. *Nature* **337**, 121–122 (1989).
  66. Maronedze, C., Thomas, L., Serrano, N. L., Lilley, K. S. & Gehring, C. The RNA-binding protein repertoire of *Arabidopsis thaliana*. *Sci. Rep.* **6**, 1–13 (2016).
  67. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341 (2018).
  68. Kiledjian, M. & Dreyfuss, G. Primary structure and binding activity of the hnRNP U protein: Binding RNA through RGG box. *EMBO J.* **11**, 2655–2664 (1992).

## 8 References

69. Castello, A., Fischer, B., Frese, C. K., Horos, R., Alleaume, A. M., Foehr, S., Curk, T., Krijgsveld, J. & Hentze, M. W. Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* **63**, 696–710 (2016).
70. Rajyaguru, P., She, M. & Parker, R. Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins. *Mol. Cell* **45**, 244–254 (2012).
71. Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y. S., Mele, A., Fraser, C. E., Stone, E. F., Chen, C., Fak, J. J., Chi, S. W., Licatalosi, D. D., Richter, J. D. & Darnell, R. B. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
72. Mazroui, R., Hout, M. E., Tremblay, S., Fillion, C., Labelle, Y. & Khandjian, E. W. Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Hum. Mol. Genet.* **11**, 3007–3017 (2002).
73. Rhoads, S. N., Monahan, Z. T., Yee, D. S. & Shewmaker, F. P. The Role of Post-Translational Modifications on Prion-Like Aggregation and Liquid-Phase Separation of FUS. *Int. J. Mol. Sci.* **19**, (2018).
74. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R. & McKnight, S. L. Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
75. Ryan, V. H., Dignon, G. L., Zerze, G. H., Chabata, C. V., Silva, R., Conicella, A. E., Amaya, J., Burke, K. A., Mittal, J. & Fawzi, N. L. Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **69**, 465-479.e7 (2018).
76. Monahan, Z., Ryan, V. H., Janke, A. M., Burke, K. A., Rhoads, S. N., Zerze, G. H., O’Meally, R., Dignon, G. L., Conicella, A. E., Zheng, W., Best, R. B., Cole, R. N., Mittal, J., Shewmaker, F. & Fawzi, N. L. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* **36**, 2951–2967 (2017).
77. Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblentz, A., Schifferer, M., Ruepp, M. D., Simons, M., Niessing, D., Madl, T. & Dormann, D. Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **173**, 706-719.e13 (2018).
78. Thandapani, P., O’Connor, T. R., Bailey, T. L. & Richard, S. Defining the RGG/RG Motif. *Molecular Cell* vol. 50 613–623 (2013).
79. Birney, E., Kumar, S. & Krainer, A. R. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* **21**, 5803–5816 (1993).
80. Phan, A. T., Kuryavyi, V., Darnell, J. C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., Darnell, R. B. & Patel, D. J. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat. Struct. Mol. Biol.* **18**, 796–804 (2011).
81. Chong, P. A., Vernon, R. M. & Forman-Kay, J. D. RGG/RG Motif Regions in RNA Binding and Phase Separation. *J. Mol. Biol.* **430**, 4650–4665 (2018).
82. Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. Amino acid-base interactions: A three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* **29**, 2860–2874 (2001).
83. Vernon, R. M. C., Chong, P. A., Tsang, B., Kim, T. H., Bah, A., Farber, P., Lin, H. & Forman-Kay, J. D. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **7**, (2018).
84. van Nocker, S. & Vierstra, R. D. Two cDNAs from *Arabidopsis thaliana* encode putative RNA binding proteins containing glycine-rich domains. *Plant Mol. Biol.* **21**, 695–699 (1993).

## 8 References

85. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
86. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., *et al.* AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
87. Schöning, J. C., Streitner, C., Page, D. R., Hennig, S., Uchida, K., Wolf, E., Furuya, M. & Staiger, D. Auto-regulation of the circadian slave oscillator component AtGRP7 and regulation of its targets is impaired by a single RNA recognition motif point mutation. *Plant J.* **52**, 1119–1130 (2007).
88. Yan, Y., Ham, B. K., Chong, Y. H., Yeh, S. D. & Lucas, W. J. A Plant SMALL RNA-BINDING PROTEIN 1 Family Mediates Cell-to-Cell Trafficking of RNAi Signals. *Mol. Plant* **13**, 321–335 (2020).
89. Leder, V., Lummer, M., Tegeler, K., Humpert, F., Lewinski, M., Schüttpehl, M. & Staiger, D. Mutational definition of binding requirements of an hnRNP-like protein in Arabidopsis using fluorescence correlation spectroscopy. *Biochem. Biophys. Res. Commun.* **453**, 69–74 (2014).
90. Xu, F., Wang, L., Li, Y., Shi, J., Staiger, D., Chen, W., Wang, L. & Yu, F. The Receptor Kinase FER Mediates Phase Separation of Glycine-Rich RNA-Binding Protein 7 to Confer Temperature Resilience in Arabidopsis. *bioRxiv* 2022.03.06.483201 (2022).
91. Zand Karimi, H., Baldrich, P., Rutter, B. D., Borniego, L., Zajt, K. K., Meyers, B. C. & Innes, R. W. Arabidopsis apoplastic fluid contains sRNA- and circular RNA-protein complexes that are located outside extracellular vesicles. *Plant Cell* 1863–1881 (2022) doi:10.1093/plcell/koac043.
92. Meyer, K., Köster, T., Nolte, C., Weinholdt, C., Lewinski, M., Grosse, I. & Staiger, D. Adaptation of iCLIP to plants determines the binding landscape of the clock-regulated RNA-binding protein AtGRP7. *Genome Biol.* **18**, 1–22 (2017).
93. Kim, J. Y., Kim, W. Y., Kwak, K. J., Oh, S. H., Han, Y. S. & Kang, H. Glycine-rich RNA-binding proteins are functionally conserved in Arabidopsis thaliana and Oryza sativa during cold adaptation process. *J. Exp. Bot.* **61**, 2317–2325 (2010).
94. Lee, H. J., Kim, J. S., Yoo, S. J., Kang, E. Y., Han, S. H., Yang, K. Y., Kim, Y. C., McSpadden Gardener, B. & Kang, H. Different roles of glycine-rich RNA-binding protein7 in plant defense against Pectobacterium carotovorum, Botrytis cinerea, and tobacco mosaic viruses. *Plant Physiol. Biochem.* **60**, 46–52 (2012).
95. Carpenter, C. D., Kreps, J. A. & Simon, A. E. Genes encoding glycine-rich Arabidopsis thaliana proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. *Plant Physiol.* **104**, 1015–1025 (1994).
96. Heintzen, C., Nater, M., Apel, K. & Staiger, D. AtGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8515–8520 (1997).
97. Streitner, C., Hennig, L., Korneli, C. & Staiger, D. Global transcript profiling of transgenic plants constitutively overexpressing the RNA-binding protein AtGRP7. *BMC Plant Biol.* **10**, 221 (2010).
98. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoykov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A., *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).

## 8 References

99. Schmidt, F., Marnef, A., Cheung, M. K., Wilson, I., Hancock, J., Staiger, D. & Lodomery, M. A proteomic analysis of oligo(dT)-bound mRNP containing oxidative stress-induced *Arabidopsis thaliana* RNA-binding proteins ATGRP7 and ATGRP8. *Mol. Biol. Rep.* **37**, 839–845 (2010).
100. Jiang, Y., Yang, B., Harris, N. S. & Deyholos, M. K. Comparative proteomic analysis of NaCl stress-responsive proteins in *Arabidopsis* roots. *J. Exp. Bot.* **58**, 3591–3607 (2007).
101. Cao, S., Jiang, L., Song, S., Jing, R. & Xu, G. AtGRP7 is involved in the regulation of abscisic acid and stress responses in *Arabidopsis*. *Cell. Mol. Biol. Lett.* **11**, 526–535 (2006).
102. Cocquerelle, C., Mascrez, B., Hétuin, D. & Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.* **7**, 155–160 (1993).
103. Zaphiropoulos, P. G. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: Correlation with exon skipping. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6536–6541 (1996).
104. Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., Evtantal, N., Memczak, S., Rajewsky, N. & Kadener, S. CircRNA Biogenesis competes with Pre-mRNA splicing. *Mol. Cell* **56**, 55–66 (2014).
105. Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L. H. & Bindereif, A. Exon circularization requires canonical splice signals. *Cell Rep.* **10**, 103–111 (2015).
106. Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. & Brown, P. O. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS One* **7**, e30733 (2012).
107. Liang, Y., Liu, N., Yang, L., Tang, J., Wang, Y. & Mei, M. A Brief Review of circRNA Biogenesis, Detection, and Function. *Curr. Genomics* **22**, 485–495 (2021).
108. Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., Marzluff, W. F. & Sharpless, N. E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *Rna* **19**, 141–157 (2013).
109. Zhang, X. O., Wang, H. Bin, Zhang, Y., Lu, X., Chen, L. L. & Yang, L. Complementary sequence-mediated exon circularization. *Cell* **159**, 134–147 (2014).
110. Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmanidis, M., Phillips, C. A., Roslan, S., Schreiber, A. W., Gregory, P. A. & Goodall, G. J. The RNA binding protein quaking regulates formation of circRNAs. *Cell* **160**, 1125–1134 (2015).
111. Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., Herzog, M., Schreyer, L., Papavasileiou, P., Ivanov, A., Öhman, M., Refojo, D., Kadener, S. & Rajewsky, N. Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol. Cell* **58**, 870–885 (2014).
112. Sun, X., Wang, L., Ding, J., Wang, Y., Wang, J., Zhang, X., Che, Y., Liu, Z., Zhang, X., Ye, J., Wang, J., Sablok, G., Deng, Z. & Zhao, H. Integrative analysis of *Arabidopsis thaliana* transcriptomics reveals intuitive splicing mechanism for circular RNA. *FEBS Lett.* **590**, 3510–3516 (2016).
113. Patop, I. L., Wüst, S. & Kadener, S. Past, present, and future of circRNAs. *EMBO J.* **38**, 1–13 (2019).
114. Rahimi, K., Venø, M. T., Dupont, D. M. & Kjems, J. Nanopore sequencing of brain-derived full-length circRNAs reveals circRNA-specific exon usage, intron retention and microexons. *Nat. Commun.* **12**, 1–15 (2021).
115. Rahimi, K., Venø, M. T., Dupont, D. M. & Kjems, J. Nanopore sequencing of full-length circRNAs in human and mouse brains reveals circRNA-specific exon usage and intron retention. *bioRxiv* 567164 (2019) doi:10.1101/567164.

## 8 References

116. Ye, C. Y., Chen, L., Liu, C., Zhu, Q. H. & Fan, L. Widespread noncoding circular RNAs in plants. *New Phytol.* **208**, 88–95 (2015).
117. Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K. & Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
118. Park, O. H., Ha, H., Lee, Y., Boo, S. H., Kwon, D. H., Song, H. K. & Kim, Y. K. Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. *Mol. Cell* **74**, 494–507.e8 (2019).
119. Ebbesen, K. K., Kjems, J. & Hansen, T. B. Circular RNAs: Identification, biogenesis and function. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859**, 163–168 (2016).
120. Wang, P. L., Bao, Y., Yee, M.-C., Barrett, S. P., Hogan, G. J., Olsen, M. N., Dinneny, J. R., Brown, P. O. & Salzman, J. Circular RNA Is Expressed across the Eukaryotic Tree of Life. *PLoS One* **9**, e90859 (2014).
121. Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S. E., Graveley, B. R. & Lai, E. C. Genome-wide Analysis of Drosophila Circular RNAs Reveals Their Structural and Sequence Properties and Age-Dependent Neural Accumulation. *Cell Rep.* **9**, 1966–1980 (2014).
122. Werfel, S., Nothjunge, S., Schwarzmayr, T., Strom, T. M., Meitinger, T. & Engelhardt, S. Characterization of circular RNAs in human, mouse and rat hearts. *J. Mol. Cell. Cardiol.* **98**, 103–107 (2016).
123. Zhao, W., Cheng, Y., Zhang, C., You, Q., Shen, X., Guo, W. & Jiao, Y. Genome-wide identification and characterization of circular RNAs by high throughput sequencing in soybean. *Sci. Rep.* **7**, 1–11 (2017).
124. Danan, M., Schwartz, S., Edelheit, S. & Sorek, R. Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.* **40**, 3131–3142 (2012).
125. Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L. L., Wang, Y., Wong, C. C. L., Xiao, X. & Wang, Z. Extensive translation of circular RNAs driven by N<sup>6</sup>-methyladenosine. *Cell Res.* **27**, 626–641 (2017).
126. You, X., Vlatkovic, I., Babic, A., Will, T., Epstein, I., Tushev, G., Akbalik, G., Wang, M., Glock, C., Quedenau, C., Wang, X., Hou, J., Liu, H., Sun, W., Sambandan, S., Chen, T., Schuman, E. M. & Chen, W. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* **18**, 603–610 (2015).
127. Ye, C. Y., Chen, L., Liu, C., Zhu, Q. H. & Fan, L. Widespread noncoding circular RNAs in plants. *New Phytol.* **208**, 88–95 (2015).
128. Darbani, B., Noeparvar, S. & Borg, S. Identification of Circular RNAs from the Parental Genes Involved in Multiple Aspects of Cellular Metabolism in Barley. *Front. Plant Sci.* **7**, 776 (2016).
129. Lu, T., Cui, L., Zhou, Y., Zhu, C., Fan, D., Gong, H., Zhao, Q., Zhou, C., Zhao, Y., Lu, D., Luo, J., Wang, Y., Tian, Q., Feng, Q., Huang, T. & Han, B. Transcriptome-wide investigation of circular RNAs in rice. *RNA* **21**, 2076–2087 (2015).
130. Wang, Z., Liu, Y., Li, D., Li, L., Zhang, Q., Wang, S. & Huang, H. Identification of Circular RNAs in Kiwifruit and Their Species-Specific Response to Bacterial Canker Pathogen Invasion. *Front. Plant Sci.* **8**, 413 (2017).
131. Zuo, J., Wang, Q., Zhu, B., Luo, Y. & Gao, L. Deciphering the roles of circRNAs on chilling injury in tomato. *Biochem. Biophys. Res. Commun.* **479**, 132–138 (2016).
132. Zhang, P., Li, S. & Chen, M. Characterization and Function of Circular RNAs in Plants. *Front. Mol. Biosci.* **7**, 1–10 (2020).
133. Frydrych Capelari, É., da Fonseca, G. C., Guzman, F. & Margis, R. Circular and Micro RNAs from

## 8 References

- Arabidopsis thaliana Flowers Are Simultaneously Isolated from AGO-IP Libraries. *Plants* **8**, 302 (2019).
134. Owens, R. A., Blackburn, M. & Ding, B. Possible involvement of the phloem lectin in long-distance viroid movement. *Mol. Plant-Microbe Interact.* **14**, 905–909 (2001).
135. Tolstyko, E. A., Lezzhov, A. A., Morozov, S. Y. & Solovyev, A. G. Phloem transport of structured RNAs: A widening repertoire of trafficking signals and protein factors. *Plant Sci.* **299**, (2020).
136. Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. V., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D. T., Gruzinov, A., Borges, C., Jeffries, C. M., Svergun, D. I. & Franke, D. ATSAS 3.0: Expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
137. Jakobi, T., Uvarovskii, A. & Dieterich, C. Circtools—a one-stop software solution for circular RNA research. *Bioinformatics* **35**, 2326–2328 (2019).
138. Panjkovich, A. & Svergun, D. I. CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics* **34**, 1944–1946 (2018).
139. Svergun, D., Barberato, C. & Koch, M. H. CRY SOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
140. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).
141. Semenyuk, A. V. & Svergun, D. I. GNOM – a program package for small-angle scattering data processing. *J. Appl. Crystallogr.* **24**, 537–540 (1991).
142. Yang, J. & Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).
143. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
144. Panjkovich, A. & Svergun, D. I. SASpy: a PyMOL plugin for manipulation and refinement of hybrid models against small angle X-ray scattering data. *Bioinformatics* **32**, 2062–2064 (2016).
145. Panjkovich, A. & Svergun, D. I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* **18**, 5707–5719 (2016).
146. Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
147. Wördehoff, M. & Hoyer, W.  $\alpha$ -Synuclein Aggregation Monitored by Thioflavin T Fluorescence Assay. *Bio-Protocol* **8**, (2018).
148. Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular interaction studies using microscale thermophoresis. *Assay and Drug Development Technologies* vol. 9 342–353 (2011).
149. Chalhouh, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corr ea, M., Da Silva, C., Just, J., Falentin, C., Koh, C. S., Le Clainche, I., Bernard, M., Bento, P., Noel, B., *et al.* Early allopolyploid evolution in the post-neolithic Brassica napus oilseed genome. *Science (80- )*. **345**, 950–953 (2014).
150. Cheng, J., Metge, F. & Dieterich, C. Specific identification and quantification of circular RNAs from sequencing data. *Bioinformatics* **32**, 1094–1096 (2016).

## 8 References

151. Kwak, K. J., Park, S. J., Han, J. H., Kim, M. K., Oh, S. H., Han, Y. S. & Kang, H. Structural determinants crucial to the RNA chaperone activity of glycine-rich RNA-binding proteins 4 and 7 in *Arabidopsis thaliana* during the cold adaptation process. *J. Exp. Bot.* **62**, 4003–4011 (2011).
152. Staiger, D., Zecca, L., Wieczorek Kirk, D. A., Apel, K. & Eckstein, L. The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. *Plant J.* **33**, 361–371 (2003).
153. Khan, F., Daniëls, M. A., Folkers, G. E., Boelens, R., Saqlan Naqvi, S. M. & Van Ingen, H. Structural basis of nucleic acid binding by *Nicotiana tabacum* glycine-rich RNA-binding protein: Implications for its RNA chaperone function. *Nucleic Acids Res.* **42**, 8705–8718 (2014).
154. Staiger, D., Zecca, L., Wieczorek Kirk, D. A., Apel, K. & Eckstein, L. The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. *Plant J.* **33**, 361–371 (2003).
155. Xue, C., Lin, T. Y., Chang, D. & Guo, Z. Thioflavin T as an amyloid dye: Fibril quantification, optimal concentration and effect on aggregation. *R. Soc. Open Sci.* **4**, (2017).
156. Koch, M. H. J., Vachette, P. & Svergun, D. I. *Small-angle scattering: A view on the properties, structures and structural changes of biological macromolecules in solution*. *Quarterly Reviews of Biophysics* vol. 36 (2003).
157. Bizien, T., Durand, D., Roblin, P., Thureau, A., Vachette, P. & Pérez, J. A Brief Survey of State-of-the-Art BioSAXS. 217–231 (2016).
158. Durand, D., Vivès, C., Cannella, D., Pérez, J., Pebay-Peyroula, E., Vachette, P. & Fieschi, F. NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* **169**, 45–53 (2010).
159. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).
160. Nagai, K., Oubridge, C., Kuglstatter, A., Menichelli, E., Isel, C. & Jovine, L. Structure, function and evolution of the signal recognition particle. *EMBO J.* **22**, 3479–3485 (2003).
161. Wild, K., Weichenrieder, O., Leonard, G. A. & Cusack, S. The 2 Å structure of helix 6 of the human signal recognition particle RNA. *Structure* **7**, 1345–1352 (1999).
162. Shiraki, K., Mimura, M., Nishinami, S. & Ura, T. Effect of additives on liquid droplets and aggregates of proteins. *Biophys. Rev.* **12**, 587–592 (2020).
163. Lin, Y., Mori, E., Kato, M., Xiang, S., Wu, L., Kwon, I. & McKnight, S. L. Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. *Cell* **167**, 789–802.e12 (2016).
164. Li, L., Nelson, C. J., Solheim, C., Whelan, J. & Millar, A. H. Determining degradation and synthesis rates of arabidopsis proteins using the kinetics of progressive 15N labeling of two-dimensional gel-separated protein spots. *Mol. Cell. Proteomics* **11**, 1–16 (2012).
165. Seyednasrollah, F., Laiho, A. & Elo, L. L. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief. Bioinform.* **16**, 59–70 (2015).
166. Wang, X., Schwartz, J. C. & Cech, T. R. Nucleic acid-binding specificity of human FUS protein. *Nucleic Acids Res.* **43**, 7535–7543 (2015).
167. Ozdilek, B. A., Thompson, V. F., Ahmed, N. S., White, C. I., Batey, R. T. & Schwartz, J. C. Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. *Nucleic Acids Res.* **45**, 7984–7996 (2017).

## 8 References

168. Reichel, M., Köster, T. & Staiger, D. Marking RNA: M6A writers, readers, and functions in Arabidopsis. *Journal of Molecular Cell Biology* vol. 11 899–910 (2019).
169. Hoell, J. I., Larsson, E., Runge, S., Nusbaum, J. D., Duggimpudi, S., Farazi, T. A., Hafner, M., Borkhardt, A., Sander, C. & Tuschl, T. RNA targets of wild-type and mutant FET family proteins. *Nat. Struct. Mol. Biol.* **18**, 1428–1431 (2011).
170. Rogelj, B., Easton, L. E., Bogu, G. K., Stanton, L. W., Rot, G., Curk, T., Zupan, B., Sugimoto, Y., Modic, M., Haberman, N., Tollervey, J., Fujii, R., Takumi, T., Shaw, C. E. & Ule, J. Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.* **2**, 1–10 (2012).
171. Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L. H., Dale, R. K., Smith, S. A., Yarosh, C. A., Kelly, S. M., Nabet, B., Mecnas, D., *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177 (2013).
172. Huelga, S. C., Vu, A. Q., Arnold, J. D., Liang, T. D., Liu, P. P., Yan, B. Y., Donohue, J. P., Shiue, L., Hoon, S., Brenner, S., Ares, M. & Yeo, G. W. Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins. *Cell Rep.* **1**, 167–178 (2012).
173. Rossbach, O., Hung, L. H., Khrameeva, E., Schreiner, S., König, J., Curk, T., Zupan, B., Ule, J., Gelfand, M. S. & Bindereif, A. Crosslinking-immunoprecipitation (iCLIP) analysis reveals global regulatory roles of hnRNP L. *RNA Biol.* **11**, 146–155 (2014).
174. Uren, P. J., Bahrami-Samani, E., de Araujo, P. R., Vogel, C., Qiao, M., Burns, S. C., Smith, A. D. & Penalva, L. O. F. High-throughput analyses of hnRNP H1 dissects its multi-functional aspect. *RNA Biol.* **13**, 400–411 (2016).
175. Yagi, R., Miyazaki, T. & Oyoshi, T. G-quadruplex binding ability of TLS/FUS depends on the  $\beta$ -spiral structure of the RGG domain. *Nucleic Acids Res.* **46**, 5894–5901 (2018).
176. Van Dusen, C. M., Yee, L., McNally, L. M. & McNally, M. T. A Glycine-Rich Domain of hnRNP H/F Promotes Nucleocytoplasmic Shuttling and Nuclear Import through an Interaction with Transportin 1. *Mol. Cell. Biol.* **30**, 2552–2562 (2010).
177. McBride, A. E., Conboy, A. K., Brown, S. P., Ariyachet, C. & Rutledge, K. L. Specific sequences within arginine-glycine-rich domains affect mRNA-binding protein function. *Nucleic Acids Res.* **37**, 4322–4330 (2009).
178. Vasilyev, N., Polonskaia, A., Darnell, J. C., Darnell, R. B., Patel, D. J. & Serganov, A. Crystal structure reveals specific recognition of a G-quadruplex RNA by a  $\beta$ -turn in the RGG motif of FMRP. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E5391–E5400 (2015).
179. Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H. M., Duhr, S., Baaske, P. & Breitsprecher, D. MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **1077**, 101–113 (2014).
180. López-Méndez, B., Baron, B., Brautigam, C. A., Jowitt, T. A., Knauer, S. H., Uebel, S., Williams, M. A., Sedivy, A., Abian, O., Abreu, C., Adamczyk, M., Bal, W., Berger, S., Buell, A. K., Carolis, C., Daviter, T., Fish, A., Garcia-Alai, M., Guenther, C., *et al.* Reproducibility and accuracy of microscale thermophoresis in the NanoTemper Monolith: a multi laboratory benchmark study. *Eur. Biophys. J.* **50**, 411–427 (2021).
181. Kwak, K. J., Park, S. J., Han, J. H., Kim, M. K., Oh, S. H., Han, Y. S. & Kang, H. Structural determinants crucial to the RNA chaperone activity of glycine-rich RNA-binding proteins 4 and 7 in Arabidopsis thaliana during the cold adaptation process. *J. Exp. Bot.* **62**, 4003–4011 (2011).
182. Kim, M. K., Jung, H. J., Kim, D. H. & Kang, H. Characterization of glycine-rich RNA-binding proteins in Brassica napus under stress conditions. *Physiol. Plant.* **146**, 297–307 (2012).

## 8 References

183. Ryan, V. H., Perdikari, T. M., Naik, M. T., Saueressig, C. F., Lins, J., Dignon, G. L., Mittal, J., Hart, A. C. & Fawzi, N. L. Tyrosine phosphorylation regulates hnRNP A2 granule protein partitioning and reduces neurodegeneration. *EMBO J.* **40**, 1–22 (2021).
184. Monahan, Z., Ryan, V. H., Janke, A. M., Burke, K. A., Rhoads, S. N., Zerbe, G. H., O’Meally, R., Dignon, G. L., Conicella, A. E., Zheng, W., Best, R. B., Cole, R. N., Mittal, J., Shewmaker, F. & Fawzi, N. L. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* **36**, 2951–2967 (2017).
185. Ghisolfi, L., Joseph, G., Amalric, F. & Erard, M. The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J. Biol. Chem.* **267**, 2955–2959 (1992).
186. Martin, E. W., Holehouse, A. S., Peran, I., Farag, M., Incicco, J. J., Bremer, A., Grace, C. R., Soranno, A., Pappu, R. V. & Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science (80-. )*. **367**, 694–699 (2020).
187. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell* **60**, 208–219 (2015).
188. Mollieux, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T. & Taylor, J. P. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163**, 123–133 (2015).
189. Linsenmeier, M., Faltova, L., Palmiero, U. C., Seiffert, C., Küffner, A. M., Pinotsi, D., Zhou, J., Mezzenga, R. & Arosio, P. The interface of condensates of the hnRNP A1 low complexity domain promotes formation of amyloid fibrils. *bioRxiv* 2022.05.23.493075 (2022).
190. Schwartz, J. C., Wang, X., Podell, E. R. & Cech, T. R. RNA Seeds Higher-Order Assembly of FUS Protein. *Cell Rep.* **5**, 918–925 (2013).
191. Muschol, M. & Rosenberger, F. Liquid-liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J. Chem. Phys.* **107**, 1953–1962 (1997).
192. Cinar, H., Fetahaj, Z., Cinar, S., Vernon, R. M., Chan, H. S. & Winter, R. H. A. Temperature, Hydrostatic Pressure, and Osmolyte Effects on Liquid–Liquid Phase Separation in Protein Condensates: Physical Chemistry and Biological Implications. *Chem. – A Eur. J.* **25**, 13049–13069 (2019).
193. Guccione, E. & Richard, S. The regulation, functions and clinical relevance of arginine methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 642–657 (2019).
194. Lorton, B. M. & Shechter, D. Cellular consequences of arginine methylation. *Cell. Mol. Life Sci.* **76**, 2933–2956 (2019).
195. Nichols, R. C., Wang, X. W., Tang, J., Hamilton, B. J., High, F. A., Herschman, H. R. & Rigby, W. F. C. The RGG domain in hnRNP A2 affects subcellular localization. *Exp. Cell Res.* **256**, 522–532 (2000).
196. Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).
197. Xiao, M. S. & Wilusz, J. E. An improved method for circular RNA purification using RNase R that efficiently removes linear RNAs containing G-quadruplexes or structured 3' ends. *Nucleic Acids Res.* **47**, 8755–8769 (2019).
198. Szabo, L. & Salzman, J. Detecting circular RNAs: bioinformatic and experimental challenges. (2016) doi:10.1038/nrg.2016.114.

## 8 References

199. Nielsen, A. F., Bindereif, A., Bozzoni, I., Hanan, M., Hansen, T. B., Irimia, M., Kadener, S., Kristensen, L. S., Legnini, I., Morlando, M., Jarlstad Olesen, M. T., Pasterkamp, R. J., Preibisch, S., Rajewsky, N., Suenkel, C. & Kjems, J. Best practice standards for circular RNA research. *Nat. Methods* (2022) doi:10.1038/s41592-022-01487-2.
200. Hansen, T. B. Improved circRNA identification by combining prediction algorithms. *Front. Cell Dev. Biol.* **6**, 1–9 (2018).
201. Allender, C. J. & King, G. J. Origins of the amphiploid species *Brassica napus* L. investigated by chloroplast and nuclear molecular markers. *BMC Plant Biol.* **10**, (2010).
202. Tang, C., Yu, T., Xie, Y., Wang, Z., McSwiggin, H., Zhang, Y., Zheng, H. & Yan, W. Template switching causes artificial junction formation and false identification of circular RNAs. *bioRxiv* 259556 (2018) doi:10.1101/259556.
203. Zand Karimi, H., Baldrich, P., Rutter, B. D., Borniego, L., Zajt, K. K., Meyers, B. C. & Innes, R. W. Arabidopsis apoplastic fluid contains sRNA- and circular RNA-protein complexes that are located outside extracellular vesicles. *Plant Cell* **34**, 1863–1881 (2022).
204. Chu, Q., Bai, P., Zhu, X., Zhang, X., Mao, L., Zhu, Q. H., Fan, L. & Ye, C. Y. Characteristics of plant circular RNAs. *Brief. Bioinform.* **21**, 135–143 (2018).
205. Conn, V. M., Hugouvieux, V., Nayak, A., Conos, S. A., Capovilla, G., Cildir, G., Jourdain, A., Tergaonkar, V., Schmid, M., Zubieta, C. & Conn, S. J. A circRNA from *SEPALLATA3* regulates splicing of its cognate mRNA through R-loop formation. *Nat. Plants* **3**, 1–5 (2017).
206. Rahimi, K., Færch Nielsen, A., Venø, M. T. & Kjems, J. Nanopore long-read sequencing of circRNAs. *Methods* **196**, 23–29 (2021).
207. Hansen, T. B. Improved circRNA identification by combining prediction algorithms. *Frontiers in Cell and Developmental Biology* vol. 6 (2018).

## 9 Supplements

## 9.1 Additional Material and Methods

## 9.1.1 Oligonucleotides

All oligonucleotides used in this work are listed in Table 33 with sequence and usage.

**Table 33: Oligonucleotides used in this work.** Name, usage and sequence of each oligonucleotide is listed.

Name	Usage	Sequence 5'-3'
pLic_fw	Colony PCR and sequencing	TGTGAGCGGATAACAATTCC
pLic_rev	Colony PCR and sequencing	AGCAGCCAACCTCAGCTTCC
M13_pUC_fw	Colony PCR and sequencing	CCCAGTCACGACGTTGTAAAACG
M13_pUC_rev	Colony PCR and sequencing	AGCGGATAACAATTCACACAGG
RuBISCO_small_sub_fw	Phloem-purity testing	TTCACCGGCTTGAAGTCATC
RuBISCO_small_cub_rev	Phloem-purity testing	CCGGTACTCCTCTTGCAT
ThioredoxinH_fw	Phloem-purity testing	CTCAAGGCAGCAAAGAATC
ThioredoxinH_rev	Phloem-purity testing	ATGGCCTGAACATCGAACTC
At2g21660_fw	Cloning of At2g21660 into expression plasmid	TTAACATATGATGGCGTCCGGTGATGTT
At2g21660_rev	Cloning of At2g21660 into expression plasmid	TTAAGCGGCCACCTAATTAAGGAATTACCATCCTCC
At4g28440_fw	Cloning of At4g28440 into expression plasmid	TTAACATATGATGGCGACTGGAAGTGG
At4g28440_rev	Cloning of At4g28440 into expression plasmid	AATTGCGGCCGCTCACTGATCACCACCAACATTGATC
At2g_Mut_Stop_fw	Stop-codon mutation of AtGRP7	GAGGCTCAGTCACGATGAAGCGGTGGCGCGGA
At2g_Mut_Stop_rev	Stop-codon mutation of AtGRP7	TCCGCCGCCACCGCTTCATCGTGACTGAGCTC
AtGRP7_RRM_fw	Amplification of AtGRP7 RRM	ttaaGGTCTCgttgcATGGCGTCCGGTGATGTT
AtGRP7_RRM_rev	Amplification of AtGRP7 RRM	ttaaGGTCTCCTCGTGACTGAGCCTCGTTAACAGTGATG
AtGRP7_T7_fw	<i>In vitro</i> transcription template amplification	TAATACGACTCACTATAGGGATGGCGTCCGGTGATGTT
AtGRP7_rev	<i>In vitro</i> transcription template amplification	ACCTAATTAAGGAATTACCATCCTCC
AtGRP7_UTR_T7_fw	<i>In vitro</i> transcription template amplification	TAATACGACTCACTATAGGGCAGAATTACGTAACCGACTCTAAACC
AtGRP7_UTR_rev	<i>In vitro</i> transcription template amplification	ttaaGGTCTCgaagcTGCGAATGAACCAAAATAGTTTCTACTT
BnGRP7_NheI_fw	Cloning of BnGRP7 into expression plasmid	TTAAGTAGCATGGCGTCCCTGATGTTGA
BnGRP7_HindIII_rev	Cloning of BnGRP7 into expression plasmid	TTAAAAGCTTTTACCAACCACCACC
BnGRP7_T7_fw	<i>In vitro</i> transcription template amplification	taatacactcactataggGGATGGCGTCCCTGATGTTGA
BnGRP7_rev	<i>In vitro</i> transcription template amplification	TTACCAACCACCACCACC
BnGRP7_UTR_T7_fw	<i>In vitro</i> transcription template amplification	TAATACGACTCACTATAGGGGCATCACACCTTCCCTCTCTTAC
BnGRP7_UTR_rev	<i>In vitro</i> transcription template amplification	ATTTGTTAAACTTAAAAAATAAACACGGCAAG
BnPARCL_fw_T7	<i>In vitro</i> transcription template amplification	taatacactcactataggGG ATGCAGTACTACGAAACCCGT
BnPARCL_rev	<i>In vitro</i> transcription template amplification	GGTACTCGCTGTGCTACCG
BnPARCL_fw	<i>In vitro</i> transcription template amplification	ATGCAGTACTACGAAACCCGT
BnPARCL_rev_T7prom	<i>In vitro</i> transcription template amplification	taatacactcactataggGGTACTCGCTGCTACCG
miRNA164_T7_fw	<i>In vitro</i> transcription template amplification	GATCTTAATACGACTCACTATAGGGTGAGAGCAGGGCACGTGCA
miRNA164_rev	<i>In vitro</i> transcription template	TGCACGTGCCCTGCTTCCACCCTATAGTGAGTCGTATTAAGATC
GRP7-3'UTR_T7_fw	<i>In vitro</i> transcription template	taatacactcactataggGGATTTTGTCTGTTCTGCTTTAGATTTGATGT
GRP7-3'UTR_rev	<i>In vitro</i> transcription template	ACATCAAATCTAAAGCAGAACCAAGACAAAATCCcctatagtgagtcgtatta
miRNA <sub>novel2</sub> _T7_fw	<i>In vitro</i> transcription template	taatacactcactataggGGtcttctctgtttggtttgaac
miRNA <sub>novel2</sub> _rev	<i>In vitro</i> transcription template	gttcaaaccaaacgagaacaagaCCcctatagtgagtcgtatta
miRNA <sub>novel106</sub> _T7_fw	<i>In vitro</i> transcription template	taatacactcactataggGGagatacagatctcttagctttaac
miRNA <sub>novel106</sub> _rev	<i>In vitro</i> transcription template	gttaaaagtaagagatcgtatctCCcctatagtgagtcgtatta
miRNA <sub>novel149</sub> _T7_fw	<i>In vitro</i> transcription template	taatacactcactataggGGattaattgtgctggtgtagacatc
miRNA <sub>novel149</sub> _rev	<i>In vitro</i> transcription template	gatgtctacaccagacaattaatCCcctatagtgagtcgtatta

## 9 Supplements

<b>BnPARCL_Gamotif_T7_fw</b>	<i>In vitro</i> transcription template amplification	GATCTTAATACGACTCACTATAGGGGAGAAGACGAAGAAGAAGAAATGCAGTACTACGAAACCCGTG
<b>BNACNNG53670D_fw_T7</b>	<i>In vitro</i> transcription template amplification	GATCTTAATACGACTCACTATAGGATGGGAAATCGTTGTGCTAGAAC
<b>BNACNNG53670D_rev</b>	<i>In vitro</i> transcription template amplification	TTTCTCTCTGCTGTTTATGGTTTTCTG
<b>BnaCNG53670D_120UTR_T7_fw</b>	<i>In vitro</i> transcription template amplification	GATCTTAATACGACTCACTATAGGGGACACAAGTCATACCTAGGTTCCGG
<b>BnaCNG53670D_120UTR_rev</b>	<i>In vitro</i> transcription template amplification	ATCCGGAAGGATAAAACATCCCAAG
<b>BnaCNG53670D_rev_ohneGA</b>	<i>In vitro</i> transcription template amplification	TTTTCTGAGTAAGAGTTCAGCAGCTG
<b>BnaC03G26040D_fw_T7</b>	<i>In vitro</i> transcription template amplification	GATCTTAATACGACTCACTATAGGGCTCGGTTGCTAGGAGATAAGAAAATC
<b>BnaC03G26040D_rev</b>	<i>In vitro</i> transcription template amplification	CGTAGATATTAGAGAAAATAATCTCAGAGTACACC
<b>BnaA03g14400D_T7_fw</b>	<i>In vitro</i> transcription template amplification	taatacgactcactataggGGttcttctctcaaaacggccaacacg
<b>BnaA03g14400D_rev</b>	<i>In vitro</i> transcription template amplification	aaaaactgcttttttattagaactatatacaaaatagcc
<b>BnaA07g32150D_UTR_T7_fw</b>	<i>In vitro</i> transcription template amplification	taatacgactcactataggGGtctctcacaccacacatacatac
<b>BnaA07g32150D_UTR_rev</b>	<i>In vitro</i> transcription template amplification	tttgataagcgactaaactccgaac
<b>BnaA09g50830D_T7_fw</b>	<i>In vitro</i> transcription template amplification	taatacgactcactataggGGATGTCCCGAGGACCAGG
<b>BnaA09g50830D_rev</b>	<i>In vitro</i> transcription template amplification	TCAGAAGCCGCTCCTTCAG
<b>BnaC05g32740D_T7_fw</b>	<i>In vitro</i> transcription template amplification	taatacgactcactataggGGATGAAAGTAAACTGCTTGCTGCACACTAGACGTGTAA
<b>BnaC05g32740D_T7_rev</b>	<i>In vitro</i> transcription template amplification	TTACACGCTAAGTGTGCAAGCAAGCAGTTTACTTTTCATCCctatagtgagtcgtatta
<b>BnaC09g46650D_T7_fw</b>	<i>In vitro</i> transcription template amplification	taatacgactcactataggGGATGGCTGACATAGCACTACTAGTGG
<b>BnaC09g46650D_rev</b>	<i>In vitro</i> transcription template amplification	TTAAGCAGAAAAGAAGCCGTTAGAGAC
<b>BnaC09g45930D_T7_fw</b>	<i>In vitro</i> transcription template amplification	taatacgactcactataggGGATGCCTCCCTCTGTCTCTTCTTCT
<b>BnaC09g45930D_rev</b>	<i>In vitro</i> transcription template amplification	CTACGACTTTTCTCAGCGGAGG
<b>circRNA1E1_lin_fw</b>	Amplification of linear transcript circRNA1	ATGATGTCGTTCCATCTCCTACTTTTTTC
<b>circRNA1E1_rev</b>	Amplification of BSJ circRNA1	AATCCGCTGGCCAGGT
<b>circRNA1E2_rev</b>	Amplification of BSJ circRNA1	TCGTAGATGAGTAGCATCTCCAAACC
<b>circRNA1E7_fw</b>	Amplification of BSJ circRNA1	GGAACCCCTTCCCTAAGCTCTATGT
<b>circRNA1E1_total_rev</b>	Amplification of whole circRNA1	GAGATGTCACGGCGACTGG
<b>circRNA1E1_total_fw</b>	Amplification of whole circRNA1	ACCTGGGCCAGCGGATT
<b>circRNA2_fw</b>	Amplification of BSJ circRNA2	GTCCATGCCACGTCAGTTTCCG
<b>circRNA2_rev</b>	Amplification of BSJ circRNA2	GCGTCTGGTCCCAGCGT
<b>circRNA2_total_fw</b>	Amplification of whole circRNA2	GTTCCAGCCTCTGGTCTTC
<b>circRNA2_total_rev</b>	Amplification of whole circRNA2	CGGAAACTGACGTGGCATGGAC
<b>circRNA3lin_fw</b>	Amplification of lin transcript circRNA3	CTATCTATGCGCTCTCTTGTGCATACG
<b>circRNA3_fw</b>	Amplification of BSJ circRNA3	AGCTCTGTTGGCTTCACC
<b>circRNA3_rev</b>	Amplification of BSJ circRNA3	GAGATCGAGAGGAGAGAAGAAAACAG
<b>circRNA3_total_fw</b>	Amplification of whole circRNA3	TACTCACCACCTTCACAGGTCA
<b>circRNA3_total_rev</b>	Amplification of whole circRNA3	GGTGAAGCCAACAGAGAGCT
<b>circRNA4E7_lin_fw</b>	Amplification of lin transcript circRNA4	ACTTTTGTGTCTCCGTGGGG
<b>circRNA4E1_fw</b>	Amplification of BSJ circRNA4	CAAGCCTCCATCGATTCTGTGAA
<b>circRNA4E5_rev</b>	Amplification of BSJ circRNA4	AGAGGAAGGCGAGGCTAAAAC
<b>circRNA4_total_fw1</b>	Amplification of whole circRNA4	GTTTTAGCCTCGCTCTCTCT
<b>circRNA4_total_rev1</b>	Amplification of whole circRNA4	TTCTTACTCTCTCACAGGAGCAG
<b>circRNA5E5_lin_fw</b>	Amplification of lin transcript circRNA5	GAGGAGAAGAGCTGAGAAGTGCA
<b>circRNA5E6_fw</b>	Amplification of BSJ circRNA5	CACTCTTCTCGGAATGAAAAAGGG
<b>circRNA5E6_rev</b>	Amplification of BSJ circRNA5	GCTTCTTTCACCTTGAAGGTTGTCT
<b>circRNA5_total_fw</b>	Amplification of whole circRNA5	AGACAACCTTCAAGGTGAAAAGAAGC
<b>circRNA5_total_rev</b>	Amplification of whole circRNA5	TTGCTGTTGCTTCCAGTAAAACG
<b>circRNA7_fw</b>	Amplification of BSJ circRNA7	GTAAACATAGCAAACATAATAAGTCTTAG
<b>circRNA7_rev</b>	Amplification of BSJ circRNA7	CGTGTTGTATCGGAAAGT

## 9 Supplements

<b>circRNA8_fw</b>	Amplification of BSJ circRNA8	CGCTCGAGTCCAACCCT
<b>circRNA8_rev</b>	Amplification of BSJ circRNA8	AAGGCTTGGGGTTCAACCTC
<b>circRNA8_lin_rev</b>	Amplification of lin transcript circRNA8	GTCGAGCTTAGTAACGCAGGC
<b>SRP_RNA1_fw</b>	Amplification of lin/circ SRP1	AACTATGGCCGATCTCATTAGGC
<b>SRP_RNA1circ_rev</b>	Amplification of circular SRP1	CCAACATCTTGTGTCTCCACTCG
<b>SRP_RNA1_rev</b>	Amplification of linear SRP1	GGCCCGTGATCCTCCACC
<b>SRP_circlin_fw</b>	Amplification of lin/circ SRP	ACCAAGTTGGGAGTTGAGGC
<b>SRP_lin_rev</b>	Amplification of linear SRP	GATGGTCCAATCGGAATCGGC
<b>SRPcirc_rev</b>	Amplification of circular SRP	ATGGGCAGCCAGAATAAAGC

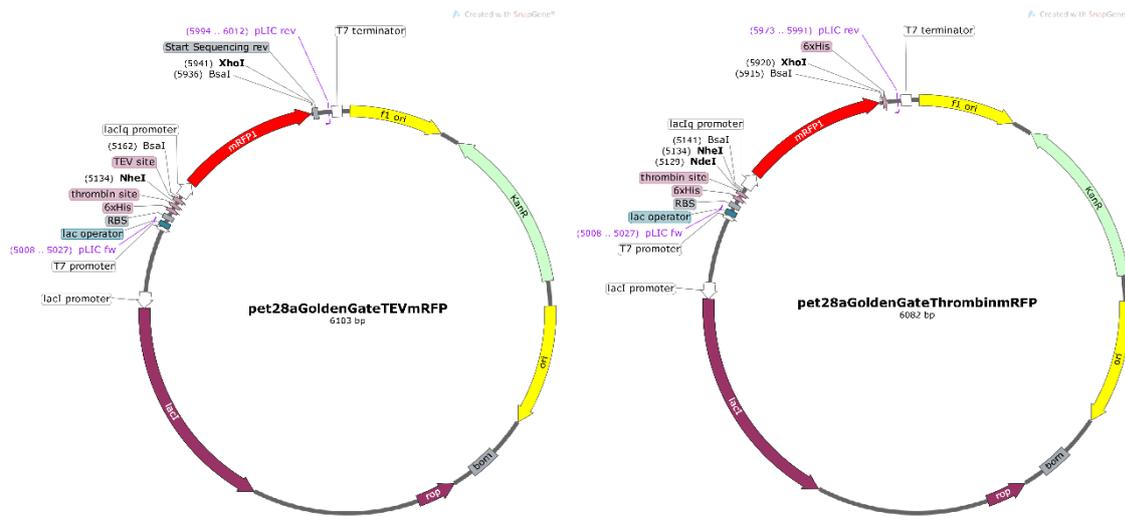
### 9.1.2 Plasmid maps

Maps of plasmids used in this thesis are shown in this chapter.

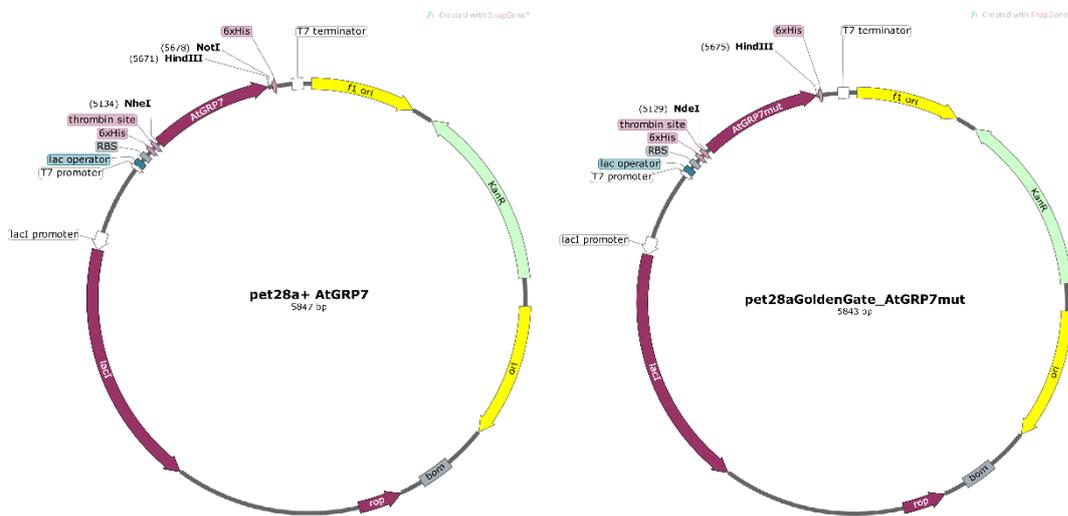


**Figure 40: pet28a+ and pUC57 cloning and expression plasmids.** Pet28a+ was used to insert genes and further expression and purification of their transcripts. pUC57 was used for blunt-end cloning of genes and PCR-amplicons via *SmaI*.

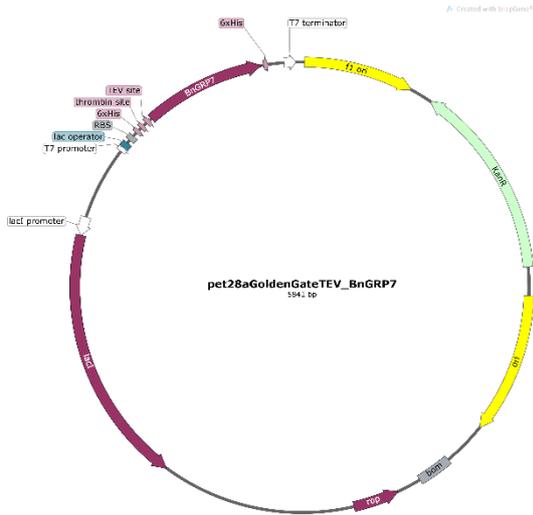
## 9 Supplements



**Figure 41: Modified pet28a+ plasmids suitable for Golden Gate cloning via *BsaI*.** For easy selection of positive colonies after transformation, an mRFP was placed in the multiple cloning site between *BsaI* recognition and cutting sequences. Six different plasmids were designed, three without and three with MBP as N-terminal solubility tag. The plasmids contain either Thrombin, TEV or PreScission protease recognition and cutting sequences.



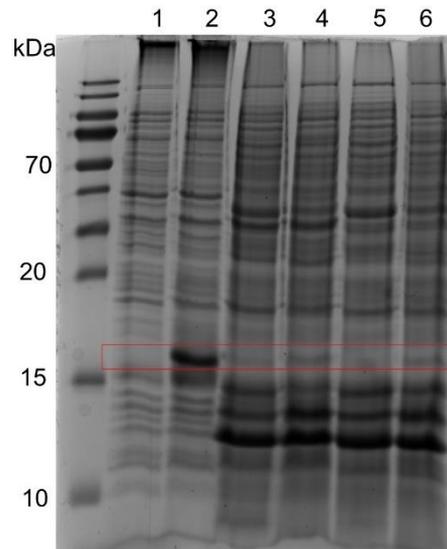
**Figure 42: Expression plasmids (pet28a+ and pet28a+GGThrombin) containing AtGRP7 and AtGRP7mut.** The N-terminal His-tag is in frame with AtGRP7 and can be cut off by Thrombin during protein purification. AtGRP7short plasmid is the same as pet28a+ AtGRP7, only containing a point mutation at amino acid 88 to a stop codon.



**Figure 43: Expression plasmid pet28a+ GGTEV containing BnGRP7.** The N-terminal His-tag is in frame with AtGRP7 and can be cut off by TEV during protein purification.

## 9.2 Additional figures for results

### 9.2.1 Solubility of AtGRP7



**Figure 44: Solubility of AtGRP7 with and without 1.6 hexandiol during lysis.** 1: before induction, 2: after induction, 3: After lysis, 4: after lysis with hexandiol 1.6, 5: Supernatant after lysis after, 6: Supernatant after lysis with hexandiol 1.6

## 9 Supplements

### 9.2.2 Degradation of AtGRP7

The degradation of AtGRP7 was analyzed by mass spectrometry Figure 45.

Full length protein            MASCOT score: 164            72 % protein sequence coverage

```
1  MASGDVEYRC  FVGGLAWATD  DRALETAFAQ  YGDVIDSKII  NDRETGRSRG
51  FGFVTFKDEK  AMKDAIEGMN  GQDLDGRSIT  VNEAQRSGSG  GGGGHRGGGG
101 GGYRSGGGGG  YSGGGGSYGG  GGGRREGGGG  YSGGGGGYSS  RGGGGGSYGG
151 GRREGGGGYG  GGEGGGYGGG  GGGGGW
```

degradation protein            MASCOT score: 110            39 % protein sequence coverage

```
1  MASGDVEYRC  FVGGLAWATD  DRALETAFAQ  YGDVIDSKII  NDRETGRSRG
51  FGFVTFKDEK  AMKDAIEGMN  GQDLDGRSIT  VNEAQRSGSG  GGGGHRGGGG
101 GGYRSGGGGG  YSGGGGSYGG  GGGRREGGGG  YSGGGGGYSS  RGGGGGSYGG
151 GRREGGGGYG  GGEGGGYGGG  GGGGGW
```

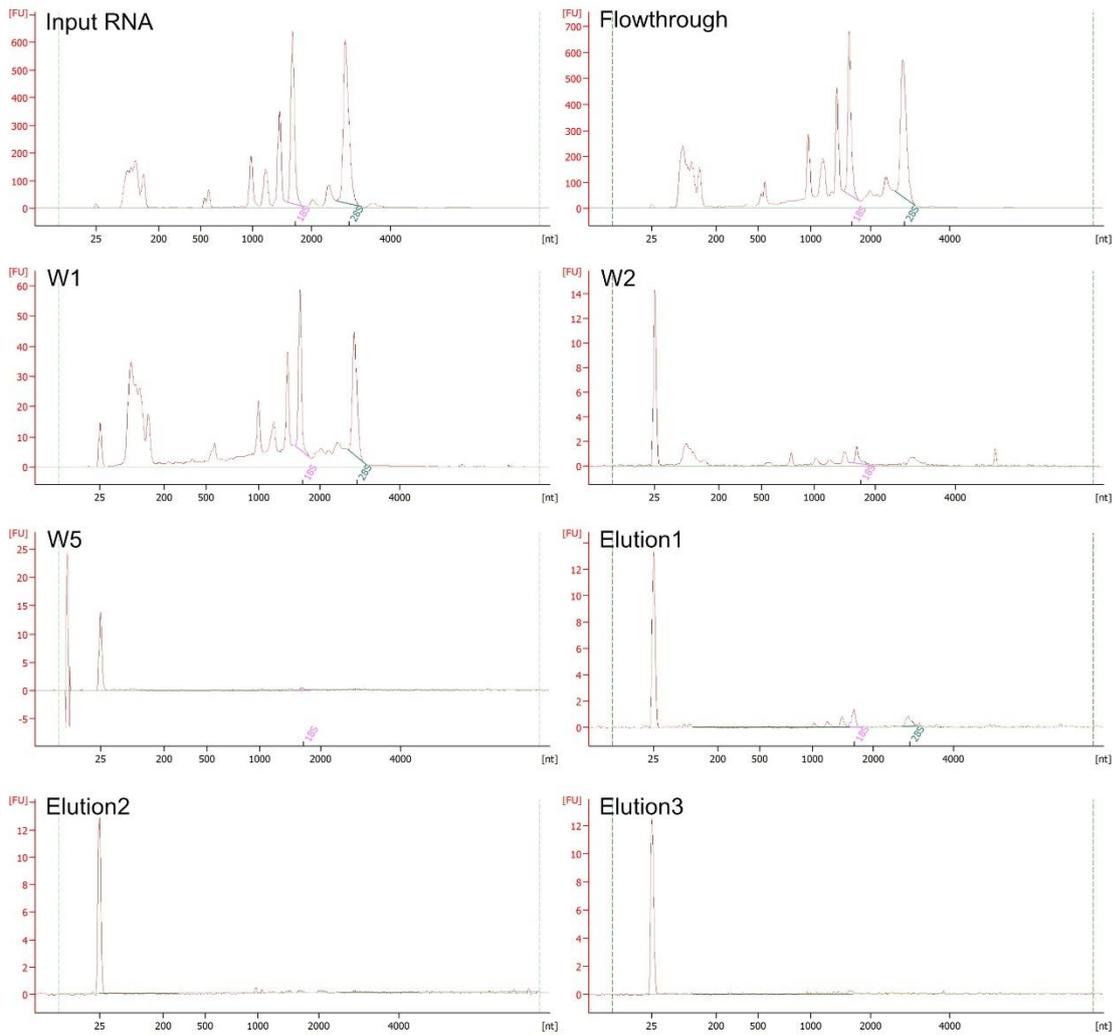
degradation protein 2            MASCOT score: 111            49 % protein sequence coverage

```
1  MASGDVEYRC  FVGGLAWATD  DRALETAFAQ  YGDVIDSKII  NDRETGRSRG
51  FGFVTFKDEK  AMKDAIEGMN  GQDLDGRSIT  VNEAQRSGSG  GGGGHRGGGG
101 GGYRSGGGGG  YSGGGGSYGG  GGGRREGGGG  YSGGGGGYSS  RGGGGGSYGG
151 GRREGGGGYG  GGEGGGYGGG  GGGGGW
```

**Figure 45: Sequence coverage of AtGRP7 and its degradation products.** The peaks were selected in mMass by *in silico* digestion of AtGRP7 sequence and peak picking. Peaks at least 4 times higher than the background were selected. The minimum score number for a significant hit at a p-value of >0.001 is 72.

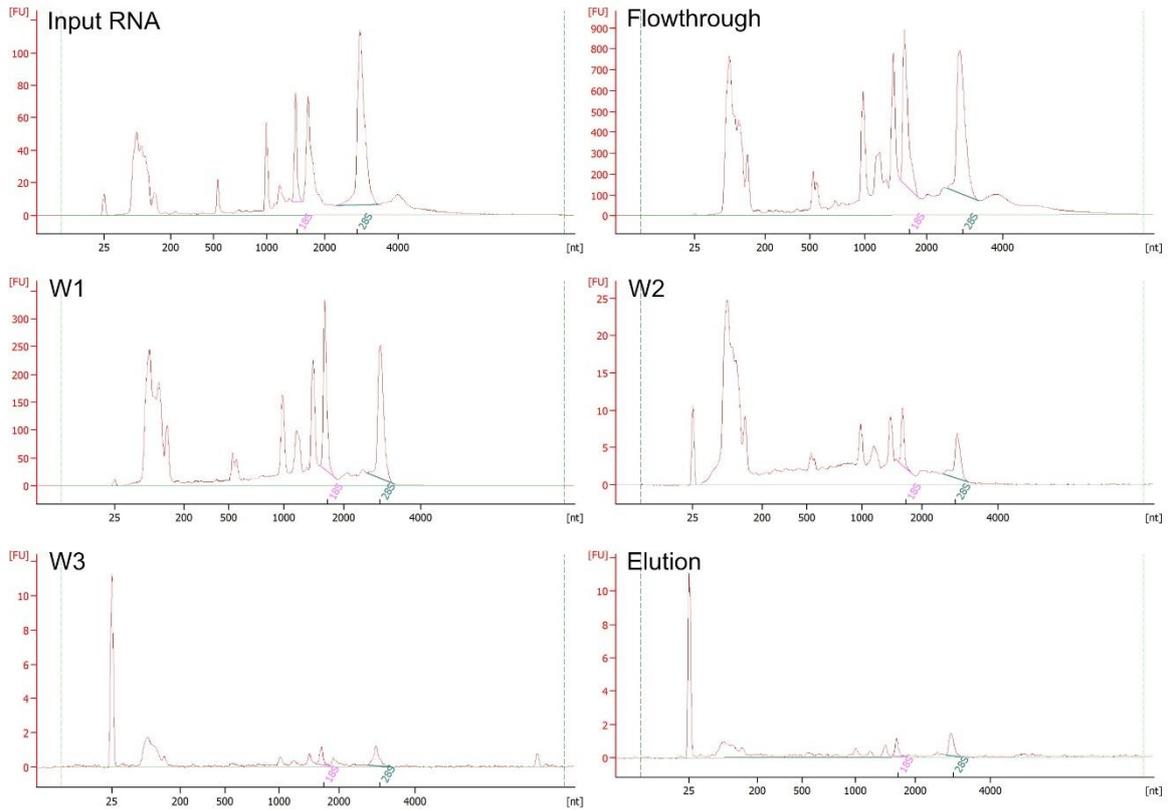
## 9 Supplements

### 9.2.3 CnBr-Sepharose beads negative control



**Figure 46: Electropherograms of CnBr-sepharose beads negative control.** Electropherograms of Input leaf RNA, flowthrough, washing step one, two and five as well as elution one to three are displayed. Bioanalyzer run with a Nanochip for total eukaryotic RNA. Y-axis: fluorescence units (FU), X-axis: length in nucleotides (nt). Due to the different amount of RNA used, the y-axis displayed varies between the different electropherograms showing different FU.

## 9 Supplements



**Figure 47: Electropherograms of CnBr-sepharose beads with AtGRP7short.** Electropherograms of Input leaf RNA, flowthrough, washing step one, two and three as well as elution are displayed. Bioanalyzer run with a Nanochip for total eukaryotic RNA. Y-axis: fluorescence units (FU), X-axis: length in nucleotides (nt). Due to the different amount of RNA used, the y-axis displayed varies between the different electropherograms showing different FU.

## 9 Supplements

### 9.2.4 Additional Illumina sequencing results

**Table 34: List with all significantly enriched RNAs by AtGRP7 CnBr column from *B. napus* leaf RNA compared to total leaf RNA.**

GRP7value	Inputvalue	log2foldchange	pvalue	qvalue	gene_id	gene_name
7.20232633	0	Inf	0.00022439	0.32843561	GSBRNA2T00049484001	BnaA09g39850D
18.375228	0	Inf	0.00077378	0.32843561	GSBRNA2T00001753001	BnaC01g26930D
1.932241	0	Inf	0.00142463	0.36562033	GSBRNA2T00048376001	BnaA05g17250D
1.86925067	0	Inf	0.00154403	0.36562033	GSBRNA2T00132370001	BnaC03g19360D
5.40448067	0	Inf	0.00304215	0.41864812	GSBRNA2T00105102001	BnaC04g41400D
7.526416	0	Inf	0.00321241	0.41864812	GSBRNA2T00154087001	BnaA10g13900D
1.77130833	0	Inf	0.00497502	0.44909916	XLOC_005721	-
1.60051733	0	Inf	0.0051784	0.44909916	GSBRNA2T00030919001	BnaA02g25850D
22.0209177	0	Inf	0.00549746	0.44909916	GSBRNA2T00098761001	BnaA07g16530D
5.99724933	0	Inf	0.0058469	0.44909916	GSBRNA2T00152258001	BnaC09g26870D
1.671269	0	Inf	0.00862326	0.46866673	GSBRNA2T00016810001	BnaCnng28740D
3.27207133	0	Inf	0.0113562	0.52107495	GSBRNA2T00091733001	BnaC02g34090D
2.06347567	0	Inf	0.01276009	0.53320818	GSBRNA2T00047515001	BnaA02g34080D
8.61317367	0	Inf	0.01492786	0.55052658	GSBRNA2T00005574001	BnaA01g26500D
49.3951307	0	Inf	0.02567701	0.58619732	GSBRNA2T00082228001	BnaCnng12930D
1.37679433	0	Inf	0.02723169	0.58619732	GSBRNA2T00153897001	BnaA05g22170D
185.76289	0	Inf	0.0311497	0.58619732	GSBRNA2T00043080001	BnaC05g32140D
9.424608	0	Inf	0.03789849	0.58619732	GSBRNA2T00043080001	BnaC05g32140D
2.737717	0	Inf	0.03833267	0.58619732	GSBRNA2T00011127001	BnaA08g30370D
2.21992333	0	Inf	0.0412534	0.58619732	GSBRNA2T00018571001	BnaC01g39340D
5.07774	0.000278	14.156814	0.011243	0.52107495	GSBRNA2T00158381001	BnaC03g43360D
2.66664833	0.00202667	10.361703	0.00265525	0.41833439	GSBRNA2T00140972001	BnaC03g11130D
3.24015833	0.00395567	9.67792774	0.04028088	0.58619732	GSBRNA2T00038928001	BnaC05g44040D
2.02018267	0.01335	7.2415022	0.01308346	0.53320818	GSBRNA2T00150475001	BnaA10g01010D
2.054733	0.01415833	7.18115568	0.03468992	0.58619732	GSBRNA2T00006316001	BnaC01g20180D
4.00410167	0.08715367	5.52177343	0.01079674	0.50987101	GSBRNA2T00011113001	BnaC09g54000D
24.837755	1.06429	4.54457156	0.0143349	0.54311929	GSBRNA2T00015388001	BnaA01g34180D
15.027747	0.699036	4.42611817	0.00517909	0.44909916	GSBRNA2T00134990001	BnaC09g08190D
1.799809	0.09981267	4.17247709	0.02050578	0.58339905	GSBRNA2T00028876001	BnaC02g39660D
2.77114033	0.18184433	3.9297039	0.00128396	0.36562033	GSBRNA2T00018477001	BnaA09g01200D
2.12693333	0.16459133	3.69181454	0.04326457	0.58619732	GSBRNA2T00085702001	BnaC04g42270D
2.161741	0.245608	3.13776422	0.01621452	0.56616348	GSBRNA2T00066147001	BnaC03g26830D
2.91375633	0.37520133	2.95714338	0.00999048	0.49481244	GSBRNA2T00142895001	BnaC07g25650D
4.94965333	0.67282267	2.87902927	0.02222257	0.58619732	GSBRNA2T00092904001	BnaA08g29920D
18.9416287	3.52914267	2.42417073	0.00541689	0.44909916	GSBRNA2T00082187001	BnaCnng12690D
2.06743033	0.45290033	2.19057321	0.02321087	0.58619732	GSBRNA2T00106254001	BnaC05g37960D
1.96714067	0.461153	2.09278274	0.02604651	0.58619732	GSBRNA2T00111546001	BnaA03g32040D
3.938827	0.92400767	2.09178933	0.00025191	0.32843561	GSBRNA2T00127186001	BnaC08g32880D

## 9 Supplements

3.487138	0.878864	1.98833162	0.04291983	0.58619732	GSBRNA2T00156706001	BnaC04g20490D
3.79425633	1.01459033	1.90291983	0.02948023	0.58619732	XLOC_012718	PSBK
2.594978	0.72843567	1.83284884	0.00817794	0.46713757	GSBRNA2T00104311001	BnaA02g20640D
2.34596433	0.65997733	1.8296927	0.02331025	0.58619732	GSBRNA2T00072672001	BnaA01g12900D
1.896157	0.549504	1.78687653	0.04684432	0.58619732	GSBRNA2T00121706001	BnaC08g20810D
2.72239667	0.80749467	1.75335265	0.0060382	0.44909916	GSBRNA2T00140087001	BnaC08g07860D
2.53639433	0.81179667	1.64358874	0.00343648	0.41864812	GSBRNA2T00012883001	BnaA02g08330D
3.77420533	1.213301	1.63723541	0.0288354	0.58619732	GSBRNA2T00014258001	BnaC07g49260D
4.082539	1.34259967	1.60443748	0.01134662	0.52107495	GSBRNA2T00129637001	BnaA03g02010D
12.0709207	4.072743	1.56746303	0.01239058	0.53004406	GSBRNA2T00129360001	BnaA09g32640D
2.47580033	0.842663	1.55486728	0.02177628	0.58619732	GSBRNA2T00066147001	BnaC03g26830D
5.84814133	2.05169933	1.51115885	0.02126066	0.58619732	GSBRNA2T00042530001	BnaC08g46660D
3.80690833	1.43718167	1.40537739	0.01959688	0.57944672	GSBRNA2T00048505001	BnaA05g12250D
2.86263133	1.09294	1.38912768	0.04939884	0.58619732	GSBRNA2T00145553001	BnaA09g22010D
2.20947333	0.86103233	1.3595632	0.00545201	0.44909916	GSBRNA2T00038110001	BnaC04g51050D
17.1072653	6.699113	1.35256717	0.00836799	0.46713757	GSBRNA2T00118848001	BnaC09g25620D
7.70630367	3.285986	1.2297127	0.00418018	0.44909916	GSBRNA2T00077487001	BnaA07g01850D
2.78952367	1.190732	1.22817005	0.02323761	0.58619732	GSBRNA2T00019122001	BnaC03g23440D
1.902023	0.829965	1.19641229	0.02735153	0.58619732	GSBRNA2T00026265001	BnaC01g37620D
2.70539433	1.186323	1.18934203	0.02612968	0.58619732	GSBRNA2T00093873001	BnaCnng18210D
3.90500367	1.77467867	1.13776608	0.0399167	0.58619732	GSBRNA2T00082208001	BnaCnng12810D
3.492589	1.59798667	1.12804151	0.021732	0.58619732	GSBRNA2T00125378001	BnaA06g16230D
4.21986633	1.942016	1.11964221	0.01311046	0.53320818	GSBRNA2T00046820001	BnaCnng47740D
27.1040147	12.5386713	1.11212208	0.03195344	0.58619732	GSBRNA2T00100826001	BnaC09g16500D
1.69236367	0.79422533	1.09141933	0.0468568	0.58619732	GSBRNA2T00123215001	BnaA04g23940D
6.785152	3.25744733	1.05863928	0.02994437	0.58619732	GSBRNA2T00147165001	BnaC06g09740D
4.12227367	1.99467267	1.04728827	0.00148973	0.36562033	GSBRNA2T00084031001	BnaA08g06520D
3.605303	1.77069833	1.02580207	0.03440027	0.58619732	GSBRNA2T00016084001	BnaC09g46780D

**Table 35: List of all through BnGRP7 enriched phloem mRNAs of *B. napus* compared to total phloem sap RNA of *B. napus*.**

gene_id	GRP7	Input	log2FoldChange	pvalue	gene_name
GSBRNA2T00052496001	35.0865381	2.78E-17	8.13680007	0.00108644	BnaA09g50830D
GSBRNA2T00016101001	33.6164523	2.78E-17	8.07673797	1.13E-07	BnaC09g46650D
GSBRNA2T00038673001	31.8118763	2.78E-17	7.99628489	0.00054664	BnaC09g28670D
GSBRNA2T00148693001	24.1425432	2.78E-17	7.60072992	0.00010494	BnaA03g09990D
GSBRNA2T00000538001	23.5118165	2.78E-17	7.56319861	1.61E-05	BnaC04g19880D
GSBRNA2T00111902001	21.8956772	2.78E-17	7.46121369	7.15E-05	BnaA03g29040D
GSBRNA2T00026930001	21.1829164	2.78E-17	7.41336331	0.00069732	BnaC05g32740D
GSBRNA2T00123887001	19.1816753	2.78E-17	7.27125569	8.26E-05	BnaC03g39700D
GSBRNA2T00033871001	18.8092601	2.78E-17	7.24312541	0.00014435	BnaC01g43230D
GSBRNA2T00135381001	18.8098458	2.78E-17	7.24303877	0.00021464	BnaA10g17620D
GSBRNA2T00006880001	17.4491361	2.78E-17	7.13538481	0.00044758	BnaCnng69410D
GSBRNA2T00035732001	16.8712536	2.78E-17	7.08873102	0.00208017	BnaC09g09220D

## 9 Supplements

GSBRNA2T00068676001	16.8046968	2.78E-17	7.08166176	0.00037048	BnaC02g42670D
GSBRNA2T00156711001	16.1938342	2.78E-17	7.02961213	0.00150349	BnaC04g20530D
GSBRNA2T00077896001	15.7947541	2.78E-17	6.99294029	0.00051378	BnaC02g23620D
GSBRNA2T00041483001	15.7789331	2.78E-17	6.99146271	0.00060128	BnaA02g10130D
GSBRNA2T00018640001	15.1239912	2.78E-17	6.93089573	0.00048392	BnaC01g38850D
GSBRNA2T00101603001	14.8394634	2.78E-17	6.90414971	0.0010073	BnaC03g25880D
GSBRNA2T00015980001	14.8336417	2.78E-17	6.90376497	0.00246105	BnaC09g47600D
GSBRNA2T00047986001	14.7868057	2.78E-17	6.89848664	0.00127292	BnaC08g48600D
ENSRNA049472815	14.7671041	2.78E-17	6.89633983	0.00169396	U1
GSBRNA2T00142066001	14.1386195	2.78E-17	6.83445695	0.00073391	BnaA03g14570D
GSBRNA2T00136426001	13.4330961	2.78E-17	6.76130064	0.00141545	BnaA10g26300D
GSBRNA2T00071552001	13.4209278	2.78E-17	6.75983508	0.00233316	BnaC03g33890D
GSBRNA2T00157603001	13.1401345	2.78E-17	6.72984412	0.00150468	BnaC01g17360D
GSBRNA2T00131595001	13.1198381	2.78E-17	6.72754688	0.00201616	BnaA01g11290D
GSBRNA2T00013777001	13.0950379	2.78E-17	6.72473677	0.00253937	BnaC01g41930D
GSBRNA2T00152643001	12.469946	2.78E-17	6.65510169	0.00216787	BnaC04g07960D
GSBRNA2T00069089001	12.102419	2.78E-17	6.61228379	0.00202302	BnaC06g05340D
GSBRNA2T00042111001	11.4439482	2.78E-17	6.53234666	0.00261211	BnaA06g25520D
GSBRNA2T00135021001	31.3005714	0.97006377	4.83979004	6.46E-05	BnaC09g08490D
GSBRNA2T00136267001	225.487277	8.74985533	4.66714485	0.00011816	BnaA10g25030D
GSBRNA2T00044735001	23.8998295	0.97006377	4.45238482	0.00059225	BnaC05g37060D
GSBRNA2T00008912001	23.8570558	0.97006377	4.44939571	0.00114321	BnaC05g08840D
GSBRNA2T00142084001	393.857204	18.474526	4.40425906	0.00036473	BnaA03g14400D
GSBRNA2T00133324001	21.8822413	0.97006377	4.32573984	0.00023019	BnaC09g34300D
GSBRNA2T00086702001	20.526605	0.97006377	4.23399306	0.00043294	BnaA03g12180D
GSBRNA2T00144575001	19.842097	0.97006377	4.18537259	0.00033053	BnaA06g31490D
GSBRNA2T00156018001	18.180494	0.97006377	4.05998897	0.000669	BnaC08g26860D
GSBRNA2T00086759001	18.1379972	0.97006377	4.0566547	0.00105139	BnaC04g13780D
GSBRNA2T00076840001	17.8256563	0.97006377	4.0317383	0.00135571	BnaA01g06950D
GSBRNA2T00113451001	17.4745466	0.97006377	4.00326105	0.00127723	BnaA03g40900D
GSBRNA2T00053646001	17.1468921	0.97006377	3.97616762	0.00118876	BnaAnng22250D
GSBRNA2T00127138001	17.1207609	0.97006377	3.97380904	0.00248316	BnaC08g32510D
GSBRNA2T00014288001	31.9068522	1.94254108	3.9516908	6.64E-05	BnaCnng27240D
GSBRNA2T00065371001	16.4800251	0.97006377	3.91935396	0.00144031	BnaC02g06730D
ENSRNA049473092	29.9251557	1.94254108	3.8596098	0.00027269	mir-393
GSBRNA2T00132634001	25.2344548	1.94254108	3.61484656	0.00039461	BnaA05g03530D
GSBRNA2T00095755001	24.8908986	1.94254108	3.59516834	0.00020225	BnaC03g65280D
GSBRNA2T00015544001	24.5627828	1.94254108	3.57610883	0.00022754	BnaC02g28750D
GSBRNA2T00153531001	23.8634034	1.94254108	3.53440305	0.00264386	BnaC03g32060D
GSBRNA2T00063732001	34.5671584	2.91502413	3.51099834	0.00033329	BnaA09g08650D
GSBRNA2T00040553001	23.1535724	1.94254108	3.49097101	0.00183185	BnaC02g19520D
GSBRNA2T00056212001	22.4984412	1.94254108	3.44992224	0.00161977	BnaA01g24330D
GSBRNA2T00035277001	22.4809876	1.94254108	3.4487803	0.0017397	BnaC09g41950D
GSBRNA2T00102482001	22.1832497	1.94254108	3.42979217	0.0013091	BnaA07g29430D
GSBRNA2T00147163001	28.6564701	2.91502413	3.24216837	0.00060588	BnaC06g09720D

## 9 Supplements

GSBRNA2T00111322001	46.4205466	4.85997558	3.22230054	2.14E-06	BnaA04g06100D
GSBRNA2T00090457001	27.8915429	2.91502413	3.20301278	0.00026028	BnaC08g40640D
GSBRNA2T00099953001	81.5886725	8.74985533	3.20232924	3.40E-08	BnaC07g44660D
GSBRNA2T00121705001	18.8550737	1.94254108	3.19678524	0.00209283	BnaC08g20800D
GSBRNA2T00051863001	18.4916108	1.94254108	3.16892773	0.00244737	BnaC01g36620D
GSBRNA2T00079400001	58.175166	6.80491749	3.072027	7.36E-05	BnaA05g25970D
GSBRNA2T00029472001	56.1368363	6.80491749	3.02081038	1.48E-06	BnaA05g19310D
GSBRNA2T00018157001	23.5254261	2.91502413	2.95860052	0.00108794	BnaC01g35940D
GSBRNA2T00046823001	23.2005275	2.91502413	2.93858507	0.00155491	BnaCnng47760D
GSBRNA2T00077926001	22.2233376	2.91502413	2.87685624	0.00191279	BnaC02g23710D
GSBRNA2T00134389001	28.9128703	3.88750167	2.85443685	0.00121132	BnaC03g02200D
GSBRNA2T00141157001	21.5317588	2.91502413	2.83149426	0.00187931	BnaC04g09040D
GSBRNA2T00095364001	41.0732708	5.83244725	2.78917742	4.58E-05	BnaA10g09110D
GSBRNA2T00011038001	26.5295412	3.88750167	2.73083063	0.00215433	BnaA05g34630D
GSBRNA2T00122439001	25.9387636	3.88750167	2.6986547	0.00176906	BnaC05g31040D
GSBRNA2T00135339001	44.8600265	6.80491749	2.6984866	0.00070028	BnaC02g19080D
GSBRNA2T00097666001	25.2583595	3.88750167	2.66043113	0.00174955	BnaCnng67540D
GSBRNA2T00021972001	54.8033626	8.74985533	2.62930022	0.00027324	BnaC04g48990D
GSBRNA2T00137925001	24.5620597	3.88750167	2.62028901	0.00183084	BnaC07g01890D
GSBRNA2T00155507001	36.2987404	5.83244725	2.6115047	0.00048383	BnaC08g30830D
GSBRNA2T00134172001	41.7140242	6.80491749	2.59343124	0.00013967	BnaC03g00470D
GSBRNA2T00005964001	29.5720986	4.85997558	2.5739495	0.00147918	BnaAnng02640D
GSBRNA2T00031031001	69.6477057	11.6672586	2.56450666	6.55E-07	BnaC02g20300D
ENSRNA049472669	61.9092655	10.6947911	2.51905699	8.04E-06	snoZ102_R77
GSBRNA2T00088620001	45.1559928	7.77738675	2.51814368	0.0002183	BnaA04g24690D
GSBRNA2T00057239001	67.3357026	11.6672586	2.51589343	1.18E-05	BnaA01g33070D
GSBRNA2T00078049001	28.2971403	4.85997558	2.51075348	0.0024737	BnaA06g32890D
GSBRNA2T00158304001	484.635722	85.5746736	2.49998859	1.10E-05	BnaC03g43740D
GSBRNA2T00095044001	132.075325	23.3368576	2.49494299	0.00217694	BnaA08g04270D
GSBRNA2T00070905001	27.9587601	4.85997558	2.49338904	0.00189321	BnaA09g17440D
GSBRNA2T00133690001	26.8999388	4.85997558	2.43795394	0.00185648	BnaC07g28760D
GSBRNA2T00076989001	31.9269797	5.83244725	2.42711117	0.00098972	BnaA04g20460D
GSBRNA2T00097727001	276.941346	52.5108365	2.39605631	7.75E-06	BnaCnng19840D
ENSRNA049447570	107.693039	20.4194588	2.39166786	2.40E-06	snoR31
GSBRNA2T00124973001	29.6030925	5.83244725	2.31846022	0.00235076	BnaC03g47950D
GSBRNA2T00112871001	130.924771	26.2542561	2.31266463	1.40E-07	BnaC06g27900D
GSBRNA2T00048262001	41.0598035	8.74985533	2.21391994	0.00119085	BnaC04g52430D
GSBRNA2T00030731001	40.999869	8.74985533	2.21175221	0.00216698	BnaC06g43720D
GSBRNA2T00112379001	30.9554595	6.80491749	2.1645696	0.00236725	BnaA07g11670D
GSBRNA2T00076283001	56.6773148	12.6397257	2.15319837	0.00209975	BnaC05g48450D
GSBRNA2T00139283001	52.1307058	11.6672586	2.14747623	0.00014369	BnaC07g23580D
GSBRNA2T00097591001	42.6741038	9.72232342	2.11941975	0.00146329	BnaC08g47940D
GSBRNA2T00022405001	67.9978597	15.5571264	2.11871329	1.22E-05	BnaAnng14770D
GSBRNA2T00055394001	80.6499019	18.474526	2.11859647	4.92E-06	BnaA02g30950D
GSBRNA2T00132617001	42.4277039	9.72232342	2.11102275	0.00067514	BnaA05g03670D

## 9 Supplements

GSBRNA2T00086962001	80.1364833	18.474526	2.10912982	3.35E-06	BnaA02g29290D
GSBRNA2T00025002001	46.483456	10.6947911	2.10656534	0.0006718	BnaCnng33540D
GSBRNA2T00080896001	91.391868	21.3919251	2.08857714	2.01E-06	BnaC09g50490D
GSBRNA2T00071217001	37.3559047	8.74985533	2.07795113	0.00208301	BnaC09g23880D
GSBRNA2T00117730001	65.9548637	15.5571264	2.07483097	2.24E-05	BnaA09g18740D
GSBRNA2T00093053001	45.3356144	10.6947911	2.07064822	0.00130288	BnaA10g06540D
GSBRNA2T00019370001	40.3393225	9.72232342	2.03853144	0.00143388	BnaC08g43310D
GSBRNA2T00037314001	106.741363	26.2542561	2.01805227	5.08E-07	BnaC04g50320D
GSBRNA2T00062672001	39.0291632	9.72232342	1.99098765	0.00167757	BnaC04g32200D
GSBRNA2T00087238001	38.3305705	9.72232342	1.96507689	0.00180018	BnaCnng14940D
GSBRNA2T00087150001	117.776079	30.1441204	1.9614895	1.20E-07	BnaAnng08870D
GSBRNA2T00114679001	49.3967236	12.6397257	1.95566444	0.00117028	BnaC02g08560D
GSBRNA2T00112552001	53.1393437	13.6121928	1.9548214	0.00103241	BnaC07g35160D
GSBRNA2T00114211001	105.606319	27.2267222	1.95062593	4.52E-07	BnaA08g01410D
GSBRNA2T00030811001	127.964426	33.0615184	1.94834474	0.00011109	BnaA02g26590D
GSBRNA2T00096994001	67.7486634	17.5020595	1.94482919	0.00019118	BnaA01g30780D
GSBRNA2T00044576001	55.8098172	14.5846596	1.92682025	0.00017031	BnaC05g09810D
GSBRNA2T00026137001	59.414917	15.5571264	1.92452158	0.0006582	BnaCnng76860D
GSBRNA2T00150749001	155.065632	40.841246	1.92147434	2.49E-08	BnaA08g17880D
GSBRNA2T00009467001	646.56631	173.09659	1.9003989	3.05E-05	BnaA09g03840D
GSBRNA2T00116032001	54.5784492	14.5846596	1.89454157	0.00037335	BnaC01g16880D
GSBRNA2T00133048001	54.4047027	14.5846596	1.89013079	0.00058194	BnaA05g00160D
GSBRNA2T00093806001	64.5005869	17.5020595	1.87424609	0.0001889	BnaCnng17990D
GSBRNA2T00097532001	42.7225184	11.6672586	1.86107233	0.00151224	BnaCnng19750D
GSBRNA2T00083128001	39.0172942	10.6947911	1.85473856	0.00256707	BnaA02g17130D
GSBRNA2T00138905001	53.0538041	14.5846596	1.85388575	0.00118183	BnaA03g21630D
GSBRNA2T00127487001	109.239411	30.1441204	1.85326958	1.57E-06	BnaC08g35330D
GSBRNA2T00064626001	66.8930972	18.474526	1.84920317	9.18E-05	BnaA04g19430D
GSBRNA2T00136443001	56.21835	15.5571264	1.84482621	0.00028799	BnaA10g26410D
GSBRNA2T00002403001	90.9877243	25.2817899	1.8424543	6.02E-05	BnaC03g30870D
GSBRNA2T00129191001	118.73174	33.0615184	1.84048535	8.30E-07	BnaA09g31350D
GSBRNA2T00111452001	86.7116099	24.3093238	1.8293637	3.76E-05	BnaC05g24590D
GSBRNA2T00143671001	58.9416698	16.529593	1.82611538	0.00033389	BnaC06g22610D
GSBRNA2T00072873001	41.3576856	11.6672586	1.81441916	0.0023331	BnaA01g31970D
GSBRNA2T00110055001	72.0861272	20.4194588	1.81322009	0.00020686	BnaC07g14100D
GSBRNA2T00094652001	299.957269	86.5471394	1.79189387	3.45E-10	BnaA09g06080D
GSBRNA2T00110412001	50.7625708	14.5846596	1.79038553	0.00066861	BnaC04g23060D
GSBRNA2T00154456001	60.5245984	17.5020595	1.78256972	0.00023561	BnaC07g39040D
GSBRNA2T00105116001	310.136599	90.4370025	1.77667192	1.59E-08	BnaC04g41300D
GSBRNA2T00103681001	80.0372873	23.3368576	1.7725008	6.77E-05	BnaC06g36360D
GSBRNA2T00053224001	46.732114	13.6121928	1.76998954	0.00174073	BnaA10g12080D
GSBRNA2T00098845001	69.0631383	20.4194588	1.75151702	0.00034122	BnaA07g17140D
GSBRNA2T00114785001	52.4718781	15.5571264	1.7456457	0.00067272	BnaC02g07820D
GSBRNA2T00045076001	84.8763887	25.2817899	1.74208811	0.00025704	BnaA04g24840D
GSBRNA2T00121240001	51.7696266	15.5571264	1.72625953	0.00166571	BnaA01g03420D

## 9 Supplements

GSBRNA2T00142080001	106.365549	32.0890524	1.72474878	5.28E-06	BnaA03g14430D
GSBRNA2T00000494001	119.08109	35.9789164	1.72314993	2.56E-06	BnaC04g19660D
GSBRNA2T00078137001	64.3840981	19.4469924	1.72046737	0.00073283	BnaA06g32110D
GSBRNA2T00019291001	83.2774723	25.2817899	1.71486397	0.00027038	BnaC08g42560D
GSBRNA2T00011042001	57.0851118	17.5020595	1.69832033	0.00183637	BnaA05g34660D
GSBRNA2T00071576001	59.9116602	18.474526	1.69033994	0.0007021	BnaC03g34080D
GSBRNA2T00012504001	212.613909	66.125358	1.68313685	0.0002261	BnaC04g06110D
GSBRNA2T00087562001	237.558949	73.9050843	1.68277938	8.55E-09	BnaAnng09020D
GSBRNA2T00065595001	68.8893362	21.3919251	1.6813818	0.00036539	BnaC05g17630D
GSBRNA2T00142136001	93.8418219	29.1716543	1.68133123	2.31E-05	BnaA03g14060D
GSBRNA2T00121333001	53.0701221	16.529593	1.67521194	0.00245976	BnaA01g02600D
GSBRNA2T00103738001	84.0116618	26.2542561	1.67330033	0.00021479	BnaC06g36840D
GSBRNA2T00026368001	52.8536001	16.529593	1.66921908	0.00128146	BnaC01g38220D
GSBRNA2T00059849001	102.162526	32.0890524	1.666838	0.00013956	BnaC03g41620D
GSBRNA2T00044866001	58.5084662	18.474526	1.65633312	0.00059965	BnaA03g49330D
GSBRNA2T00059912001	73.7854398	23.3368576	1.65520088	0.00042876	BnaA08g19700D
GSBRNA2T00056373001	70.2573704	22.3643914	1.64584132	0.00101156	BnaA01g14370D
GSBRNA2T00096956001	60.5874672	19.4469924	1.63297177	0.00099182	BnaA01g30470D
GSBRNA2T00091437001	63.4818202	20.4194588	1.63039626	0.00100237	BnaAnng31140D
GSBRNA2T00100334001	48.0577246	15.5571264	1.61924916	0.00261274	BnaC07g47570D
ENSRNA049467007	56.8045968	18.474526	1.61382648	0.00110646	snoR86
GSBRNA2T00109531001	61.8325445	20.4194588	1.59246401	0.00120014	BnaC09g02570D
GSBRNA2T00158003001	134.086085	44.7311096	1.58124152	7.21E-06	BnaC07g09910D
GSBRNA2T00019947001	67.1098513	22.3643914	1.57986209	0.00257843	BnaC01g38700D
GSBRNA2T00015694001	154.273158	51.5383707	1.5795566	1.04E-06	BnaA09g30860D
GSBRNA2T00087829001	52.4271733	17.5020595	1.57585588	0.00223884	BnaA01g13490D
GSBRNA2T00016054001	63.2211235	21.3919251	1.55762836	0.00239639	BnaC09g47040D
GSBRNA2T00077487001	165.319724	56.4006999	1.54929858	0.00068889	BnaA07g01850D
GSBRNA2T00124712001	8561.26288	2926.14697	1.54879286	8.87E-19	BnaA03g18970D
GSBRNA2T00123386001	2077.90056	712.815054	1.54339575	3.65E-20	BnaA04g22480D
GSBRNA2T00039891001	93.3985089	32.0890524	1.53771778	0.00018527	BnaA01g06730D
GSBRNA2T00088707001	127.096052	43.7586437	1.53556654	5.90E-05	BnaA09g23070D
GSBRNA2T00152838001	75.9976576	26.2542561	1.52883318	0.00066391	BnaC04g06680D
GSBRNA2T00125165001	2412.11399	836.318197	1.52808638	1.62E-17	BnaC03g35680D
GSBRNA2T00135223001	55.7816088	19.4469924	1.51416182	0.00251844	BnaC02g18420D
GSBRNA2T00149321001	58.5152931	20.4194588	1.51301924	0.00140124	BnaA04g00200D
GSBRNA2T00110574001	1415.45559	497.900135	1.50711062	1.32E-18	BnaC03g37060D
GSBRNA2T00101646001	55.4999676	19.4469924	1.50677589	0.00194751	BnaC03g26210D
GSBRNA2T00028682001	55.4813473	19.4469924	1.50634164	0.00190926	BnaC03g22720D
GSBRNA2T00026201001	63.247231	22.3643914	1.49444223	0.00105611	BnaCnng35070D
GSBRNA2T00004531001	542.717601	192.545904	1.49434895	1.74E-14	BnaA08g11760D
GSBRNA2T00155160001	2022.20507	718.649848	1.49241363	1.69E-18	BnaC04g46860D
GSBRNA2T00079660001	62.6013529	22.3643914	1.47960237	0.00131668	BnaC01g22320D
GSBRNA2T00128345001	65.3021293	23.3368576	1.4793336	0.00107409	BnaA07g06160D
GSBRNA2T00030956001	99.1685574	35.9789164	1.4595763	0.00016418	BnaC02g19940D

## 9 Supplements

GSBRNA2T00005797001	66.9451507	24.3093238	1.45661661	0.00113389	BnaA09g06820D
GSBRNA2T00100301001	74.6677142	27.2267222	1.45114789	0.00168551	BnaC07g47300D
GSBRNA2T00047470001	119.849837	43.7586437	1.45093278	0.00151347	BnaA02g34380D
GSBRNA2T00053093001	332.995812	122.528372	1.44162969	1.06E-08	BnaC02g11980D
GSBRNA2T00032693001	2954.27789	1088.18681	1.44080457	9.93E-11	BnaC09g39910D
GSBRNA2T00144323001	84.3514859	31.1165864	1.43508867	0.00089546	BnaA06g29450D
GSBRNA2T00036893001	86.8213604	32.0890524	1.43226798	0.00031094	BnaA02g10950D
GSBRNA2T00078540001	1195.05039	443.442056	1.43007222	1.23E-15	BnaC02g01330D
GSBRNA2T00107829001	94.013758	35.0064504	1.42187596	0.00164254	BnaC01g10760D
GSBRNA2T00039025001	552.443054	206.160424	1.42167192	2.14E-11	BnaC05g44770D
ENSRNA049470168	75.3585101	28.1991883	1.41399534	0.00198999	Plant_U3
GSBRNA2T00144398001	83.1000398	31.1165864	1.41342245	0.00042898	BnaA06g30030D
GSBRNA2T00103524001	1510.24627	569.862596	1.40595563	5.40E-17	BnaC06g35200D
GSBRNA2T00023744001	4290.67286	1624.01541	1.40158867	1.93E-17	BnaC05g40380D
GSBRNA2T00121978001	115.456489	43.7586437	1.39695734	7.35E-05	BnaA09g36760D
GSBRNA2T00118778001	61.5552056	23.3368576	1.39435643	0.00227251	BnaA07g14070D
GSBRNA2T00040689001	107.334208	40.841246	1.39116455	0.00016745	BnaC09g32450D
GSBRNA2T00008121001	86.7271034	33.0615184	1.3879606	0.00089355	BnaA04g29430D
GSBRNA2T00109532001	116.94119	44.7311096	1.38399815	0.00038679	BnaC09g02560D
GSBRNA2T0007725001	179.475212	69.0427554	1.37670823	1.96E-05	BnaA03g51520D
GSBRNA2T00001498001	106.236262	40.841246	1.37646074	0.00050036	BnaA02g06700D
GSBRNA2T00034770001	98.6113815	37.9238483	1.37556283	0.00032671	BnaC06g10690D
GSBRNA2T00009741001	138.908098	53.4833024	1.37488396	1.86E-05	BnaA09g03140D
GSBRNA2T00072871001	103.536439	39.8687801	1.37408756	0.00022279	BnaA01g31980D
GSBRNA2T00102075001	95.9590986	36.9513823	1.37360687	0.00077539	BnaA07g26930D
GSBRNA2T00122772001	131.129547	50.5659048	1.37259894	5.28E-05	BnaC02g13470D
GSBRNA2T00085176001	132.939172	51.5383707	1.36480273	0.00033269	BnaC03g44220D
GSBRNA2T00074936001	75.2716631	29.1716543	1.36375534	0.00233587	BnaC04g21820D
GSBRNA2T00043692001	135.232896	52.5108365	1.36260286	3.81E-05	BnaC05g31880D
GSBRNA2T00102807001	132.438067	51.5383707	1.35955143	5.68E-05	BnaA08g21500D
GSBRNA2T00058350001	75.0134986	29.1716543	1.35870282	0.00183289	BnaA04g14740D
GSBRNA2T00043883001	296.914231	115.721112	1.35851662	2.11E-08	BnaA05g37090D
GSBRNA2T00074423001	92.410538	35.9789164	1.35792648	0.00068317	BnaA01g25840D
GSBRNA2T00127150001	218.673442	85.5746736	1.35219733	8.39E-07	BnaC08g32600D
GSBRNA2T00096991001	1171.05799	459.973972	1.34805055	8.15E-13	BnaA01g30750D
GSBRNA2T00094565001	1462.51572	579.587253	1.33517749	2.84E-16	BnaA09g20840D
GSBRNA2T00090601001	105.601015	41.8137119	1.33393685	0.00046602	BnaA09g28940D
GSBRNA2T00087818001	124.293966	49.593439	1.32350053	0.00021809	BnaA01g13550D
GSBRNA2T00032831001	496.152212	198.380698	1.32210258	1.17E-10	BnaC09g11950D
GSBRNA2T00055533001	99.7119215	39.8687801	1.31964378	0.00248048	BnaAnng22690D
GSBRNA2T00112685001	106.715573	42.7861779	1.31582099	0.00028963	BnaC07g36190D
GSBRNA2T00040549001	108.904038	43.7586437	1.31303151	0.00035718	BnaC02g19480D
GSBRNA2T00108933001	113.642625	45.7035755	1.31180251	0.00025562	BnaC08g25670D
GSBRNA2T00136884001	598.820716	241.16919	1.3116463	0.00028063	BnaC08g36790D
GSBRNA2T00145263001	256.00968	103.079057	1.31137124	3.64E-06	BnaC08g00900D

## 9 Supplements

GSBRNA2T00089234001	287.062516	115.721112	1.30992777	9.50E-08	BnaCnng64430D
GSBRNA2T00014258001	594.271099	242.141655	1.29499173	4.59E-09	BnaC07g49260D
GSBRNA2T00125378001	124.115091	50.5659048	1.29330204	8.50E-05	BnaA06g16230D
GSBRNA2T00100306001	73.9898829	30.1441204	1.29184842	0.00191327	BnaC07g47350D
GSBRNA2T00116606001	169.160974	69.0427554	1.2912048	0.0013249	BnaA05g04020D
GSBRNA2T00081912001	99.9065193	40.841246	1.28787816	0.00041872	BnaA07g31820D
ENSRNA049468396	329.228103	135.170427	1.28357027	7.75E-08	snoZ103
GSBRNA2T00092302001	149.272146	61.263029	1.28319888	4.69E-05	BnaC06g12400D
GSBRNA2T00129553001	87.4404502	35.9789164	1.27816845	0.00098925	BnaA09g34250D
GSBRNA2T00146665001	4551.35074	1879.77389	1.27570291	2.00E-09	BnaA07g32150D
GSBRNA2T00070774001	176.559249	72.9326186	1.27409363	1.25E-05	BnaC07g03030D
GSBRNA2T00115235001	82.4696266	34.0339844	1.2735866	0.00181696	BnaC04g25340D
GSBRNA2T00096888001	150.583966	62.2354948	1.2731859	4.58E-05	BnaA10g06360D
GSBRNA2T00065926001	93.6542546	38.8963142	1.26473402	0.00154446	BnaA10g16790D
GSBRNA2T00012129001	134.153068	56.4006999	1.24824395	0.0005788	BnaAnng36550D
GSBRNA2T00137206001	108.726167	45.7035755	1.24786038	0.00070233	BnaC02g21410D
GSBRNA2T00093632001	105.964563	44.7311096	1.24183844	0.0005707	BnaCnng65950D
GSBRNA2T00124048001	464.301298	196.435767	1.24054148	9.16E-08	BnaC04g19210D
GSBRNA2T00013888001	233.597813	99.1891942	1.2347301	0.00176643	BnaCnng26720D
GSBRNA2T00108932001	918.561171	390.928908	1.23232932	5.21E-10	BnaC08g25680D
GSBRNA2T00097596001	125.741344	53.4833024	1.23141391	0.0002812	BnaC08g47990D
GSBRNA2T00099185001	748.320492	318.966446	1.22996402	6.18E-12	BnaA07g20130D
GSBRNA2T00031865001	186.847561	79.739879	1.22728029	0.00018056	BnaC08g43830D
GSBRNA2T00130029001	122.42898	52.5108365	1.21924866	0.00162011	BnaA03g05270D
GSBRNA2T00017630001	1959.18058	843.125457	1.21632494	7.93E-15	BnaA09g56740D
GSBRNA2T00119945001	219.247087	94.3268655	1.21591344	0.0001152	BnaC05g46550D
GSBRNA2T00128317001	113.077064	48.6209731	1.21539731	0.00038774	BnaA07g05920D
GSBRNA2T00013988001	106.228087	45.7035755	1.21459481	0.00099843	BnaCnng26880D
GSBRNA2T00057073001	2860.05525	1234.05667	1.21256462	5.80E-06	BnaC08g41800D
GSBRNA2T00143777001	90.1132279	38.8963142	1.2094812	0.00165476	BnaC06g21960D
GSBRNA2T00129360001	258.647685	111.831249	1.20889645	4.93E-05	BnaA09g32640D
GSBRNA2T00049542001	94.5388129	40.841246	1.20827621	0.00112975	BnaA09g40370D
GSBRNA2T00053229001	148.316368	64.1804264	1.20689434	6.94E-05	BnaA10g12130D
GSBRNA2T00014357001	98.8926181	42.7861779	1.20627383	0.00106785	BnaAnng37060D
GSBRNA2T00123790001	280.312471	121.555906	1.20448177	2.82E-07	BnaC03g39130D
GSBRNA2T00044079001	4502.62345	1960.48854	1.19952234	1.35E-10	BnaC07g47800D
GSBRNA2T00072247001	307.986862	134.197961	1.19784919	2.31E-07	BnaA02g27670D
GSBRNA2T00152426001	138.274947	60.2905632	1.19578194	0.00028499	BnaC02g26450D
GSBRNA2T00140451001	1001.1325	437.607261	1.19373817	2.28E-12	BnaC03g06910D
GSBRNA2T00120481001	104.654233	45.7035755	1.19289591	0.00077719	BnaA02g31520D
GSBRNA2T00104934001	166.680722	72.9326186	1.19108114	0.0014493	BnaA05g23900D
GSBRNA2T00065649001	5270.74844	2312.52112	1.18850504	5.04E-14	BnaA10g11350D
GSBRNA2T00060804001	90.5598915	39.8687801	1.18086525	0.00194974	BnaC03g13810D
GSBRNA2T00137917001	109.765252	48.6209731	1.17253395	0.00189962	BnaC07g01970D
GSBRNA2T00127500001	96.5044338	42.7861779	1.17111954	0.00170297	BnaC08g35450D

## 9 Supplements

GSBRNA2T00016078001	354.401646	157.537138	1.1689383	8.17E-08	BnaC09g46840D
GSBRNA2T00099743001	142.211672	63.2079606	1.16836058	0.00020476	BnaC07g42930D
GSBRNA2T00035436001	109.298418	48.6209731	1.16656309	0.00067137	BnaA09g47080D
GSBRNA2T00006749001	166.041425	73.9050843	1.16651066	0.00049068	BnaA01g36210D
GSBRNA2T00058695001	894.676248	398.708633	1.16583956	9.44E-10	BnaCnng53310D
GSBRNA2T00017218001	93.8855957	41.8137119	1.16439897	0.00168002	BnaC04g37630D
GSBRNA2T00054303001	562.105301	250.893847	1.16351505	1.27E-07	BnaA07g04280D
GSBRNA2T00091599001	113.314995	50.5659048	1.16216157	0.000689	BnaCnng16860D
GSBRNA2T00028976001	356.99807	159.48207	1.16184649	7.98E-07	BnaC06g18520D
GSBRNA2T00030080001	574.471146	257.701107	1.15623915	7.02E-06	BnaCnng36920D
GSBRNA2T00103042001	106.344377	47.6485073	1.15600614	0.0007759	BnaA08g23520D
GSBRNA2T00001303001	502.652108	225.609738	1.15526203	1.41E-07	BnaA02g05180D
GSBRNA2T00014958001	4928.71309	2216.24702	1.15305931	1.28E-14	BnaCnng27910D
GSBRNA2T00143187001	121.196038	54.4557682	1.15216097	0.00057742	BnaA02g03260D
GSBRNA2T00118704001	135.964843	61.263029	1.14840851	0.00062698	BnaA07g13580D
GSBRNA2T00067281001	187.137339	84.6022079	1.14433722	0.00015284	BnaA01g29400D
GSBRNA2T00020415001	120.520911	54.4557682	1.14413854	0.00106615	BnaA04g27380D
GSBRNA2T00016789001	126.721019	57.3731657	1.14156744	0.0009289	BnaC04g52010D
GSBRNA2T00000400001	145.831846	66.125358	1.13928974	0.00040813	BnaA05g28420D
GSBRNA2T00077509001	2257.52576	1025.94901	1.1377353	9.88E-11	BnaA05g15290D
GSBRNA2T00119975001	147.56786	67.0978238	1.1356372	0.00079037	BnaC05g46300D
GSBRNA2T00070154001	157.314453	71.9601528	1.126804	0.00075223	BnaA02g13090D
GSBRNA2T00071042001	163.623937	74.8775501	1.12629636	0.0003171	BnaC04g26600D
GSBRNA2T00091561001	146.5764	67.0978238	1.125951	0.0003009	BnaC02g09970D
GSBRNA2T00138957001	201.523441	92.381934	1.12416136	5.84E-05	BnaA03g21160D
GSBRNA2T00009783001	112.382396	51.5383707	1.12269874	0.00082873	BnaA09g02840D
GSBRNA2T00074807001	132.975577	61.263029	1.11629452	0.00057111	BnaCnng59220D
GSBRNA2T00144194001	379.454295	175.041521	1.11549789	5.73E-07	BnaA06g28560D
GSBRNA2T00101379001	1131.53951	522.211777	1.11546721	2.85E-08	BnaC01g34360D
GSBRNA2T00052901001	124.32884	57.3731657	1.11415686	0.00117921	BnaC08g02390D
GSBRNA2T00115393001	189.144786	87.5196052	1.11056989	4.73E-05	BnaC07g21680D
GSBRNA2T00121060001	3293.45644	1526.76884	1.10907887	6.71E-12	BnaA01g04920D
GSBRNA2T00052344001	307.241021	142.950152	1.10310214	4.42E-06	BnaAnng21880D
GSBRNA2T00148727001	215.257115	100.16166	1.1025506	0.00020323	BnaA03g09690D
GSBRNA2T00128106001	150.233768	70.0152212	1.10022282	0.00052177	BnaC08g03730D
GSBRNA2T00150057001	133.196311	62.2354948	1.09617649	0.00104921	BnaC03g45790D
GSBRNA2T00145642001	227.795139	106.96892	1.08957512	1.48E-05	BnaA09g22620D
GSBRNA2T00049055001	347.206479	163.371933	1.0870175	6.85E-06	BnaC06g23410D
GSBRNA2T00130738001	192.179013	90.4370025	1.08627386	7.58E-05	BnaC01g04640D
GSBRNA2T00099229001	274.320965	129.335632	1.08375597	8.91E-05	BnaA07g20510D
GSBRNA2T00009742001	1419.60427	670.026563	1.08312747	4.06E-09	BnaA09g03130D
GSBRNA2T00124438001	212.295104	100.16166	1.08273912	4.54E-05	BnaA07g25220D
GSBRNA2T00050840001	643.784534	304.37946	1.08042032	1.19E-09	BnaC06g15890D
GSBRNA2T00135986001	121.36352	57.3731657	1.07931278	0.00106321	BnaA10g22620D
GSBRNA2T00013624001	203.572139	96.271797	1.07929645	0.00014152	BnaC08g49820D

## 9 Supplements

GSBRNA2T00064228001	129.482889	61.263029	1.07813203	0.00066464	BnaA02g19760D
GSBRNA2T00071025001	139.743069	66.125358	1.07784357	0.00086322	BnaC04g26730D
GSBRNA2T00139038001	991.368036	470.671095	1.07450165	9.77E-11	BnaA03g20460D
GSBRNA2T00067672001	143.363264	68.0702896	1.07308807	0.00161111	BnaAnng25560D
GSBRNA2T00013975001	1156.62156	550.413282	1.0712304	8.66E-09	BnaAnng13110D
GSBRNA2T00138209001	118.343112	56.4006999	1.06759047	0.00128834	BnaC01g13050D
GSBRNA2T00139085001	179.304107	85.5746736	1.06619556	0.0016522	BnaA03g20090D
GSBRNA2T00145360001	411.859549	197.408233	1.06044079	1.26E-06	BnaC08g00240D
GSBRNA2T00055727001	286.075655	137.115358	1.06041823	4.63E-05	BnaCnng52300D
GSBRNA2T00044620001	456.262524	218.802478	1.05996545	6.30E-05	BnaC05g09510D
GSBRNA2T00129062001	204.452086	98.2167285	1.05683533	6.88E-05	BnaC03g53440D
GSBRNA2T00090937001	285.313887	137.115358	1.05666332	1.66E-05	BnaA02g12030D
GSBRNA2T00139843001	113.367542	54.4557682	1.05607598	0.00145893	BnaC08g16180D
GSBRNA2T00136995001	656.543327	316.049049	1.05447687	3.23E-09	BnaC08g35860D
GSBRNA2T00121704001	542.790039	261.590969	1.05278251	2.04E-08	BnaC08g20790D
GSBRNA2T00150459001	304.658424	146.840015	1.05233635	0.00018584	BnaA10g01150D
GSBRNA2T00141082001	256.197954	123.500838	1.05220539	0.00021068	BnaC03g12050D
GSBRNA2T00004610001	153.377714	73.9050843	1.05207657	0.00036331	BnaC06g40750D
GSBRNA2T00071285001	111.011562	53.4833024	1.05174838	0.00193239	BnaC06g14570D
GSBRNA2T00032030001	203.747597	98.2167285	1.0515452	0.00013393	BnaC05g04510D
GSBRNA2T00003604001	108.916539	52.5108365	1.05083947	0.00232325	BnaC06g43480D
GSBRNA2T00145159001	235.292364	113.77618	1.04728345	4.08E-05	BnaC08g01600D
ENSRNA049451265	104.579188	50.5659048	1.0465397	0.00249203	SNORD14
GSBRNA2T00135966001	946.133737	458.029041	1.04649551	1.66E-05	BnaA10g22460D
GSBRNA2T00145985001	289.057898	140.032755	1.04508156	2.10E-05	BnaC09g04040D
GSBRNA2T00119640001	146.540353	70.987687	1.04446799	0.00110743	BnaAnng01000D
GSBRNA2T00071664001	1143.3991	554.303145	1.04441805	1.70E-10	BnaAnng06270D
GSBRNA2T00003103001	114.440373	55.4282341	1.04406726	0.00229789	BnaA06g23470D
GSBRNA2T00059731001	198.518551	96.271797	1.04304774	7.32E-05	BnaA06g38750D
GSBRNA2T00059427001	224.548012	108.913852	1.04279807	5.91E-05	BnaA07g01020D
GSBRNA2T00031290001	140.224622	68.0702896	1.04131454	0.00104817	BnaC07g34170D
GSBRNA2T00064738001	651.929553	317.021515	1.03984257	3.02E-07	BnaC04g00250D
GSBRNA2T00004277001	2515.2237	1228.22187	1.03408142	1.27E-09	BnaC04g53170D
GSBRNA2T00139054001	460.079706	224.637273	1.03403903	1.15E-06	BnaA03g20310D
GSBRNA2T00144411001	137.190331	67.0978238	1.03052035	0.00102665	BnaA06g30150D
GSBRNA2T00012481001	667.22108	326.746172	1.02985541	1.26E-06	BnaC02g23980D
GSBRNA2T00123924001	452.790564	221.719875	1.02983673	3.91E-07	BnaC03g39940D
GSBRNA2T00080010001	186.467219	91.4094682	1.02768745	0.00052842	BnaC02g03270D
GSBRNA2T00157470001	331.118406	162.399467	1.02715571	2.83E-06	BnaC01g18280D
GSBRNA2T00135951001	562.456217	276.177955	1.02581124	2.01E-08	BnaA10g22330D
GSBRNA2T00141946001	2383.4216	1172.79133	1.02301714	0.00088266	BnaA03g15610D
GSBRNA2T00142013001	614.345794	303.406995	1.01751264	0.00195949	BnaA03g15020D
GSBRNA2T00069607001	1251.27411	618.485881	1.01650231	4.83E-09	BnaA02g27180D
GSBRNA2T00054474001	249.917273	123.500838	1.01618601	0.00021967	BnaAnng22410D
GSBRNA2T00104894001	112.108572	55.4282341	1.01434818	0.00266002	BnaA05g24200D

## 9 Supplements

GSBRNA2T00089627001	463.290634	229.499601	1.01313267	8.83E-07	BnaA08g27430D
GSBRNA2T00086709001	519.489265	257.701107	1.01100642	1.15E-07	BnaA03g12130D
GSBRNA2T00155549001	1434.23912	711.842588	1.01049073	1.50E-10	BnaC08g30470D
GSBRNA2T00127552001	170.276651	84.6022079	1.0081997	0.00218716	BnaC04g37900D
GSBRNA2T00138925001	158.515822	78.7674132	1.00768687	0.00114519	BnaA03g21440D
GSBRNA2T00095296001	183.878702	91.4094682	1.00746044	0.00043469	BnaA05g09640D
GSBRNA2T00019691001	254.193122	126.418235	1.00684537	0.00014467	BnaCnng74170D
GSBRNA2T00057870001	287.147336	142.950152	1.00574207	1.82E-05	BnaC06g13590D
GSBRNA2T00048661001	265.594392	132.253029	1.00545365	0.00039566	BnaA05g11280D
GSBRNA2T00148744001	1026.77808	511.514654	1.00514706	0.0010474	BnaA03g09530D
GSBRNA2T00076808001	374.512473	186.71111	1.00393538	3.07E-05	BnaA01g07200D
GSBRNA2T00125563001	1893.72733	944.261889	1.00386906	3.28E-11	BnaA05g13800D
GSBRNA2T00090178001	3118.21936	1554.97035	1.00375263	3.15E-11	BnaC02g34580D
GSBRNA2T00034909001	272.745682	136.142892	1.00182604	1.65E-05	BnaA06g11780D
GSBRNA2T00075668001	253.208844	126.418235	1.00146161	0.00013017	BnaC05g47210D
GSBRNA2T00103355001	179.15811	89.4645367	1.00060622	0.00058974	BnaC06g33780D

### 9.2.5 Normalized MST-time traces and binding curves of AtGRP7

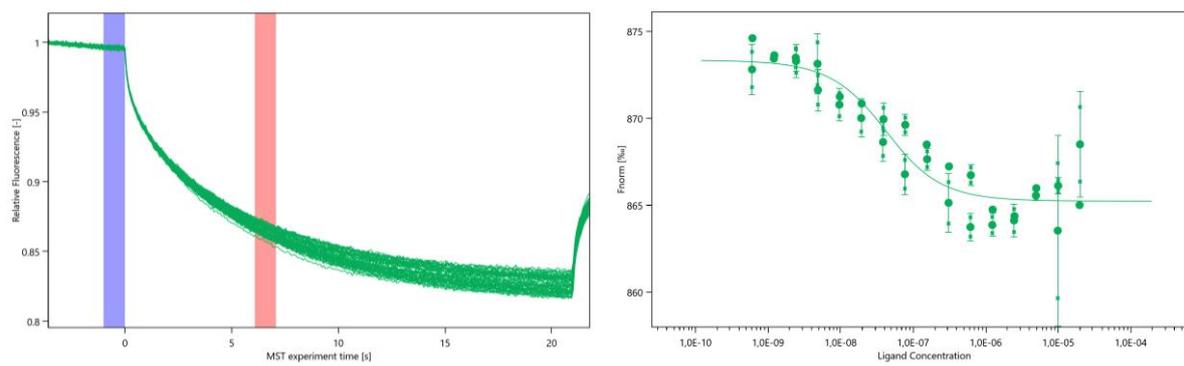


Figure 48: AtGRP7 - *miRNA164* MST-traces and binding curve.

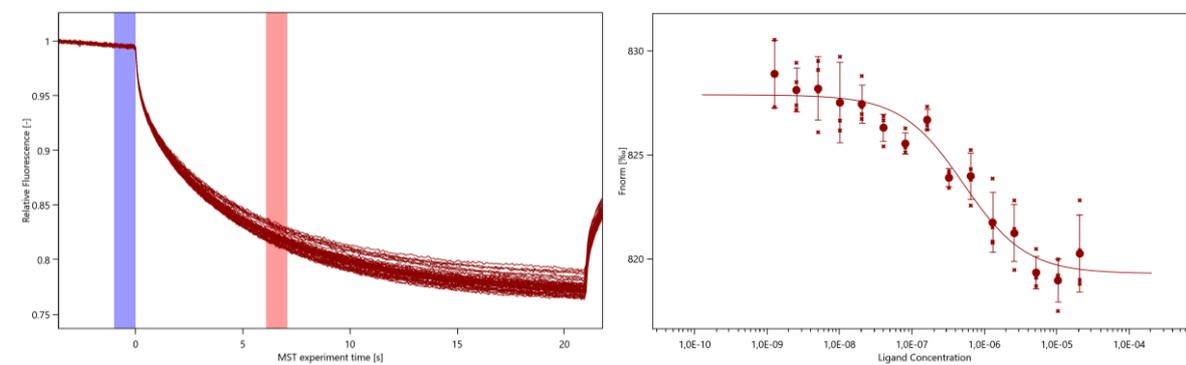


Figure 49: AtGRP7 - *dsmiRNA164* MST-traces and binding curve.

## 9 Supplements

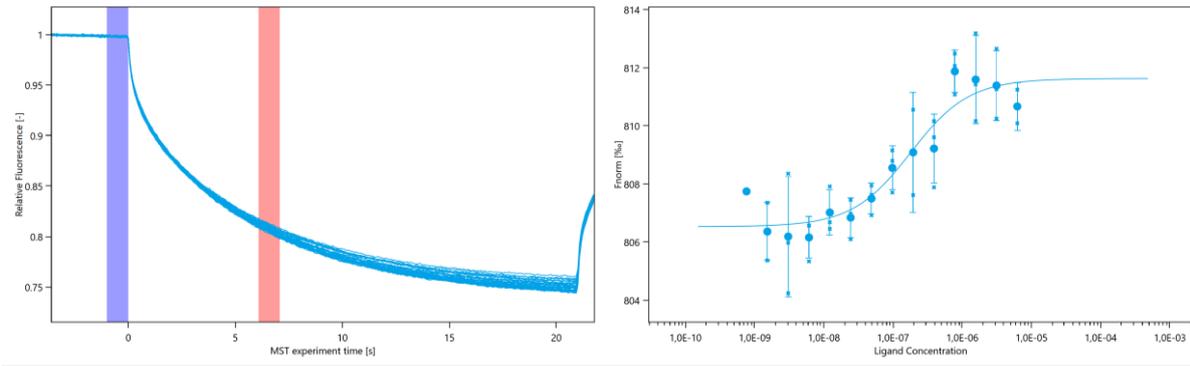


Figure 50: AtGRP7 - 21R- MST-traces and binding curve.

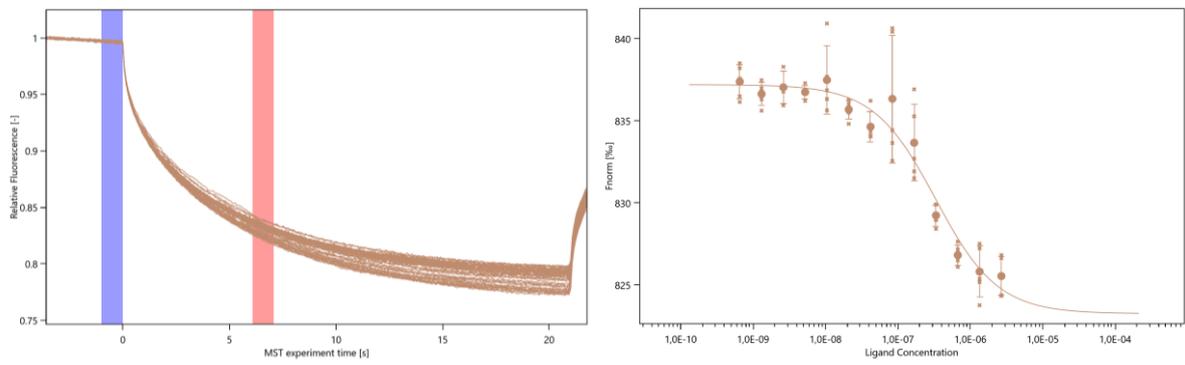


Figure 51: AtGRP7 - *tRNAmet* MST-traces and binding curve.

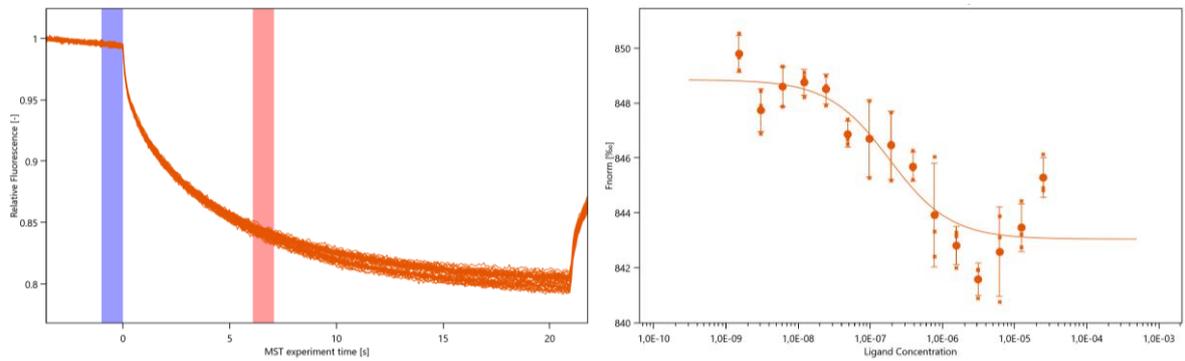


Figure 52: AtGRP7 - *GRP7 3'UTR* MST-traces and binding curve.

## 9 Supplements

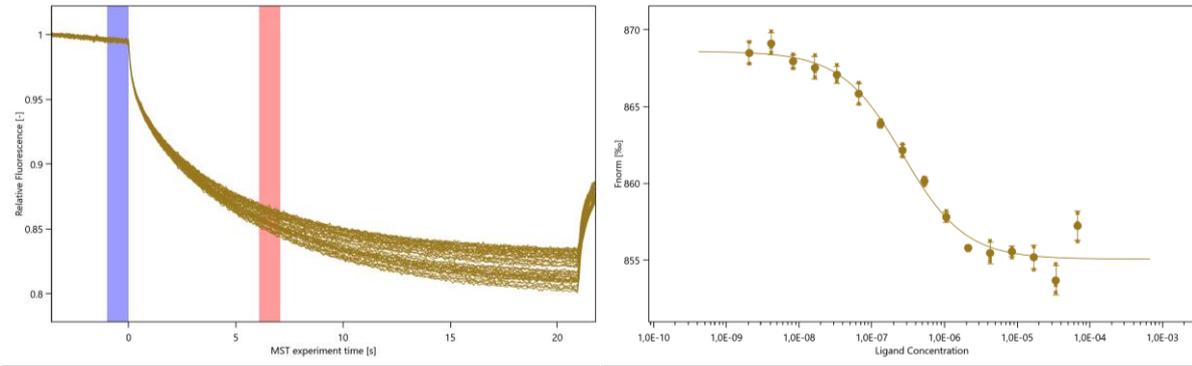


Figure 53: AtGRP7 - *BnmiRnovel2* MST-traces and binding curve.

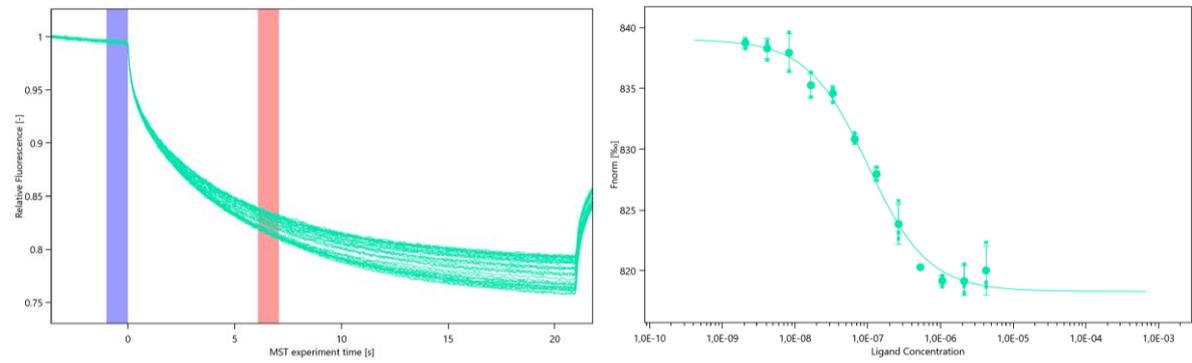


Figure 54: AtGRP7 - *BnmiRnovel106* MST-traces and binding curve.

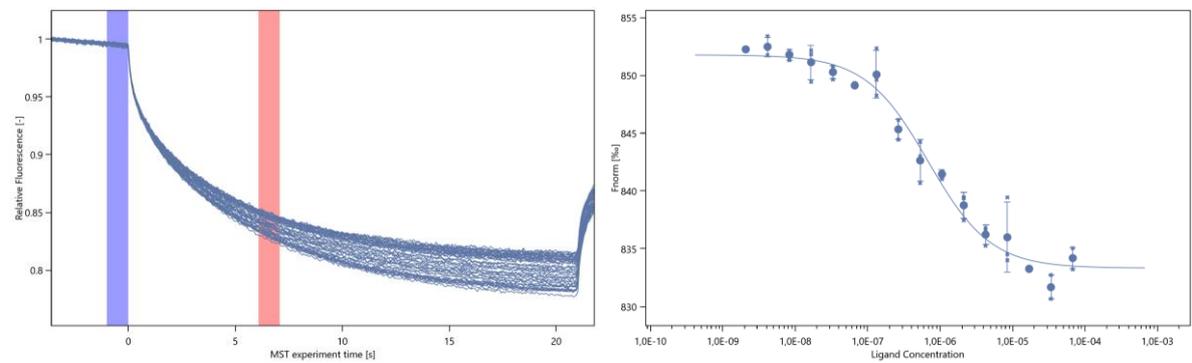


Figure 55: AtGRP7 - *BnmiRnovel149* MST-traces and binding curve.

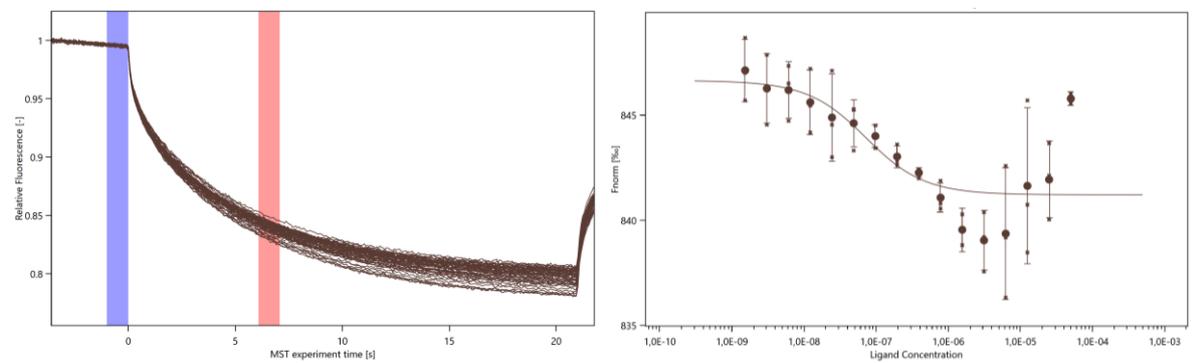


Figure 56: AtGRP7 - *miR164 m6A* MST-traces and binding curve.

## 9 Supplements

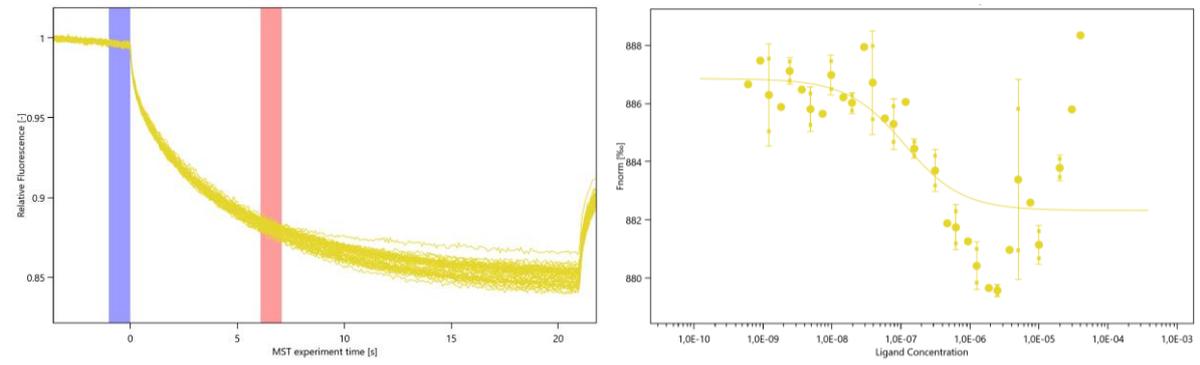


Figure 57: AtGRP7 - *miR164* m5C MST-traces and binding curve.

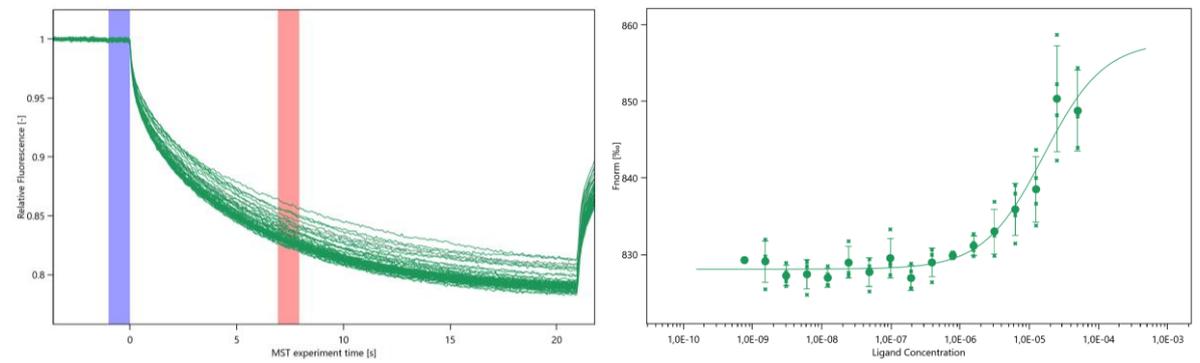


Figure 58: AtGRP7 - *GFP-RNA* MST-traces and binding curve.

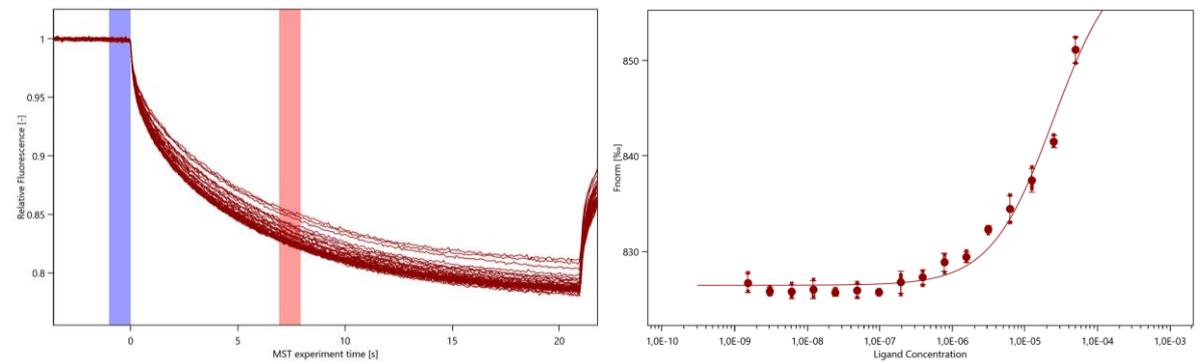


Figure 59: AtGRP7 - *GFP+GRP7UTR-RNA* MST-traces and binding curve.

## 9 Supplements

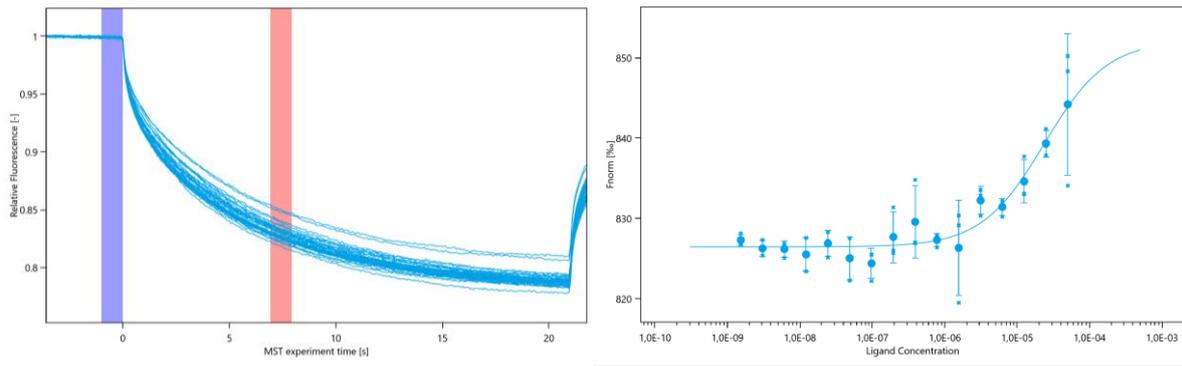


Figure 60: AtGRP7 - *GFP-TCTP-RNA* MST-traces and binding curve.

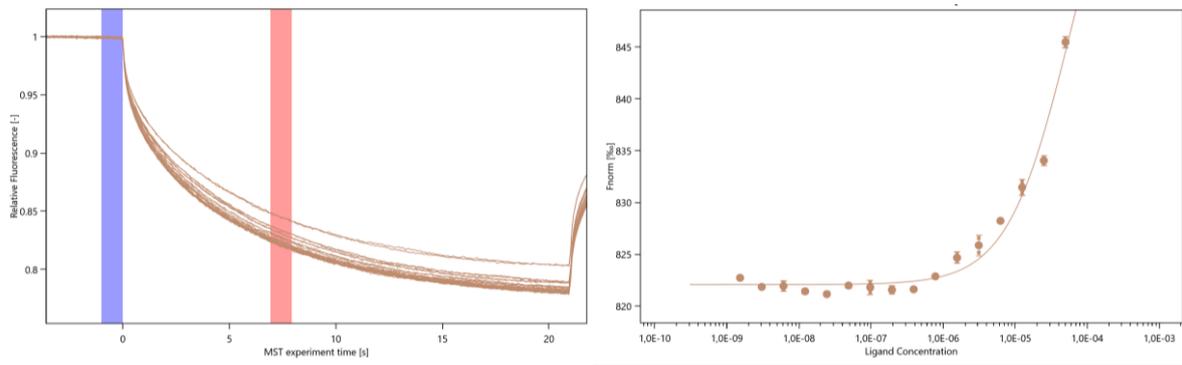


Figure 61: AtGRP7 - *GFP-TCTP-GRP7-UTR-RNA* MST-traces and binding curve.

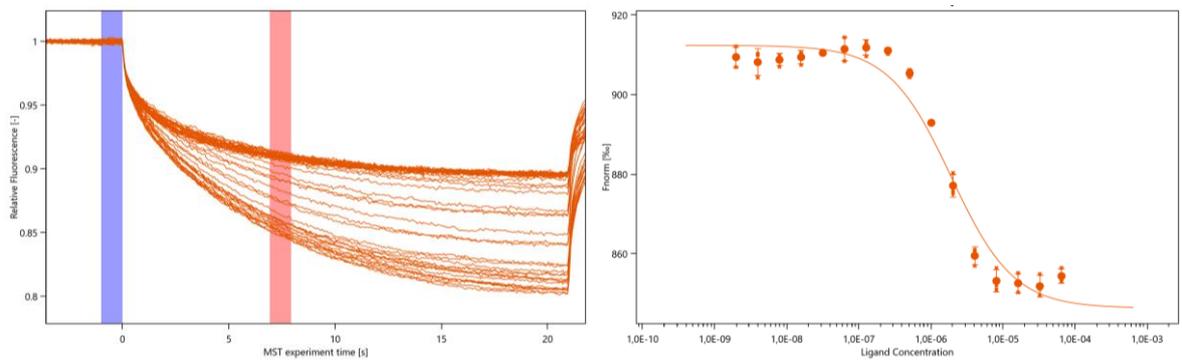


Figure 62: AtGRP7 - *BnaCNG53670D\_w/o GA* MST-traces and binding curve.

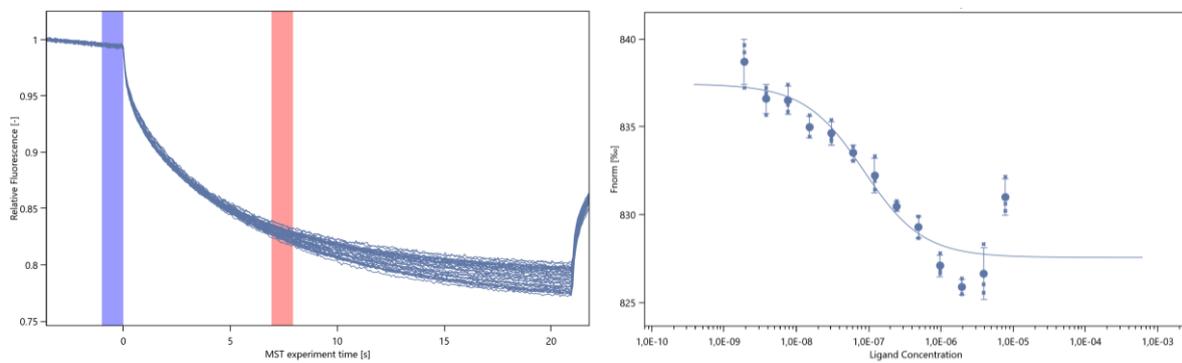


Figure 63: AtGRP7 - *BnPARCL* MS- traces and binding curve.

9 Supplements

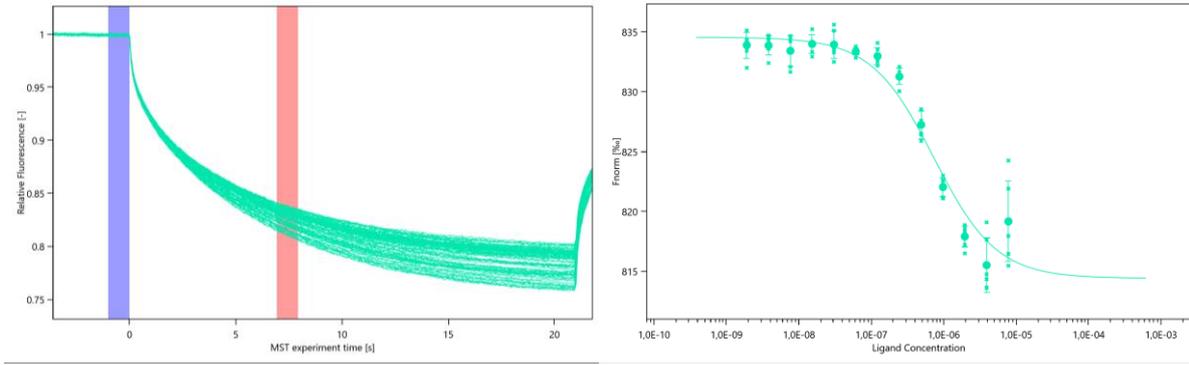


Figure 64: AtGRP7 – *BnaCNGG53670D* MST-traces and binding curve.

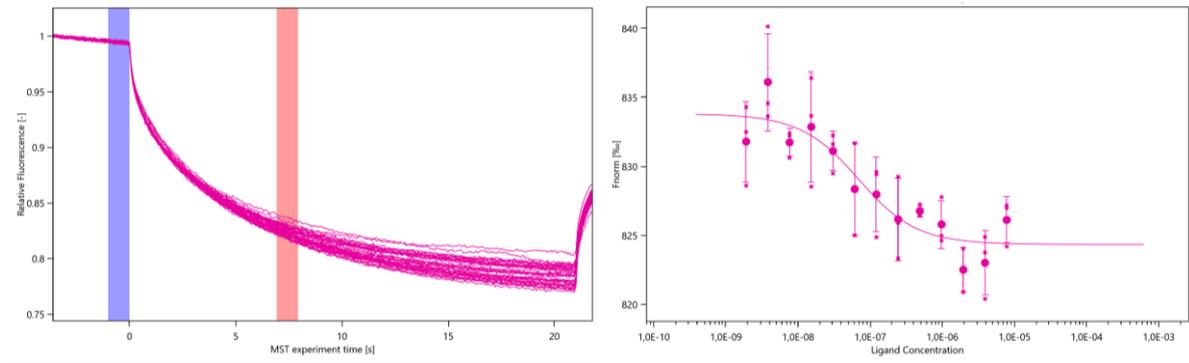


Figure 65: AtGRP7 – *BnPARCL\_withGAmotif* MST-traces and binding curve.

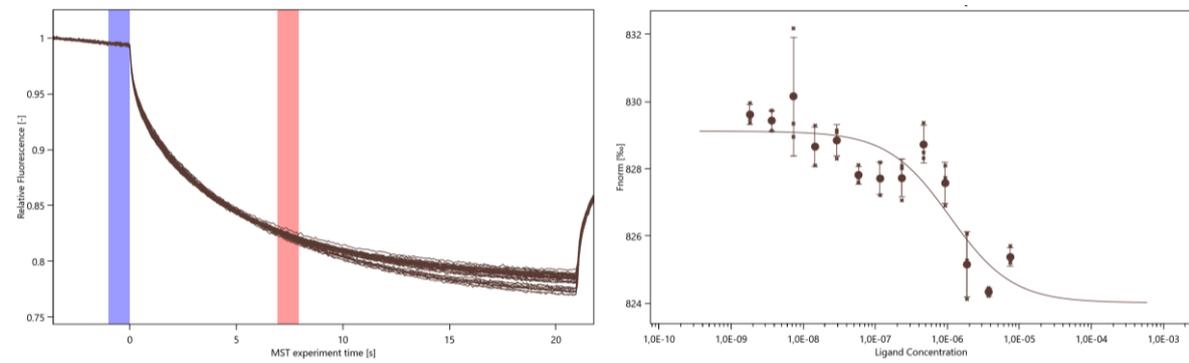


Figure 66: AtGRP7 – *BnaC03g26040D* MST-traces and binding curve.

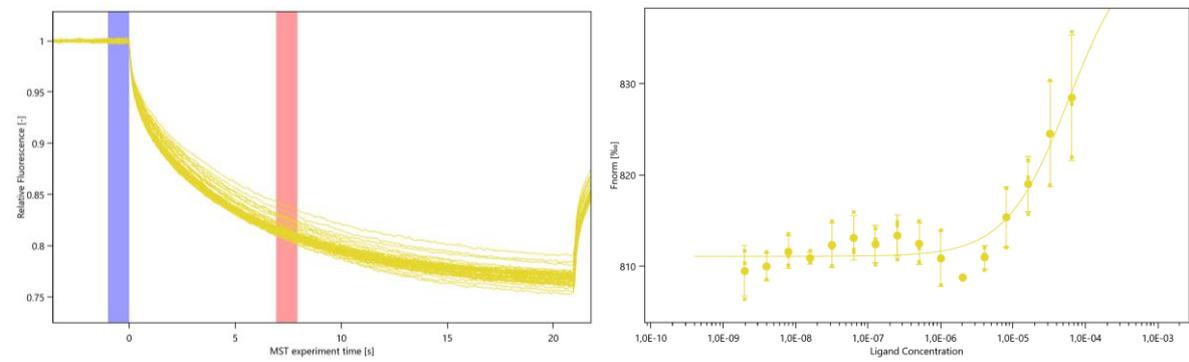


Figure 67: AtGRP7 - *BnaCNGG53670D\_120ntUTR* MST-traces and binding curve.

## 9 Supplements

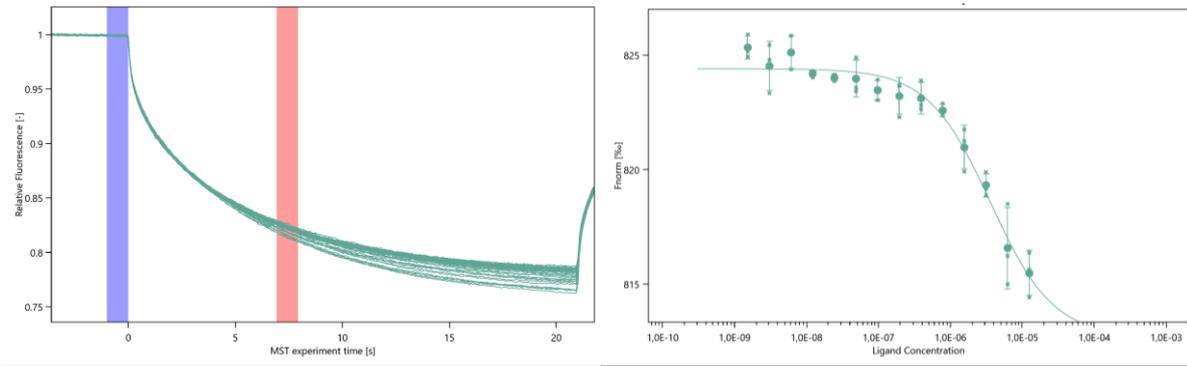


Figure 68: AtGRP7 - *BnaA03g14400D* MST-traces and binding curve.

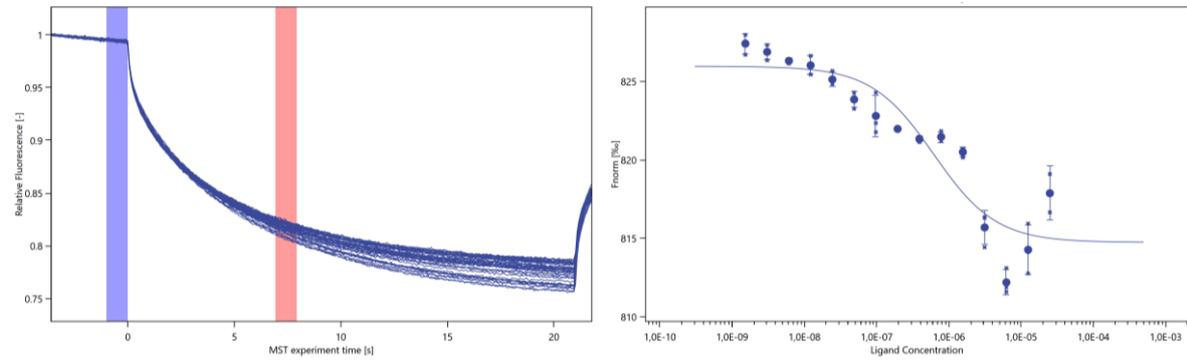


Figure 69: AtGRP7 - *BnaC09g46650D* MST-traces and binding curve.

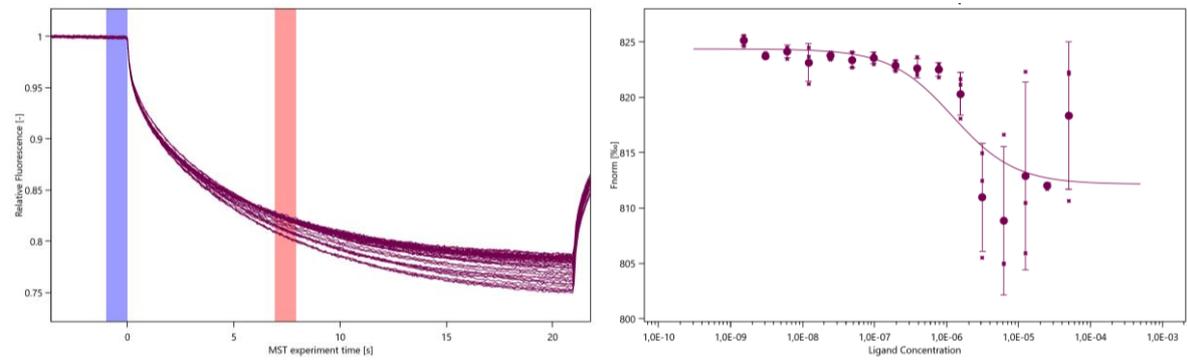


Figure 70: AtGRP7 - *BnaA07g32150D* MST-traces and binding curve.

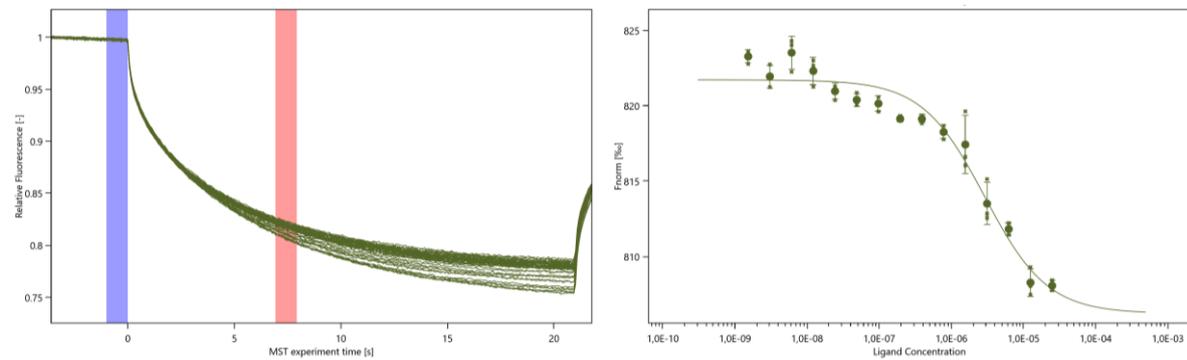


Figure 71: AtGRP7 - *BnGRP7* MST-traces and binding curve.

## 9 Supplements

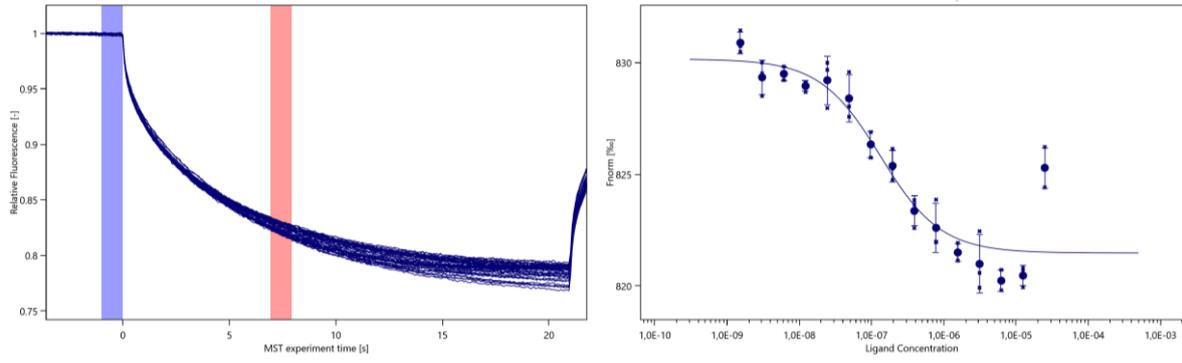


Figure 72: AtGRP7 - *BnGRP7+UTR* RNA MST-traces and binding curve.

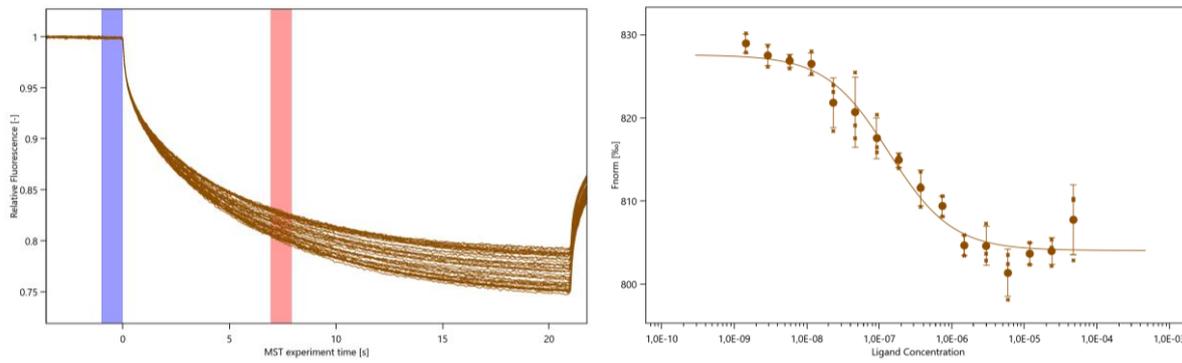


Figure 73: AtGRP7 - *BnGRP7+UTR+Intron* RNA MST-traces and binding curve

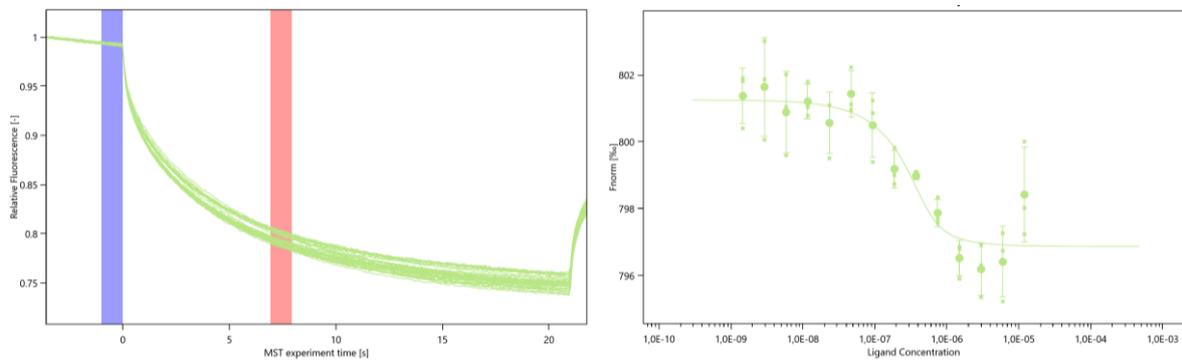


Figure 74: AtGRP7 - *dsBnPARCL* MST-traces and binding curve.

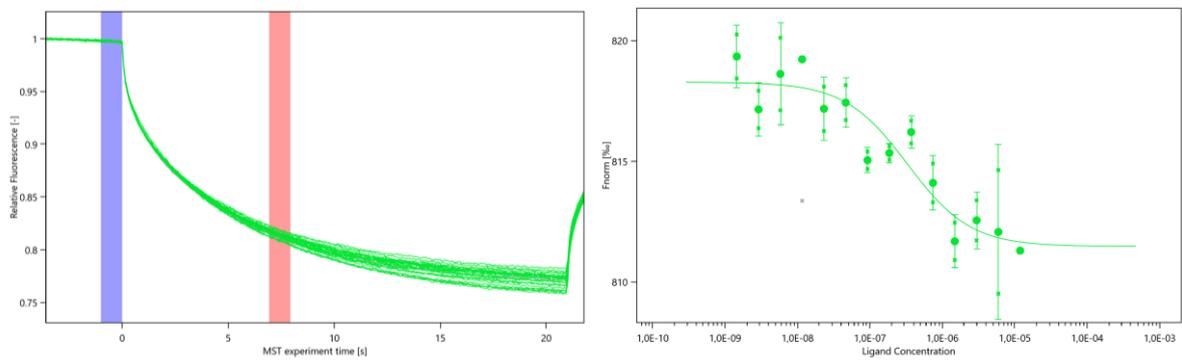


Figure 75: AtGRP7 - *AtGRP7-RNA* MST-traces and binding curve.

9 Supplements

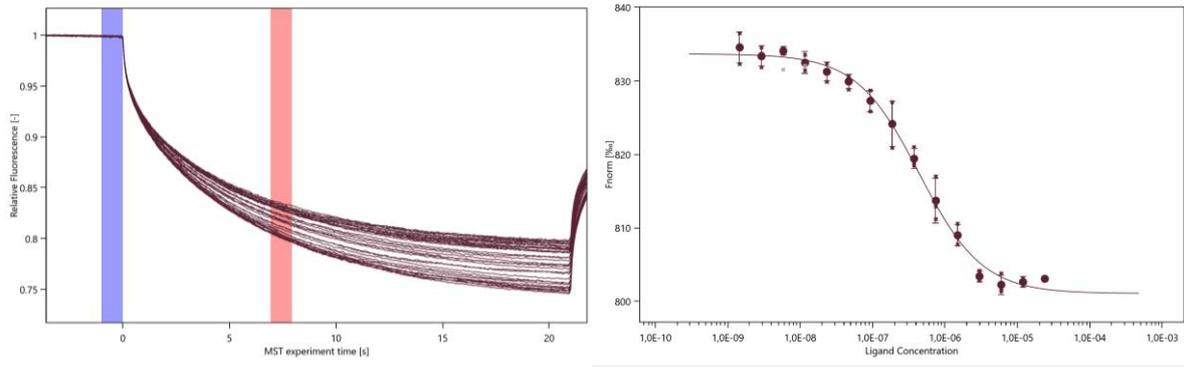


Figure 76: AtGRP7 – AtGRP7\_UTR+Intron MST-traces and binding curve.

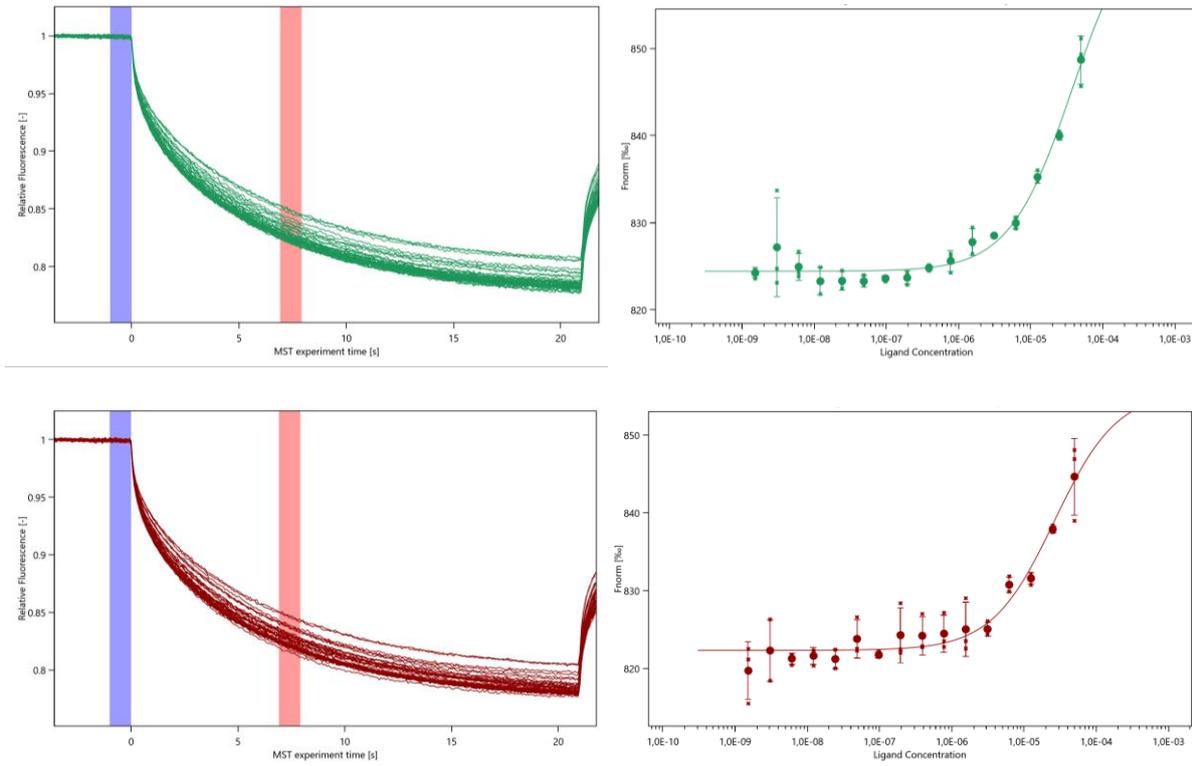


Figure 77: AtGRP7 - GFP-TCTP-RNA m5C MST-traces and binding curve

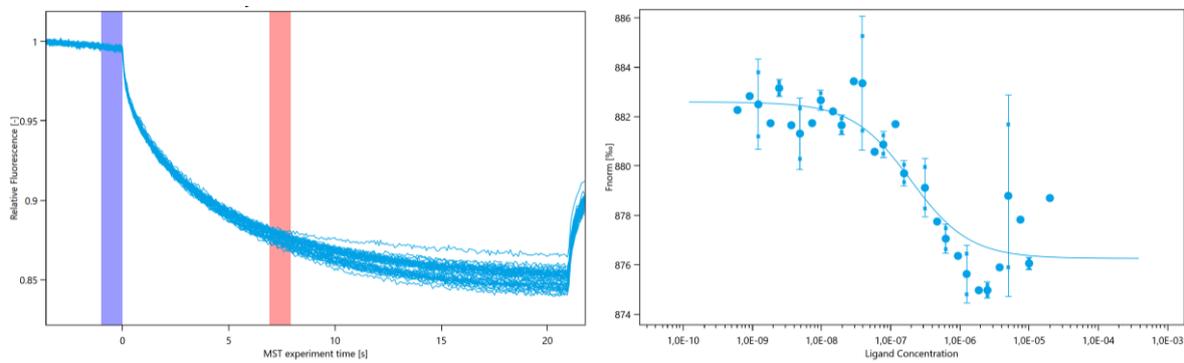


Figure 78: AtGRP7 - miR164 m5C MST-traces and binding curve.

## 9 Supplements

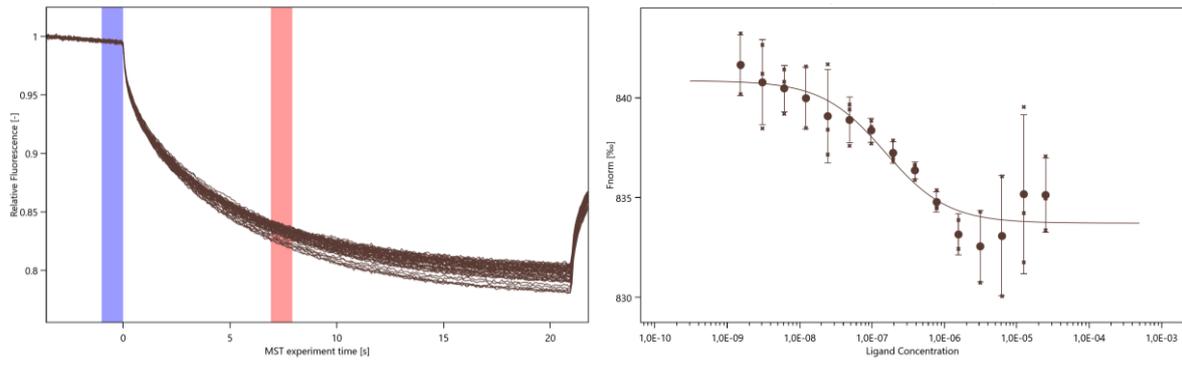


Figure 79: AtGRP7 - *miR164* m6A MST-traces and binding curve.

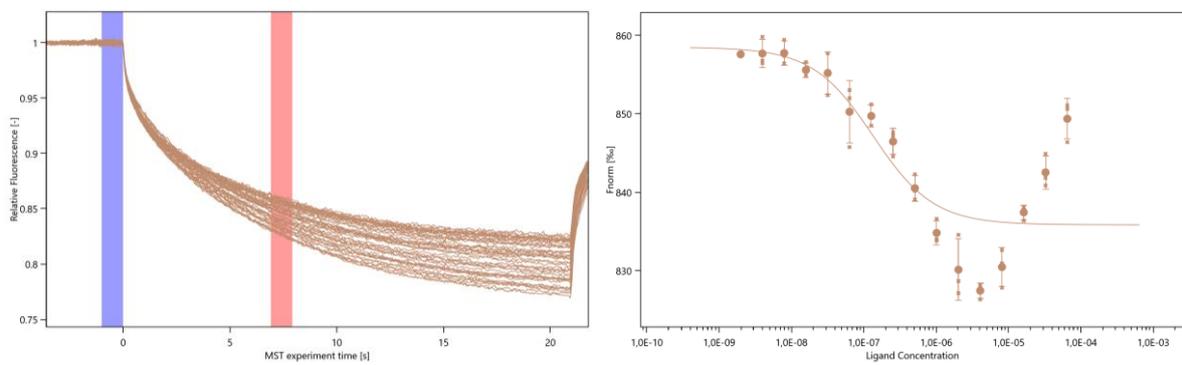


Figure 80: AtGRP7 - *BnaCNNG53670D* m6A MST-traces and binding curve.

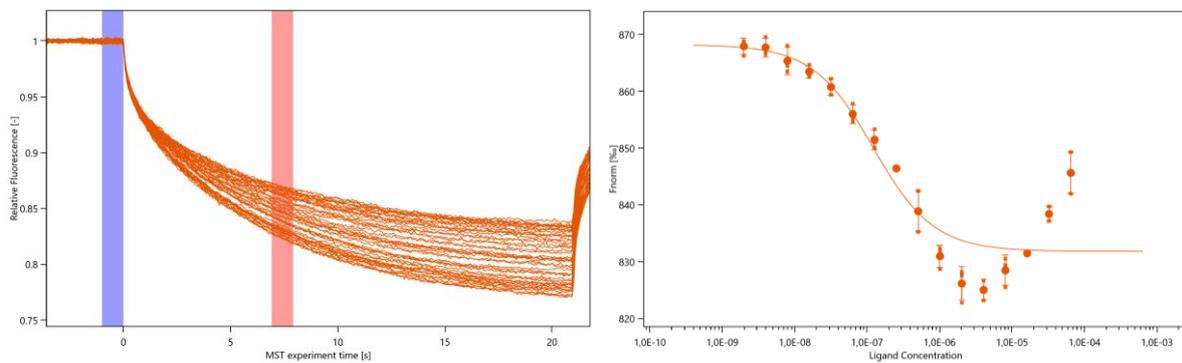


Figure 81: AtGRP7 - *BnaCNNG53670D* m6A without GA-motif MST-traces and binding curve.

9 Supplements

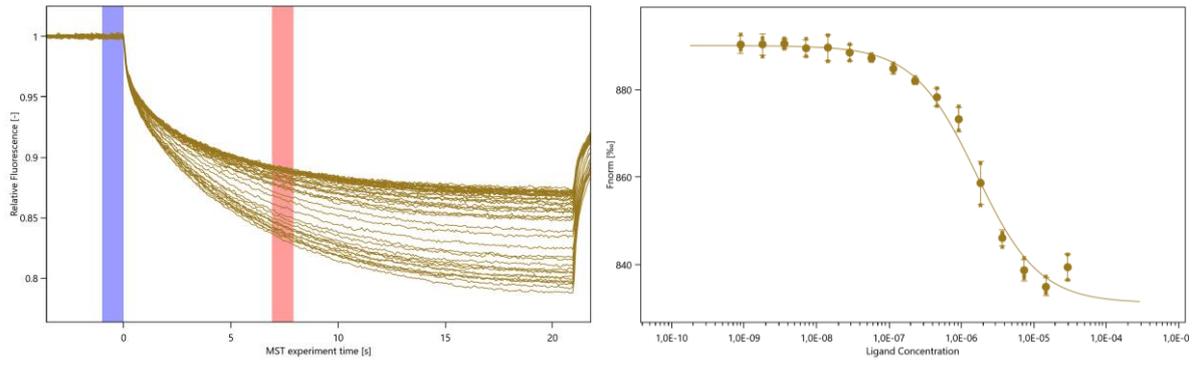


Figure 82: AtGRP7 - *BnaCNG53670D* m5C MST-traces and binding curve.

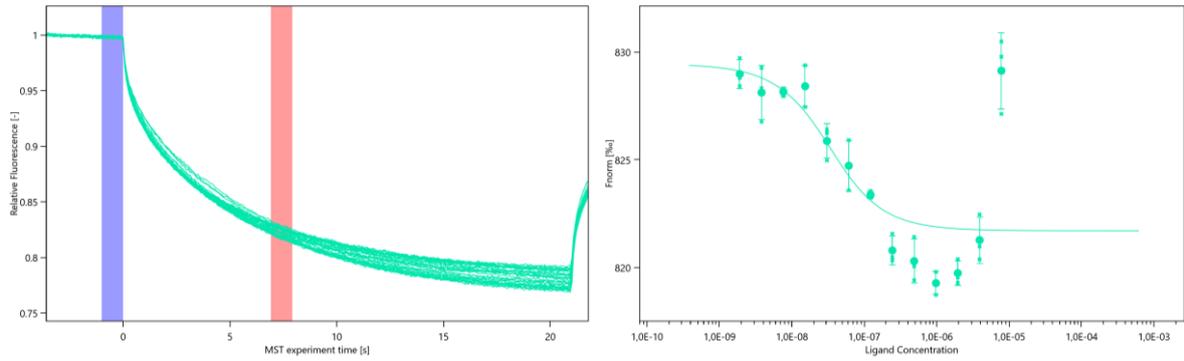


Figure 83: AtGRP7 - *BnPARCL* m6A MST-traces and binding curve.

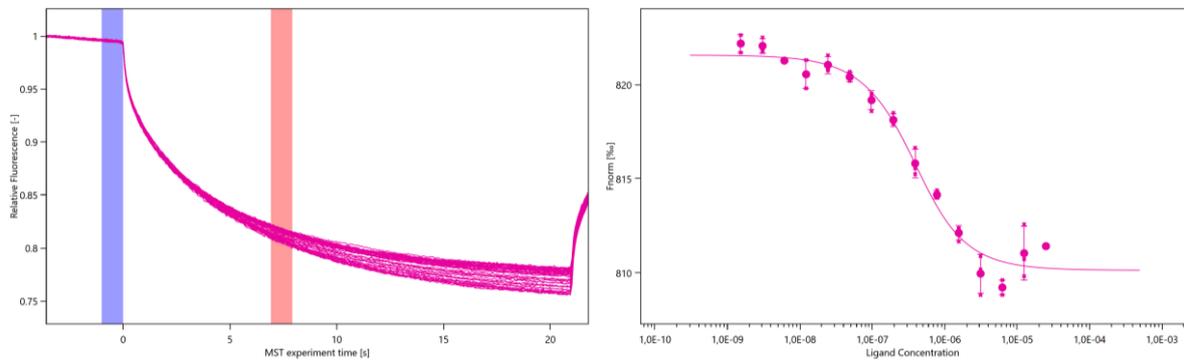


Figure 84: AtGRP7 - *BnPARCL* m5C MST-traces and binding curve.

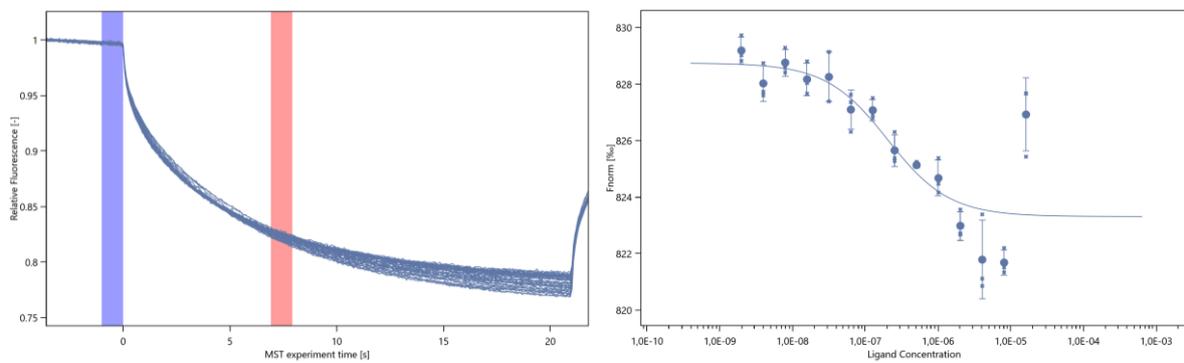


Figure 85: AtGRP7 - *BnaC03g26040D* m5C MST-traces and binding curve.

## 9 Supplements

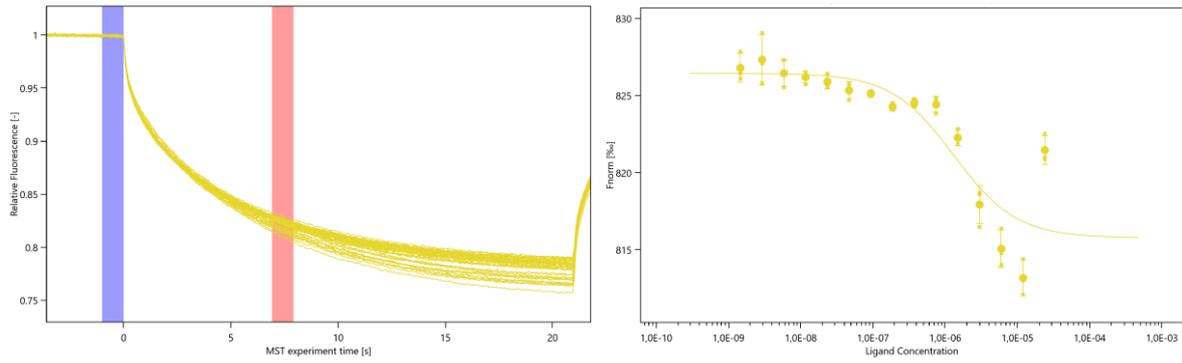


Figure 86: AtGRP7 - *BnaC03g26040D* m6A MST-traces and binding curve.

### 9.2.6 Normalized MST-time traces and binding curves of AtGRP7short

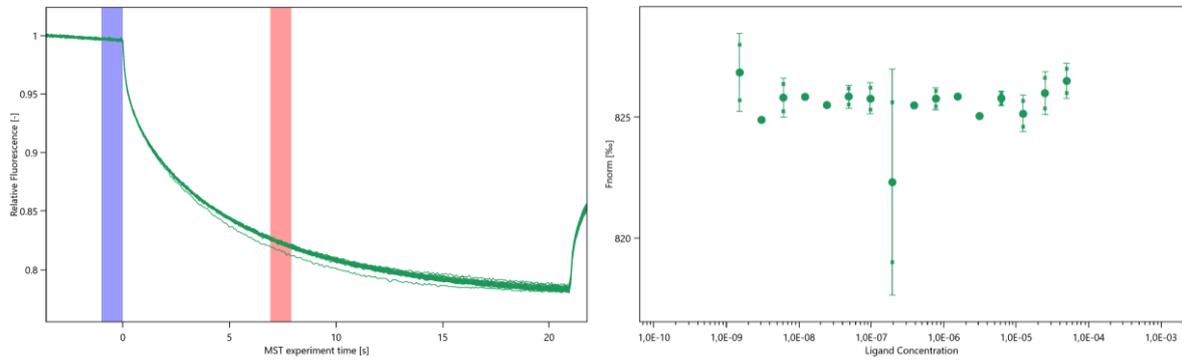


Figure 87: AtGRP7short - *BnPARCL* RNA MST-traces and binding curve.

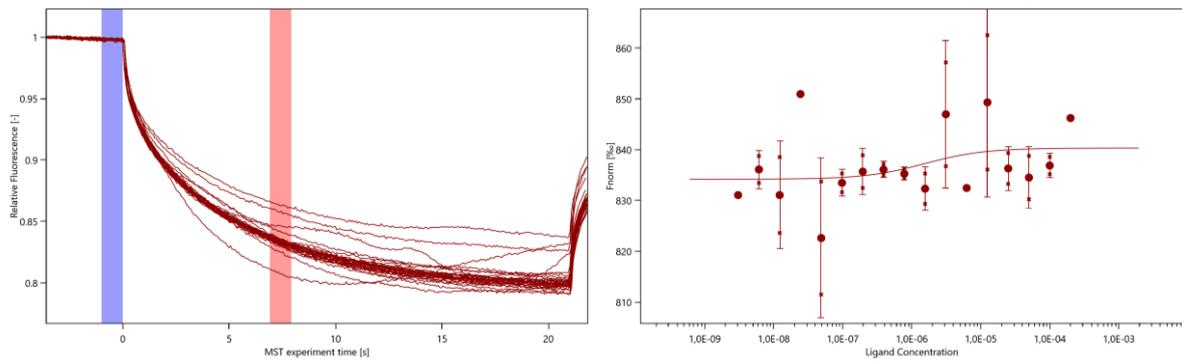


Figure 88: AtGRP7 short – *tRNAmets* MST-traces and binding curve.

## 9 Supplements

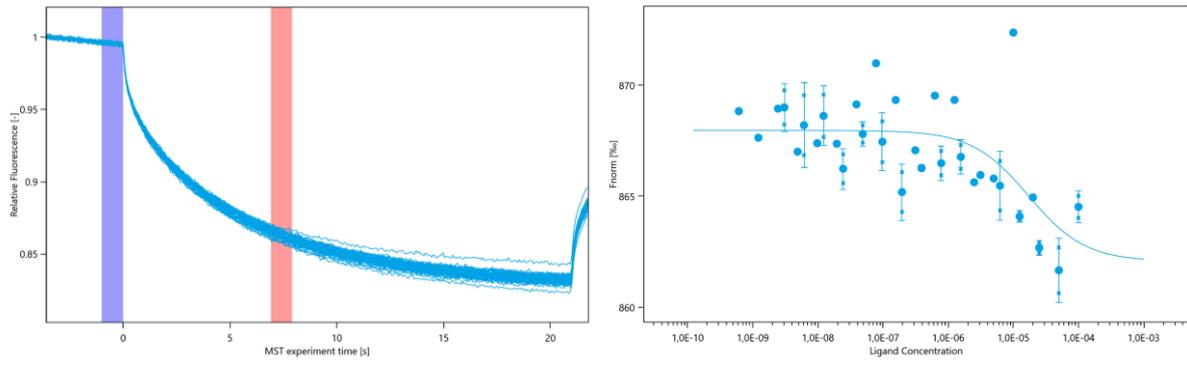


Figure 89: AtGRP7short - *miR164* MST-traces and binding curve.

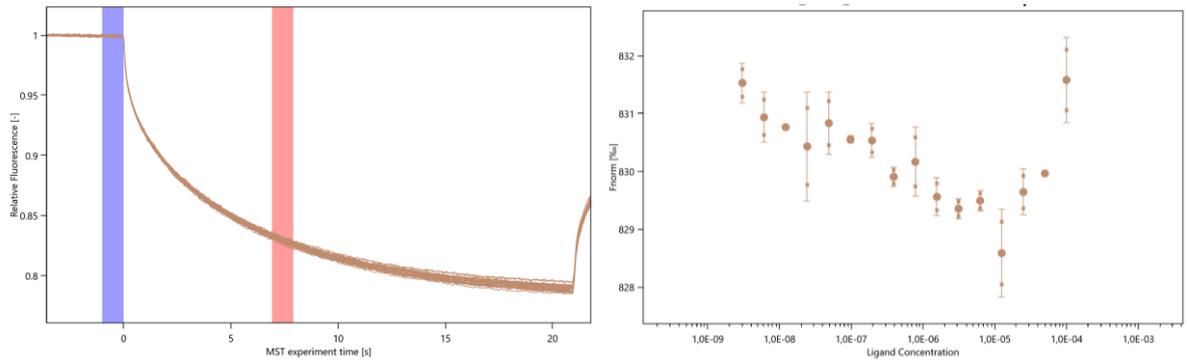


Figure 90: AtGRP7short - *BnaCNG53670D* MST-traces and binding curve.

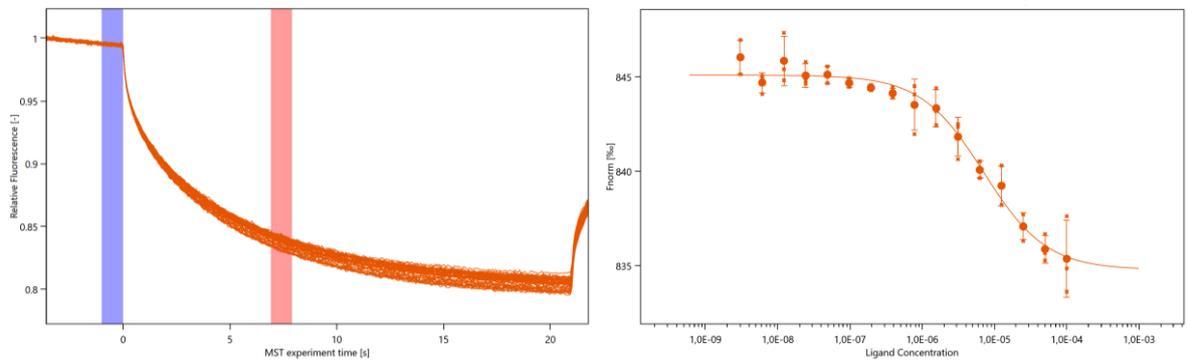


Figure 91: AtGRP7short - *BnmirNovel2* MST-traces and binding curve.

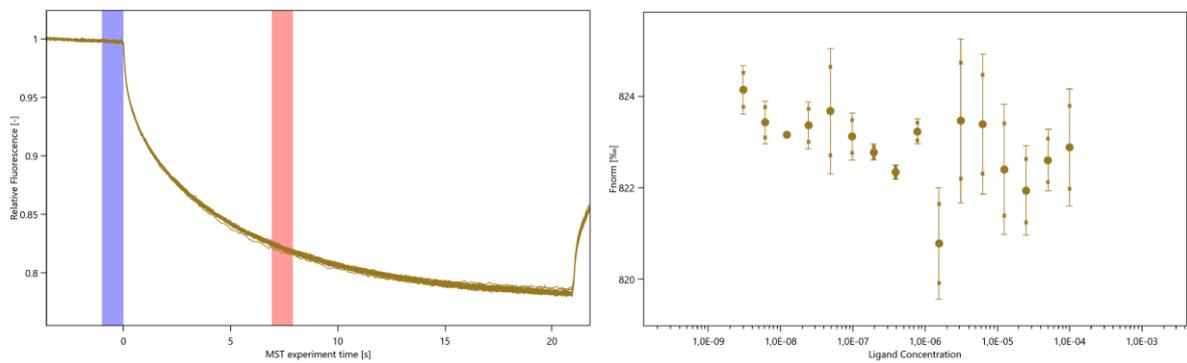


Figure 92: AtGRP7short - *BnGRP7 RNA* MST-traces and binding curve.

## 9 Supplements

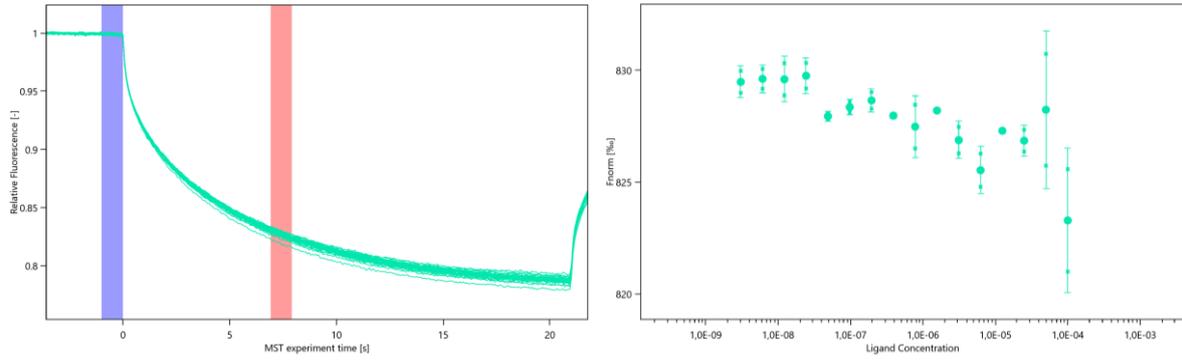


Figure 93: AtGRP7short - *BnGRP7+UTR* RNA MST-traces and binding curve.

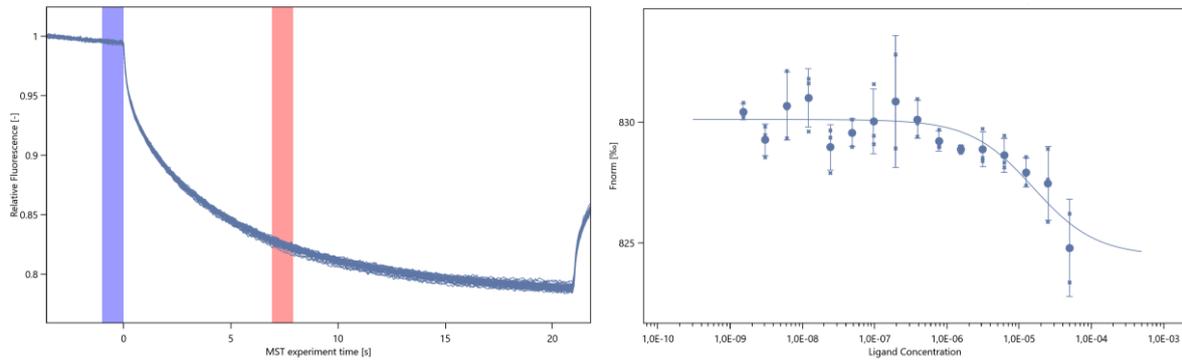


Figure 94: AtGRP7short - *BnmiRnovel106* MST-traces and binding curve.

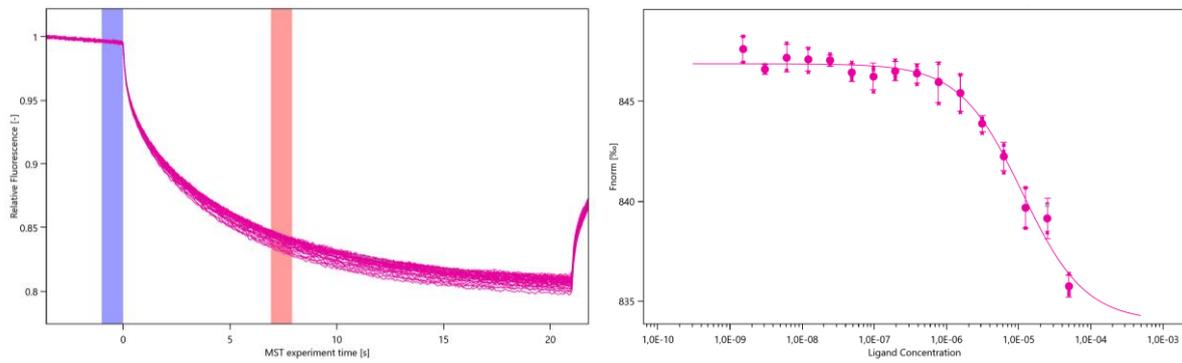


Figure 95: AtGRP7short - *GRP7 3'UTR* MST-traces and binding curve.

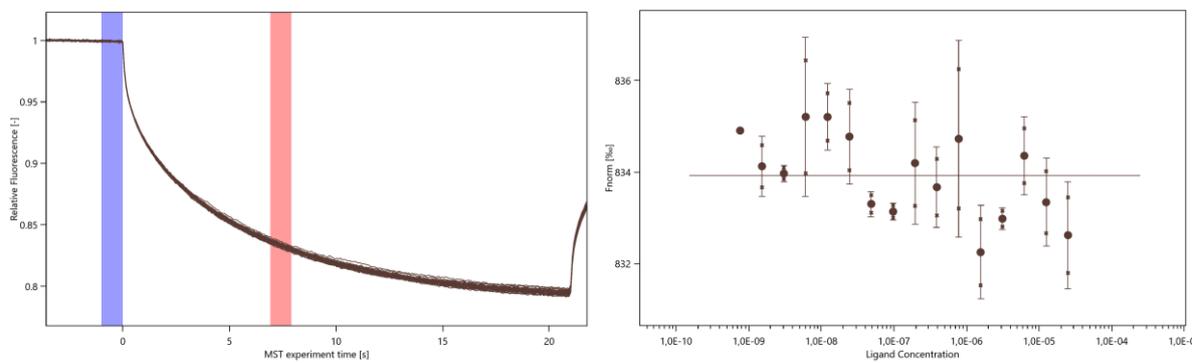


Figure 96: AtGRP7short - *AtGRP7* RNA MST-traces and binding curve.

## 9 Supplements

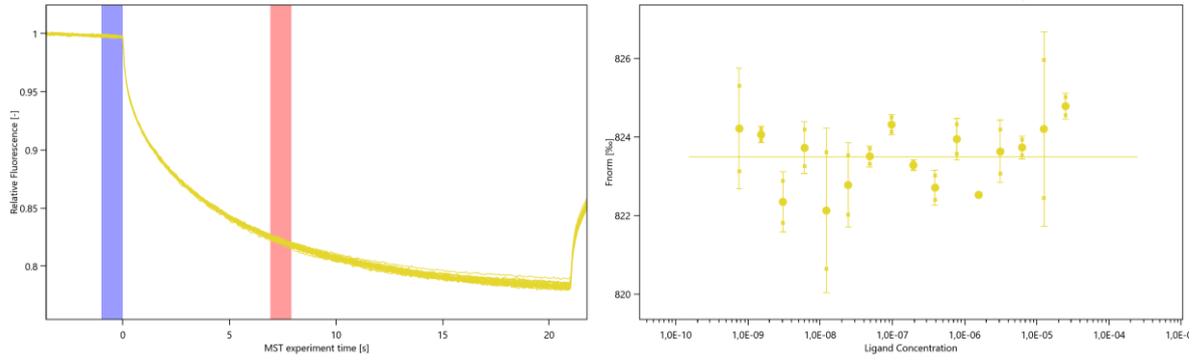


Figure 97: AtGRP7short – AtGRP7+UTR+Intron MST-traces and binding curve.

### 9.2.7 Normalized MST-time traces and binding curves of AtGRP7mut

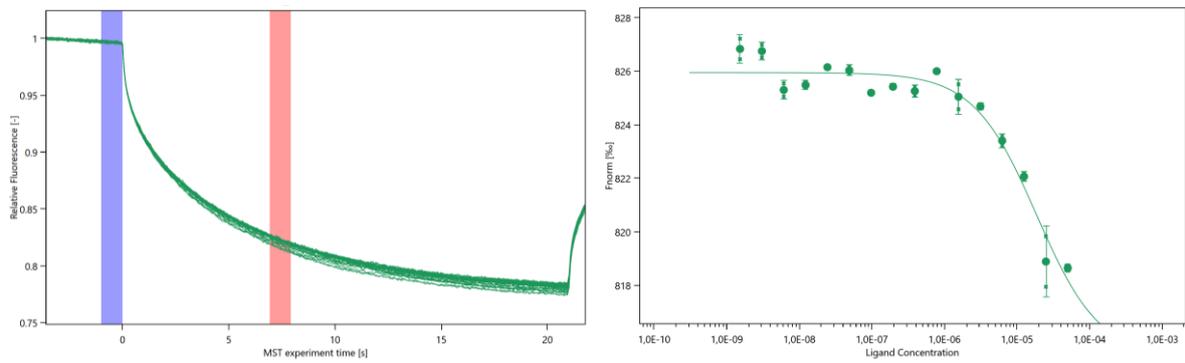


Figure 98: AtGRP7mut – BnPARCL MST-traces and binding curve.

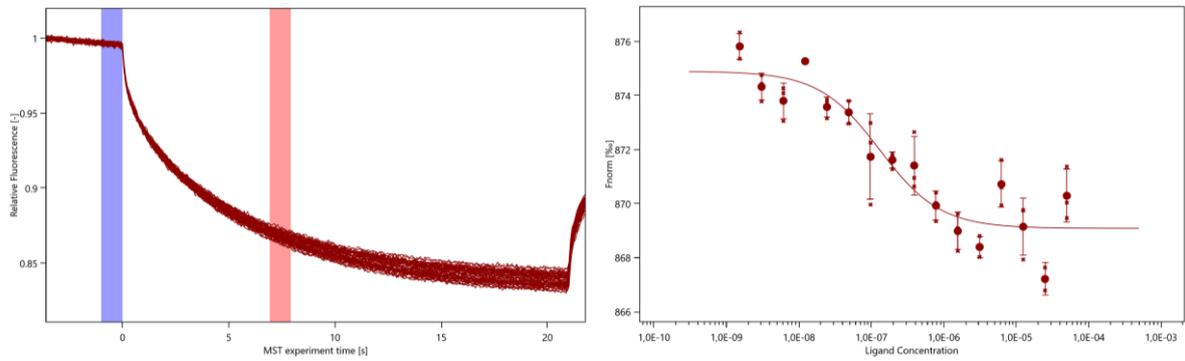


Figure 99: AtGRP7mut - miR164 MST-traces and binding curve.

## 9 Supplements

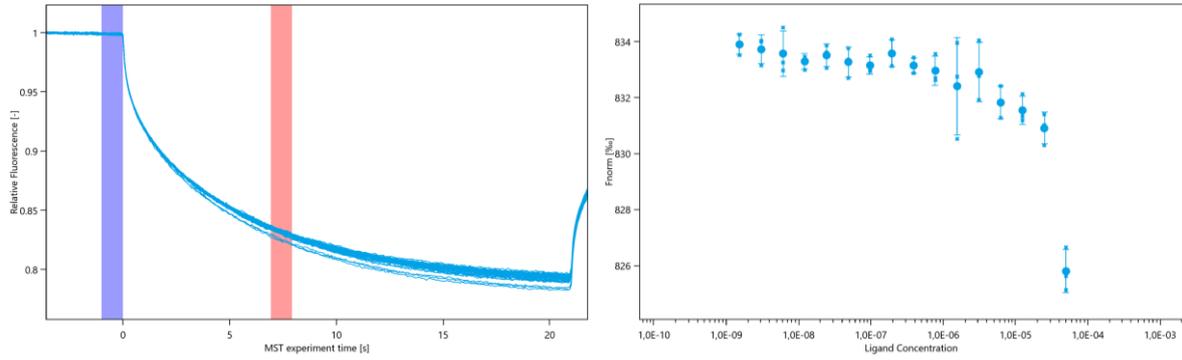


Figure 100: AtGRP7mut - *BnaCNG53670D* MST-traces and binding curve.

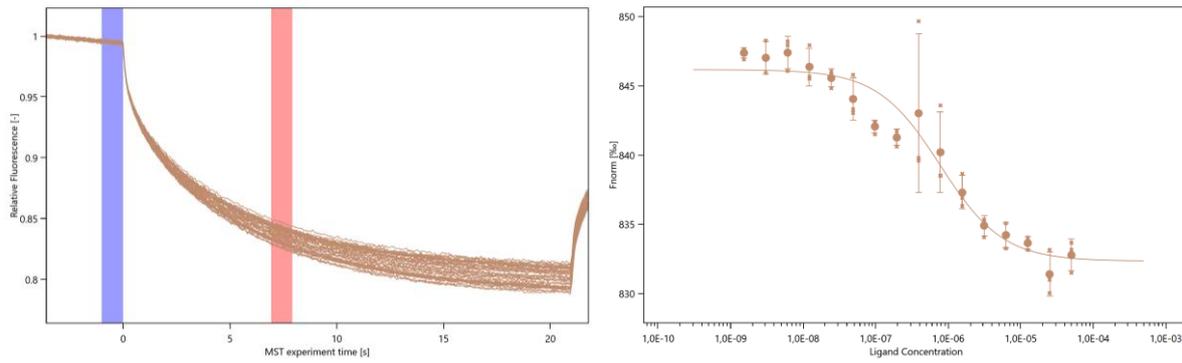


Figure 101: AtGRP7mut - *BnmiRnovel2* MST-traces and binding curve.

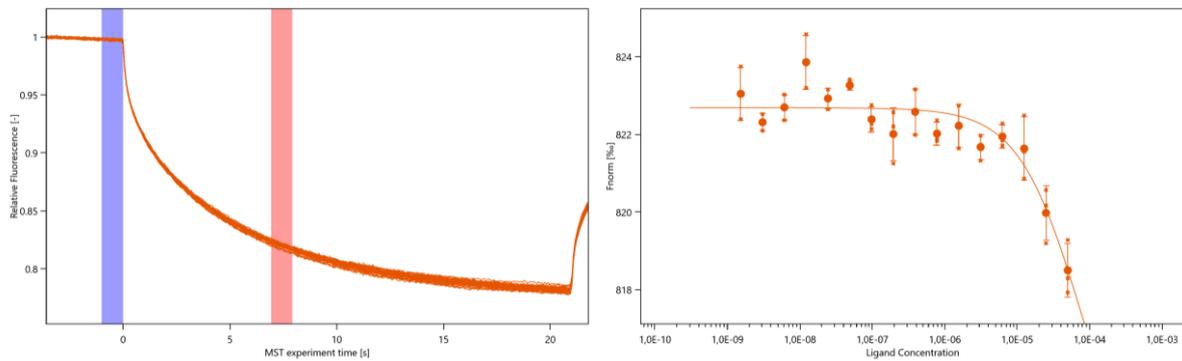


Figure 102: AtGRP7mut - *BnGRP7 RNA* MST-traces and binding curve.

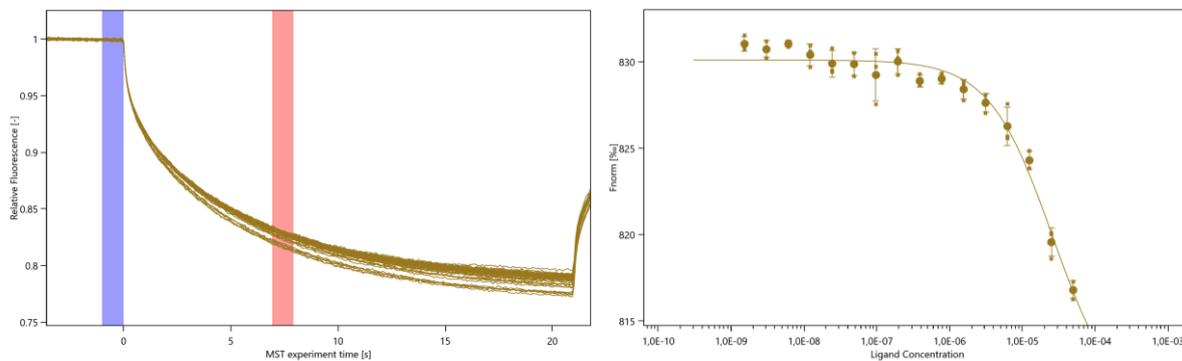


Figure 103: AtGRP7mut - *BnGRP7+UTR RNA* MST-traces and binding curve.

## 9 Supplements

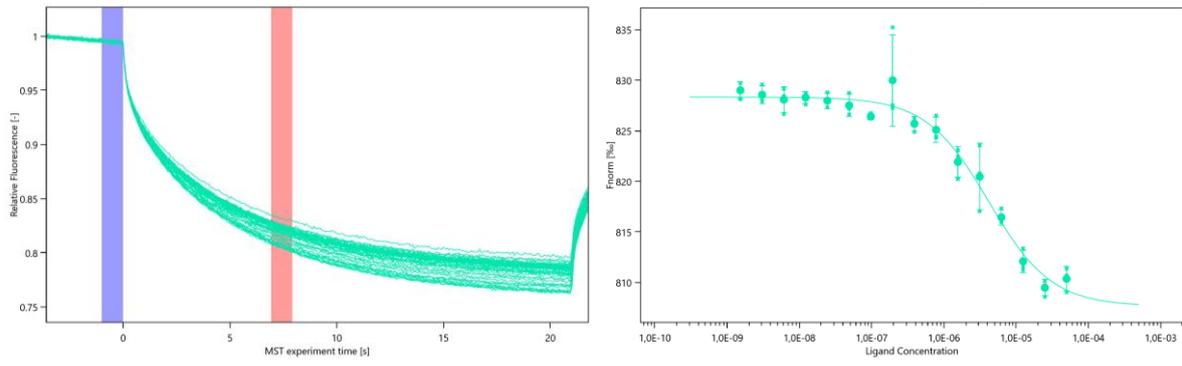


Figure 104: AtGRP7mut - *BnmiRnovel106* MST-traces and binding curve.

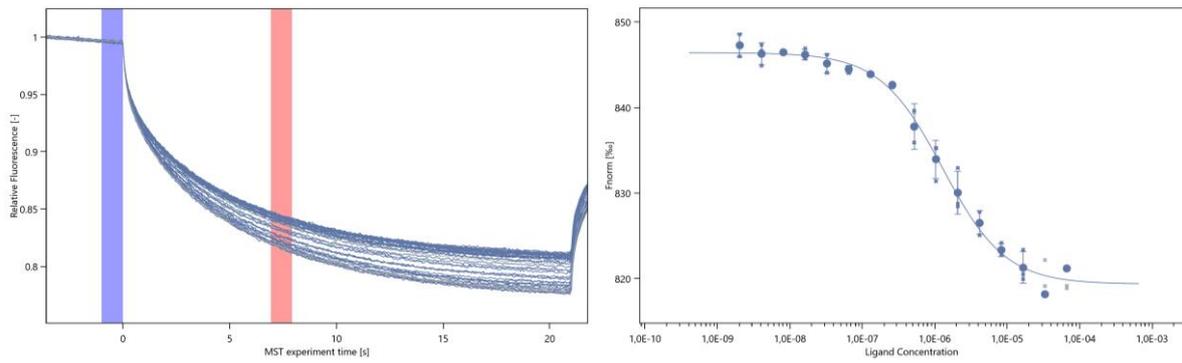


Figure 105: AtGRP7mut - *GRP7 3'UTR* MST-traces and binding curve.

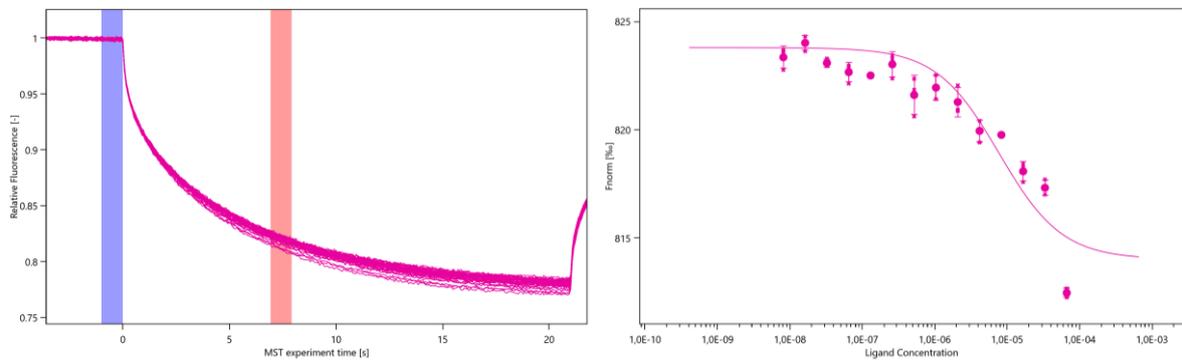


Figure 106: AtGRP7mut - *AtGRP7 RNA* MST-traces and binding curve.

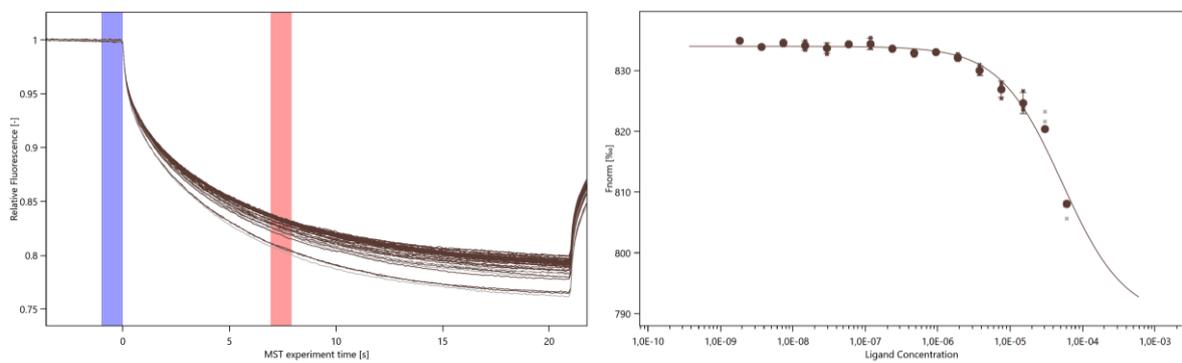


Figure 107: AtGRP7mut - *AtGRP7+UTR+Intron RNA* MST-traces and binding curve.

## 9 Supplements

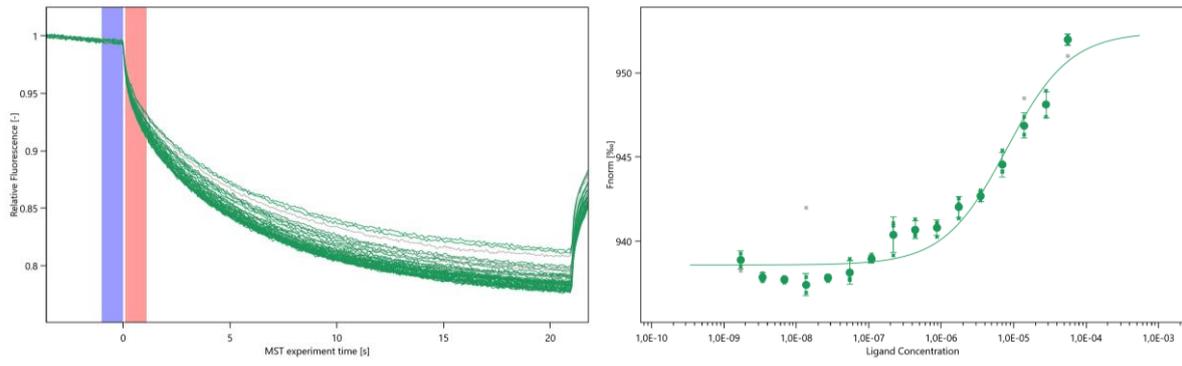


Figure 108: BnGRP7 - BnaC05g32740D MST-traces and binding curve.

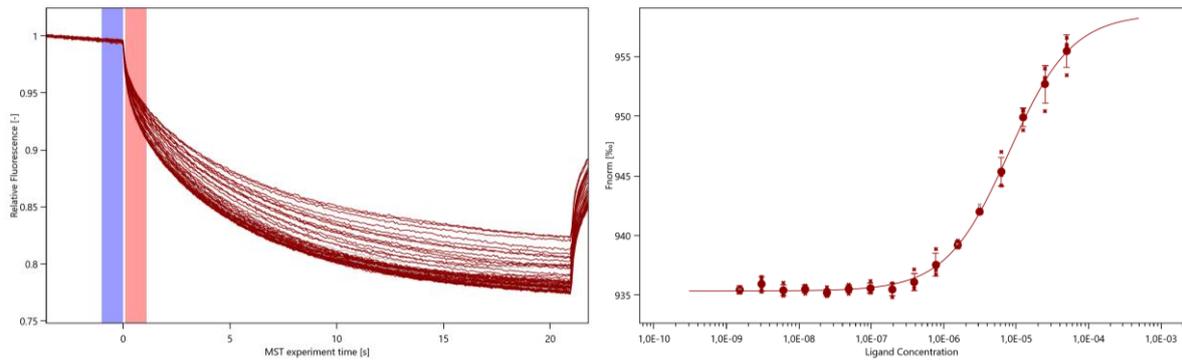


Figure 109: BnGRP7 - BnaC09g46650D MST-traces and binding curve.

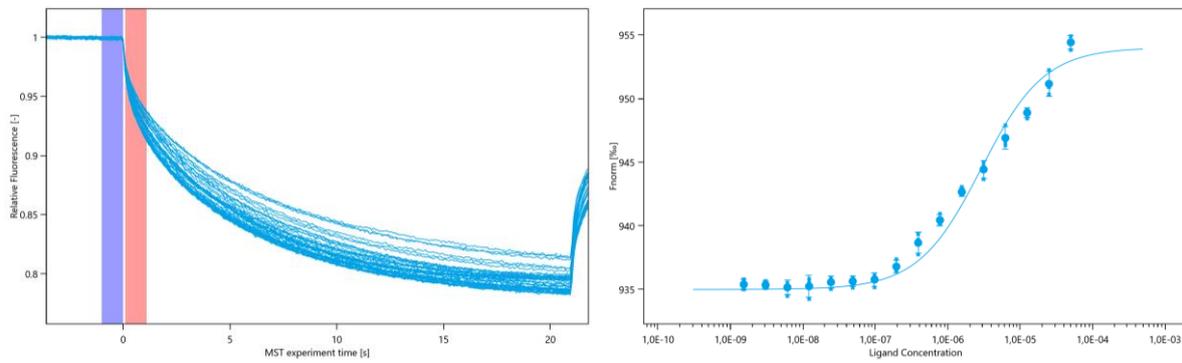


Figure 110: BnGRP7 - BnaA07g32150D MST-traces and binding curve.

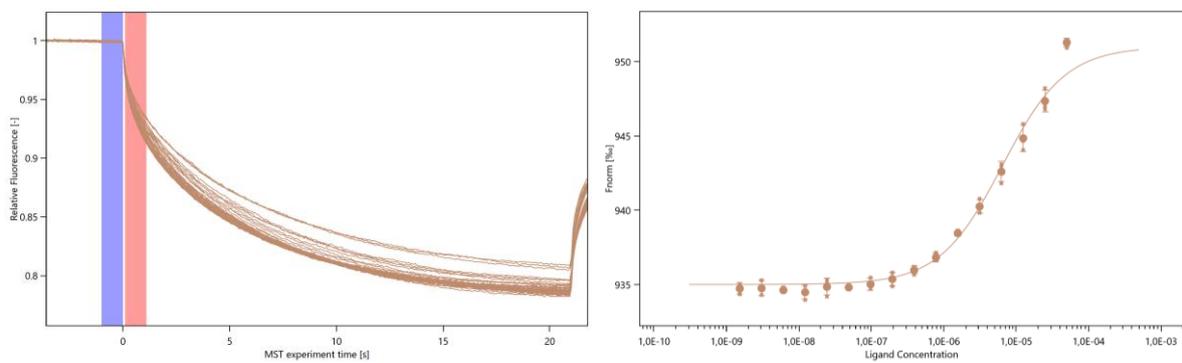


Figure 111: BnGRP7 - BnaA03g14400D MST-traces and binding curve.

## 9 Supplements

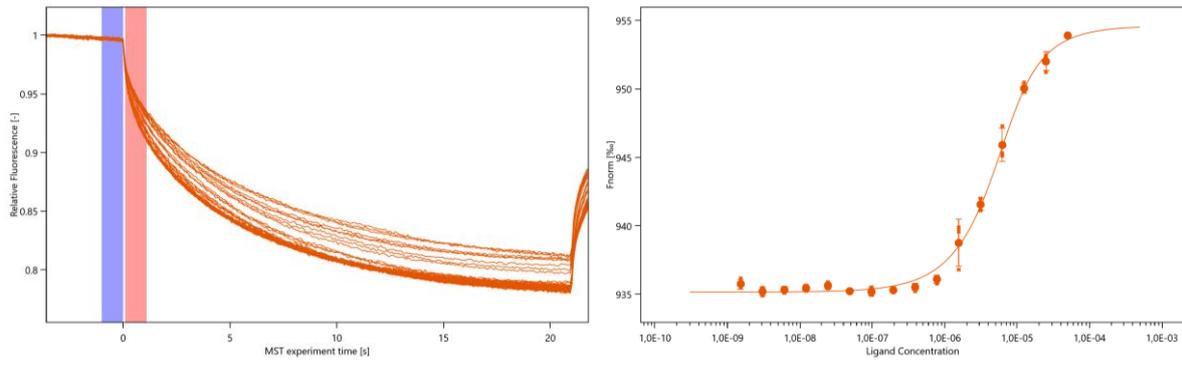


Figure 112: BnGRP7 - BnaA09g50830D MST-traces and binding curve.

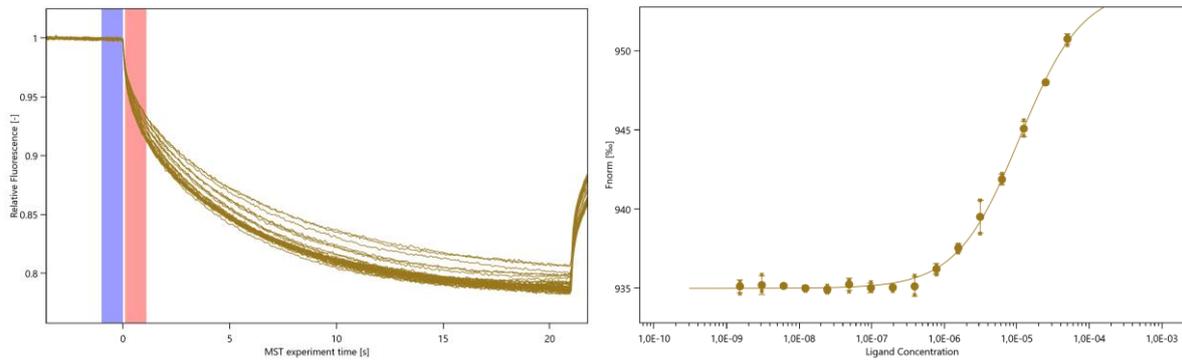


Figure 113: BnGRP7 - BnaC09g45930D MST-traces and binding curve.

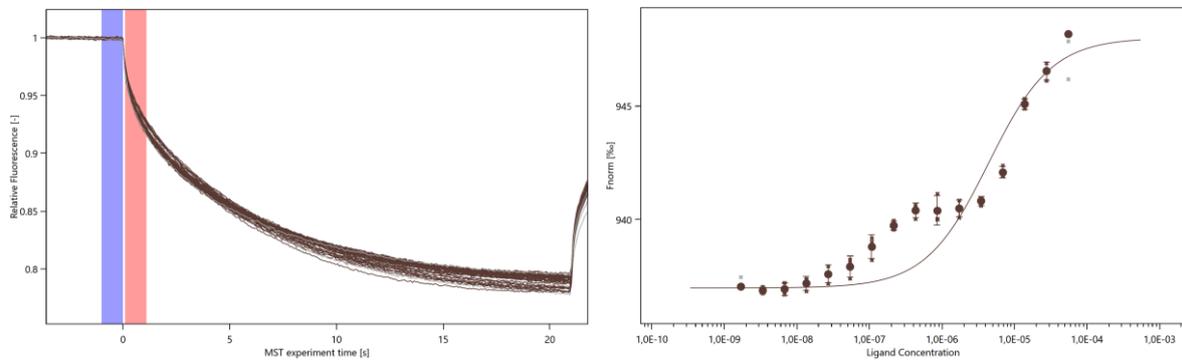


Figure 114: BnGRP7 - BnaCNG53670D MST-traces and binding curve.

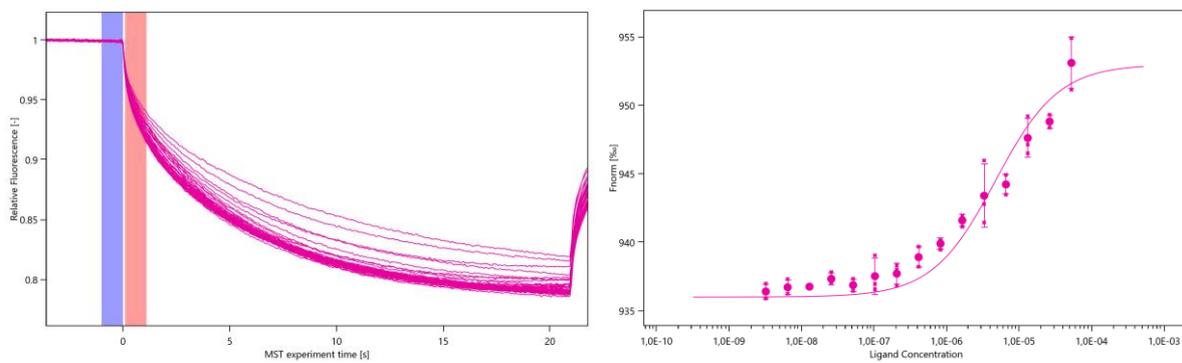


Figure 115: BnGRP7 - BnaC03g26040D MST-traces and binding curve.

## 9 Supplements

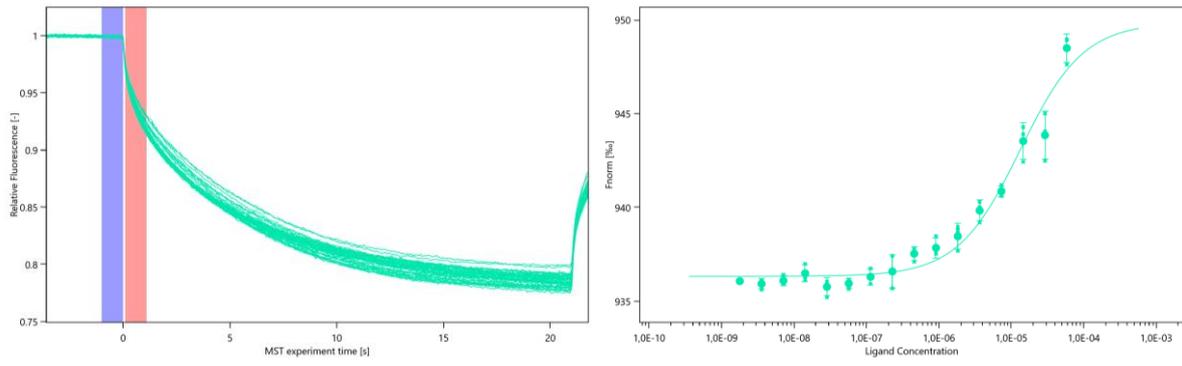


Figure 116: BnGRP7 - BnaC03g26040D m6A MST-traces and binding curve.

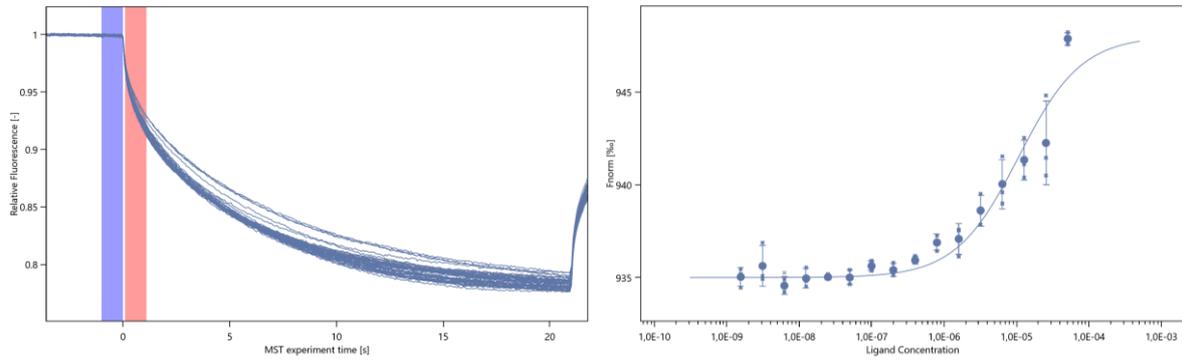


Figure 117: BnGRP7 - BnaC03g26040D m5C MST-traces and binding curve.

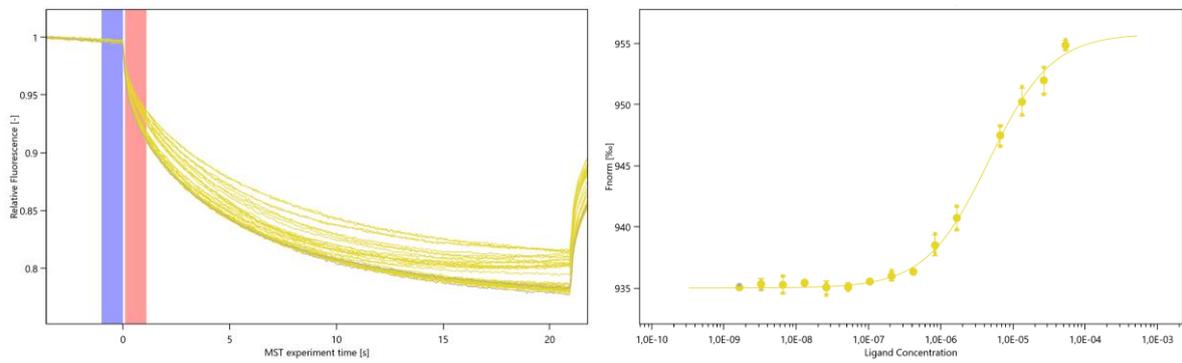


Figure 118: BnGRP7 - BnPARCL MST-traces and binding curve.

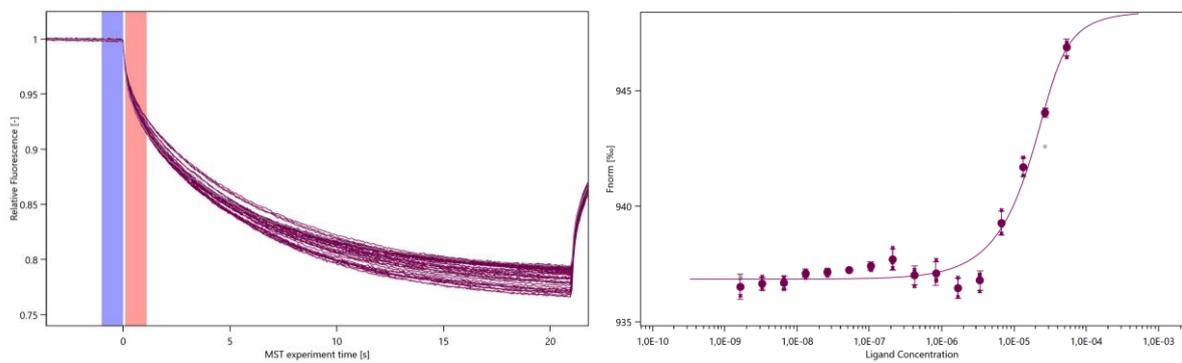


Figure 119: BnGRP7 - BnaCNG53670D m6A MST-traces and binding curve.

## 9 Supplements

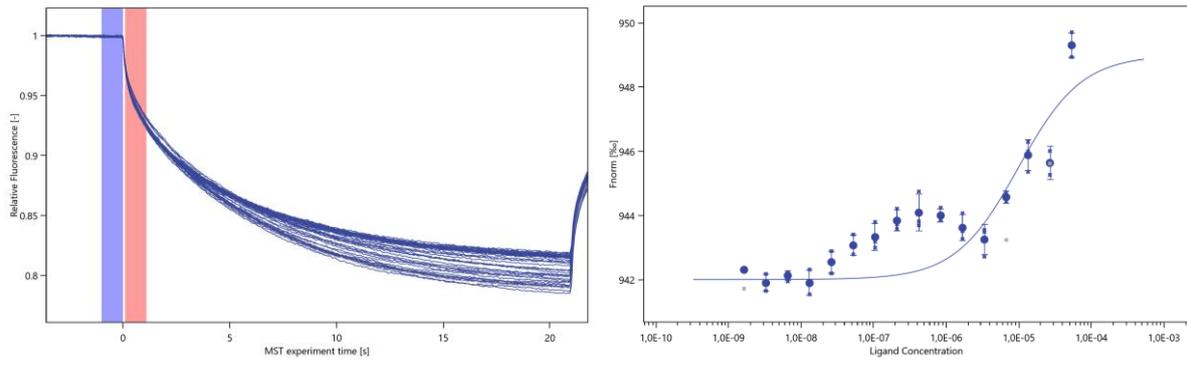


Figure 120: BnGRP7 - BnaCNG53670D m5C MST-traces and binding curve.

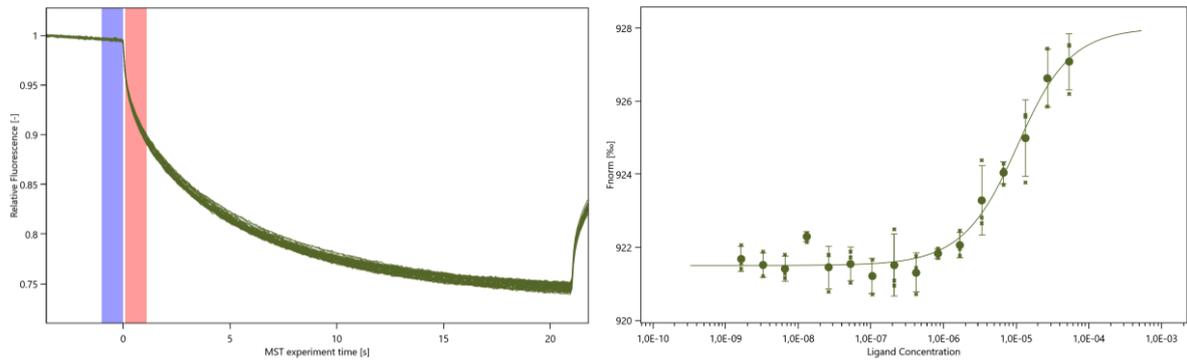


Figure 121: BnGRP7 - dsBnPARCL MST-traces and binding curve.

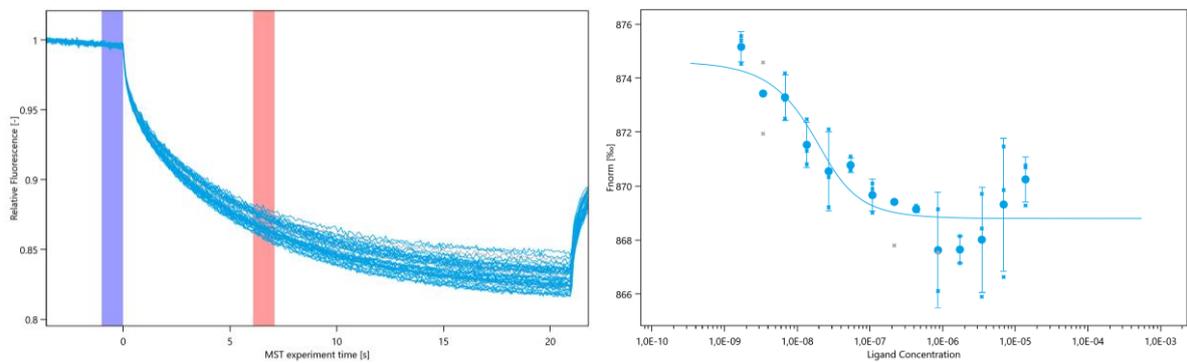


Figure 122: BnGRP7 - miR164 MST-traces and binding curve.

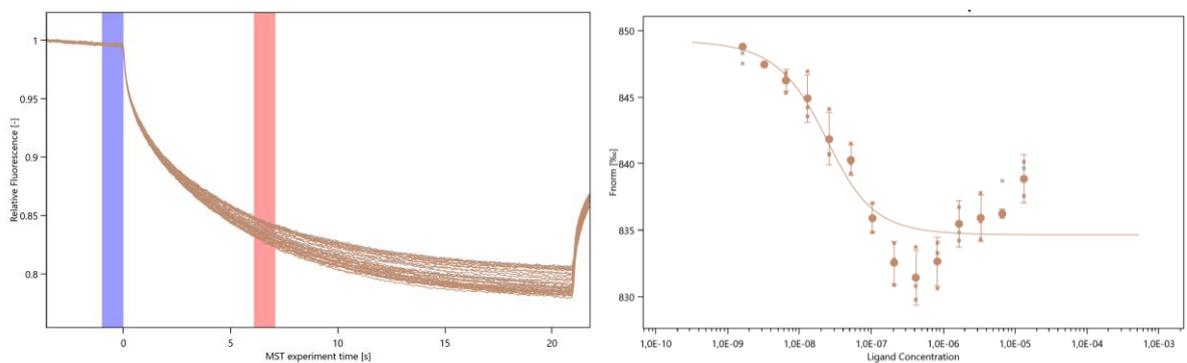


Figure 123: BnGRP7 - GRP7 3'UTR MST-traces and binding curve.

## 9 Supplements

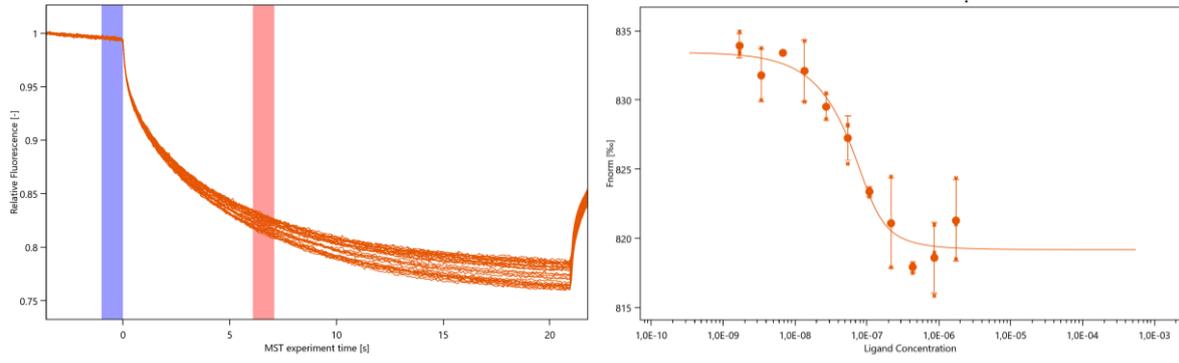


Figure 124: BnGRP7 - BnmiRnovel106 MST-traces and binding curve.

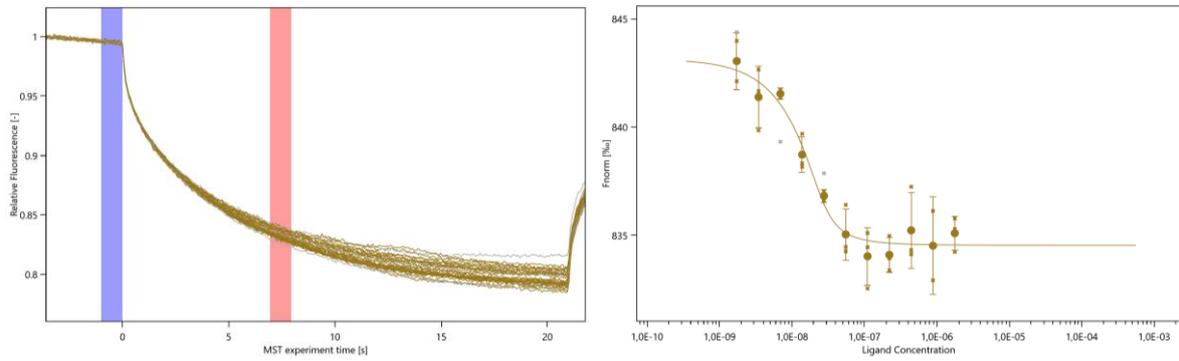


Figure 125: BnGRP7 - BnmiRnovel149 MST-traces and binding curve.

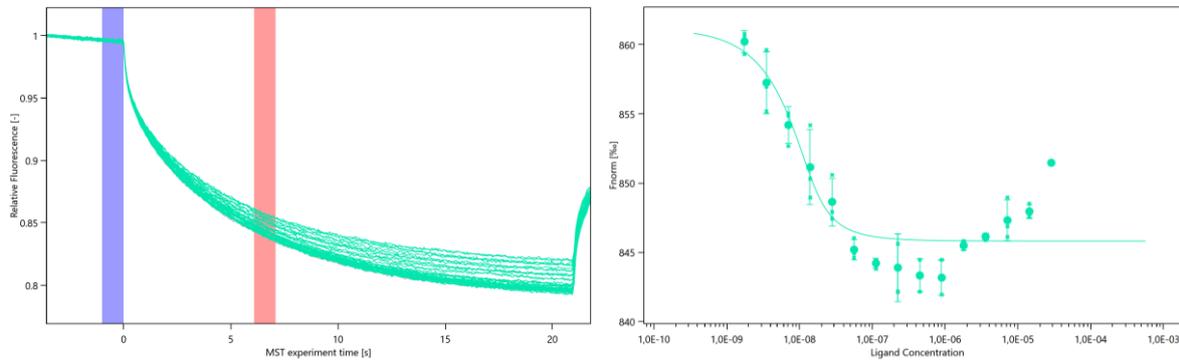


Figure 126: BnGRP7 - BnmiRnovel2 MST-traces and binding curve.

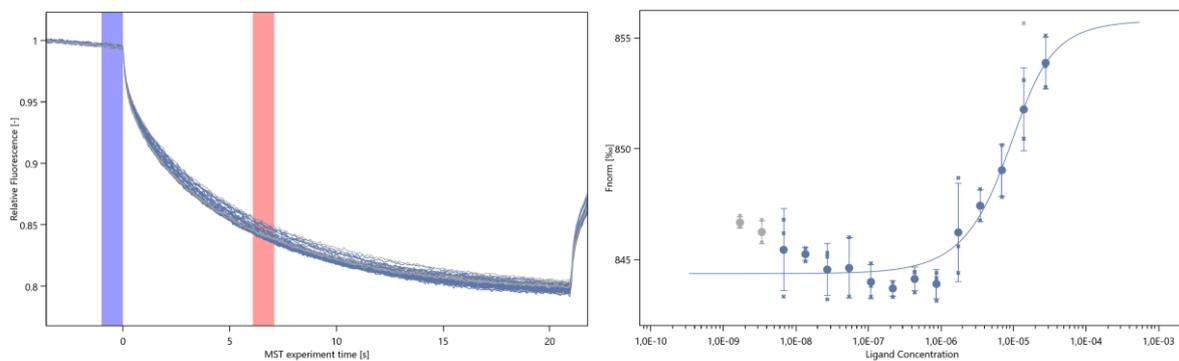


Figure 127: BnGRP7 - miR164 m5C MST-traces and binding curve.

## 9 Supplements

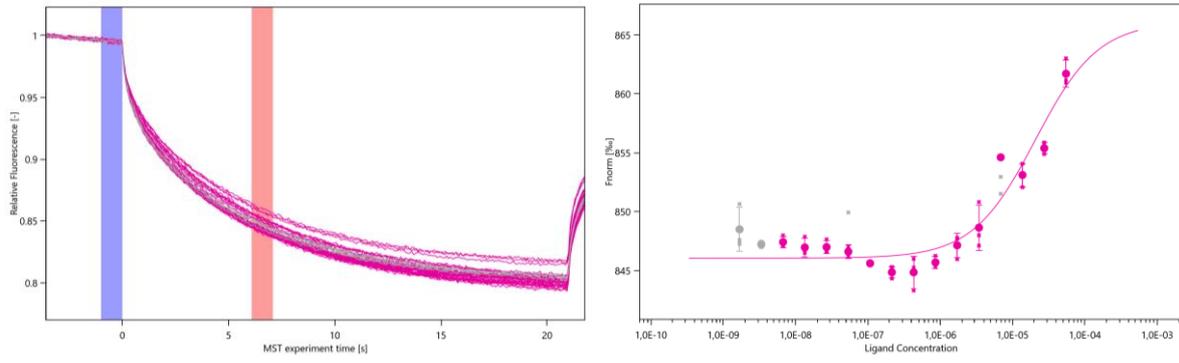


Figure 128: BnGRP7 - miR164 m6A MST-traces and binding curve.

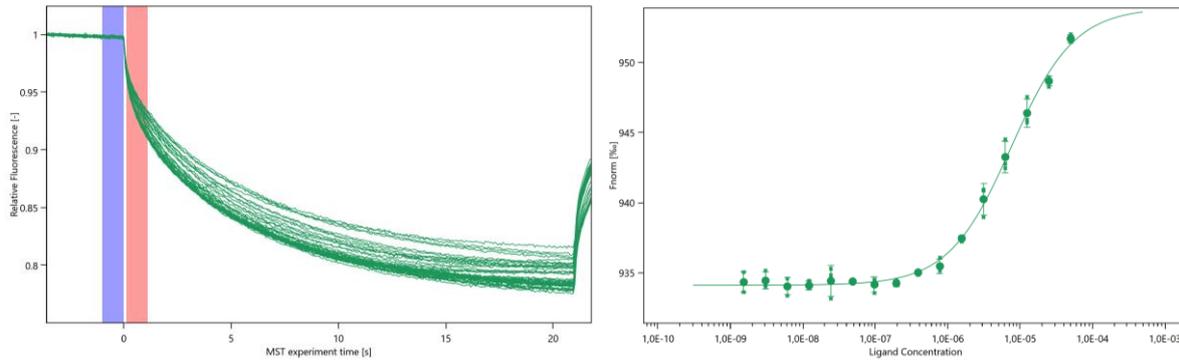


Figure 129: BnGRP7 - BnPARCL m6A MST-traces and binding curve.

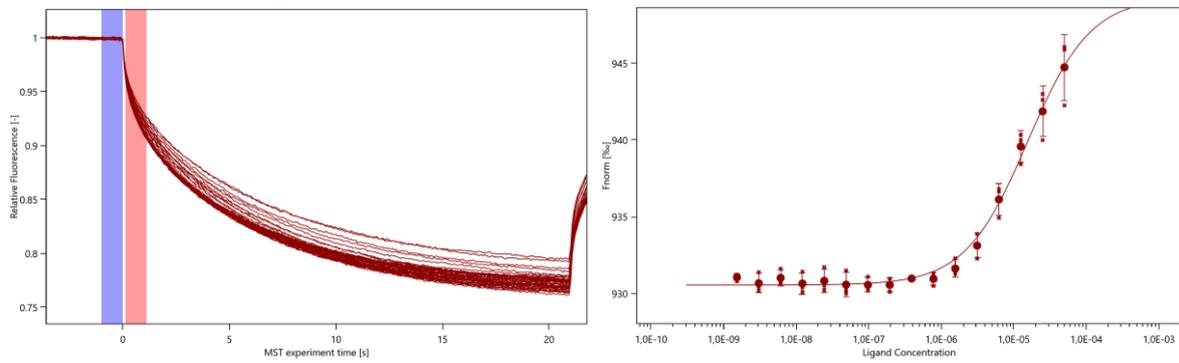


Figure 130: BnGRP7 - BnPARCL m5C MST-traces and binding curve.

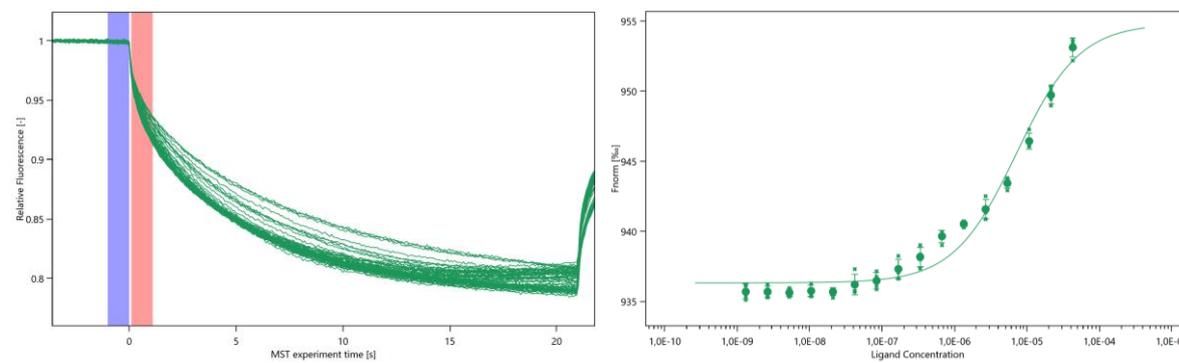


Figure 131: BnGRP7 - BnGRP7 RNA MST-traces and binding curve.

9 Supplements

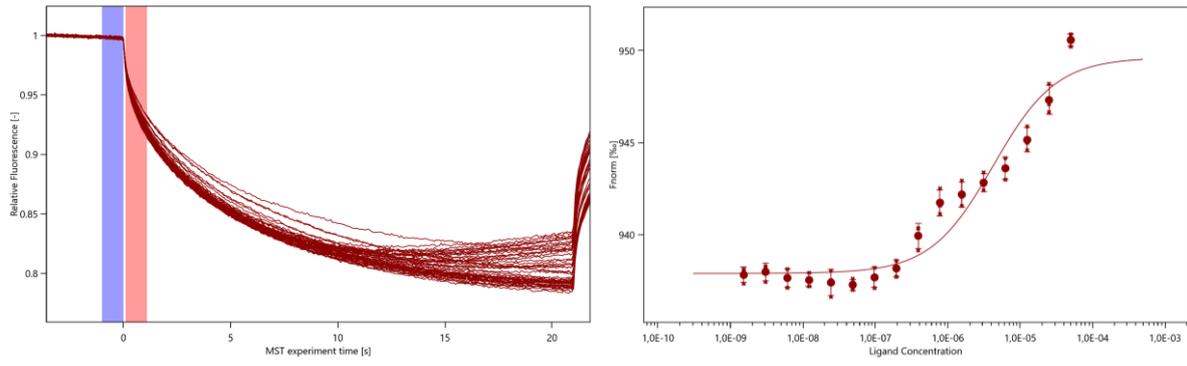


Figure 132: BnGRP7 - *BnGRP7+UTR* MST-traces and binding curve.

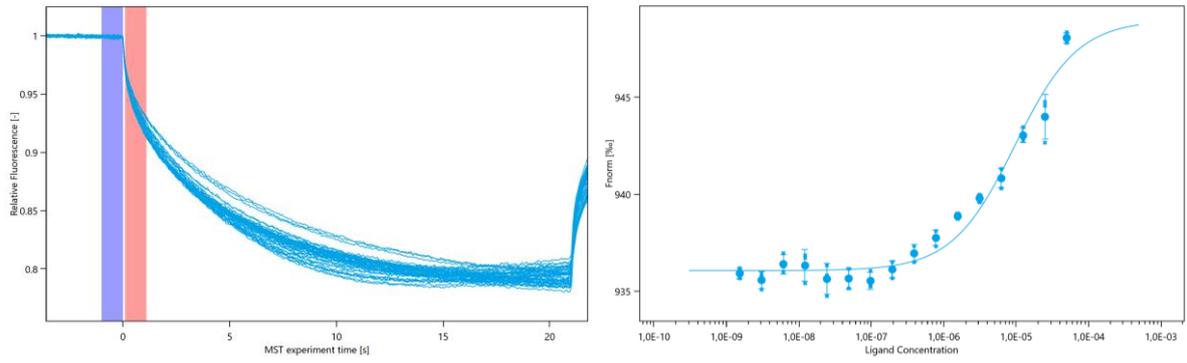


Figure 133: BnGRP7 - *BnGRP7+UTR+Intron RNA* MST-traces and binding curve.

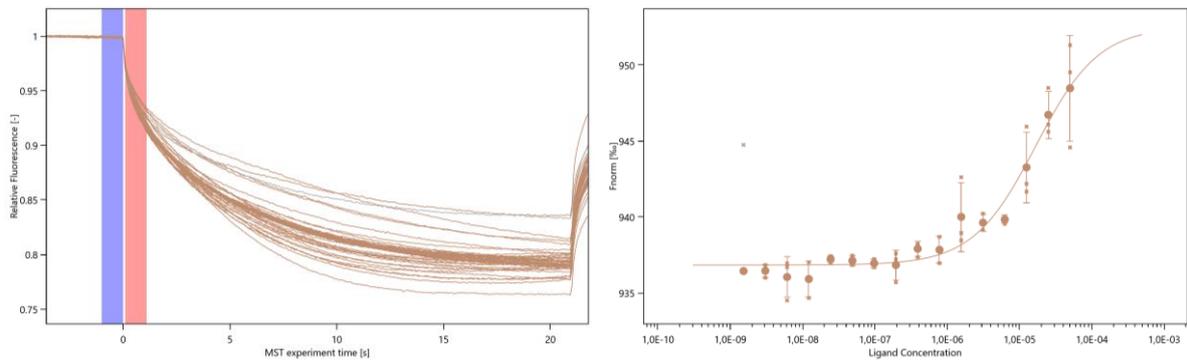


Figure 134: BnGRP7 - *AtGRP7+UTR+Intron RNA* MST-traces and binding curve.

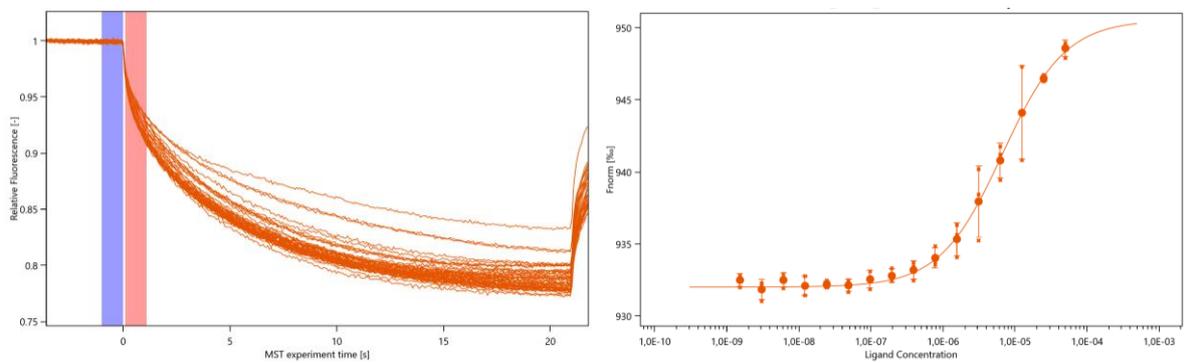


Figure 135: BnGRP7 + *AtGRP7 RNA* MST-traces and binding curve.

## 9 Supplements

### 9.2.8 Tables with dissociation constants of AtGRP7mut and AtGRP7short

**Table 36: Overview of dissociation constants ( $K_d$ ) measured for AtGRP7mut for different RNAs.**

RNA	Protein name	$K_d$ mean [ $\mu\text{M}$ ]	$K_d$ SD [ $\mu\text{M}$ ]	N
BnmiRnovel2	/	0.1341	0.0713	3
GRP7 3'UTR	AtGRP7	0.1467	0.7830	3
miR164	/	0.32	0.0563	3
BnmiRnovel106	/	3.60	0.4993	3
AtGRP7	AtGRP7	5.4037	1.4565	3
BnGRP7	BnGRP7	9.44	1.8951	2
BnPARCL mRNA	BnPARCL	16.6200	7.2973	2
AtGRP7+UTR+Intron	AtGRP7	30.1117	24.1476	3
BnGRP7+UTR	BnGRP7	30.6267	16.8817	3
BnaCNGG53670D	Putative thioredoxin H10	/	/	3

**Table 37: Overview of dissociation constants ( $K_d$ ) measured for AtGRP7short for different RNAs.**

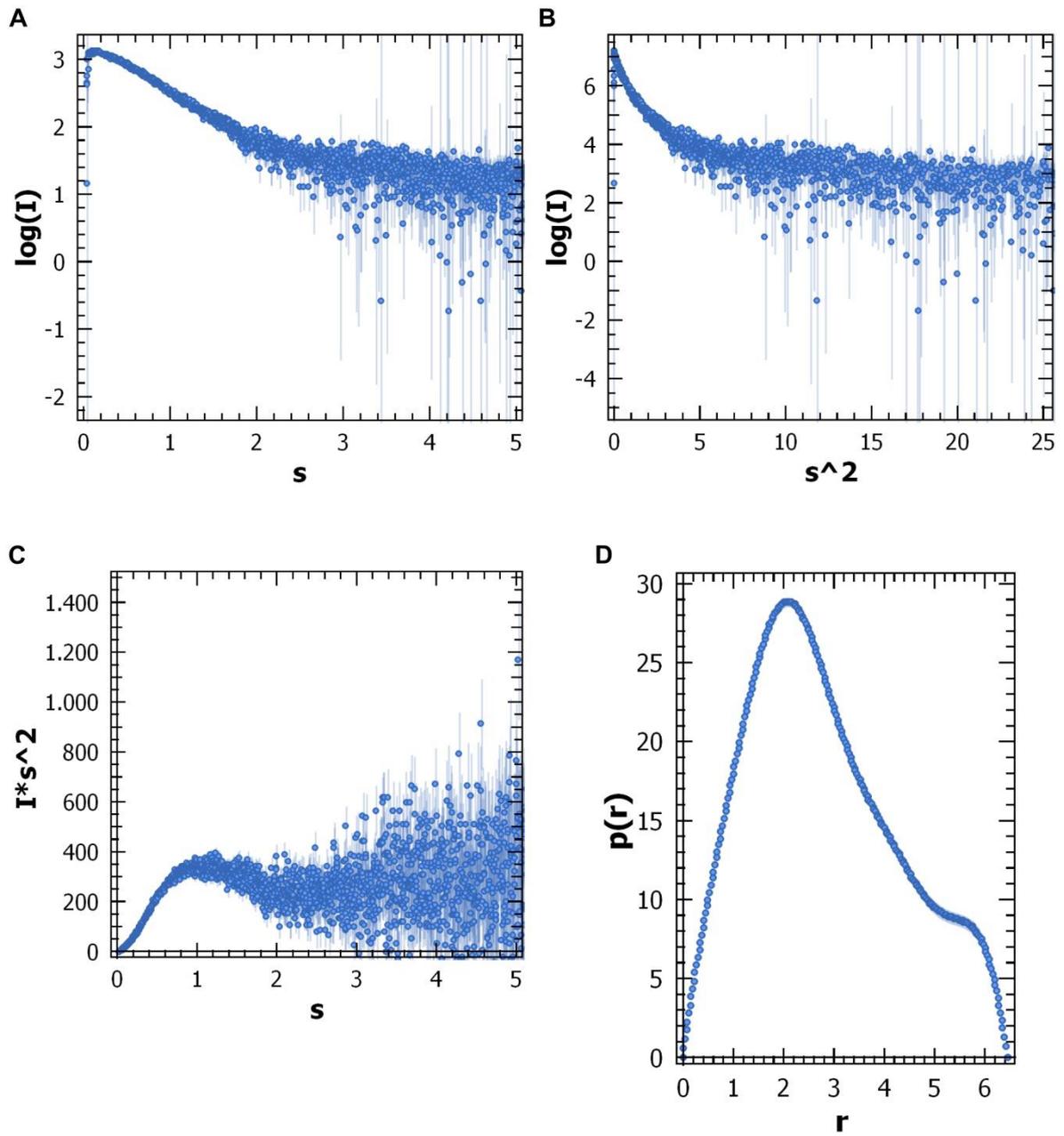
RNA	Protein name	$K_d$ mean [ $\mu\text{M}$ ]	$K_d$ SD [ $\mu\text{M}$ ]	N
miR164	/	4.49	2.3666	3
BnmiRnovel2	/	5.93	3.7217	3
GRP7 3'UTR	AtGRP7	12.1767	3.8608	3
BnmiRnovel106	/	12.3667	10.0462	3
BnPARCL mRNA	BnPARCL	/	/	3
BnaCNGG53670D	Putative thioredoxin H10	/	/	3
BnGRP7	BnGRP7	/	/	3
BnGRP7+UTR	BnGRP7	/	/	3
AtGRP7	AtGRP7	/	/	3
AtGRP7+UTR+Intron	AtGRP7	/	/	3

## 9.2.9 SEC-SAXS measurement of AtGRP7

Additional information and plots for AtGRP7 SEC-SAXS measurements.

**Table 38: Structural parameters of SEC-SAXS measurements of AtGRP7.**

Structural parameters	
<b>Guinier analysis</b>	GRP7 23 °C
<b>I(0) (cm<sup>-1</sup>)</b>	1286.46 ± 4.82
<b>Rg (Å)</b>	23.06 ± 0.16
<b>s min (Å<sup>-1</sup>)</b>	0.23
<b>sRg max (Å<sup>-1</sup>)</b>	1.30
<b>Fidelity (<i>AutoRg</i>)</b>	0.13
<b>MW</b>	
<b>from I(0) (Da)</b>	12199
<b>from Vc (Da)</b>	15200
<b>from MoW (Da)</b>	14717
<b>From Bayesian-interference (Da)</b>	14825
<b>P(r)</b>	
<b>I(0) (cm<sup>-1</sup>)</b>	1286.46 ± 4.73
<b>Rg (Å)</b>	23.06 ± 0.01
<b>D max (Å)</b>	85.33
<b>s range</b>	0.00 – 0.343
<b>Chi<sup>2</sup></b>	1.43
<b>Porod volume (Å<sup>3</sup>)</b>	26061.10
<b>Total quality estimate</b>	0.87



**Figure 136: Plots of SEC-SAXS measurements.** A) Scattering plot of AtGRP7. B) Guinier plot of AtGRP7. C) Kratky plot of AtGRP7. D) Distance distribution function of AtGRP7 generated with Gnom.

## 9 Supplements

### 9.2.10 circRNAs identified in lnc and circRNA sequencing

**Table 39: Predicted circRNAs in lnc-sequencing of *B. napus* leaf RNA.**

Location	First replicate	Second replicate	Third replicate	total read count	Gene	Junction Type	Strand	Start-End Region
LK034926 12893 3099	134	232	191	557	not_annotated	1	-	intergenic-intergenic
LK032891 85002 20613	104	102	139	345	not_annotated	2	+	intergenic-intergenic
LK032223 268479 168146	20	43	29	92	not_annotated	2	-	intergenic-intergenic
LK032105 210538 469641	12	21	36	69	not_annotated	1	-	intergenic-intergenic
LK032116 501707 197745	14	20	18	52	not_annotated	1	-	intergenic-intergenic
LK032334 301970 302466	17	21	12	50	not_annotated	2	+	intergenic-intergenic
LK032116 501707 501870	12	15	9	36	BnaA01g34300D	2	+	intergenic-exon
LK032144 174755 268532	18	7	11	36	not_annotated	2	-	intergenic-intergenic
LK032349 117157 117272	12	11	13	36	not_annotated	2	+	intergenic-intergenic
LK032349 117157 102445	12	11	13	36	BnaC03g59520D	1	-	exon-exon
LK031862 1147499 1147641	11	5	14	30	not_annotated	1	+	intergenic-intergenic
LK031862 1147499 975390	11	5	14	30	BnaA08g16820D	2	-	exon-exon
LK032334 302417 117272	8	13	9	30	not_annotated	1	+	intergenic-intergenic
LK032045 141379 53425	9	9	10	28	BnaC02g25110D	2	-	intergenic-exon
LK031822 1193072 1193636	10	7	8	25	not_annotated	1	+	intergenic-intergenic
LK032872 18226 85037	6	10	8	24	BnaC03g76110D	2	+	intergenic-intergenic
LK031822 1193072 764385	7	8	7	22	not_annotated	2	-	intergenic-intergenic
LK034680 3566 12926	2	5	15	22	not_annotated	2	-	intergenic-intergenic
LK031915 975258 975390	5	7	9	21	not_annotated	1	+	intergenic-intergenic
LK031915 975258 255398	5	7	9	21	not_annotated	2	-	intergenic-intergenic
LK044562 1621 712	6	5	9	20	not_annotated	2	-	intergenic-intergenic
LK032104 439794 210579	6	4	9	19	not_annotated	2	-	intergenic-intergenic
LK034639 8859 9046	2	14	3	19	BnaC07g47800D	2	+	exon-exon
LK034639 8859 3646	2	14	3	19	not_annotated	1	-	intergenic-intergenic
LK032591 210782 143349	7	5	6	18	not_annotated	2	-	intergenic-intergenic
LK032045 141379 143249	5	4	8	17	not_annotated	1	+	intergenic-intergenic
LK032084 380999 300286	4	6	7	17	not_annotated	2	+	intergenic-intergenic
LK034324 20103 20209	5	6	6	17	not_annotated	2	+	intergenic-intergenic
LK032713 30913 75593	3	4	8	15	not_annotated	1	-	intergenic-intergenic
LK031850 1530258 1147641	6	3	5	14	not_annotated	2	-	intergenic-intergenic
LK031928 255187 915163	2	8	4	14	not_annotated	2	-	intergenic-intergenic
LK032237 235107 302005	3	4	7	14	not_annotated	1	-	intergenic-intergenic
LK032627 143260 97939	4	7	3	14	not_annotated	2	+	intergenic-intergenic
LK032690 97847 30994	9	2	3	14	not_annotated	2	+	intergenic-intergenic
LK031787 23357 1952271	2	4	7	13	not_annotated	2	+	intergenic-intergenic

## 9 Supplements

LK031818 819003 1193636	2	9	2	13	BnaC03g18980D	1	-	exon-exon
LK032225 166944 168146	2	6	5	13	BnaCnng05580D	1	+	exon-exon
LK032548 102330 102445	2	4	7	13	not_annotated	2	+	intergenic-intergenic
LK032548 102330 210825	2	4	7	13	BnaA07g07570D	1	-	exon-exon
LK034324 20103 9046	5	4	4	13	BnaUnng00850D	1	-	exon-intergenic
LK032001 913795 915163	3	5	4	12	not_annotated	1	+	intergenic-intergenic
LK049179 674	5	2	5	12	not_annotated	2	+	intergenic-intergenic
LK031792 1952218 2037192	2	3	6	11	not_annotated	2	-	intergenic-intergenic
LK031842 762825 1530322	5	3	3	11	BnaA02g02290D	1	-	exon-exon
LK032059 371760 381063	2	4	5	11	not_annotated	1	-	intergenic-intergenic
LK032099 300241 439832	3	5	3	11	not_annotated	2	-	intergenic-intergenic
LK032116 500320 501870	2	2	7	11	BnaA01g34300D	2	+	intergenic-exon
LK036416 2907 3099	4	5	2	11	not_annotated	1	+	intergenic-intergenic
LK036416 2907 6151	4	5	2	11	BnaUnng02400D	2	-	exon-exon
LK032051 53346 371840	3	2	5	10	not_annotated	2	-	intergenic-intergenic
LK032116 469424 469641	2	5	3	10	not_annotated	1	+	intergenic-intergenic
LK032116 469424 502116	2	5	3	10	not_annotated	2	-	intergenic-intergenic
LK033288 19883 61048	2	4	4	10	not_annotated	2	-	intergenic-intergenic
LK033330 61002 36110	4	4	2	10	not_annotated	2	-	intergenic-intergenic
LK033540 36037 20209	3	3	4	10	not_annotated	2	+	intergenic-intergenic
LK031818 819003 819166	2	5	2	9	not_annotated	2	+	intergenic-intergenic
LK031809 2036223 819166	2	3	2	7	not_annotated	1	+	intergenic-intergenic
LK032225 166944 235194	2	2	3	7	not_annotated	2	-	intergenic-intergenic
LK032962 6591 19937	2	2	3	7	BnaC06g07550D	2	+	intergenic-intergenic
LK032001 913795 143249	2	2	2	6	not_annotated	2	-	intergenic-intergenic
LK036802 5926 4047	2	2	2	6	not_annotated	2	-	intergenic-intergenic
LK038596 3969 1834	2	2	2	6	not_annotated	1	-	intergenic-intergenic

**Table 40: Predicted circRNAs in lnc-sequencing of *B. napus* phloem RNA.**

Location	First replicate	Second replicate	Third replicate	total read count	Gene	Junction Type	Strand	Start-End Region
LK034926 12893 12926	459	342	376	1177	not_annotated	1	-	intergenic-intergenic
LK034680 3566 3646	106	96	101	303	not_annotated	2	-	intergenic-intergenic
LK032891 85002 85037	86	91	67	244	not_annotated	2	+	intergenic-intergenic
LK034680 3566 3646	68	56	64	188	BnaCnng45920D	1	+	exon-exon
LK032045 141379 143249	50	45	41	136	BnaC02g25110D	2	-	intergenic-exon
LK032045 141379 143249	47	46	39	132	not_annotated	1	+	intergenic-intergenic
LK032144 174755 197745	53	36	30	119	not_annotated	2	-	intergenic-intergenic
LK032104 439794 439832	37	16	29	82	not_annotated	2	-	intergenic-intergenic

## 9 Supplements

LK031850 1530258 1530301	23	20	13	56	not_annotated	2	-	intergenic-intergenic
LK033217 59750 61637	22	22	12	56	BnaC02g48340D	2	-	exon-exon
LK032105 210538 210579	12	23	20	55	not_annotated	1	-	intergenic-intergenic
LK033217 59750 61637	21	22	9	52	not_annotated	1	+	intergenic-intergenic
LK031850 1530258 1530322	15	21	15	51	not_annotated	2	-	intergenic-intergenic
LK032627 143260 143349	18	15	15	48	not_annotated	2	+	intergenic-intergenic
LK032872 18226 75593	14	16	11	41	BnaC03g76110D	2	+	intergenic-intergenic
LK032382 220221 220383	10	17	10	37	not_annotated	1	+	intergenic-intergenic
LK032962 6591 20613	8	14	15	37	BnaC06g07550D	2	+	intergenic-intergenic
LK032382 220221 220383	9	16	10	35	BnaC02g31350D	2	-	intergenic-intergenic
LK032627 143260 143349	14	9	12	35	BnaCnng12800D	1	-	exon-exon
LK032334 302417 302466	10	13	11	34	not_annotated	1	+	intergenic-intergenic
LK032028 793949 794026	16	9	8	33	not_annotated	1	-	intergenic-intergenic
LK031790 1450399 1450585	4	15	13	32	not_annotated	2	+	intergenic-intergenic
LK031790 1450399 1450585	4	15	13	32	not_annotated	1	-	intergenic-intergenic
LK031836 1118209 1118604	15	2	13	30	BnaC04g10070D	2	+	intergenic-exon
LK032697 169151 169201	9	10	8	27	not_annotated	2	+	intergenic-intergenic
LK032223 268479 268532	5	14	7	26	not_annotated	2	-	intergenic-intergenic
LK033218 21174 21230	12	8	4	24	not_annotated	1	-	intergenic-intergenic
LK031985 805439 805896	8	7	8	23	BnaC09g12120D	1	+	exon-exon
LK032001 913795 918788	9	10	3	22	not_annotated	1	+	intergenic-intergenic
LK032001 913795 915163	8	4	9	21	not_annotated	1	+	intergenic-intergenic
LK031823 1210324 1210506	10	2	8	20	not_annotated	2	-	intergenic-intergenic
LK032011 272517 272561	4	8	8	20	not_annotated	2	-	intergenic-intergenic
LK031823 1210324 1210506	10	2	6	18	BnaC03g02510D	1	+	exon-exon
LK032011 272520 272561	2	6	8	16	not_annotated	2	-	intergenic-intergenic
LK032946 7393 7466	9	4	3	16	not_annotated	2	+	intergenic-intergenic
LK032948 54376 54417	5	3	8	16	not_annotated	2	-	intergenic-intergenic
LK031818 819003 819166	8	2	5	15	BnaC03g18980D	1	-	exon-exon
LK032679 39161 39242	7	3	5	15	not_annotated	1	-	intergenic-intergenic
LK031816 348641 351847	8	3	3	14	BnaA01g09040D	1	+	exon-intergenic
LK031935 224259 225080	9	2	3	14	not_annotated	1	+	intergenic-intergenic
LK032017 753843 753977	6	4	4	14	BnaC03g60470D	2	+	exon-exon
LK032017 753843 753977	6	4	4	14	not_annotated	1	-	intergenic-intergenic
LK032028 793949 794026	7	4	3	14	BnaA09g13230D	2	+	exon-exon
LK032334 301970 302005	5	4	5	14	not_annotated	2	+	intergenic-intergenic

## 9 Supplements

LK035188 10464 10508	2	2	10	14	not_annotated	2	-	intergenic-intergenic
LK035188 10467 10508	2	3	9	14	not_annotated	2	-	intergenic-intergenic
LK044162 1839 1908	4	7	3	14	not_annotated	1	+	intergenic-intergenic
LK031823 1895973 1903835	6	3	4	13	BnaC03g03920D	2	-	intergenic-intergenic
LK032001 913795 918788	6	5	2	13	not_annotated	2	-	intergenic-intergenic
LK032274 210092 210205	7	4	2	13	not_annotated	2	+	intergenic-intergenic
LK032274 210092 210205	7	4	2	13	not_annotated	1	-	intergenic-intergenic
LK032977 72233 72483	5	2	6	13	BnaA03g06980D	1	+	intergenic-exon
LK032977 72233 72483	5	2	6	13	not_annotated	2	-	intergenic-intergenic
LK034324 20103 20209	5	4	4	13	not_annotated	2	+	intergenic-intergenic
LK034324 20103 20209	5	4	4	13	BnaUnng00850D	1	-	exon-intergenic
LK031789 2803353 2803392	4	4	4	12	not_annotated	2	-	intergenic-intergenic
LK031789 2891693 2891764	3	2	7	12	not_annotated	2	-	intergenic-intergenic
LK031818 819003 819166	6	2	4	12	not_annotated	2	+	intergenic-intergenic
LK032960 6840 6928	5	5	2	12	BnaA06g02960D	1	-	exon-exon
LK031810 672404 672481	3	3	5	11	not_annotated	2	+	intergenic-intergenic
LK032439 318419 318579	2	4	5	11	BnaA02g17620D	2	+	exon-exon
LK032439 318419 318579	2	4	5	11	not_annotated	1	-	intergenic-intergenic
LK031795 597157 597207	5	2	3	10	not_annotated	1	+	intergenic-intergenic
LK031906 786263 786407	2	5	3	10	Plant_U3	1	+	exon-exon
LK031906 786263 786407	2	5	3	10	not_annotated	2	-	intergenic-intergenic
LK031916 15785 15837	2	6	2	10	not_annotated	2	+	intergenic-intergenic
LK031917 343334 343422	4	4	2	10	not_annotated	1	+	intergenic-intergenic
LK031923 1127884 1127930	4	3	3	10	not_annotated	2	-	intergenic-intergenic
LK032028 788272 788382	5	2	3	10	not_annotated	2	+	intergenic-intergenic
LK032028 788272 788382	5	2	3	10	not_annotated	1	-	intergenic-intergenic
LK031818 819167 819282	2	5	2	9	not_annotated	2	+	intergenic-intergenic
LK031818 819167 819282	2	5	2	9	BnaC03g18980D	1	-	exon-exon
LK031929 5744 5932	4	2	3	9	not_annotated	2	+	intergenic-intergenic
LK031929 5744 5932	4	2	3	9	BnaC06g27870D	1	-	exon-exon
LK031935 224259 225080	4	2	3	9	not_annotated	2	-	intergenic-intergenic
LK032116 105778 105834	3	3	3	9	not_annotated	2	+	intergenic-intergenic
LK032548 102330 102445	5	2	2	9	not_annotated	2	+	intergenic-intergenic
LK032548 102330 102445	5	2	2	9	BnaA07g07570D	1	-	exon-exon
LK032948 54370 54417	3	2	4	9	not_annotated	2	-	intergenic-intergenic
LK031850 704103 704192	3	2	3	8	not_annotated	2	+	intergenic-intergenic
LK032097 624142 624186	2	4	2	8	not_annotated	2	-	intergenic-intergenic

## 9 Supplements

LK032253 328739 328796	2	3	3	8	not_annotated	2	+	intergenic-intergenic
LK032613 105563 105701	4	2	2	8	BnaA02g04940D	2	+	exon-exon
LK032613 105563 105701	4	2	2	8	not_annotated	1	-	intergenic-intergenic
LK033781 17206 17256	3	2	3	8	not_annotated	2	-	intergenic-intergenic
LK032028 788515 788562	3	2	2	7	not_annotated	2	-	intergenic-intergenic
LK032432 218637 218690	2	3	2	7	not_annotated	1	-	intergenic-intergenic
LK032908 40503 40559	2	2	3	7	not_annotated	2	-	intergenic-intergenic
LK031807 134983 135057	2	2	2	6	not_annotated	2	-	intergenic-intergenic
LK031818 1588660 1588862	2	2	2	6	BnaC03g20240D	2	+	exon-exon
LK031818 1588660 1588862	2	2	2	6	not_annotated	1	-	intergenic-intergenic
LK031850 452816 454048	2	2	2	6	BnaA07g34720D	1	+	exon-exon

**Table 41: Novogene-predicted circRNAs in total leaf RNA from circRNA enriched sequencing.**

Location	leaf circ total RC	found by circtools	strand
LK032910:68256-68831	13.8007276	yes	+
LK032970:16047-16675	15.9493964	yes	+
LK032212:228634-230189	18.4048494	yes	+
LK032110:230540-230848	19.9423926	yes	-
LK031822:1193071-1193636	95.0414784	yes	+
LK031947:224263-224984	6.56669096	yes	+
LK032228:311453-312041	5.39228346	yes	-
LK032482:287533-288111	5.83684711	yes	+
LK031850:452815-454048	5.18786307	no	+
LK032578:86847-87311	4.03778495	yes	+
LK031813:2017073-2018095	3.46272283	yes	+
LK032212:228634-232236	2.9180616	yes	+
LK032212:228639-232236	2.18877144	yes	+
LK032091:485236-485981	1.16188185	no	-
LK032416:57834-58008	1.46004333	no	-
LK032372:242286-323778	5.18786307	no	-

**Table 42: Novogene-predicted circRNAs in total phloem sap RNA from enriched circRNA sequencing.**

location	phloem circ total RC	found with circtools	strand
LK031862:570939-571436	36.5696067	no	-
LK031817:1087665-1087961	57.7771026	no	-
LK032225:9186-9529	0.81413875	no	-
LK033075:72990-75706	3.78338379	no	-
LK031902:1054924-1055373	1.05445409	no	-
LK032372:242286-323778	5.27008789	no	-

## 9 Supplements

**Table 43: DCC-predicted circRNAs from leaf circRNA enriched sequencing**

Location	first	second	total RC	Gene	Junction Type	Strand	Start-End Region
LK032891 85002 85037	1929	1893	3822	not_annotated	2	+	intergenic-intergenic
LK032334 302417 302466	118	163	281	not_annotated	1	+	intergenic-intergenic
LK032334 301970 302005	93	137	230	not_annotated	2	+	intergenic-intergenic
LK031822 1193072 1193636	57	34	91	not_annotated	1	+	intergenic-intergenic
LK032223 268479 268532	41	34	75	not_annotated	2	-	intergenic-intergenic
LK031822 1193072 1193636	33	27	60	not_annotated	2	-	intergenic-intergenic
LK032104 439794 439832	27	12	39	not_annotated	2	-	intergenic-intergenic
LK031915 991716 991907	18	20	38	BnaC09g27520D	1	-	intergenic-exon
LK031915 991716 991907	17	20	37	not_annotated	2	+	intergenic-intergenic
LK032067 408766 408844	20	13	33	not_annotated	2	-	intergenic-intergenic
LK032067 408766 408844	17	13	30	not_annotated	1	+	intergenic-intergenic
LK032970 16048 16675	17	12	29	BnaA09g51780D	1	+	exon-exon
LK031809 2036223 2037192	11	17	28	not_annotated	1	+	intergenic-intergenic
LK031865 390401 391316	14	13	27	BnaC02g05730D	1	-	exon-exon
LK032225 166944 168146	11	16	27	BnaCnng05580D	1	+	exon-exon
LK032116 501707 502081	16	10	26	BnaA01g34300D	2	+	intergenic-exon
LK031915 964776 964834	7	15	22	not_annotated	1	+	intergenic-intergenic
LK032057 436723 436887	17	5	22	BnaC04g15560D	1	+	exon-intergenic
LK032057 436723 436887	17	5	22	not_annotated	2	-	intergenic-intergenic
LK032212 228635 230189	18	4	22	BnaCnng05120D	1	+	intergenic-exon
LK032225 166944 167815	14	7	21	BnaCnng05580D	1	+	exon-exon
LK032970 16048 16675	9	12	21	not_annotated	2	-	intergenic-intergenic
LK031819 189546 190258	13	7	20	BnaA05g03610D	1	+	exon-exon
LK032057 437003 437223	6	13	19	BnaC04g15560D	1	+	intergenic-intergenic
LK032057 437003 437223	6	13	19	not_annotated	2	-	intergenic-intergenic
LK032110 230541 230848	6	12	18	BnaC08g14490D	1	-	exon-exon
LK032116 501707 502081	10	8	18	not_annotated	1	-	intergenic-intergenic
LK032212 228635 230189	14	4	18	tRNA-Pro	2	-	intergenic-intergenic
LK032198 241686 242417	7	10	17	BnaC06g12160D	1	+	exon-exon
LK031889 298418 299059	9	6	15	not_annotated	1	-	intergenic-intergenic
LK032225 166944 167815	8	7	15	not_annotated	2	-	intergenic-intergenic
LK032045 141379 143249	4	10	14	not_annotated	1	+	intergenic-intergenic
LK032045 141379 143249	4	10	14	BnaC02g25110D	2	-	intergenic-exon
LK032110 230541 230848	6	8	14	not_annotated	2	+	intergenic-intergenic
LK032910 68257 68831	2	12	14	BnaC07g38670D	1	+	exon-exon
LK032947 65992 66306	3	10	13	BnaCnng19950D	1	+	exon-intergenic
LK034324 762 866	3	10	13	not_annotated	1	+	intergenic-intergenic
LK031842 762825 764385	4	8	12	not_annotated	2	+	intergenic-intergenic
LK031842 762825 764385	4	8	12	BnaA02g02290D	1	-	exon-exon
LK032334 301970 302050	3	8	11	not_annotated	2	+	intergenic-intergenic
LK031819 189546 190258	4	6	10	not_annotated	2	-	intergenic-intergenic
LK031872 1328 1881	3	7	10	not_annotated	1	-	intergenic-intergenic

## 9 Supplements

LK031872 199036 199094	3	7	10	not_annotated	1	+	intergenic-intergenic
LK031875 464792 465762	4	6	10	BnaC04g46650D	1	-	exon-exon
LK031915 991914 992095	6	4	10	not_annotated	2	+	intergenic-intergenic
LK031915 991914 992095	6	4	10	BnaC09g27520D	1	-	exon-exon
LK032320 376111 376169	8	2	10	not_annotated	1	-	intergenic-intergenic
LK032578 86848 87311	3	7	10	BnaA02g14340D	1	+	exon-exon
LK032627 9972 10271	6	4	10	not_annotated	2	-	intergenic-intergenic
LK032627 9530 9627	5	4	9	not_annotated	1	-	intergenic-intergenic
LK034789 8788 9266	6	3	9	not_annotated	2	+	intergenic-intergenic
LK034789 8788 9266	6	3	9	not_annotated	1	-	intergenic-intergenic
LK031813 2017074 2018095	6	2	8	BnaC03g53070D	1	+	exon-exon
LK031885 797638 797687	5	3	8	not_annotated	1	+	intergenic-intergenic
LK031947 224267 224984	5	3	8	BnaC07g10790D	1	+	intergenic-exon
LK032116 469424 469641	5	3	8	not_annotated	1	+	intergenic-intergenic
LK032116 469424 469641	5	3	8	not_annotated	2	-	intergenic-intergenic
LK032116 501707 501870	3	5	8	not_annotated	1	-	intergenic-intergenic
LK032334 302006 302050	4	4	8	not_annotated	2	+	intergenic-intergenic
LK034926 12893 12926	2	6	8	not_annotated	1	-	intergenic-intergenic
LK031872 982911 983730	2	5	7	BnaA10g15450D	1	-	exon-exon
LK031915 991920 992009	2	5	7	not_annotated	1	+	intergenic-intergenic
LK031915 991920 992009	2	5	7	BnaC09g27520D	2	-	exon-exon
LK031947 224267 224984	4	3	7	not_annotated	2	-	intergenic-intergenic
LK032079 484491 484556	3	4	7	not_annotated	2	+	intergenic-intergenic
LK032116 501707 501870	3	4	7	BnaA01g34300D	2	+	intergenic-exon
LK032198 241686 242417	2	5	7	not_annotated	2	-	intergenic-intergenic
LK032212 228640 230189	4	3	7	BnaCnng05120D	1	+	intergenic-exon
LK032228 311454 312041	2	5	7	BnaA02g25170D	1	-	exon-exon
LK032791 112834 112904	5	2	7	not_annotated	1	-	intergenic-intergenic
LK032840 20094 20608	4	3	7	not_annotated	1	+	intergenic-intergenic
LK032320 384070 384141	3	3	6	not_annotated	2	+	intergenic-intergenic
LK032482 287534 288111	2	4	6	BnaC09g25960D	1	+	exon-intergenic
LK031865 384271 391316	3	2	5	BnaC02g05730D	1	-	exon-exon
LK031967 7043 7475	3	2	5	BnaA09g38170D	1	+	exon-exon
LK032351 206367 206977	3	2	5	BnaA03g19530D	1	+	exon-exon
LK031850 1640119 1640623	2	2	4	BnaA07g32220D	1	-	exon-exon
LK032026 17855 18527	2	2	4	BnaC04g29310D	1	-	exon-exon
LK032079 616660 619264	2	2	4	BnaC04g02920D	1	+	exon-exon
LK032805 32586 33025	2	2	4	BnaA02g28180D	1	-	exon-exon

## 9 Supplements

**Table 44: DCC-predicted circRNAs from phloem circRNA enriched sequencing.**

location	First	Second	total RC	Gene	JunctionType	Strand	Start-End Region
LK032891 85002 85037	479	752	1231	not_annotated	2	+	intergenic-intergenic
LK032960 6840 6928	152	108	260	BnaA06g02960D	1	-	exon-exon
LK032334 301970 302005	34	44	78	not_annotated	2	+	intergenic-intergenic
LK034926 12893 12926	22	43	65	not_annotated	1	-	intergenic-intergenic
LK032960 6840 6928	18	12	30	Plant_SRP	2	+	exon-exon
LK031813 2017074 2018095	8	19	27	BnaC03g53070D	1	+	exon-exon
LK031813 2017074 2018095	8	19	27	not_annotated	2	-	intergenic-intergenic
LK033027 67763 67844	10	17	27	not_annotated	2	+	intergenic-intergenic
LK031797 1982759 1983857	4	20	24	BnaC06g35330D	1	+	exon-exon
LK034680 3566 3646	3	18	21	not_annotated	2	-	intergenic-intergenic
LK032020 1792 2028	10	4	14	not_annotated	1	+	intergenic-intergenic
LK032020 1792 2028	10	4	14	BnaA06g20290D	2	-	exon-exon
LK033445 40218 40304	8	3	11	not_annotated	1	+	intergenic-intergenic
LK034680 3566 3646	3	5	8	BnaCnng45920D	1	+	exon-exon
LK033921 17256 18908	3	4	7	BnaA01g34800D	2	+	exon-intergenic