Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Masterthesis

Faculty of Life Sciences
Department of Biotechnology

Reverse Engineering of a Ribosome Profiling Workflow

Silke Machauer

27.02.2015

Hochschule für Angewandte Wissenschaften Hamburg
Fakultät Life Sciences
Department Biotechnologie
Lohbrügger Kirchstraße 65
21033 Hamburg

In cooperation with:

University of Glasgow
Institute of Molecular Cell and Systems Biology
Davidson Building
Glasgow G12 8QQ

Author: Silke Machauer
Submission Date: 27.02.2015

1. Examiner: Prof. Dr. Ernst A. Sanders
2. Examiner:Prof Neil Bulleid

Acknowledgments

I would firstly like to thank Prof Neil Bulleid for giving me the chance to carry out the work for my Master Thesis in his research group. The project he had thought of for me was not only a great challenge but also helped me to gain valuable experience and skills.

I would also like to thank Marie-Anne and Ojay for fruitful discussions and their help, and of course everybody else in the group for their friendship and for welcoming me as a new member of the team.

Gratitude also goes to my dear friends and family back home in Germany and to new friends I made in Scotland.

Most of all, I would like to thank my partner in life and in love, Graeme – Hasi – Morris, without whom I would not be where I am today, literally. Thank you for your help and support, for enduring me in my bad moods and for enjoying the good times with me. I love you!

Statutory Declaration

I hereby assure that this thesis is a result of my personal work and that no other than the indicated aids have been used for its completion. Furthermore I assure that all quotations and statements that have been inferred literally or in a general manner from published or unpublished writings are marked as such. Beyond this I assure that the work has not been used, neither completely nor in parts, to pass any previous examination.

_____

Glasgow, 27.02.2015
Silke Machauer

Abstract

Ribosome profiling is a relatively new method that was only developed in 2009, but it has already enabled researches all over the world to obtain remarkable findings. By isolating the ribosome protected fragments and converting them into a cDNA library, the position on the mRNA of all ribosomes that were engaged in the process of translation at the time of cell lysis can be determined with the help of a high throughput sequencing technique. Because the method is still not well established and not yet very prevalent, the aim of this work was to reverse engineer the workflow, examining every single step closely and introducing an optimised ribosome profiling protocol for our laboratory with the future prospects of developing it further into a protocol for selective ribosome profiling that will hopefully help understanding the role of protein disulfide isomerases in co-translational folding processes.

Contents

Abbreviations

AP	Affinity Purification
cDNA	complementary DNA
DSP	Dithiobis [Succinimidyl Propionate]
E. coli	*Escherichia coli*
ER	Endoplasmic Reticulum
Ero1	ER Oxidoreductin 1
ESP	1-Ethyl-3-[3-Dimethylaminopropyl]
FBS	Foetal Bovine Serum
IMAC	Immobilised Ion Affinity Chromatography
IP	Immunoprecipitation
miRNA	microRNA
NGS	Next Generation Sequencing
nt	Nucleotides
ORF	Open Reading Frame
PEG	Polyethylene Glycol
PDI	Protein Disulfide Isomerase
PNK	Polynucleotide Kinase
RPF	Ribosome Protected Fragment
rRNA	Ribosomal RNA
RT	Reverse Transcription
TF	Trigger Factor
uORF	upstream Open Reading Frame

List of Figures

1. Introduction

The research group of Prof Neil Bulleid at the Institute of Molecular Cell and Systems Biology, University of Glasgow has an interest in how proteins fold in mammalian cells. The main interest of the group lies in a family of proteins primarily found in the endoplasmic reticulum (ER) called protein disulfide isomerase (PDI) oxidoreductases; enzymes that catalyse the formation and degradation of disulfide bonds between cysteine amino acids in proteins, thus supporting protein folding. The exact reaction that is taking place between PDI and its substrate protein is yet to be determined, however, it is thought that the enzyme in its oxidised form binds to the unfolded protein creating a mixed disulfide, which prevents the protein from forming non-native disulfides and allows it to fold. After folding, the thiols that stabilise the correct conformation are in close proximity to each other and a disulfide exchange reaction takes place that reduces PDI and thereby releases it from its binding partner. The oxidoreductase ER oxidoreductin 1 (Ero1) oxidises PDI after this reaction by transferring electrons from the enzyme to molecular oxygen, thus producing hydrogen peroxide (Oka & Bulleid 2013).

Disulfide bonds are in many cases essential for the correct folding and function of a protein and if deficient they can lead to a number of diseases. Neurodegenerative misfolding diseases, such as Alzheimer's, Parkinson's, or Huntington's disease are examples where misfolding of the protein leads to its inability to be secreted, causing it to accumulate in the cell (Mossuto 2013). In order to eventually find a cure for these diseases, it is necessary to understand the mechanism of those enzymes that are involved in the process of protein folding, one of them being PDI.

In 2012, 21 members of the human PDI family were known, varying in structure, abundance, and function (Galligan & Petersen 2012). Not all members are capable of rearranging disulfide bonds, some lack catalytic cysteines and others only act as reductases. One aim is to

characterise the different members of the PDI family and to identify their role in protein folding, particularly with regard to co-translational events.

To date the exact function of each PDI is not known. Conceivably, different PDIs engage in the folding of different proteins; alternatively, different PDIs may bind to the nascent amino acid chain during different stages of the process or after translation is finished; a combination of both options is also possible. In a less complex yet not unlikely scenario, the members of the PDI family are partly or even completely redundant.

By crosslinking the nascent chain of a newly synthesised polypeptide to PDIs that are interacting with it at a given moment during translation, the proteins that are interacting with the respective PDI, can be specified. However, the exact determination of the individual protein that is crosslinked to a particular PDI is difficult, if not impossible by classical biochemical techniques, which is why a more powerful method is needed.

The research group of Prof Jonathan Weissman at the University of California developed such a tool, called Ribosome Profiling and published it for the first time in 2009 (Ingolia et al. 2009). The technique offers genome-wide information on proteins being synthesised *in vivo* at the time point of cell lysis by isolating mRNA fragments that are currently associated with a ribosome and thus belong to a mRNA sequence being translated into an amino acid sequence of a specific protein (Ingolia 2010; Michel & Baranov 2013).

In collaboration with Prof Weissman, the research group of Prof Bernd Bukau at the German Cancer Research Center, Heidelberg developed the method further into what was named Selective Ribosome Profiling (Oh et al. 2011; Becker et al. 2013). In this technique, co-translationally active proteins that are associated with the nascent polypeptide during protein synthesis, such as chaperons and targeting factors, are stabilised within their ribosome-nascent chain complexes via crosslinking, and the associated monosomes are

isolated according to the ribosome profiling protocol. With the help of this modified ribosome profiling approach, it is possible to detect interactions of co-translationally acting factors with nascent proteins and early folding and modification events can be monitored. The protocol by Becker et al. (2013) describes the technique for chaperons and targeting factors, however, it can be modified to allow the interaction of PDIs with nascent polypeptides to be studied.

However, before this protocol could be introduced into the laboratory, it was necessary to study the principles of the ordinary ribosome profiling method and to make sure that it flows smoothly.

This work is about the reverse engineering of a ribosome profiling workflow and the creation of our own protocol that allowed us to carry out the method in local facilities with the available materials and instruments. In addition, each step of the procedure was studied intensely with the aim of understanding the method thoroughly, providing a starting point for troubleshooting in case of problems arising when developing the protocol further.

2. Theoretical Background

Ribosome profiling is a relatively new method that is not yet commonly used in many laboratories. This may be due to the fact that the process itself is complex and contains numerous steps that can be error-prone and should, therefore, be studied carefully before carrying out the actual experimental work. This chapter describes the development of the strategy and the details of each step of the protocol.

2.1 Development of the Ribosome Profiling Strategy

When starting the Human Genome Project in 1990, the aim was not only to map and sequence the human genome for the very first time, but this ambitious proposition also laid the foundation for the development of new technologies, which allowed making DNA sequencing more affordable and accurate (Hubbard 2005). This development led to a shift in the general research focus towards the entity of genes being expressed and proteins being synthesised in a cell, rather than concentrating on individual molecules or pathways (Michel & Baranov 2013).

An easy and common way to access information about genes being expressed in a cell at a given time point, e.g. after treatment with a certain drug or after being exposed to stress, is to extract the total mRNA, which is equivalent to the transcriptome of the cell at this time point. Reverse transcription to complementary DNA (cDNA) and subsequent analysis by quantitative real-time PCR allows the selective investigation of gene expression in a cell (Nolan et al. 2006). A more general approach is the application of microarrays, where the obtained mRNA is converted into fluorescently labelled cDNA, which is then hybridised with matching DNA single strands immobilised on a microarray (Brown & Botstein 1999). Both techniques are valuable methods that are widely applied when it comes to the analysis of the transcriptome; however the quantity of a certain transcript does not necessarily correlate to the transcription level, as RNA transcripts are liable to reduced stability depending on their length

and composition (Nolan et al. 2006; Michel & Baranov 2013). Furthermore, neither of these techniques take into account that more than one protein can be encoded in one gene, depending on which open reading frame (ORF) is used for the translation (Ribrioux et al. 2008; Michel & Baranov 2013), nor can they mirror the abundance of proteins derived from the mRNA pool, as each mRNA transcript is usually translated by a varying number of ribosomes (Ingolia 2014).

A different approach is to examine the protein expression directly at the level of proteomics, for example by performing 2D electrophoresis, which separates the extracted proteins by size and charge, or by mass spectrometry, which allows identification of a protein according to its peptide fingerprint (Gupta et al. 2007). Both techniques, however, are relatively laborious and difficult to automate (Hinkson & Elias 2011); furthermore, it is difficult to draw conclusions about the gene expression levels based on the amount of proteins being present in the cell. For example, some proteins might occur in such low quantities that they can simply not be recognised during analysis; additionally, variable protein stability can lead to a false impression of the level of protein being expressed (Michel & Baranov 2013).

A further difficulty that arises from these techniques, whether they are based on transcriptome or proteome analysis, is that neither of them allows drawing conclusions about translational regulations and co-translational modifications happening during protein synthesis. Although, these questions may be even more relevant in certain areas other than studies on protein abundance or gene expression levels alone.

An alternative approach would be to determine the levels of transcripts present on actively translating polysomes; polysome analysis was repeatedly reported in literature during the late 1990s and early 2000s (Arava et al. 2003; Sagliocco et al. 1996; Zong et al. 1999). The technique is based on the size fractionating of polysomes by means of a sucrose density gradient centrifugation. Based on the number of ribosomes being attached to a mRNA

transcript, it can be concluded to what extend a transcript is being translated into an amino acid sequence; in combination with DNA microarrays, the technique has also been used to generate genome-wide polysome profiles (Arava et al. 2005). Polysome analysis, polysome profiling and ribosome density mapping are useful tools in the analysis of protein synthesis; however, they can be comparatively imprecise in relation to the effort they require (Ingolia 2010). Additionally, they do not offer the possibility to obtain information about the exact position of each ribosome on the mRNA transcript, which is of interest in some research areas, as ribosomes can be accumulated on certain sections of the transcript (Gerashchenko et al. 2012; Liu et al. 2013). Furthermore, they do not discriminate between ribosomes translating mRNA into an amino acid sequence and ribosomes localised on upstream open reading frames (uORFs) (Arava et al. 2005), which are located within the 5' untranslated region of mRNA transcripts and, therefore, do not take part in the translation of the encoded gene (Vilela & McCarthy 2003).

In 2009, Jonathan Weissman's the research group published a method, which for the first time built a bridge over the technological gap between methods of transcriptome and proteome analysis by developing the polysome analysis further and thus implementing a technique they called Ribosome Profiling (Ingolia et al. 2009).

Translation of a mRNA transcript is initialised by assembly of both ribosome subunits in a complex with the initiation site on the 5' end of the mRNA molecule (Lodish et al. 2008). From this moment on until termination, the ribosome encloses a region of the mRNA, which has a length of approximately 28 to 30 nucleotides (nt) and moves through the gap between the two ribosome subunits during the elongation process. Hence, the ribosome encloses a small part of the mRNA it is translating at every time point of the translation. Takanami et al. could show in 1965 that this region on the mRNA is protected from nuclease digestion by the ribosome enclosing it, which is taken advantage of in the method. By inhibiting the elongation

process, each ribosome is halted at its current position, whilst protecting the respective mRNA fragment. After cell lysis and isolation of polysomes, unprotected mRNA can be digested, leaving only the footprints of the ribosomes, the so called ribosome protected fragments (RPFs) (Ingolia et al. 2009). This approach was combined with a sequencing technique called deep sequencing, or next generation sequencing, which allows parallel sequencing of a multitude of small DNA fragments (Metzker 2010). By comparing the obtained sequences with a reference database via bioinformatics, it is possible to draw conclusions about the locations on mRNA transcripts of all ribosomes carrying out protein synthesis at the time point of cell lysis.

In addition to giving access to valuable information about protein synthesis levels, other less obvious fields of research can be addressed with the help of ribosome profiling. O'Connor et al. showed that interactions between messenger and ribosomal RNA in *Escherichia coli* (*E. coli*) leads to a "caterpillar-like" movement of the ribosome along the mRNA (O'Connor et al. 2013). To obtain these findings, the group used data from Li et al., who in turn investigated translational pausing in *E. coli* and *Bacillus subtilis* (Li et al. 2012). Thus, with the help of ribosome profiling, both studies came to the conclusion that so called Shine-Dalgarno sequences play a significant role in translation rates of the investigated prokaryotes. Other research groups have used the technique in order to investigate the effects that stress conditions have on the translation process (Gerashchenko et al. 2012; Liu et al. 2013).

An approach called selective ribosome profiling combines the technique with a selective purification, using for instance an antibody that specifically recognises, for example, enzymes that are involved in the folding process of nascent chains, which helps studying co-translational events. Oh et al. used this technique to investigate the substrates of the chaperone trigger factor in *E. coli* and found that the protein interacts with the nascent chain

after it had already been partially synthesised, rather than being bound to the ribosome (Oh et al. 2011). Han et al. developed the method further to folding-associated co-translational sequencing; this allowed them to investigate intermediates that are formed during folding of the fusion protein Flag-FRB-GFP, by using antibodies and protein binding partners to pull down different intermediates (Han et al. 2012). In 2013, a Nature protocol was published that gave accurate instructions for selective ribosome profiling to investigate the interactions taking place between the ribosome-nascent chain complexes and proteins potentially being involved in co-translational folding events (Becker et al. 2013) giving other research groups the opportunity to partake in the possibilities the technique offers. The latest work carried out by Prof Weissman's group concentrates on glycosylation of nascent polypeptides in the endoplasmic reticulum and has the aim of identifying the roles different oligosaccharyltransferases play in the process (Costa et al. 2014).

Only five years have passed since the ribosome profiling method was first published, but it has already widely spread as the previously mentioned examples illustrate. The Illumina company Epicentre® recognised this potential and developed the ARTSeq™ Ribosome Profiling Kit that contains some alterations compared to the original protocol (Epicentre 2013; Ingolia et al. 2012).

2.2 Workflow of the Ribosome Profiling Strategy

The complete workflow of the ribosome profiling strategy can be divided into two stages. Firstly, RPFs are isolated from the remaining cell debris and mRNA that is not covered by ribosomes at the moment of cell lysis, the second step consists of the generation of a cDNA library suitable for deep sequencing.

In Figure 1 an overview of the complete workflow as proposed by Ingolia et al. (2012) is given, starting with a living cell, in which translation is taking place. Ribosomes are associated with mRNA molecules, usually more than one at a time, forming so called polyribosomes or polysomes. Each ribosome encloses a mRNA section of approximately 28 to 30 nucleotides, which is being translated at that moment and is thusly protected from digestion by nucleases (Takanami et al. 1965). This characteristic is taken advantage of in the following procedure, when the RPFs are isolated from the remaining mRNA.

Before lysis, the cells are treated with the antibiotic cycloheximide, which inhibits protein synthesis by interfering with the translocation of the mRNA molecule from the ribosome, thus preventing translation elongation (Schneider-Poetsch et al. 2010; Godchaux et al. 1967), which ensures the ribosome remains in its original position. The lysis buffer also contains DNase, which serves to remove the genomic DNA that would otherwise represent perturbing impurities in the following steps. RNase I is an endoribonuclease and proved to produce constant footprint sizes (Ingolia 2010); unlike other ribonucleases, it cleaves any phosphodiester bond in RNA (Nicholson 1997), which allows the complete removal of mRNA that is not protected by ribosomes, a process referred to as nuclease footprinting. In this step, however, care needs to be taken not to incubate the sample with the nuclease for too long, as this may degrade the ribosomes, which mostly consist of RNA themselves. The reaction is stopped by adding an RNase inhibitor.

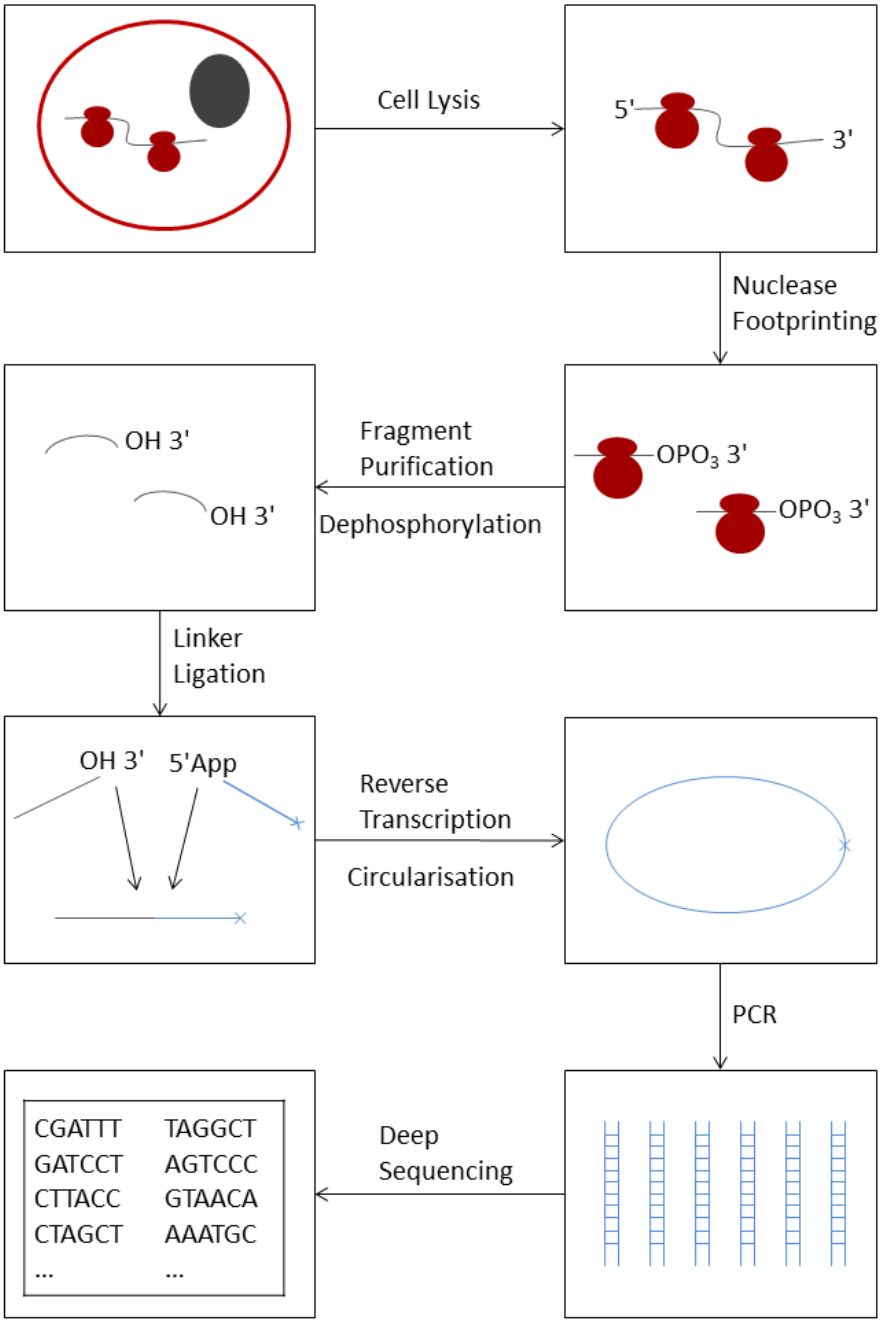

Figure 1 Overview of the ribosome profiling strategy. RNA fragments are shown in black, whereas DNA sequences are coloured blue. Based on (Ingolia et al. 2012)

Subsequently, RPFs are purified, which was initially carried out by a sucrose density gradient ultracentrifugation, a highly specialised technique that originally aims to separate polysomes of different sizes from each other. Accordingly, the method is relatively elaborate and requires a certain amount of skill and special apparatus (Ingolia 2010). However, the ribosome

profiling approach ideally produces only monosomes and therefore does not require such a complex procedure. It is indeed sufficient to carry out ultracentrifugation through a sucrose cushion rather than a gradient in order to successfully purify monosomes (Ingolia et al. 2011). The obtained ribosome-RNA complexes are denatured by the addition of detergents such as SDS and phenol and the RPFs are purified further in preparation for the subsequent steps.

Each ribosome-RPF complex contains only one molecule of mRNA, but several kilobases of ribosomal RNA (rRNA), which make up a substantial part of the obtained RNA during this process and should be removed in order to receive less ambiguous sequencing data. As the sequences of the most common human rRNAs are known, these molecules can be removed effectively by incubating the sample with a pool of single stranded DNA molecules that are complementary to the rRNA sequences. The Ribo-Zero rRNA Removal Kit (Epicentre) can be used for this purpose. It consists of the mentioned DNA molecules as well as magnetic beads. The precise procedure is not described any further in the company's protocol, however, it is probable that the DNA molecules are tagged in a way that allows them to bind to the magnetic beads, thus enabling their removal.

In the second part of the protocol, RPFs are converted into a cDNA library, which can be deep sequenced. The RPFs are ligated with a linker and then transcribed into cDNA, which after circularisation can be used to set up a deep sequencing library; Figure 2 describes these steps in more detail.

In order to be suitable for the ligation to a cloning linker, the RPF needs a 3' hydroxyl terminus, however, the previous digestion with the ribonuclease leaves the molecule with either a 3' phosphoester or a 2', 3' phosphodiester terminus (DelCardayré & Raines 1995), which means that the molecule needs to be dephosphorylated. Bacteriophage T4 Polynucleotide Kinase (PNK) catalyses the phosphorylation of a 5' hydroxyl terminus of

deoxyribonucleic acids as well as ribonucleic acids in the presence of a nucleoside triphosphate (Richardson 1981). In 1977, Cameron and Uhlenbeck described the enzyme's additional ability to dephosphorylate the 3' termini of nucleic acids in the absence of ATP. This ability is utilised in this step of the protocol, where each RPF is dephosphorylated on its 3' end prior to ligation to a universal microRNA (miRNA) cloning linker. Other phosphatases, such as alkaline phosphatase could be used, however, PNK was shown to catalyse the reaction more efficiently (Ingolia 2010).

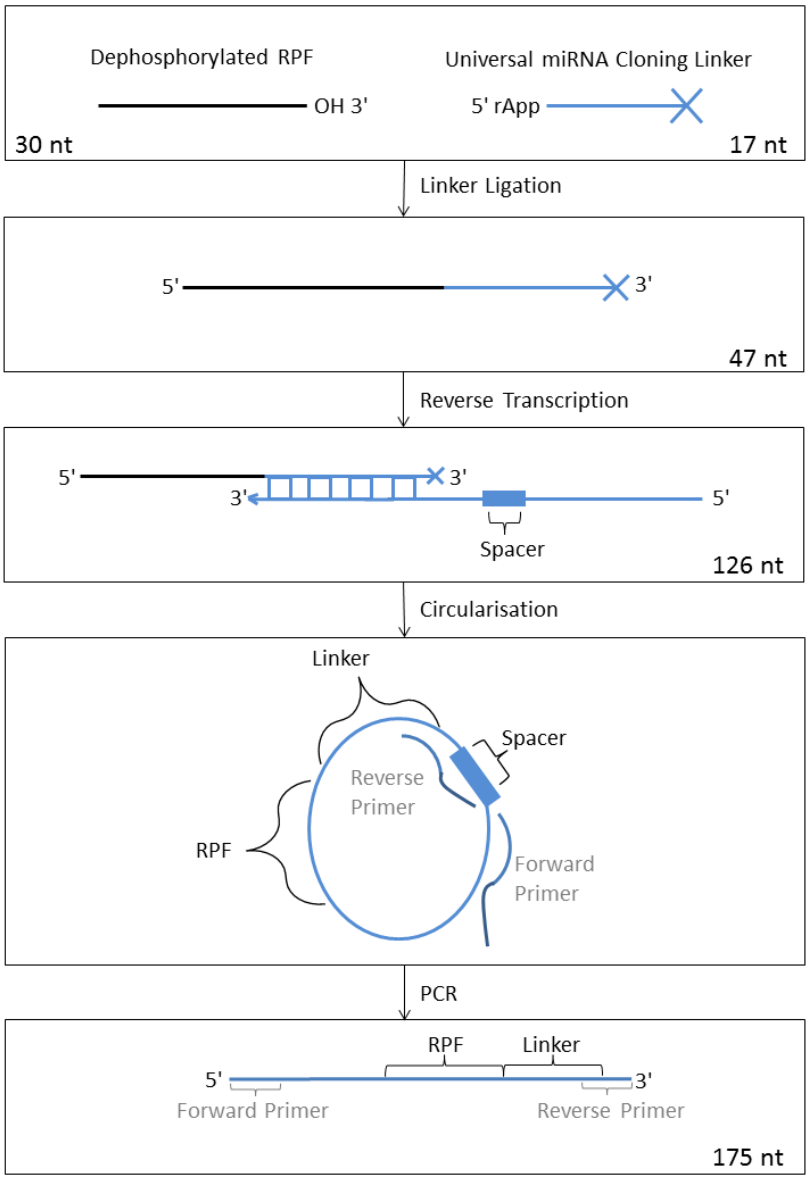

Figure 2 Overview of the preparation of RPFs for sequencing. RNA fragments are shown in black, whereas DNA sequences are coloured blue. Based on (Ingolia et al. 2012)

In previous versions of the protocol, the next step was to add a tail of 25 – 30 adenines to the 3' end of the RPF, which allowed base pairing to an oligonucleotide containing a poly-d(T) sequence on its 3' terminus and that is otherwise composed as the reverse transcription primer described below. The oligo-d(T) also served as a reverse transcription priming site. Unfortunately, the process of polyadenylation does not lead to a uniformly sized poly-d(A) sequence on each RPF, but results in a size distribution of poly-d(A) tails. The RPFs obtained by footprinting are of slightly different lengths and may end in one or more adenines, which cannot be distinguished from adenines added during the polyadenylation process. For this reason, ambiguities can occur when mapping a RPF sequence to its origin in the genome, especially because deep sequencing is more error-prone at the end of a fragment (Ingolia et al. 2009; Ingolia 2010). In a modified process published in 2012, RPFs were, therefore, ligated to a cloning linker with a sequence complementary to a reverse transcription primer, in order to avoid these problems (Ingolia et al. 2012).

The linker is blocked on its 3' end, which prevents circularisation and contains a polyadenylate on its 5' terminus; this polyadenylation is crucial for the subsequent ligation catalysed by the truncated T4 RNA ligase 2. This modified version of the enzyme lacks its N-terminus and, therefore, consists of only 249 instead of 334 amino acid residues. Furthermore, it contains an adenyltransferase/RNA ligase domain with a bound AMP molecule, which is why the enzyme, unlike the full-length protein, can catalyse the ligation without additional ATP added to the reaction, as long as the 5' end of the nucleotide is polyadenylated (Ho et al. 2004). Viollet et al. (2011) carried out experiments investigating different reaction conditions on the enzyme's efficiency. According to their results, it can be improved by increasing the reaction time to up to 24 hours. The reaction can be carried out at room temperature; however, decreasing the temperature to 16 °C can be beneficial, as the reduced Brownian motion gives the ligation products more time to form. According to the

manufacturer of the ligase (New England Biolabs) even ligations on ice for 24 hours have been reported. The concept of molecular crowding is another principle that can influence the formation of the ligation product positively by – amongst other effects, that are still not fully understood – increasing the sample viscosity, which has the same effect on Brownian motion (Miyoshi & Sugimoto 2008). To simulate molecular crowding, the inert polymer polyethylene glycol (PEG) can be added to the reaction at a concentration of up to 25 %. Higher percentages do not further influence the reaction and in addition the handling of the sample becomes more difficult due to increased viscosity (Viollet et al. 2011).

Subsequently, the ligation product is transcribed into cDNA. For this reaction a primer is used that pairs with the cloning linker; additionally, it contains a hexaethylene glycol spacer and two sequences that act as forward and reverse priming sequences in the following PCR amplification. By circularising the obtained single stranded DNA molecule, using the enzyme CircLigase, which catalyses the intramolecular ligation in the presence of ATP, the RPF is now flanked by a forward and a reverse priming sequence, which allows amplification of the sequence by PCR. During this reaction the spacer still remains in the molecule, preventing the DNA polymerase from rolling circle amplification and also ensuring torsional flexibility, which enables the conversion from flaccid single-stranded DNA into a rigid double strand. The resulting product is a library of linear DNA, which corresponds to the initially purified RPFs. All oligonucleotides that are involved in the generation of a cDNA library as well as the approximately 175 nt long PCR product that is obtained at the end of the process are summarised in Figure 3, showing complementary sequences in matching colours.

Linker

5'- CTGTAGGCAC CATCAAT - 3'

Reverse Transcription Primer

5'- AGATCGGAAG AGCGTCGTGT AGGGAAAGAG TGTAGATCTC TGGTGGTCGC-(SpC18)-CACTCA-(SpC18)-TTCA GACGTGTGCT CTTCCGATCT ATTGATGGTG CCTACAG - 3'

Forward PCR Primer

5'- AATGATACGG CGACCACCAG ATCTACAC - 3'

Reverse PCR Primer

5'- CAAGCAGAAG ACGGCATACG AGATNNNNNN GTGACTGGAG TTCAGACGTG TGCTCTTCCG ATCT - 3'

PCR product

5'- AATGATACGG CGACCACCAG ATCTACACTC TTTCCCTACA CGACGCTCTT CCGATCT - RPF - CTG TAGGCACCAT CAATAGATCG GAAGAGCACA CGTCTGAACT CCAGTCACNN NNNNATCTCG TATGCCGTCT TCTGCTTG - 3'

Figure 3 Sequences of all oligonucleotides that are involved in the creation of a cDNA library suitable for NGS. Complimentary sequences are shown in different colours: The linker sequence is green, complimentary parts of the forward PCR primer are shown in blue, and the complimentary sequences for the reverse PCR primer are shown in purple.

2.3 Deep Sequencing of RPFs

Deep sequencing or next generation sequencing (NGS) is a technique that allows parallel sequencing of different DNA fragments by solid-phase amplification, and is used to determine the sequences of the multitude of RPFs obtained in each experiment. For NGS, the DNA molecules to be sequenced have to be ligated to an adaptor with known sequence, which is complementary to sequences immobilised on a sequencing platform (Grada & Weinbrecht 2013). This step however can be omitted in this case, since the adapter sequence is identical to the used reverse primer.

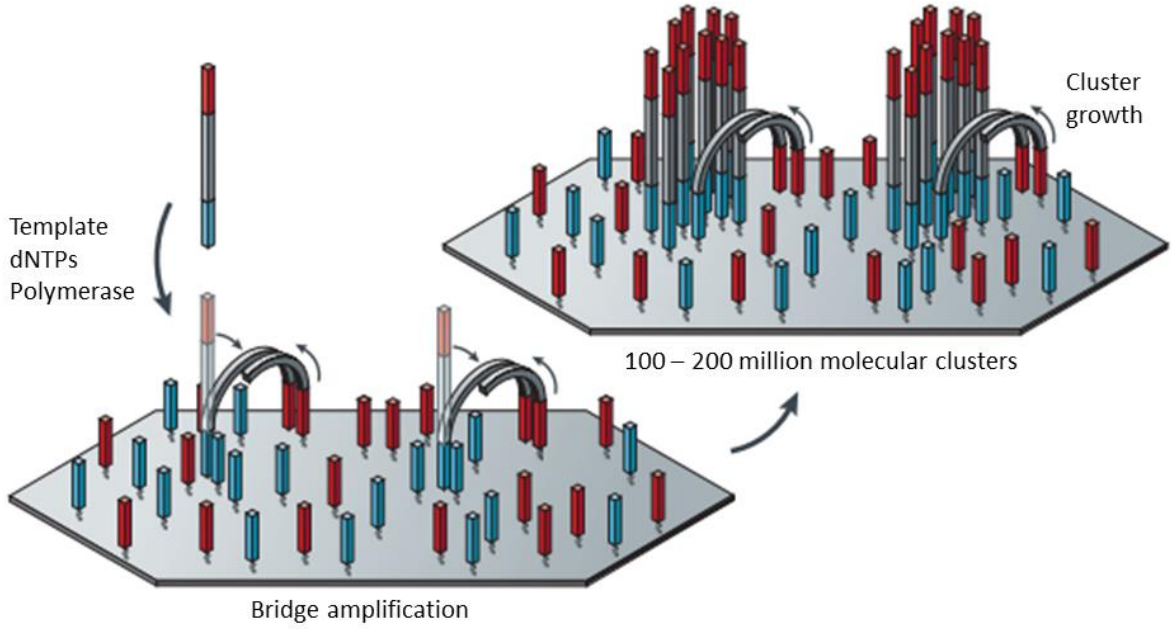

Figure 4 Schematic of bridge amplification as performed by the Illumina sequencing platform. Templates are immobilised with their 3' and 5' terminus on the sequencing platform and form clusters through amplification. (Metzker 2010), modified

There are various approaches of NGS, using different strategies for template preparation, sequencing, imaging and data analysis. The Illumina platform works on the basis of solid phase amplification, as shown in Figure 4.

Primers that base pair with the known forward and reverse primers of the template are immobilised on a solid phase. During the initial priming, the templates bind with their 3' as well as with their 5' end to the primers and build bridges, which are then amplified, producing approximately 100 - 200 million molecular clusters. The clusters are then amplified again, this time by solid-phase amplification, using the so called four-colour cyclic reversible termination method, in which one nucleotide at a time is added to the sequence. All four nucleotides are labelled with a fluorescent dye, allowing imaging of each amplification step; for the next step, the dye is cleaved from the growing sequence, enabling the next nucleotide to be added (Grada & Weinbrecht 2013; Metzker 2010). Due to the nature of NGS, it generates an enormous amount of data and its storage and analysis requires expensive

computers and algorithms. The procedure itself can be summarised in a very simplified manner by the two steps processing and mapping (Gogol-Döring & Chen 2012). Especially the Illumina sequencing platform is prone to producing artefacts such as mixed clusters, phasing (a base has failed to be incorporated during one cycle), and prephasing (more than one base has been incorporated per cycle), which have to be excluded from the data analysis (Ledergerber & Dessimoz 2011). Furthermore, it can be necessary to remove adaptor sequences that are included in the read if the insert was too short (Gogol-Döring & Chen 2012). An additional problem specifically concerning the deep sequencing of ribosome profiling samples is the remaining contamination with other RNAs, such as rRNA that escaped the rRNA depletion procedure during cDNA library preparation. Ingolia et al. (2012) recommend the application of a bioinformatics filter for this case. Other contaminating RNAs can be tRNAs and small nuclear RNAs, which can usually be distinguished from RPFs by their shorter lengths (Ingolia et al. 2012).

The second step of data analysis is referred to as read mapping, in which the obtained sequences are aligned to a reference sequence, in this case the total transcriptome. Where sequences overlap within a certain tolerance due to sequencing errors, the position of the RPF in the transcriptome and thus the polypeptide that was synthesised at the moment of cell lysis can be determined (Gogol-Döring & Chen 2012).

2.4 Outline of the Project Aims

After a detailed literature review that was presented in chapters 2.1 to 2.3, it was decided to validate the single steps of the procedure, whilst mostly referring to the original ribosome profiling protocol by Ingolia et al. (2012) and the protocol provided by the ARTSeq™ Ribosome Profiling Kit (Epicentre 2013). Key elements of the workflow are the linker ligation, the reverse transcription and the PCR amplification, which should be focused on. Each step of the workflow was to be validated carefully, including the preparation of

ribosome protected fragments from real cells as well as the generation of the cDNA library, which was initially carried out with control oligonucleotides. After validation of each component of the protocol, the gained knowledge was to be combined and the complete ribosome profiling workflow was carried out using real samples. Figure 5 summarises the planned experiments of the project. Due to a lack of time and financial reasons, deep sequencing was not carried out during this work.

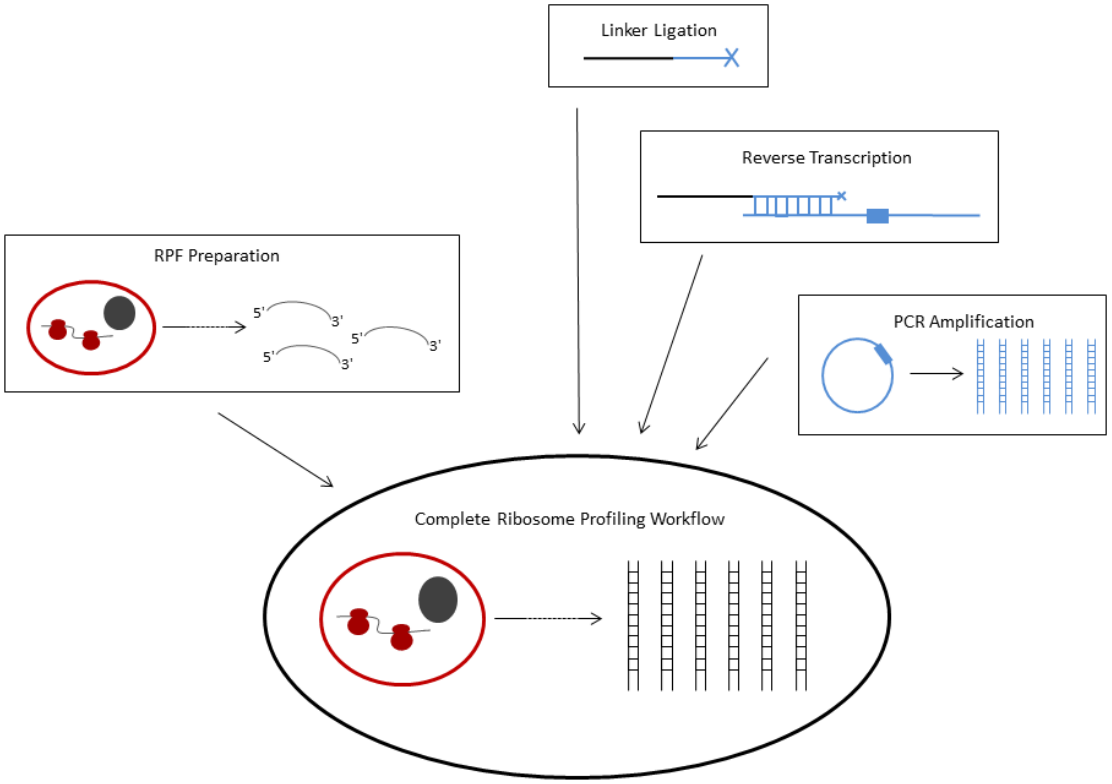

Figure 5 Outline of the project aims. Preparation of ribosome protected fragments and generation of a cDNA library were to be carried out separately with adequate controls, before the complete ribosome profiling workflow was conducted using real samples.

3. General Materials and Methods

Temperature specific incubations as well as PCR reactions were carried out in either the 2720 Thermal Cycler (Applied Biosystems), the TC-312, or the ^{5}Prime (both Techne). Electrophoresis was performed using 18 % TBE-Urea gels for RNA separation and 10 % TBE gels for the separation of PCR products. 18 % TBE-Urea gels consisted of TBE buffer (pH 8.2) supplemented with 0.44 g ml^{-1} Urea (Fisher Scientific, cat# U/0500/53), 0.45 ml ml^{-1} Acrylamide:Bisacrylamide (Sigma Aldrich, cat# A2792), 0.2 µg ml^{-1} Ammoniumpersulfate (Sigma Aldrich, cat# 31117) and 1.1 µl ml^{-1} TEMED (Sigma Aldrich, cat# T9281). 10 % TBE gels consisted of TBE buffer (pH 8.2) supplemented with 0.3 ml ml^{-1} Acrylamide:Bisacrylamide (Sigma Aldrich, cat# A2792), 1.7 µg ml^{-1} Ammoniumpersulfate (Sigma Aldrich, cat# 31117) and 0.8 µl ml^{-1} TEMED (Sigma Aldrich, cat# T9281). 5x TBE buffer consisted of 10.8 g l^{-1} Tris (Fisher Scientific, cat# BP152-1), 5.5 g l^{-1} orthoboric acid (BDH, cat# 100583R), and 0.6 g l^{-1} EDTA (Fisher Scientific, cat# D/0700/53) and was made up with demineralised water fresh from the filter module. During the optimisation, MicroSpin S400 columns (GE Healthcare, cat# 27-5140-01) for nuclease footprinting and the RNA Clean & Concentrator kit (The Epigenetics Company, cat# R1015) for RNA purification were also used, but were considered unfavourable for the process.

3.1 Cell Cultures

Two different cell lines were routinely cultured for this work. HT 1080 is an adherent cell line that was derived in the 1970s from a fibrosarcoma of a 35 year old Caucasian male without having received any form of radio- or chemotherapy (Rasheed et al. 1974). The cell line WT-PDI is the same cell line stably overexpressing PDI.

Both, HT 1080 (CCL-121, ATCC, USA) and WT-PDI (generated in local facilities) were routinely cultured in tissue culture flasks using Dulbecco's Modified Eagle Medium (Life Technologies, cat# 21969-035) supplemented with 10 % (v/v) foetal bovine serum (FBS) (Life Technologies, cat# 10500-064), 2 mM L-glutamine (Sigma Aldrich, cat# G7513), 50 U ml^{-1} penicillin, and 50 µg ml^{-1} streptomycin (both Life Technology, cat# 15070-063). Growth medium of WT-PDI was furthermore supplemented with 500 µg ml^{-1} of the antibiotic G-418 Sulfate (Promega, cat# V7983) to maintain selection pressure. The cells were constantly kept in a humidified incubator at 37 °C and 5 % CO_2 concentration and were passaged at 70-80 % confluence, using trypsin-EDTA (Life Technologies, cat# 25300-054) to detach the cells from the tissue culture flask and PBS (Life Technologies, cat# 14190) for washing the cells. Cells were tested for contamination with mycoplasma in regular intervals (Lonza, cat# LT07-705) and were found to be free of the contaminant.

3.2 Ribosome Profiling

Ribosome profiling was carried out based on a protocol by Ingolia et al. (2012); however, where modified procedures yielded improved results, these were adapted instead.

In order to protect the RNA fragments obtained during the process, all solutions and buffers were made up in nuclease-free water (Quiagen, cat# 129115) if not stated otherwise. Workbenches as well as pipettes and racks were treated with RNase Zap (Ambion, cat# AM9780) prior to use. Furthermore, non-stick RNase-free microfuge tubes (Ambion, cat# AM12450 1.5 ml and AM12350 0.5 ml) and RNase-free filter pipette tips (Mettler Toledo, cat# 17002414 10 µl, 17002420 200 µl and 17002410 1000 µl) were used for handling samples containing RNA.

3.2.1 Cell Lysis

After cells had reached approximately 80 % confluence, they were incubated for 2 min with 100 µg ml^{-1} cycloheximide (Sigma, cat# C7698) in fresh medium at 37 °C in order to halt

mRNA translation. Medium was then aspired, the cells were washed with PBS (Life Technologies, ca# 14190) and subsequently lysed using a polysome buffer made up of a 20 mM Tris Cl buffer (pH 7.4) (Fisher Scientific, cat# BP152-1) containing 150 mM NaCl (VWR, cat# 27810), 5 mM $MgCl_2$ (Fisher Scientific, cat# M/0600/53), 1 mM DTT (Melford, cat# MB1015), and 100 µg ml^{-1} cycloheximide (Sigma, cat# C7698). 1 % Triton X-100 (Sigma, cat# T8532) and 25 U ml^{-1} DNase I (Ambion, cat# AM2238) were added to obtain a lysis buffer. Cells were removed from the dish with the help of a cell scraper and after incubation on ice for 10 min they were mechanically disrupted by syringing. The lysate was cleared from cell debris by centrifugation at 20,000 x *g* for 10 min.

3.2.2 Ribosome Footprinting and rRNA Depletion

In the next step, mRNA, which was not protected by ribosomes, was digested by adding 2.4 U $µl^{-1}$ RNase I (Ambion, cat# AM2295) and the reaction was stalled after 45 min by adding 0.6 U $µl^{-1}$ RNasin (Promega, cat# N2111). Ribosome footprint complexes were separated by centrifuging the lysate in the presence of a 1 M sucrose cushion (Sigma, cat# S0389) made up in polysome buffer (see above) in an ultracentrifuge TLA100.3 rotor (Beckman Coulter, cat# 349490) at 70,000 rpm and 4 °C for 4 h. The recovered pellet contained the ribosomes and was resuspended in 600 µl TriZol (part of the miRNeasy kit, see below). The obtained RNA was purified using the miRNeasy kit (Quiagen, cat# 217004) following the manufacturer's protocol, and thereafter precipitated by adding 38.5 µl water, 1.5 µl glycoblue (Ambion, cat# AM9516), 10 µl of 3 M sodium acetate (BDH, cat# 102364Q), as well as 150 µl isopropanol (Sigma Aldrich, cat# 24137). Precipitation was carried out for 30 min on dry ice and the RNA was pelleted by centrifugation at 20,000 x *g* and 4 °C for 30 min. The pellet was resuspended in 10 mM Tris Cl buffer (pH 8) (Fisher Scientific, cat# BP152) and rRNA depletion was conducted using the Ribo-Zero Magnetic Kit (Epicentre, cat# MRZH116) according to the manufacturer's protocol. The rRNA depleted

samples were precipitated by adding 1 µl glycoblue and 10 µl of 3 M sodium acetate, as well as 150 µl isopropanol. Precipitation was carried out as previously described.

The RNA pellet was resuspended in 5 µl of a 10 mM Tris Cl buffer (pH 8) and diluted 1:1 with a denaturing loading buffer consisting of 98 % (vol/vol) formamide (Sigma Aldrich, cat# 47671), 10 mM EDTA (Fisher Scientific, cat# D/0700/53), and 300 µg ml^{-1} bromophenol blue (Sigma Aldrich, cat# 114391). Samples were denatured at 105 °C for 90 s and footprint fragments were subsequently purified by gel electrophoresis on an 18 % polyacrylamide TBE-Urea gel (see above) in TBE running buffer. Together with the samples, a 20/100 DNA size marker (IDT, cat# 51-05-15-02) as well as a 30 nt and a 28 nt control oligonucleotide (both synthesised by Sigma Aldrich, sequences see chapter 3.2.8) were loaded on the gel. Electrophoresis was performed for approximately 50 min at 200 V, until the running front had visibly reached the end of the gel. It was thereafter stained with SYBR Gold nucleic acid gel stain (Life Technologies, cat# S11494) in 1x TBE running buffer for 3 min. With the help of the control oligonucleotide, the region that was expected to contain RPFs was localised under UV light and subsequently excised. RNA was extracted from the gel slice by disrupting the gel slice mechanically, adding RNA gel extraction buffer composed of a 300 mM sodium acetate buffer (pH 5.5) (BDH, cat# 102364Q) containing 1 mM EDTA (Fisher Scientific, cat# D/0700/53), as well as 0.25 % (wt/vol) SDS (VWR, cat# 442444H), freezing on dry ice for 30 min and incubating the samples overnight at room temperature with gentle mixing. Extracted RNA was precipitated by adding 1.5 µl glycoblue and 500 µl isopropanol, following the procedure described above.

3.2.3 Dephosphorylation

The dephosphorylation reaction was set up by adding T4 PNK buffer (New England Biolabs, cat# M0236S) to a final 1x concentration, 0.4 U $µl^{-1}$ RNasin, as well as 0.2 U $µl^{-1}$ T4 PNK (New England Biolabs, cat# M0201S). The reaction was carried out at 37 °C for 1 h followed

by heat inactivation of the enzyme at 70 °C for 10 min. After the reaction, RNA was precipitated by adding 39 µl water, 1 µl glycoblue, and 10 µl of 3 M sodium acetate, as well as 150 µl isopropanol following the procedure described above, with the difference that the obtained RNA pellet was resuspended in only 8.5 µl of a 10 mM Tris Cl buffer (pH 8).

3.2.4 Linker Ligation

In addition to RPFs, linker ligation was carried out with an *in vitro* transcript (see chapter 3.3) and with a 5'-FAM labelled synthetic oligonucleotide (synthesised by Sigma Aldrich, sequence see chapter 3.2.8).

Linker ligation was carried out by adding 0.5 nmol Universal miRNA Cloning Linker (New England Biolabs, cat# S1315S) to the dephosphorylated RNA; the reaction mix was completed by adding T4 Rnl2 Buffer (New England Biolabs, cat# B0216L) to a final 1x concentration, 15 % (wt/vol) PEG 8000 (New England Biolabs, cat# B1004A), 1 U μl^{-1} RNasin, and 10 U μl^{-1} T4 Rnl2(tr) (New England Biolabs, cat# M0242S). The reaction took place at room temperature for 2.5 h. During optimisation the reaction was also carried out overnight at 16 °C. RNA was precipitated by adding 338 µl water, 1.5 µl glycoblue, and 40 µl of 3 M sodium acetate, as well as 500 µl isopropanol. Precipitation was carried out as previously described and the RNA pellet was resuspended in 5 µl of a 10 mM Tris Cl buffer (pH 8). The ligation product was separated from unreacted RPFs and linkers by gel electrophoresis, excised and extracted from the gel following the procedure defined above.

3.2.5 Reverse Transcription

The reverse transcription reaction was set up by adding 2.5 pmol of the reverse transcription primer (synthesised by Sigma Aldrich, sequence see chapter 3.2.8), first-strand buffer (Invitrogen, cat# P/N Y02321) to a final 1x concentration, 0.5 mM dNTPs (New England Biolabs, cat# N0440 (dATP), N0441 (dCTP), N0442 (dGTP), N0443 (dTTP)), 5 mM DTT (Invitrogen, cat# P/N Y00147), 1 U μl^{-1} RNasin, as well as 10 U μl^{-1} SuperScript III

(Invitrogen, cat# 18080-093) to the ligation product. The reaction mix was incubated at 48 °C for 30 min and the RNA was subsequently hydrolysed by the addition of 100 mM NaOH (Fisher Scientific, cat# S/4920/53) and incubation at 98 °C for 20 min. Obtained cDNA was precipitated by adding 156 µl water, 2 µl glycoblue, and 20 µl of 3 M sodium acetate, as well as 300 µl isopropanol. Precipitation was carried out as previously described and the RNA pellet was resuspended in 5 µl of a 10 mM Tris Cl buffer (pH 8). The reverse transcribed cDNA was separated from unreacted reverse transcription primers by gel electrophoresis on an 18 % TBE-Urea gel, excised and extracted from the gel following the procedure defined above, with the difference that DNA gel extraction buffer was used. DNA gel extraction buffer was made up of a 10 mM Tris Cl buffer (pH 8), containing 300 mM NaCl (VWR, cat# 27810) and 1 mM EDTA (Fisher Scientific, cat# D/0700/53).

3.2.6 Circularisation

Circularisation was carried out by resuspending the obtained single stranded DNA in 15 µl of a 10 mM Tris Cl buffer (pH 8) and adding CircLigase buffer to a final 1x concentration, 50 mM ATP, 2.5 mM $MnCl_2$, and 10 U μl^{-1} CircLigase. Components were all part of the CircLigase Kit (Epicentre, cat# CL4111K). The reaction mix was incubated at 60 °C for 1 h and the enzyme was subsequently heat-inactivated at 80 °C for 10 min. Circularised DNA was precipitated by adding 14 µl water, 2 µl glycoblue, and 60 µl of 0.5 M NaCl (VWR, cat# 27810), as well as 300 µl isopropanol. Precipitation was carried out as previously described and the DNA pellet was resuspended in 7 µl of a 10 mM Tris Cl buffer (pH 8).

3.2.7 PCR

PCR was carried out using the Phusion® High-Fidelity PCR Master Mix (New England Biolabs, cat# M0531S) at a final 1x concentration together with 0.5 µM of each forward and reverse primer (synthesised by Sigma Aldrich, sequences see chapter 3.2.8). The protocol consisted of 1 stage at 98 °C for 30 s, 6 - 14 cycles of 98 °C for 10 s (denaturation), 65 °C for

5 s (primer annealing), and 72 °C for 5 s (elongation). Amplification was conducted for 5 samples à 1 µl simultaneously for 6, 8, 10, 12, and 14 cycles, leaving the negative controls (minus enzyme and minus template) in the thermal cycler for 14 cycles. PCR products were merged 1:6 with a non-denaturing sample buffer consisting of a 10 mM Tris buffer (pH 8) containing 1 mM EDTA (Fisher Scientific, cat# D/0700/53), 15 % Ficoll 400 (Sigma Aldrich, cat# F2637), and 0.25 % bromophenol blue (Sigma Aldrich, cat# 114391) and were loaded on a 10 % TBE gel (see above) without previous denaturation at 105 °C. In addition to the PCR amplified samples and the negative control samples, a 100 bp size marker (New England Biolabs, cat# N3231S) was loaded on the gel to enable size comparison of the obtained bands. Electrophoresis was performed for approximately 50 min at 200 V in TBE running buffer, until the running front had visibly reached the end of the gel. The obtained gel was stained with SYBR Gold nucleic acid gel stain (Life Technologies, cat# S11494) in TBE running buffer for 3 min. The gel was visualised on a UV-light screen and the band with a clear amplification product but little unspecific amplification was selected. The respective band was excised and DNA was extracted as specified above.

3.2.8 Sequences

All oligonucleotides were synthesised by Sigma Aldrich, apart from the miRNA Cloning Linker (New England Biolabs, cat# S1315S) and the *in vitro* transcript (see chapter 3.3). The sequences below are depicted in 5' → 3' direction

Control oligonucleotide (28 nt):

NNGUACACGG AGUCGACCCG CAACGCNN

Control oligonucleotide (30 nt):

NNGUACACGG AGUCAAGACC CGCAACGCNN

5'-FAM labelled control oligonucleotide (31 nt):

AGUCGUAGCC UUUAUCCGAG AUUCAGCAAU A

PPL-pGEM 4 cut with Nco I *in vitro* transcribed (83 nt):

AUUAUGCUGA GUGAUAUCCC UCUGGCCUUC GAACGAACAA GAAAAACGUC UUCGAGUCUU AUUUGCGAGU UAAACCGUCU AUG

miRNA cloning linker (17 nt):

CTGTAGGCAC CATCAAT

Reverse transcription primer (97 nt):

AGATCGGAAG AGCGTCGTGT AGGGAAAGAG TGTAGATCTC TGGTGGTCGC-(SpC18)-CACTCA-(SpC18)-TTCA GACGTGTGCT CTTCCGATCT ATTGATGGTG CCTACAG

Forward PCR primer (28 nt):

AATGATACGG CGACCACCAG ATCTACAC

Reverse PCR primer (64 nt):

CAAGCAGAAG ACGGCATACG AGATNNNNNN GTGACTGGAG TTCAGACGTG TGCTCTTCCG ATCT

NNNNNN stands for an index that is used for deep sequencing. 12 different indexes were available as recommended by Epicentre (2013), whereas in this work, index 6 was used with the sequence 5' ATTGGC 3'.

3.3 *In vitro* Transcription

For *in vitro* transcription, 10 µg of the plasmid PPL-pGEM 4 (from laboratory stock) was digested with 2 U μl^{-1} of the restriction enzyme Nco I (New England Biolabs, cat# R3193S) in 1x Buffer H (Roche, cat# 11064 900) for 2 h at 37 °C. The digested plasmid was precipitated with 2.5 volumes of ethanol (VWR, cat# 20821.330) and 0.5 volumes of 3 M sodium acetate (BDH, cat# 102364Q).

In vitro transcription was carried out in 1x transcription buffer (see below) with 40 nmol μl^{-1} DTT (Melford, cat# MB1015), 0.8 U μl^{-1} RNasin (Promega, cat# N2111), 2.4 µM ribonucleotide triphosphate mix (Roche, cat# 11277057001), and 0.72 U μl^{-1} T7 RNA polymerase (Promega, cat# P207E) for 2 h at 37 °C. 5x transcription buffer consisted of a 400 mM HEPES KOH buffer (pH 7.5) (VWR, cat# 441485H) containing 60 mM $MgCl_2$ (Fisher Scientific, cat# M/0600/53) and 10 mM spermidine (Sigma Aldrich, cat# S2626). The *in vitro* transcript was purified by phenol/chloroform extraction using phenol:chloroform:isoamyl alcohol (Sigma Aldrich, cat# P3803) and chloroform (Fisher Scientific, cat# C/4960/17). The obtained RNA was precipitated with 2.5 volumes of ethanol and 0.5 volumes of 3 M sodium acetate and was stored in H_2O at -80 °C prior to further use.

4. Experimental Results and Discussion

Before any experiments could be carried out with regards to selective ribosome profiling, it was necessary to establish the ribosome profiling technique itself and to carefully examine each of the individual steps in terms of efficiency and viability. Two different protocols were available, of which each used individual methods and reagents. The protocol by Ingolia et al. (2012) explains each individual step in great detail, and provides exact information about reagents and materials that were used during the development of the method. The protocol included in the ARTseq^{TM} Ribosome Profiling Kit (Epicentre 2013) suggested the use of different methods or commercially available kits in some instances, which differed from the original protocol (Ingolia et al. 2012). As most enzymes were components of the kit, no or only very restricted information was given about their origin, mode of action or concentration and composition of the buffers they were supplied with. Therefore, most procedures were carried out according to Ingolia et al. (2012), but kits suggested by Epicentre (2013) that promised to be more time-saving and potentially more efficient were also tested in the progression of this project.

4.1 General Procedures

Throughout the workflow, there were two procedures that had to be carried out repeatedly: purification and concentration of RNA and extraction of reaction products after gel electrophoresis.

4.1.1 Nucleic Acid purification

After each step of the protocol the obtained product needed to be purified in order to remove enzymes and salts from the previous reaction, furthermore it was necessary to provide the oligonucleotides in a small volume of buffer so that reagents for the next reaction could be added or so that the product could be purified by gel electrophoresis. Both protocols (Ingolia et al. 2012; Epicentre 2013) recommend different approaches; while Ingolia et al. suggest a

classical precipitation with sodium acetate and isopropanol, Epicentre recommends the use of the Zymo Clean & Concentrator kit, which promises to purify and concentrate even low concentrations of small oligonucleotides in only a few minutes. This alternative approach would be very beneficial compared to the conventional method, which takes up a minimum of 30 min freezing and a 30 min centrifugation. Both procedures were compared to each other in terms of recovery efficiency using 0.1 µmol of a 30 nucleotide control oligonucleotide. Figure 6 shows the recovered quantities as determined spectrophotometrically using a NanoDrop instrument, relative to the RNA concentration in the starting material. The graph shows the superiority of the precipitation method compared to the kit, with an average of more than 80 % recovery, whereas the kit only recovers less than a quarter of the starting material.

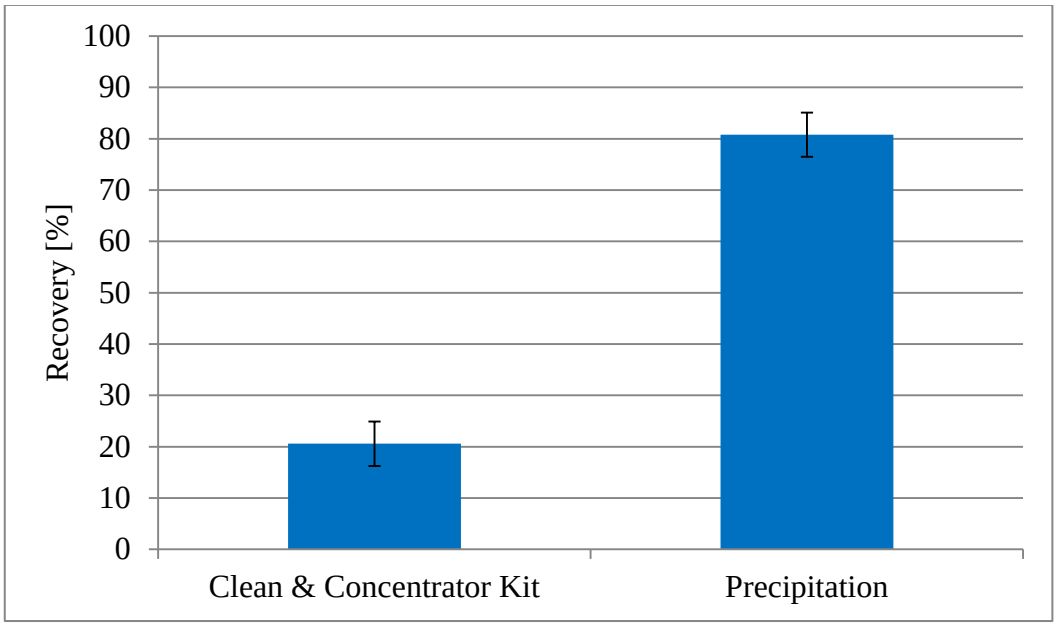

Figure 6 100 nmol of the control oligonucleotide were purified by means of an isopropanol precipitation and the Zymo Clean & Concentrator kit. Both purifications were carried out in duplicates. The concentrations of the purified products as well as the concentration of the starting material were measured spectrophotometrically. The recovery was calculated in relation to the measured concentration of the control oligonucleotide.

Results obtained by NanoDrop concentration measurements can be deceiving, especially when only low concentrations are measured, as the method does not distinguish between absorbing nucleotides and impurities.

Other impurities derived from extraction procedures, e.g. phenol or glycogen, can lead to a shift of the spectra (Thermo Scientific n.d.). During the process of this work it was found that especially low RNA concentrations in high buffer volumes lead to unrealistic results with apparent RNA amounts much higher than initially purified. Because of this unreliability the recovery of both purification methods was also investigated visually by the means of a gel electrophoresis, which is illustrated in Figure 7. The gel shows 10 nmol of the control oligonucleotide next to a tenth of each purification product, which corresponded to 10 nmol if 100 % recovery could be achieved. It confirms the results of the NanoDrop measurement, as both precipitation products show up as very bright bands that are nearly as intense as the control band.

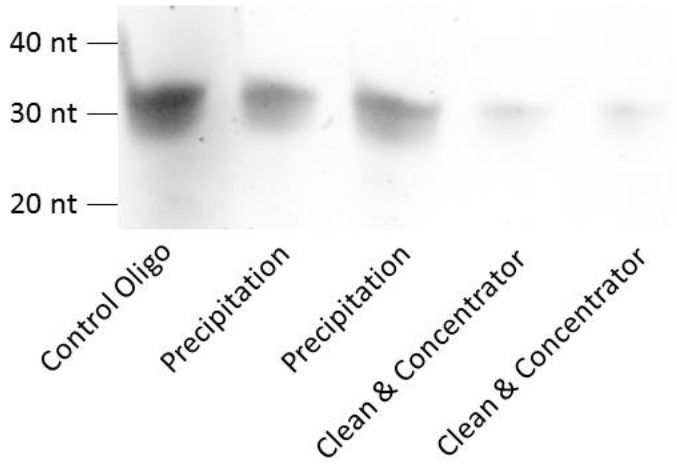

Figure 7 100 nmol of the control oligonucleotide were purified by means of an isopropanol precipitation and the Zymo Clean & Concentrator kit. Both purifications were carried out in duplicates. 10 % of each purified product (in ddH_2O) was loaded onto an 18 % TBE-Urea gel along with 10 nmol of the control oligonucleotide. Separation took place at 200 V and the gel was stained with SYBR Gold nucleic acid gel stain.

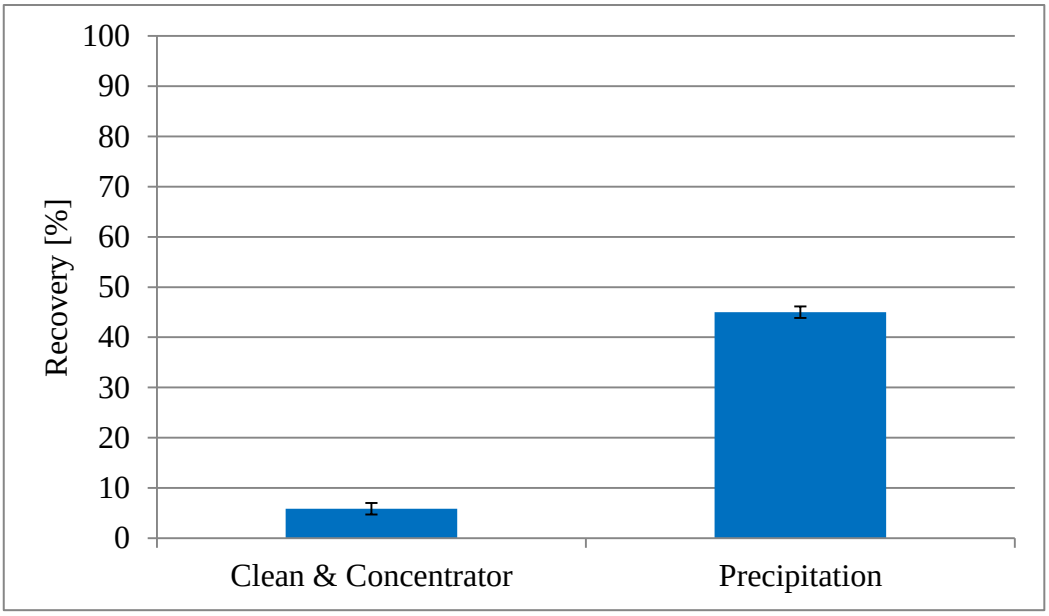

Figure 8 The band profiles from the gel in Figure 7 were analysed by means of an imaging software and their areas were integrated, allowing to quantify the amount of RNA on the gel. The recovery was calculated in relation to the measured area of the control oligonucleotide.

The kit however did not lead to a successful concentration of starting material, as the bands are a lot fainter despite the high concentration that was initially purified. Figure 8 shows a plot of the quantified bands from Figure 7, confirming the result. Although the precipitation method involves a substantial loss of material, it is with almost 50 % recovery still more reliable than the kit that only recovers a very small percentage of the initially used material.

For actual RPFs this would mean that after purification with the Zymo Clean & Concentrator kit, there will be only marginal amounts of RNA left that will hardly or not at all be visible on a gel, which would impede the process unnecessarily. Taking into account that this purification step has to be carried out after each reaction, this may lead to a large loss of sample material that cannot even be balanced with the PCR amplification at the end of the protocol as the group of recovered RPFs may not be representable anymore. This is too high a risk to be taken, which is why purifications were decided to be conducted according to Ingolia et al. (2012).

4.1.2 Gel extraction

After each separation of reaction mixes via gel electrophoresis, the obtained product band had to be excised and extracted from the gel. The Ingolia protocol (2012) suggested two different extraction methods. Both methods contain the excision of the gel slice on a UV screen. The further procedure is optional depending on the work flow of the experiment. The overnight extraction can be applied when the step is reached at the end of a work day, whereas the fast gel extraction method is useful when subsequent experiments can be conducted afterwards on the same day.

For the overnight gel extraction procedure the gel slice is simply incubated in RNA or DNA gel extraction buffer respectively with gentle agitation for a couple of hours (preferably overnight) after being frozen on dry ice for 30 minutes. Freezing the gel is not necessary, however, it supports the elution process, as ice crystals forming within the gel help loosening the polyacrylamide mesh. If sufficient time is allowed for incubation after freezing, RNA or DNA molecules can diffuse out of the polymer mesh of the gel into the buffer. For the fast gel extraction, the gel slice is forced through a needle sized hole in a microfuge tube by centrifugation, which leads to mechanical disruption. Due to the increased surface area of the gel, the molecules now have a much shorter diffusion way out of the gel. By incubating the disrupted gel slice in water for 10 minutes at 70 °C, the diffusion process is accelerated and RNA or DNA can be recovered after removing the gel slurry by means of centrifugation through a tube spin filter.

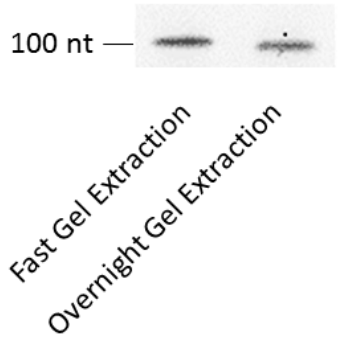

Figure 9 Two lanes of the 97 nt reverse transcription primer were separated on an 18 % TBE-Urea gel. The respective lanes were excised and extracted according to the fast and the overnight gel extraction method. The extracted and precipitated primers were loaded on another 18 % TBE-Urea gel and separated at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain.

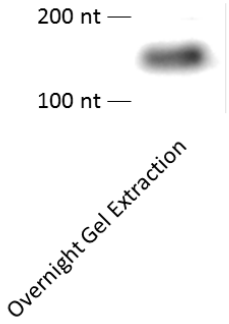

Figure 10 A 140 nt PCR product was separated on a 10 % TBE gel. The product was excised and extracted according the overnight gel extraction method. The gel slice was visualised under UV light after the procedure.

Both the fast and the overnight gel extraction method were tested by separating two lanes of the reverse transcription primer on an 18 % TBE-Urea gel and extracting each band with one of the two methods. The extracted primer was separated once more by gel electrophoresis on another 18 % TBE-Urea gel, which can be seen in Figure 9. As both bands are visible and there is no or only a marginal difference in the intensity of the bands, it was concluded that both gel extraction methods worked equally well from an 18 % TBE-Urea gel. Ingolia et al. (2012) recommended using only the overnight gel extraction method for PCR products that were obtained in the last step of the protocol. This was tested by carrying out a PCR reaction with the circularised reverse transcription primer and extracting the DNA as described in the protocol overnight. After an incubation time of approximately 15 hours, the gel slice was visualised under UV light, which is shown in Figure 10, the DNA in the gel extraction buffer was precipitated and loaded on another 10 % TBE gel (data not shown). Despite the long incubation time, the band was still well visible, whereas the precipitate contained so little DNA that a band could hardly be seen (not shown). The reason for this failed gel extraction may lie in the nature of the extracted molecules, while in Figure 9 a single stranded oligonucleotide of 97 nt length was extracted from an 18 % TBE-Urea gel,

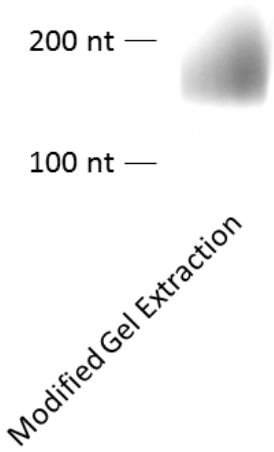

Figure 11 The 140 nt PCR product shown in Figure 10 underwent a modified gel extraction method in which the gel slice was mechanically disrupted before incubation in gel extraction buffer overnight. After extraction, DNA was precipitated and separated on a 10 % TBE gel at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain.

in Figure 10 the extraction of a 140 nt double-stranded PCR product from a 10 % TBE gel was attempted. Logic suggests that extraction from a 10 % gel should be easier than from an 18 % gel, as the mesh generated by crosslinking of the polyacrylamide is less dense. However, the double stranded DNA that is also 50 % longer than the extracted reverse transcription primer from Figure 9 will have much more difficulties diffusing out of the close meshed polymer network of a polyacrylamide gel. For this reason, a modified gel extraction method was developed, which facilitates the migration of the DNA through the mesh. The relative surface area was increased by mechanically disrupting the gel slice as in the fast gel extraction method before freezing and overnight incubation in DNA gel extraction buffer. Figure 11 shows the PCR product precipitate after this procedure separated on another 10 % TBE gel.

From these results it can be concluded that both gel extraction methods proposed by Ingolia et al. (2012) can be applied for the extraction of small oligonucleotides. Larger molecules, however, such as the double stranded PCR product of more than 140 nt should be extracted using the modified gel extraction method to ensure maximal product yield.

4.2 Cell Lysis and Nuclease Footprinting

As a first step, both protocols suggest an in-dish lysis, where cells are grown in a petri dish until confluent and are then scraped from the surface after addition of the lysis buffer (see chapter 3.2.1). This method, however, has the disadvantage that it produces larger volumes of cell lysate than intended, as there is still remaining growth medium and PBS that cannot be removed efficiently before lysis. As a consequence, the volume of obtained lysate is far higher than the 400 µl lysis buffer; usually around 800 µl lysate were recovered, in some cases up to 1200 µl. Such an increased volume requires equally increased amounts of RNAse and RNAse inhibitor during the further treatment, which makes the method more costly.

In a modified approach, lysis was carried out after trypsination of confluent cells and centrifugation. The obtained pellet was treated with the same amount of lysis buffer, which lead to only slightly higher lysate volumes, but smaller amounts of reagents necessary in the subsequent steps. Additionally, cells can be grown in culture flasks rather than petri dishes and lysate can be generated from residual cell broth after splitting.

Both procedures require the addition of the protein synthesis inhibitor cycloheximide in every step of the protocol, which ensures that the ribosomes stay attached to the mRNA. For the in dish lysis this means that the drug only needs to be added to the growth medium shortly before lysis and to the PBS used for washing the cells. The pellet lysis however requires higher amounts of the drug as it needs to be added in each step of the harvesting process, making this alternative somewhat more laborious than the in dish lysis. As will be presented later in this chapter, the numerous steps of the ribosome profiling workflow are accompanied by a loss of sample. This is not a major issue, as ribosome profiling is not aiming at recovering RPFs quantitatively, but it is thought that even after the loss of a considerable amount the relative quantities of recovered RPFs stay the same. Furthermore, the small amount of DNA left in the last step of the protocol will be amplified by PCR and, therefore,

provides a sufficient amount for deep sequencing. Generating more starting material, however, may be desirable as processed RNA and DNA will be easier to visualise after separation by gel electrophoresis.

4.3 Ribosome Recovery

After harvest, cell lysis, and monosome generation by nuclease footprinting, the ribosomes together with their RPFs had to be recovered and separated from other residual cell debris. Ingolia et al. (2012) recommend ultracentrifugation through a sucrose cushion, whereas the Epicentre kit (2013) gives the user a second option by suggesting the use of MicroSpin S-400 columns as an alternative to the ultracentrifugation. These spin columns are designed for the purification of DNA, e.g. for primer removal after a PCR reaction, and work on the base of a size exclusion chromatography. They contain a Sephacryl^{TM} resin with a pore size distribution that allows greater DNA molecules to be excluded from the pore volume, so they can elute unhindered, whereas smaller molecules, such as primers, enter the pores and do not elute immediately from the column (GE Healthcare 2006). Compared to the 4 h ultracentrifugation, this method is much more time effective, although the columns proved to be somewhat inconvenient in their handling, as they can only purify a maximum of 100 µl cell lysate each and are not allowed to dry out during the equilibration procedure, which is difficult to obtain (GE Healthcare 2006). Both procedures were carried out using a total of 300 µl lysate of both cell lines cultivated during this project. Purification of RNA was carried out in both cases with the miRNeasy kit and a subsequent precipitation as described in chapter 3.2.2 rather than using the RNA purification kit recommended by Epicentre (2013) (see chapter 4.1.1). The result of this comparison is shown in Figure 12 and shows an immense difference between both techniques.

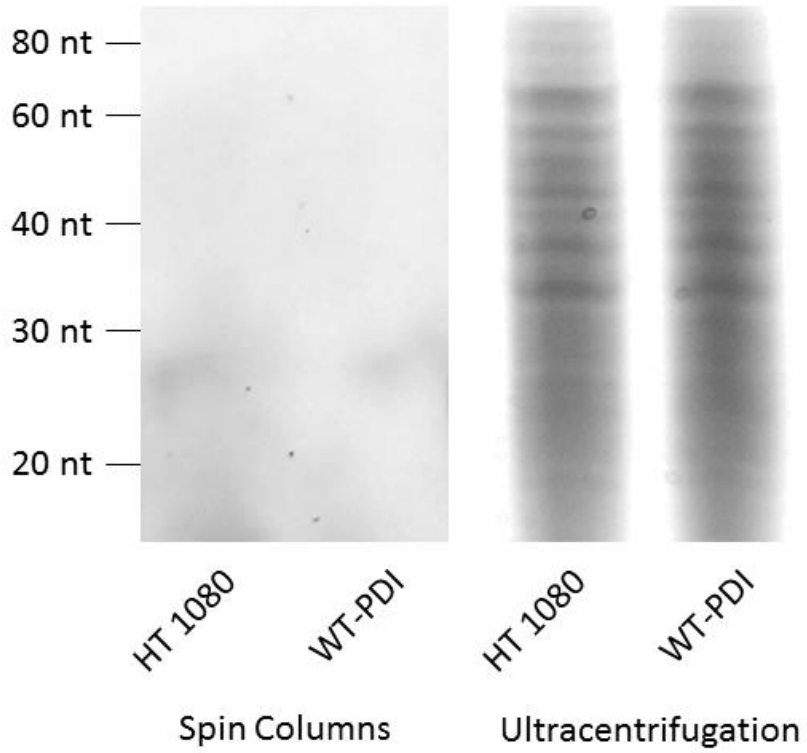

Figure 12 Ribosomes from 300 µl cell lysate of both HT 1080 and WT-PDI cells were recovered using the MicroSpin S-400 columns (Epicentre 2013) or by ultracentrifugation through a 1 M sucrose cushion (Ingolia et al. 2012). After purification using the miRNeasy kit and precipitation (see chapter 3.2.2), RNA was separated on an 18 % TBE-Urea gel at 200 V and stained with SYBR Gold nucleic acid gel stain.

The ultracentrifugation through the sucrose cushion led to the typical band pattern that was also observed by Ingolia et al. (2012), showing RNA bands of different sizes, amongst them the region around 30 nt that is believed to contain RPFs. Ribosome recovery performed with the spin columns, however, did hardly yield any visible bands at all, apart from two faint streaks below the 20 nt and the 30 nt region. Although the protocol provided by Epicentre mentions that RPF amounts may be too low to be visible on the gel and that a gel slice should still be excised in the 30 nt region, it was decided to not use the spin columns for future experiments, as the alternative approach yielded a much more satisfactory and less ambiguous result without being too laborious a procedure either. The reasons for the inefficacy of the MicroSpin S-400 columns can only be speculated, however, the proposed rationale is that RNA fragments generated during nuclease footprinting adsorb to the Sephacryl^{TM} resin and thus cannot be recovered. Ultracentrifugation experiments were repeated using smaller

volumes of cell lysate with the aim of obtaining a clearer band pattern that can confirm this step of the protocol is carried out correctly.

The obtained gel is shown in Figure 13, displaying purified sample volumes of 150 µl, 200 µl, and 300 µl. Both cell lines, but particularly HT 1080, feature the typical band pattern observed by Ingolia et al. (2012) with one band at 30 nt, which might represent the RPFs and that cannot be seen clearly when 300 µl cell lysate were applied as in Figure 12. The cell line WT-PDI exhibits a fainter band pattern as HT 1080, which is most likely due to the fact that these cells grow slower than the wild type leading to lower cell yields and, therefore, lower RNA yields at the point of cell lysis if both cultures were passaged in the same ratio at the same time.

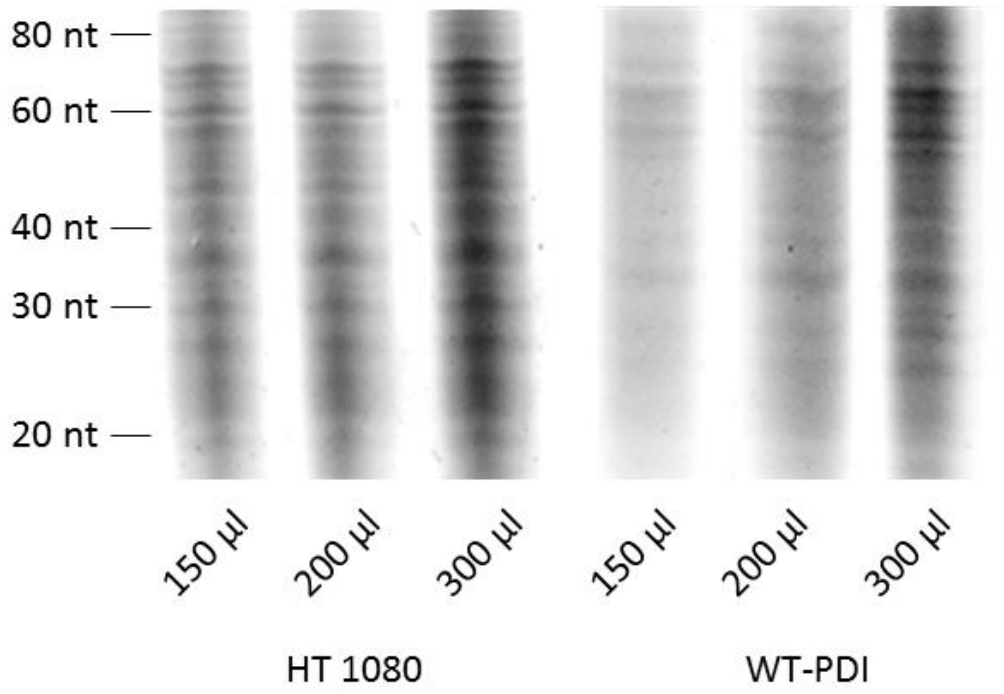

Figure 13 Ribosomes from different volumes of cell lysate (150 µl, 200 µl, 300µl) of both HT 1080 and WT-PDI cells were recovered by ultracentrifugation through a 1 M sucrose cushion (Ingolia et al. 2012). After purification using the miRNeasy kit and precipitation (see chapter 3.2.2), RNA was separated on an 18 % TBE-Urea gel at 200 V and stained with SYBR Gold nucleic acid gel stain.

4.4 Dephosphorylation and Linker Ligation

A critical step in the ribosome profiling strategy is the ligation of the miRNA cloning linker to the RPF, which is crucial for the generation of a cDNA library. As previously stated, it can be difficult to obtain sufficient amounts of RPFs that are clearly visible on a gel, which is why the following steps were decided to be carried out with a synthetic oligonucleotide first, taking the molecule through the exact same procedure as an actual RPF, starting with the dephosphorylation. Although this first reaction was not necessary as the synthetic oligonucleotide was already available in dephosphorylated form on its 3' terminus, the additional reaction was still carried out in order to keep the experimental set-up as close to the original protocol as possible. Additionally, it was not thought that the dephosphorylation step negatively influences the further procedure. The subsequent linker ligation, however, proved to be far more difficult to reach than expected, as a ligation product could not be formed even under improved reaction conditions as suggested by Viollet et al. (2011). As the reason for the failure of this reaction could not yet be determined, it was decided to attempt the reaction with an *in vitro* RNA transcript rather than a synthetic oligonucleotide. In addition, the reaction was conducted with a synthetic 5'-FAM labelled 31-mer RNA molecule identical to the one used by Viollet et al. (2011) under the same reaction conditions that were successful for this research group. Reactions were carried out at 16 °C overnight (Viollet et al. 2011) and at room temperature for 2.5 hours as suggested by Ingolia et al. (2012). The amount of miRNA cloning linker was reduced from 150 nmol to 10 nmol per reaction for economic reasons and because this should not have a negative effect on the reaction according to Viollet et al. (2012), who used substantially smaller amounts of linker.

The ligation products were separated from unreacted RNA and linker by gel electrophoresis. The Epicentre protocol (2013) omitted this step and instead removed the unwanted molecules by the addition of an enzyme that is not clearly defined. Logic suggests that this enzyme must

specifically cleave RNA and DNA molecules but does not have any affinity for RNA/DNA hybrids, such as the ligation product. However, such an enzyme could not be found anywhere in literature and it is a reaction that is not crucial for the success of the method, which is why ligation products were purified by gel electrophoresis as recommended by Ingolia et al. (2012).

4.4.1 Linker Ligation to *in vitro* Transcript

The sequence of the natural oligonucleotide is not of importance, as the ligation is supposed to work with any RNA oligonucleotide as long as it is dephosphorylated on its 3' end. The only requirement is its size; while the ligation will most probably still work with a long RNA molecule, this cannot be visualised on a gel due to its lack of resolution for higher molecular weight, as the linker is only 17 nucleotides long. The maximal acceptable size of the oligonucleotide that still allowed seeing a size shift in the gel was estimated to be 100 nt, any longer and the ligation product could not be clearly distinguished anymore from its reactant. It was decided to use the plasmid PPL-pGEM 4, which contains a T7 promotor as well as a restriction site for the restriction enzyme Nco I 59 nt downstream of the promotor as shown in Figure 14. After cutting the plasmid with Nco I the now linear plasmid could be transcribed *in vitro* using the T7 RNA polymerase. This enzyme is promotor specific and only transcribes the DNA downstream of the promotor into RNA (Sousa & Mukherjee 2003), producing a 83 nt long RNA molecule.

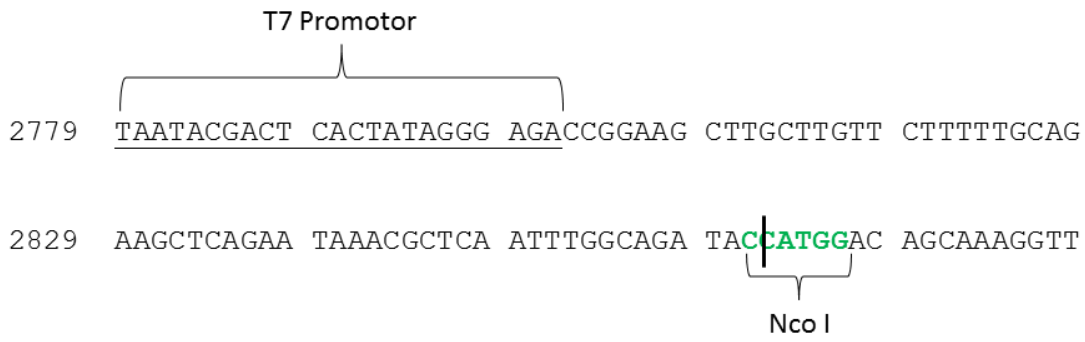

Figure 14 Section of the plasmid PPL-pGEM 4 containing a T7 promotor (underlined, position 2779-2801) as well as a Nco I restriction site (green, Nco I cuts after position 2861).

Figure 15 shows the digested plasmid as well as the obtained *in vitro* transcript. The linear plasmid is visible at the top of the gel still sitting in the sample wells as it is too big to penetrate the tight polymer mesh. Further down, there are another two bands visible, which represent two different forms of supercoiled circular plasmid DNA, which can run further than the linear DNA depending on their degree of interlacement. In addition to these DNA bands, the second lane shown in Figure 15 contains a band that represents the RNA transcript and has the predicted size of 83 nt.

The obtained RNA transcript was purified according to chapter 3.3 and could then directly undergo the process of dephosphorylation and linker ligation. Figure 16 shows the RNA transcript that can also be seen in Figure 15 along with the dephosphorylated product in the ligation reaction mix without the enzyme, which served as a negative control. As expected, the linker is visible as a band at 20 nt before the reaction and is much fainter after the reaction, as it was used up in the process. The most obvious indication for the successful reaction, however, is a clear shift of the band at 80 nt, which represents the *in vitro* transcript towards the 20 nt bigger ligation product at 100 nt. This band is visible neither in the RNA transcript, nor in the negative control and therefore clearly represents the ligation product.

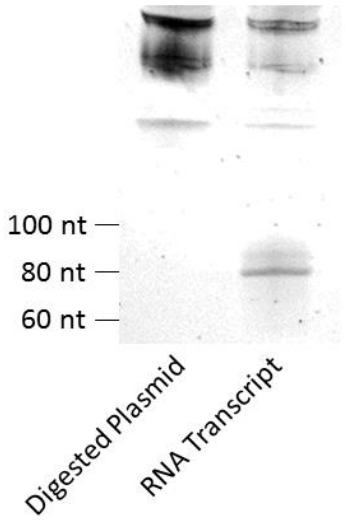

Figure 15 The plasmid PPL-pGEM 4 was digested with Nco I and an *in vitro* transcription was carried out as described in chapter 3.3. The digested plasmid and the RNA transcript were separated on an 18 % TBE-Urea gel at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain.

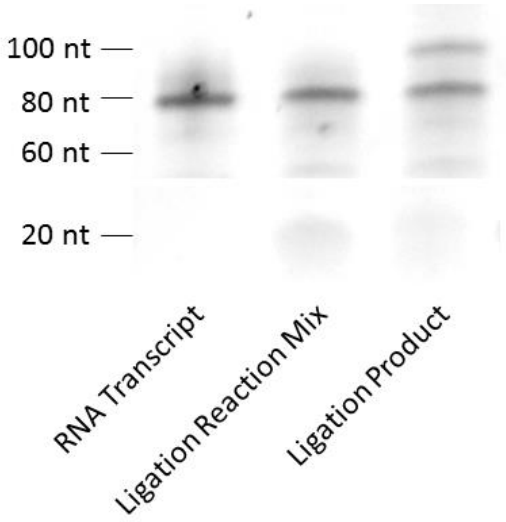

Figure 16 Ligation of the *in vitro* transcript to the miRNA cloning linker. The first lane shows the RNA transcript from Figure 15, followed by a negative control consisting of the ligation reaction mix minus the ligase and the ligation product after the reaction for 2.5 h at room temperature. Samples were separated on an 18 % TBE-Urea gel at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain.

As depicted earlier, the reaction efficiency can be improved by increasing the reaction time and/or decreasing the reaction temperature. For this reason, the same reaction was repeated at 16 °C overnight (15-18 h) instead of carrying out the reaction at room temperature for 2.5 h; concentrations of reagents, linker and transcript were not altered. As shown in Figure 17 both reactions were successful, leading to a band at 100 nt; however at 16 °C overnight, the reaction could be carried out more completely, leading to a stronger band for the ligation product. In contrast, the shorter reaction time of 2.5 h combined with the higher temperature showed a much fainter band for the ligation product, whereas the *in vitro* transcript band at 80 nt is was very prominent. Longer reaction times at lower temperatures are hence beneficial for the formation of the ligation product; however, they might prolong the process unnecessarily and a complete turnover is not mandatory in this case.

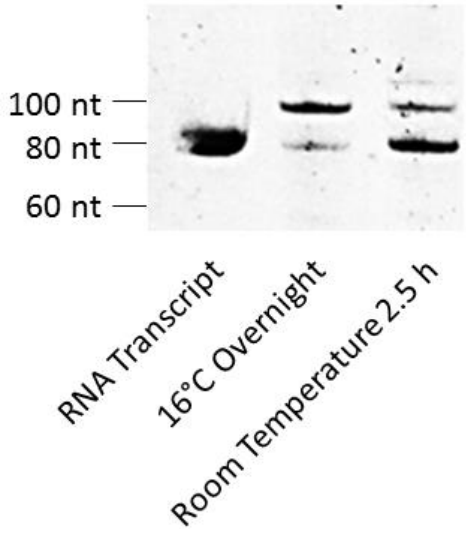

Figure 17 Ligation of the *in* vitro transcript to the miRNA cloning linker carried out at 16 °C overnight as well as at room temperature for 2.5 h. Ligation products were loaded together with the RNA transcript as a negative control on an 18 % TBE-Urea gel and electrophoresis was carried out at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain.

As ribonucleic acids are very prone to degradation even in the presence of ribonuclease inhibitors, a long incubation time can lead to stronger breakdown of the material, which was also witnessed during this work (data not shown).

In conclusion, the linker ligation can be successfully carried out after dephosphorylation of a natural RNA molecule obtained by *in vitro* transcription when carried out according to Viollet et al. (2011), using a smaller amount of miRNA cloning linker, which makes this step of the protocol more economic without hindering the reaction. A more complete turnover can be yielded by increasing the reaction time and decreasing the reaction temperature, although this is not crucial for the further steps of the protocol and may, therefore, be omitted for the sake of time and in order to keep RNA degradation to a minimum.

4.4.2 Linker Ligation to a Labelled Synthetic Oligonucleotide

As the reasons for the unsuccessful reaction with a synthetic oligonucleotide were still not clear, it was suggested to attempt the linker ligation with a 5'-FAM labelled 31-mer RNA oligonucleotide in addition to the linker ligation to a natural RNA oligonucleotide. This

oligonucleotide represented the size of an actual RPF better and, therefore, served as a better control whether the reaction can be successfully conducted. The second advantage of using a labelled oligonucleotide as a control is that correct ligation can be monitored more effectively. Furthermore, it was known that the reaction definitely works with this particular oligonucleotide (Viollet et al. 2011).

The most pivotal difference between the linker ligation to RPFs as suggested by Ingolia et al. (2012) and the ligation to the 5'-FAM labelled synthetic oligonucleotide described by Viollet et al. (2011) is the dephosphorylation, which has to be carried out with RNase I digested RNA due to the nature of the nuclease (DelCardayré & Raines 1995). However, when ligation was attempted with a synthetic oligonucleotide that had previously undergone the reaction with PNK, ligation did not succeed (data not shown). Although it was not considered to have such a major influence, it was decided to perform the linker ligation without the previous dephosphorylation in order to repeat the experiment in the exact same way as Viollet et al. (2011). The result from this experiment is depicted in Figure 18, showing the 5'-FAM labelled 31-mer as a negative control as well the ligation product. The apparent increased size if the molecule (35-38 nt instead of 31 nt) is due to the FAM label, which adds an additional 376 g mol^{-1} to the molecule.

Surprisingly, the ligation was successful when performed without the previous dephosphorylation, which can be seen in Figure 18, showing a shift of the 5'-FAM labelled oligo by 20 nt, which is the length of the miRNA cloning linker.

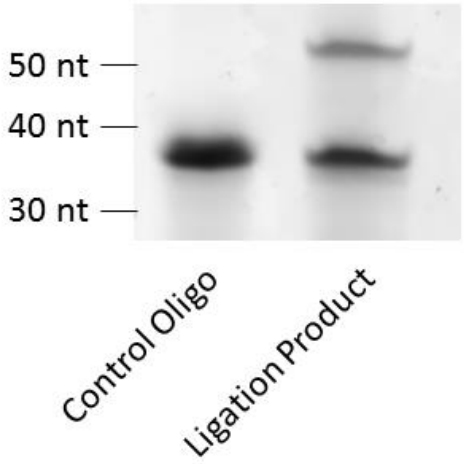

Figure 18 Linker Ligation was carried out with a 5'-FAM labelled oligonucleotide at room temperature for 2.5 h. The ligation product was loaded together with the unreacted oligonucleotide on an 18 % TBE-Urea gel and electrophoresis was performed at 200 V. The gel did not need to be stained but the labelled oligonucleotide could be visualised on a phosphorimager.

It was known before that dephosphorylation of synthetic oligonucleotides is not necessary for the success of a subsequent ligation; what is very surprising is that the previous dephosphorylation even appears to prevent the reaction and even more surprising that the same procedure does not have any effect on the *in vitro* transcript although this molecule does not have to be dephosphorylated either. Synthetic oligonucleotides are synthesised in 3' - 5' direction and are composed of nucleoside phosphoramidites rather than conventional nucleosides, as they are more prone to form internucleosidic linkages (Beaucage & Iyer 1992). It is possible that this characteristic provides a point of attack for PNK, introducing alterations into the molecule and thus making the subsequent linker ligation impossible. PNK, however, is a well characterised kinase that was already described in 1977 by Uhlenbeck and Cameron and it is unlikely that the enzyme has another catalytic ability not yet discovered. Why it does not appear to be possible to ligate a synthetic oligonucleotide that has previously underwent a reaction with PNK is an interesting question to ask, but needs to be neglected, as it does not contribute to the aim of this work.

4.5 Reverse Transcription

As the linker ligation had worked well with the *in vitro* RNA transcript, it was decided to attempt the following reverse transcription (RT) with the excised ligation product from the gel in Figure 16. Both protocols available by Epicentre (2013) and Ingolia et al. (2012) follow a very similar approach, adding the reverse transcription primer to the RPF together with a reverse transcriptase and other components, whereby the enzyme used by Epicentre (2013) is not further characterised. In this work the enzyme SuperScript III was used as recommended by Ingolia et al. (2012), which is a reverse transcriptase with reduced RNase H activity and higher thermostability; two features that promise to increase the product yield compared to other reverse transcriptases, as less RNA gets destroyed by ribonuclease activity and the reaction is more effective at higher temperatures (LifeTechnologies n.d.). In Figure 19 the reverse transcription product is shown along with the negative reaction that did not contain the enzyme as well as the reverse transcription primer.

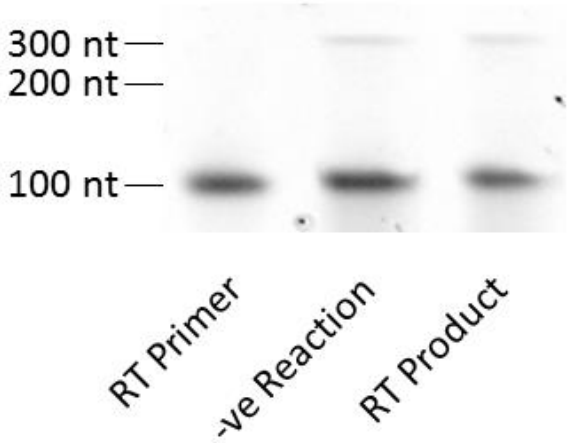

Figure 19 The ligation product from Figure 16 was reverse transcribed as described in chapter 3.2.5 and the RT product was separated together with the RT primer alone and a negative reaction that lacked the RT enzyme on an 18 % TBE-Urea gel at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain.

As described in chapter 2.2, the RT primer contains sequences that will act as priming sites in the PCR reaction later in the protocol, which explains the unusual size of 97 nucleotides, which can also be seen as a prominent band in Figure 19. The reverse transcription product is suspected to have a size of 180 nt, however, it appears to be larger than 200 nt in Figure 19, which could not be explained. Surprisingly, the negative control shows the same band as the reaction, although the product cannot have formed without the reverse transcriptase being present in the reaction mix. This can be explained by the fact that the RT primer and the ligation product are partly complementary to each other, which allows them to hybridise given the temperature during reverse transcription does not exceed the melting temperature T_m of the double strand. The approximate melting temperature can be calculated with the Wallace rule for short oligonucleotides (14-20 nt):

$$T_m \approx 2\,°\mathrm{C} \cdot (\mathrm{A} + \mathrm{T}) + 4\,°\mathrm{C} \cdot (\mathrm{G} + \mathrm{C})$$

with T_m = Approximate melting temperature

(A+T) = Number of adenines and thymines

(G+C) = Number of guanines and cytosines

which estimates the melting temperature on the basis of hydrogenbonds that have to be cleaved during the process of denaturing a DNA double strand. Adenine and thymine develop two hydrogenbonds between each other, contributing about 2 °C each to the melting temperature, while guanine and cytosine are connected by three hydrogenbonds and thereby add 4 °C each to the melting temperature.

The complementary sequence between the ligation product and the RT primer, which is effectively the sequence of the miRNA cloning linker, consists of 7 adenine and thymine bases and 10 guanine and cytosine bases; hence, according to the Wallace rule it denatures at approximately 54 °C. As the RT reaction was carried out at 50 °C, this provides good

conditions for the formation of the hybrid between RT primer and ligation product. In conclusion, what can be seen in Figure 19 is not the same oligonucleotide, although both bands migrate at the same length. It shows a hybrid which is double stranded in a section of 17 nucleotides in the centre of the molecule in the negative control lane and the reverse transcribed product, which is completely composed of single stranded DNA.

In order to remove the remaining RNA from the reaction mix, both available protocols use different methods. The Epicentre kit treats the reaction with an exonuclease followed by a mixture of ribonucleases, while the protocol by Ingolia et al. (2012) suggests an alkaline hydrolysis, a method that allows denaturing RNA on the mechanism that the ribose is prone to degradation due to its 2'-hydroxil group, which is not present in the deoxyribose of DNA (Hubert 2006). As the Epicentre protocol (2013) does not give any further information about the nature of the RNase mix, it was decided to follow the protocol by Ingolia et al. (2012) in this case. Furthermore the addition of an exonuclease as proposed by Epicentre (2013) was believed to be unfavourable, as this class of enzymes degrades DNA, which is not wanted in this case. However, the manufacturer does not specify the kind of exonuclease that is provided in the kit, which is why the purpose of this enzyme can only be speculated.

4.6 Circularisation and PCR Amplification

Both, Ingolia et al. (2012) and Epicentre (2013), use a ligase that is probably the same and that was also used in this work (see chapter 3.2.6). The enzyme catalyses intramolecular ligation of single stranded DNA and its ligation efficiency can be influenced by the sequence of the DNA molecule. To increase ligation efficiency, the manufacturer recommends extension of the reaction time by another hour. In further experiments, both the unextended RT primer, as well as the reverse transcribed *in vitro* transcript were subjected to circularisation, allowing the reaction to take place for either one or two hours, before PCR was carried out as described in chapter 3.2.7. The aim of the experiment was to amplify either

of the molecules successfully and subsequently sequence the PCR product with the help of the used PCR primers. However, this proved to be very difficult, as the PCR protocol had already been optimised for templates of a certain length and realistic concentration by Ingolia et al. (2012), resulting in very ambiguous gels that were mostly smeared and contained a huge amount of overamplified PCR products (data not shown). The control molecules that were used in this reaction were either too short (RT primer), too long (*in vitro* transcript) or too concentrated to yield appropriate results during this very specific PCR protocol and sequencing of the products could not be carried out. Optimisation of the reaction with these controls would have meant to carry out a vast number of additional experiments, in which parameters such as temperature and time for denaturing, annealing, and extension were altered; additionally, experiments with varying concentrations of starting material, but also the addition of reagents such as DMSO (Chester & Marshak 1993) might have been necessary. This would have meant an enormous effort that was considered to not be worth the expected results, since the experiment should serve a mere validation whether the provided protocol needed modification or not. It was decided to carry out the circularisation and the PCR reaction with RPFs.

4.7 The Process with Ribosome Protected Fragments

Having most steps of the protocol closely examined and validated, it was decided to advance with authentic samples from cell cultures to optimise the remaining incongruities and to finally confirm the feasibility of the ribosome profiling workflow. All procedures were carried out as previously described in chapters 3 and 4.

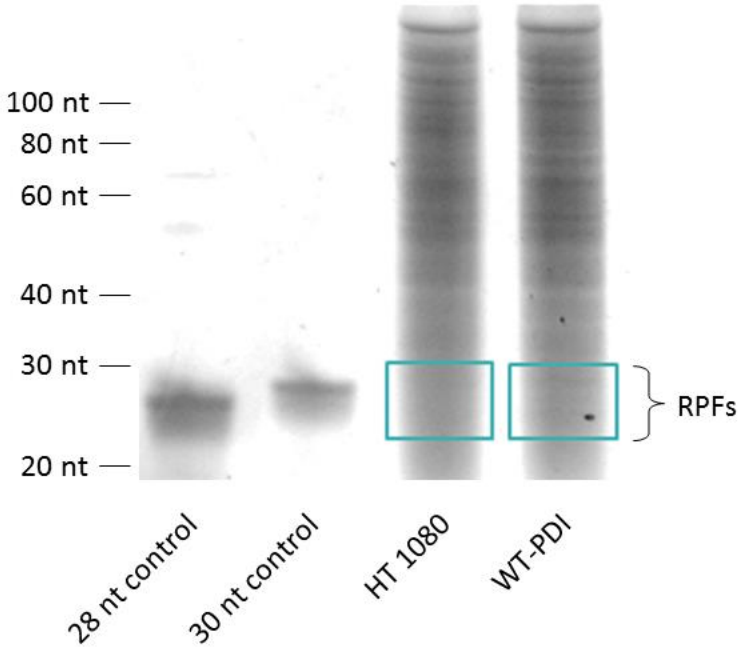

Figure 20 Cell lysis, nuclease footprinting, and ribosome recovery were carried out as described in chapter 3, without conducting the rRNA removal procedure. RNA was purified and separated on an 18 % TBE-Urea gel together with a 28 nt and a 30 nt control oligonucleotide at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain and RPF containing regions (blue squares) were excised.

Figure 20 shows the size selection of the RPFs after nuclease footprinting and ribosome recovery via ultracentrifugation. Both, the 28 nt and the 30 nt control oligonucleotide were loaded on the gel next to the RPFs containing samples, which facilitated visualisation and excision of the correct section. The blue squares in Figure 20 show which region was excised and underwent further treatment. Due to the high sample volume (300 µl), distinct bands were difficult to spot; higher impurities probably consisting of mainly ribosomal RNA were anticipated but they were accepted for the sake of better visibility of discrete bands in the following steps. Contamination with ribosomal RNA was accounted for in chapter 4.8.

RPFs were extracted from the gel and dephosphorylation and linker ligation was performed as previously described. Ingolia et al. (2012) included another precipitation step of the ligation product before separation by gel electrophoresis, which leads to a gel as shown in Figure 21.

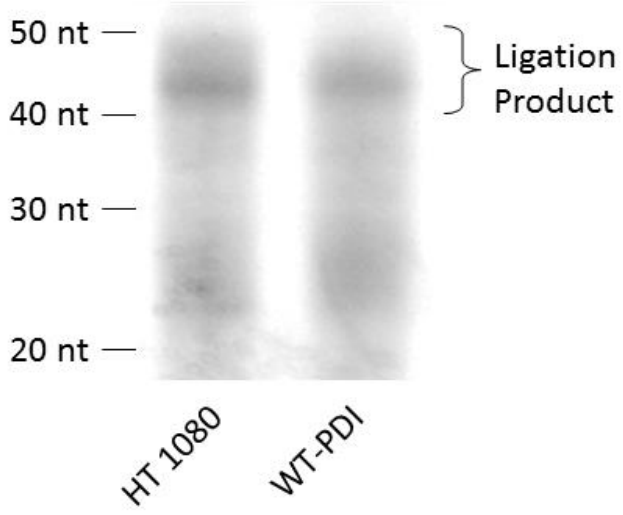

Figure 21 Linker ligation was performed as described in chapter 3.2.4 and RNA was subsequently precipitated before separation on an 18 % TBE-Urea gel at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain and the ligation product indicated in the figure was excised and extracted.

RPFs were still visible in the region between 25 nt and 30 nt, in addition, a band with the size between 40 nt and 50 nt appeared that was most probably the ligation product, while the residual linker itself was still visible as a very faint band of 20 nt. In order to save time and to omit one precipitation step in which sample material can get lost, the ligation reaction mix can be directly loaded onto an 18 % TBE-Urea gel. The reaction mix only comprises 20 µl per sample, which can be easily distributed into 4 wells, allowing two samples to be separated per gel as shown in Figure 22. The gel features the same bands that can be seen in Figure 21, with the difference that the bands are much fainter and more difficult to spot. However, as the reaction had proved to work previously, this disadvantage can be accepted, as sharp, bright bands are not required for the preparative nature of this gel electrophoresis. Because the concentration of the ligation product is lower in this case, it is advised to perform the modified gel extraction method described in chapter 4.1.2 to maximise the product yield.

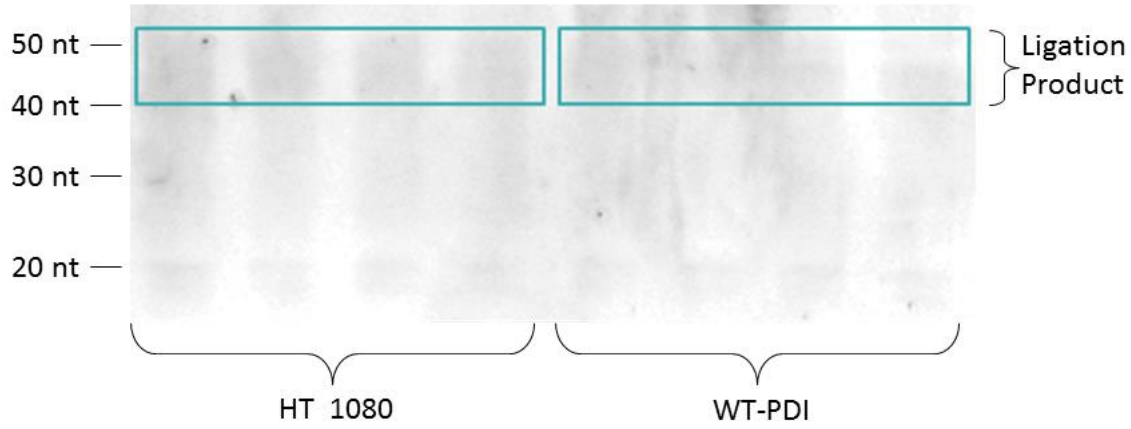

Figure 22 Linker ligation was performed as described in chapter 3.2.4 and the entire ligation reaction mix was subsequently separated on an 18 % TBE-Urea gel at 200 V without prior precipitation. The gel was stained with SYBR Gold nucleic acid gel stain and the ligation products indicated in the figure (blue squares) were excised and extracted.

Regardless if the ligation product was separated with or without prior precipitation, the extracted RNA underwent reverse transcription as described in chapter 3.2.5, which resulted after separation in a gel as shown in Figure 23. The reverse transcription adds another 79 nucleotides to the sequence, resulting in a single stranded DNA molecule of 126 nt length, which can be clearly distinguished from the 97 nt long reverse transcription primer.

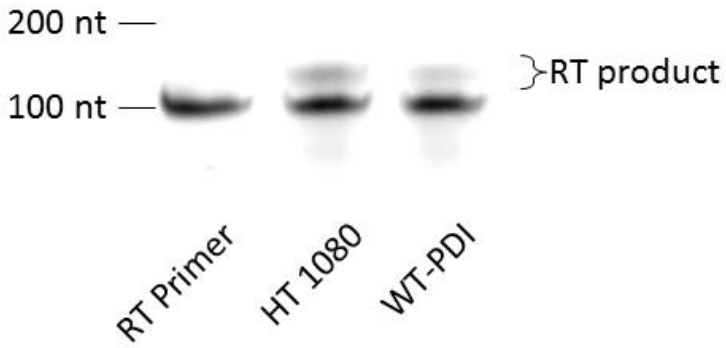

Figure 23 The obtained ligation product from Figure 22 underwent reverse transcription as described in chapter 3.2.5. The RT product was separated from the unreacted RT primer on an 18 % TBE-Urea gel at 200 V together with the RT primer alone that served as a negative control. The gel was stained with SYBR Gold nucleic acid gel stain and the RT product indicated in the figure was excised and extracted.

Circularisation and PCR amplification was carried out as recommended by Ingolia et al. (2012) and as described in chapters 3.2.6 and 3.2.7 without attempting to modify the protocol.

The success of the PCR amplification is highly dependent on template abundance, i.e. depending on how much starting material undergoes the reaction, a higher or lower number of PCR cycles is necessary for the amplification. The process is furthermore very prone to overamplification, leading to smeared PCR bands and the occurrence of bands with a molecular weight that is much higher than expected. For these reasons the reaction was carried out for differing numbers of cycles between 6 and 14, the result for HT 1080 cells is shown in Figure 24, for WT-PDI cells in Figure 25.

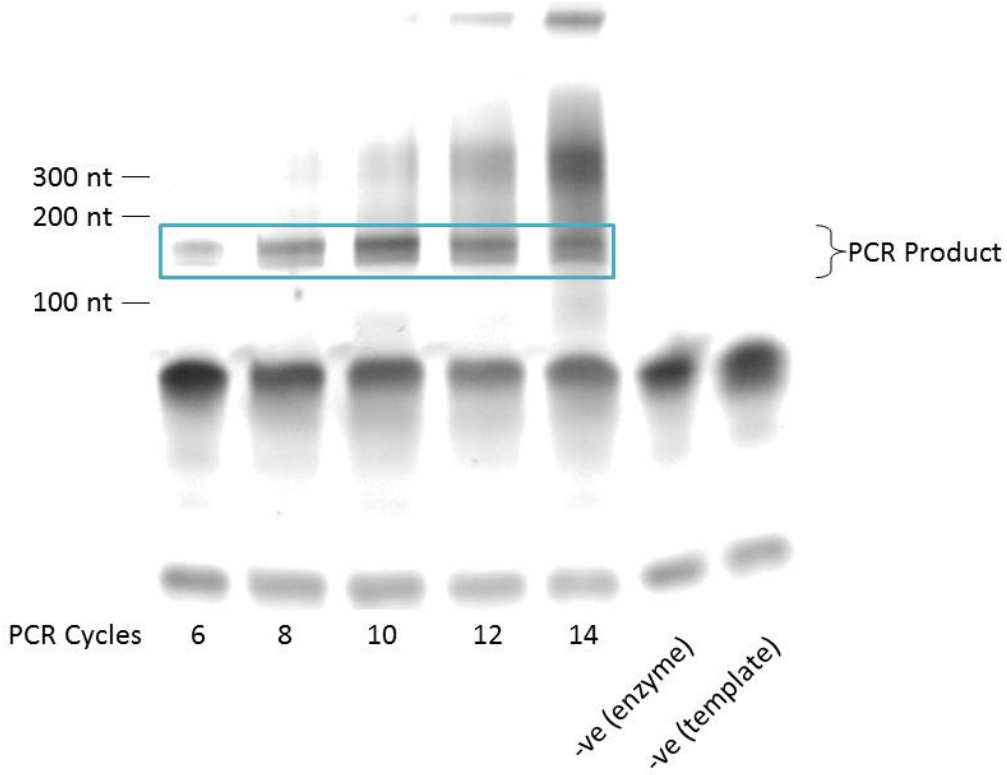

Figure 24 PCR amplification of the circularised RT product derived from HT 1080 cells was carried out as described in chapter 3.2.7 for 6, 8, 10, 12, and 14 cycles. In addition, two negative controls (no enzyme and no template) were incubated for 14 cycles and all samples were loaded onto a 10 % TBE gel without previous purification. Electrophoresis was performed at 200 V and the gel was stained with SYBR Gold nucleic acid gel stain. The expected PCR product had a size of 176 nt (blue square).

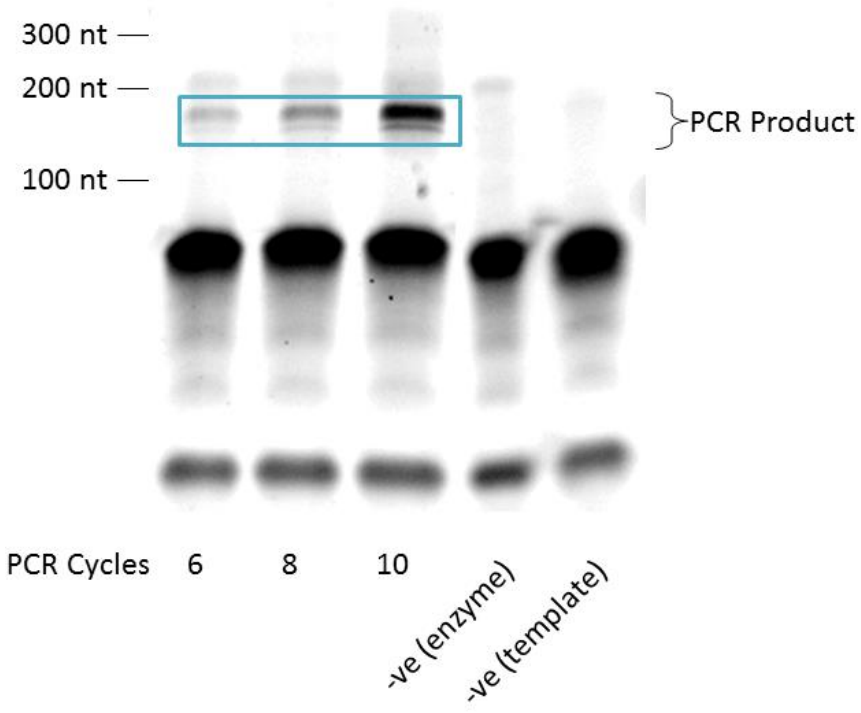

Figure 25 PCR amplification of the circularised RT product derived from WT-PDI cells was carried out as described in chapter 3.2.7 for 6, 8, and 10 cycles. In addition, two negative controls (no enzyme and no template) were incubated for 14 cycles and all samples were loaded onto a 10 % TBE gel without previous purification. Electrophoresis was performed at 200 V and the gel was stained with SYBR Gold nucleic acid gel stain. The expected PCR product had a size of 176 nt (blue square).

Due to an unexpected sample loss, PCR amplification of WT-PDI samples could only be carried out for 6, 8, and 10 cycles.

PCR amplification was successful for both samples, leading to a band of the expected size of 176 nt (blue squares). As predicted, the amount of PCR product increased with the number of PCR cycles, however, this also resulted in an increased production of unspecific PCR products that had a much higher apparent molecular weight, migrating above 200 nt. Difficult to see in Figure 24, but much clearer in Figure 25, a band just below the PCR product occured and became stronger with the number of PCR cycles conducted. This 144 nt band was derived from the amplified unextended RT primer and needed to be carefully avoided when excising the PCR product from the gel. The two negative controls show none of these artefacts apart from a faint adaptor-dimer band in the negative template reaction, but mostly the primers migrating at what is suspected to be 28 nt and 64 nt. Moreover, Figure 25 shows a faint band

at 200 nt in the negative enzyme control. This band is most likely the circularised RT product, which only has a size of 126 nt, but cannot penetrate the polyacrylamide mesh of the gel as such due to its bulkiness after circularisation. This assumption is supported by the band not occurring in the negative template control.

For these experiments and the amount of template that was introduced into the PCR reaction, a cycle number of eight or ten was appropriate to generate a sufficient amount of PCR product without overamplifying the template excessively. The respective bands could now be excised, DNA could be extracted and subjected to deep sequencing. However, the DNA samples will still contain a considerable amount of ribosomal RNA, which would unnecessarily increase the quantity of generated data.

4.8 Depletion of Ribosomal RNA

As previously stated, ribosomal RNA makes up a significant part of the RNA that is obtained after nuclease footprinting. Most contamination is in fact derived from very specific sequences that probably occur by cleavage of the rRNA inside the ribosome at those sites that are accessible for the ribonuclease (Ingolia et al. 2012). Ingolia et al. (2012) concentrate on these specific sequences and remove them with the help of a biotinylated substraction oligo pool in combination with magnetic streptavidin coupled DynaBeads. Particularly the synthesis of the oligo pool is very expensive, as a 5'-biotin-TEG has to be added to each of the 14 oligonucleotides, furthermore they need to be purified elaborately by HPLC, in order to exclude non-modified oligonucleotides. Epicentre (2013) suggests the use of an rRNA removal kit that more generally eliminates most of human rRNA and therefore needs to be applied before conversion of the RPFs into a cDNA library, e.g. after monosome purification. The separation of RNA after ribosome recovery and subsequent rRNA depletion is shown in Figure 26. Compared to Figure 20, where this step was omitted, the gel featured a substantially decreased amount of RNA bands, although the same sample volumes were used,

indicating that the rRNA removal was successful. The removal of rRNA at this early stage of the protocol has the disadvantage that RNA amounts remain very low throughout the entire cDNA library generation, which makes it much more difficult to visualise reaction products on the gel, as can be seen in Figure 27 and Figure 28, in which the ligation and the reverse transcription product are barely visible compared to their counterparts in Figure 21 and Figure 23.

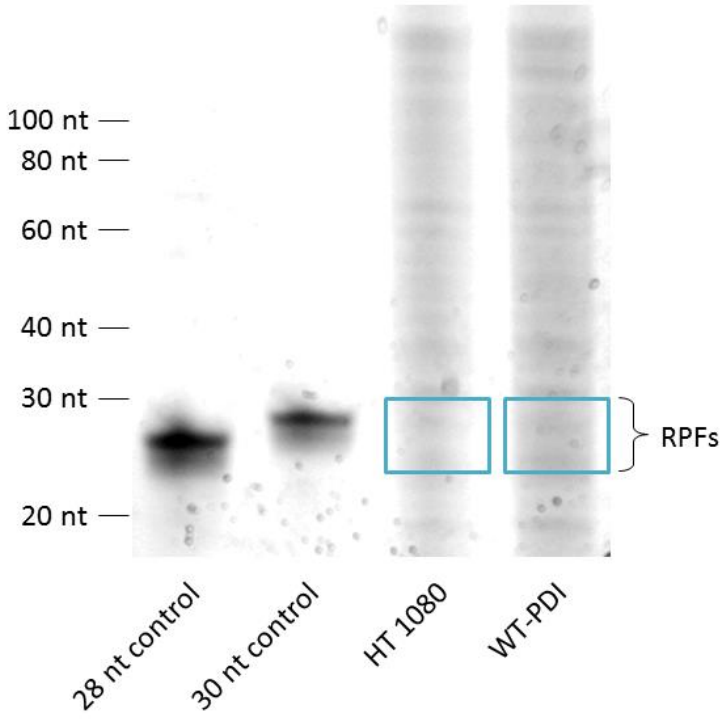

Figure 26 Cell lysis, nuclease footprinting, and ribosome recovery were carried out as described in chapter 3. RNA was purified and separated on an 18 % TBE-Urea gel together with a 28 nt and a 30 nt control oligonucleotide at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain and RPF containing regions (blue squares) excised.

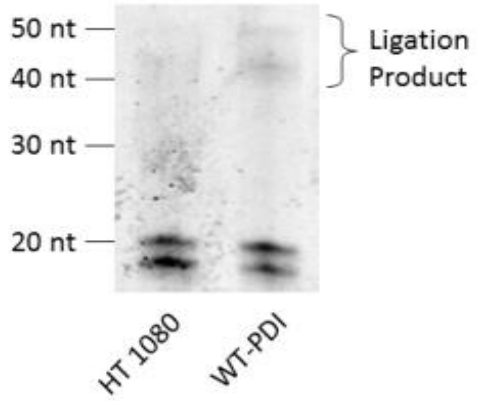

Figure 27 Linker ligation was performed as described in chapter 3.2.4 and RNA was subsequently precipitated before separation on an 18 % TBE-Urea gel at 200 V. The gel was stained with SYBR Gold nucleic acid gel stain and the ligation product indicated in the figure was excised and extracted.

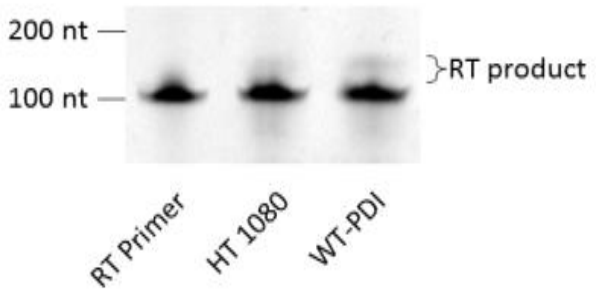

Figure 28 The obtained ligation product from Figure 27 underwent reverse transcription as described in chapter 3.2.5. The RT product was separated from the unreacted RT primer on an 18 % TBE-Urea gel at 200 V together with the RT primer alone that served as a negative control. The gel was stained with SYBR Gold nucleic acid gel stain and the RT product indicated in the figure was excised and extracted.

Despite the low sample amounts that were generated during this experiment, circularisation and PCR could successfully be performed for the HT 1080 sample, the resulting gel is shown in Figure 29. WT-PDI, however, yielded no detectable PCR product, which is probably due to a loss of the sample after extraction of the RT product (data not shown), since it was still visible in Figure 28. Figure 29 exhibits the same characteristics as its counterparts Figure 24 and Figure 25, showing the 176 nt PCR product after 8 and 10 PCR cycles, whilst 6 cycles was not enough for sufficient amplification and more than 12 cycles resulted in severe overamplification with unspecific product bands larger than 300 nt. Interestingly, the unextended RT primer band at 144 nt was much more prominent than in previous experiments, which can be explained by a likely contamination of the sample with the unreacted RT primer during excision of the RT product from the gel shown in Figure 28, as the bands migrated very close to each other, resulting in the RT primer band merging partly with the RT product band.

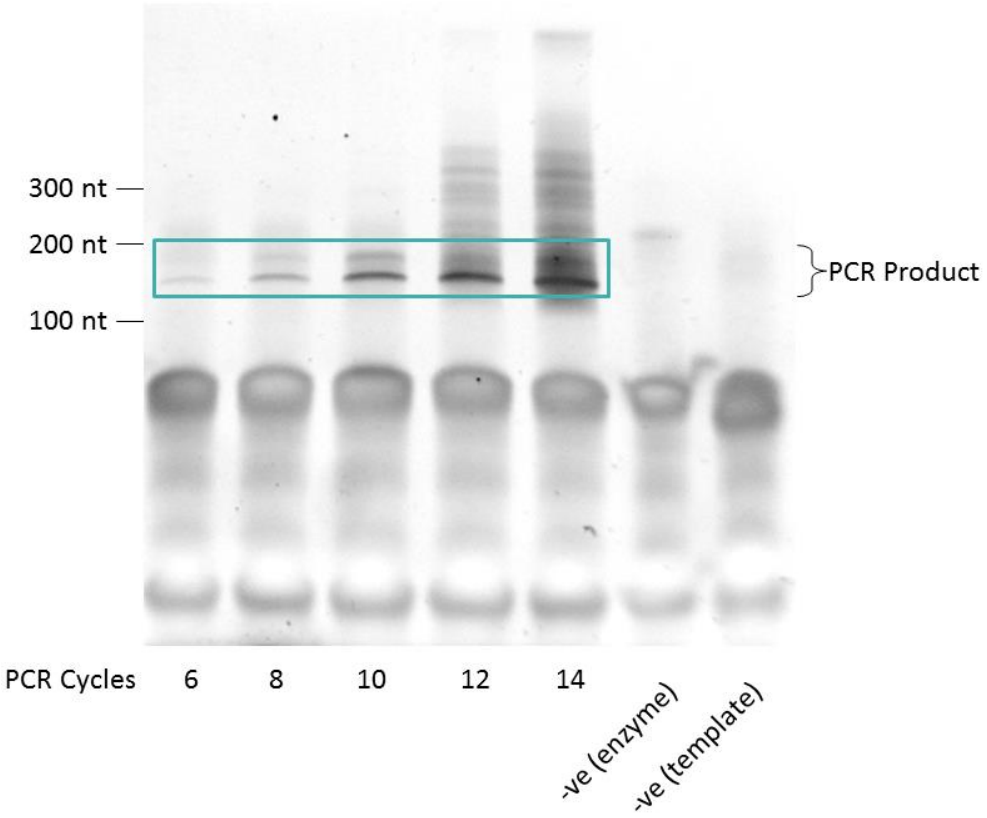

Figure 29 PCR amplification of the circularised RT product derived from HT 1080 cells was carried out as described in chapter 3.2.7 for 6, 8, 10, 12, and 14 cycles. In addition, two negative controls (no enzyme and no template) were incubated for 14 cycles and all samples were loaded onto a 10 % TBE gel without previous purification. Electrophoresis was performed at 200 V and the gel was stained with SYBR Gold nucleic acid gel stain. The expected PCR product had a size of 176 nt (blue square).

After carefully excising the PCR product and extracting it overnight as described in chapter 4.1.2, the now completed cDNA library could be subjected to deep sequencing, which was not carried out in this work due to financial and time reasons.

4.9 Bioanalyzer Assay

Deep sequencing of a cDNA library generated from RPFs requires a certain amount of DNA (approximately 4 µg), which has to be available in a very pure form. For this reason, the obtained sample needs to be subjected to a Bioanalyzer assay that measures the concentration of the 176 nt amplification product and the extent of contamination with unextended RT primer, that appears as a peak at approximately 144 nt.

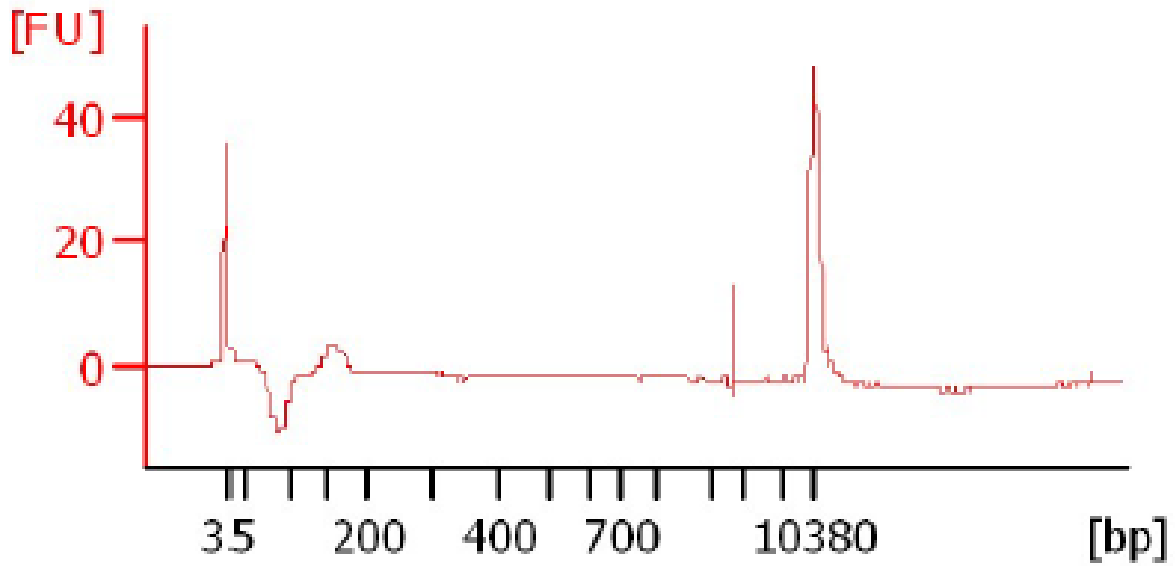

Figure 30 Ribosome profiling was carried out as described in chapter 4.7 and an excised PCR product was subjected to a high sensitivity DNA assay on the Bioanalyzer. A peak was expected at a size of 176 bp.

Samples from a different experiment that is not shown in this report, but that yielded the same result, were analysed with a high sensitivity DNA assay on the Bioanalyzer; the obtained electropherogram is shown in Figure 30. As expected, the electropherogram showed a peak at 176 nt, but also at 144 nt, which was most probably derived from the unextended RT primer. The prominent peaks at < 35 nt and 10380 nt are derived from a lower and a higher marker peak and appear in every electropherogram.

This kind of contamination represents the biggest problem and it is difficult to avoid. As can be seen in Figure 23, the 126 nt long reverse transcription product migrates only marginally higher in an 18 % TBE-Urea gel than does the unextended RT primer and the bands can even partially merge with each other. As a result, small amounts of the RT primer are extracted from the gel together with the product, that show up again as a band of 144 nt after PCR amplification. Hence, excision of the gel slice has to be carried out very carefully in order to avoid contamination. Another helpful measure is to decrease the amount of RT primer during the reverse transcription, as it is provided in excess with the protocol as it is now, although less amounts are probably sufficient.

Apart from the purity of the PCR product, the Bioanalyzer gives an accurate concentration measurement. For deep sequencing, approximately 4 µg of DNA have to be provided. In this case, only less than 1 µg could be generated, although rRNA depletion had not been carried out in this experiments, meaning that the concentration of RPFs must be even lower.

The ribosome profiling protocol contains many steps in which RNA or DNA is precipitated or extracted from a gel. As could be shown in chapter 4.1 these steps are very susceptible to sample loss, resulting in a low product yield, which apparently cannot be compensated by the PCR reaction at the end of the protocol. For this reason, it is recommended to increase the amount of starting material substantially by either setting up more dishes for in-dish lysis of cells, or as described in chapter 4.2 by performing the lysis on a cell pellet, rather than in the dish, as this allows to use more cells, that grew for example in a flask with multiple layers. Furthermore, PCR reactions at the end of the protocol can be carried out more specifically for 8 and 10 cycles and more than one product band can be excised and und processed further.

5. Future Prospects

The ribosome profiling technique is a very interesting and powerful method that generates valuable information about protein synthesis that cannot be obtained by classical approaches that only investigate either the transcriptome or the proteome (Ingolia et al. 2009). By further developing the method into selective ribosome profiling and publishing their protocol, Becker et al. (2013) equipped researches in the fields of biochemistry with another valuable tool that allows to investigate co-translational events.

With the help of selective ribosome profiling, we hope to identify nascent polypeptides that are bound by different PDIs and the exact point where they engage in co-translational folding. Ideally, the interactome of each PDI will be revealed using this technique. The *E. coli* chaperone trigger factor (TF) was the first folding factor to be characterised by this method (Oh et al. 2011) and the approach was successfully modified for other factors that take part in co-translational folding events in prokaryotes (Becker et al. 2013). The protocol provided by Becker et al. (2013) is specific for selective ribosome profiling in bacteria, which requires a few changes to adapt it to experimenting with mammalian cell cultures.

Figure 31 gives an overview of the selective ribosome profiling procedure, which differs from conventional ribosome profiling in certain stages between cell growth and generation of the cDNA library. One main difference is the crosslinking of the folding factor, in this case PDI, with the ribosome-nascent chain complex. This step can either be carried out *in vivo* before cell harvest and lysis as is shown in Figure 31 or *ex vivo* after cell lysis. Both crosslinking methods yielded comparable results for Becker et al. (2013), however, they reveal some disadvantages of the *in vivo* alternative, for instance, crosslinkers that penetrate living cells can stress the cells and thus have a potential influence on the translatome. For this reason, cells need to be pretreated with a reagent that freezes the translational status.

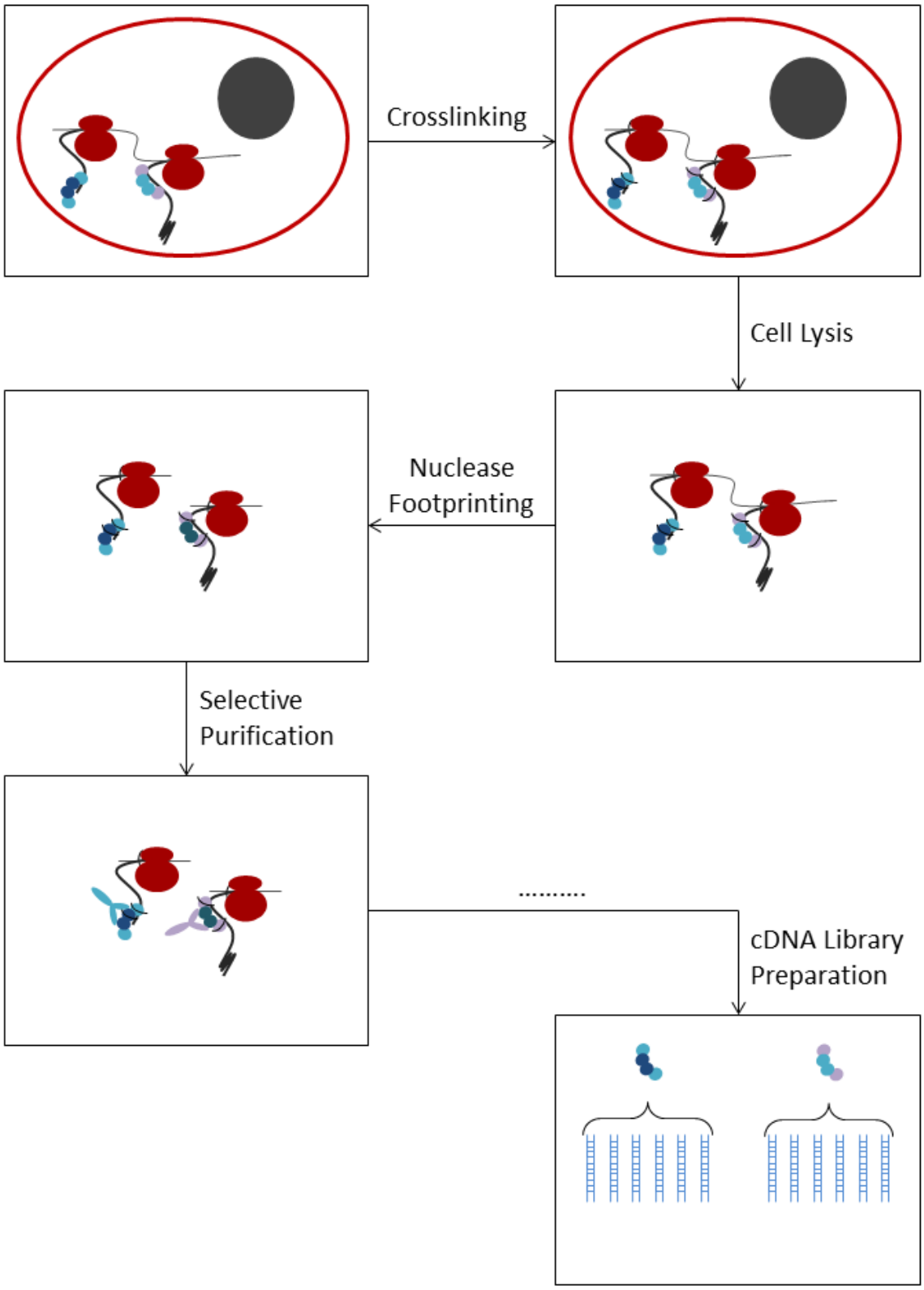

Figure 31 Overview of the selective ribosome profiling workflow. Based on (Becker et al. 2013)

For bacteria, the antibiotic chloramphenicol can be used as it inhibits the ribosome's peptidyl-transferase activity by binding to the 50S subunit of the ribosome (Goldberg 1965), for mammalian cells, a different drug would be required. Additional problems that may occur during *in vivo* crosslinking include interactions of the crosslinker with amino acids in the growth medium, which is why cell growth has to take place in minimal medium prior to the reaction; furthermore, the number of crosslinkers that can penetrate the cell membrane is limited and higher concentrations of the reagent are needed (Becker et al. 2013). Different from the schematic in Figure 31, crosslinking can also be carried out *ex vivo* after cell lysis, which omits the problems that may arise during *in vivo* crosslinking, but comes with the risk of denaturing complexes during cell lysis. Two different crosslinkers were previously used to stabilise the interaction between TF and the nascent chain (Oh et al. 2011; Becker et al. 2013). Dithiobis [succinimidyl propionate] (DSP) induces a covalent bond between primary amines, for instance between a lysine and the N-terminus of the nascent polypeptide; adding a reducing agent reverses the crosslink, which can be exploited during the subsequent purification. Another crosslinker that proved to yield good results is 1-ethyl-3-[3-dimethylaminopropyl] (ESP), which connects primary amines and carboxyl groups, for example a lysine residue with aspartic acid or glutamic acid (Becker et al. 2013). In the case of PDI it may be beneficial to use a crosslinker that interacts with cysteine residues, such as bismaleimidohexane (BMH), dibromoacetone, 1,5-difluoro-2,4-dinitrobenzene (DFDNB) or 1,4-di-(3',2'-pyridyldithio)propionamido butane (DPDPB) (Mattson et al. 1993). The crosslinking reaction with one of these reagents may be more specific than with ESP or DSP, however, they don't have the advantage of being thoroughly tested and validated yet and since PDI-polypeptide complexes are selectively purified in a later step, such highly specific crosslinking might not be necessary.

After crosslinking and cell lysis, regardless in which order, nuclease footprinting is performed and monosomes are purified according to the ribosome profiling protocol, using a sucrose cushion centrifugation (see chapters 3.2.2 and 4.3). Different from the original protocol, Becker et al. (2013) recommend the addition of up to 1 M NaCl to the sucrose cushion, which breaks up non-covalent complexes that may have formed after the crosslinking. Unfortunately, they also observed that high salt concentrations can lead to loss of data especially close to the start codon (Becker et al. 2013), which is why salt has to be added carefully and only if considered absolutely necessary. The last step that distinguishes selective ribosome profiling from the original approach is the selective purification of crosslinked factor-nascent chain complexes. Figure 31 implies the purification by immunoprecipitation (IP), using antibodies that specifically recognise the factors under examination. With the help of protein A-Sepharose beads it is then possible to pull down the complexes bound by the antibody. IP is a very simple and still effective method, provided an antibody is available that binds the factor while it is linked to the nascent chain. Another possibility is to pull down the complexes by means of affinity purification (AP), which is much more specific than IP but also more challenging. For this approach, the factor needs to be expressed as a fusionprotein together with an affinity tag, however, without influencing the factor's function or its interaction with the nascent chain (Becker et al. 2013). In the simplest case, this could be a polyhistidine-tag, which allows the purification by so called immobilised metal ion affinity chromatography (IMAC). Becker et al. (2013), however, advise against this possibility, as the positively charged tag interacts with the negatively charged surface of the ribosome. Instead, they used an AviTag for the pulldown of TF, which binds very strongly to streptavidin and can therefore withstand even sturdy washing, which is beneficial for the purification.

The following preparation of the cDNA footpring library does not differ from the protocol that was discussed in this work.

6. Summary and Conclusions

Ribosome profiling is a powerful technique that, although very young, has already helped addressing many interesting questions that could not be answered by means of classical molecular biology methods. So far, changes in gene expression and protein synthesis could be monitored for example by real-time PCR (Nolan et al. 2006), microarrays (Brown & Botstein 1999), 2D electrophoresis or mass spectrometry (Gupta et al. 2007), but each of these techniques comes with its own limitations and there is no satisfactory correlation between mRNA and protein abundance. Thanks to the development of ribosome profiling in 2009 (Ingolia et al.), these limitations could be circumvented for the first time and new perspectives were opened up. Since then, many laboratories have established the technique and generated interesting results (e.g. Gerashchenko et al. 2012; Han et al. 2012; Liu et al. 2013).

The aim of this work was to develop our own ribosome profiling workflow based on the Nature protocol published by Ingolia et al. in 2012. This required a thorough understanding of every step of the procedure, including the mode of action of each reaction carried out in the process and the respective enzymes and reagents taking part in it. The workflow can be divided into two parts, one being the preparation of RPFs and the second being the generation of a cDNA footprint library suitable for deep sequencing. While the first part could easily be carried out using real samples from cell lysates, the second part, starting with the ligation of a miRNA cloning linker to the obtained RPF was decided to be simulated with an adequate substitute consisting of a synthetic 30 nt long oligonucleotide. For reasons that could not entirely be resolved, however, the use of a synthetic RNA molecule proved to be problematic, apparently due to issues occurring during dephosphorylation. For this reason, a RNA transcript was produced *in vitro* that should mimic a natural oligonucleotide and the linker ligation and subsequent procedures could successfully be carried out with this control molecule. An exception was the PCR reaction at the end of the workflow, which was too

elaborate to be optimised for this kind of molecule that was much longer and more concentrated than RPF samples. Finally, the complete workflow from cell lysis to PCR amplification could be successfully performed, and Bioanalyzer analysis confirmed the correct length of the PCR product. However, concentration and purity were still not satisfactory for deep sequencing which requires approximately 4 µg of the extremely pure PCR product. Therefore, more work has to be done to increase product yield and purity, which can be achieved by significantly increasing the volume of starting material and improved reverse transcription.

A general problem that occurred throughout the experimental procedure was that repetition of experiments was extremely laborious. Each step only yielded a very small amount of reaction product that had to be used completely for the subsequent step, in order to obtain a visible band after purification via gel electrophoresis. In conclusion, this meant that whenever an experiment had to be repeated, the complete, or at least a large part of the previous experiments had to be conducted once more as well.

Despite these difficulties and minor incompleteness, the project aim is considered to be reached. A detailed protocol for ribosome profiling was created that has already been used by other members of the laboratory, proving its practicality and the research group is one step closer to its main goal, namely the characterisation of the members of the PDI family and their role in co-translational folding, which will hopefully be achieved with the help of selective ribosome profiling.

References

Arava, Y. et al., 2005. Dissecting eukaryotic translation and its control by ribosome density mapping. *Nucleic acids research*, 33(8), pp.2421–32.

Arava, Y. et al., 2003. Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), pp.3889–94.

Beaucage, S.L. & Iyer, R.P., 1992. Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron*, 48(12), pp.2223–2311.

Becker, A.H. et al., 2013. Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. *Nature protocols*, 8(11), pp.2212–39.

Brown, P.O. & Botstein, D., 1999. Exploring the new world of the genome with DNA microarrays. *Nature genetics*, 21(1 Suppl), pp.33–7.

Chester, N. & Marshak, D.R., 1993. Dimethyl Sulfoxide-Mediated Primer Tm Reduction: A Method for Analyzing the Role of Renaturation Temperature in the Polymerase CHain Reaction. *Analytical biochemistry*, 209, pp.284–290.

Costa, E.A., Oh, E. & Weissman, J.S., 2014. Global analysis of co-translational N-linked glycosylation in the endoplasmatic reticulum.

DelCardayré, S.B. & Raines, R.T., 1995. The Extend to which Ribonucleases Cleave Ribonucleic Acid. *Analytical biochemistry*, 225, pp.176–178.

Epicentre, 2013. *ARTseq ™ Ribosome Profiling Kit (Mammalian) Catalog Number: RPHMR12126*,

Galligan, J.J. & Petersen, D.R., 2012. The human protein disulfide isomerase gene family. *Human genomics*, 6(1), pp.1–15.

GE Healthcare, 2006. illustra MicroSpin Columns.

Gerashchenko, M. V, Lobanov, A. V & Gladyshev, V.N., 2012. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proceedings of the National Academy of Sciences of the United States of America*, 109(43), pp.17394–9.

Godchaux, W., Adamson, S.D. & Herbert, E., 1967. Effects of Cycloheximide on Polyribosome Function in Reticulocytes. *Journal of Molecular Biology*, 27, pp.57–72.

Gogol-Döring, A. & Chen, W., 2012. An Overview of the Analysis of Next Generation Sequencing Data. In J. Wang, A. C. Tan, & T. Tian, eds. *Next Generation Microarray Bioinformatics*. Humana Press, pp. 247–257.

Goldberg, I.H., 1965. Mode of Action of Antibiotics - Drugs Affecting Nucleic Acid and Protein Synthesis. *American Journal of Medicine*, 39, pp.722–752.

Grada, A. & Weinbrecht, K., 2013. Next-generation sequencing: methodology and application. *The Journal of investigative dermatology*, 133(8), p.e11.

Gupta, N. et al., 2007. Whole proteome analysis of post-translational modifications : Applications of mass-spectrometry for proteogenomic annotation. *Genome Research*, 17(9), pp.1362–1377.

Han, Y. et al., 2012. Monitoring cotranslational protein folding in mammalian cells at codon resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 109(31), pp.12467–72.

Hinkson, I. V & Elias, J.E., 2011. The dynamic state of protein turnover: It's about time. *Trends in cell biology*, 21(5), pp.293–303.

Ho, C.K. et al., 2004. Structure and mechanism of RNA ligase. *Structure (London, England : 1993)*, 12(2), pp.327–39.

Hubbard, T., 2005. *The Human Genome Project. Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics.*,

Hubert, R., 2006. Die Struktur und Funktion Biologischer Makromoleküle. In N. Campbell & J. Reece, eds. *Biologie*. München: Markl, Jürgen, pp. 75–101.

Ingolia, N.T. et al., 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science (New York, N.Y.)*, 324(5924), pp.218–23.

Ingolia, N.T., 2010. *Genome-wide translational profiling by ribosome footprinting.* 2nd ed., Elsevier Inc.

Ingolia, N.T., 2014. Ribosome profiling: new views of translation, from single codons to genome scale. *Nature reviews. Genetics*, 15(3), pp.205–13.

Ingolia, N.T. et al., 2012. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature protocols*, 7(8), pp.1534–50.

Ingolia, N.T., Lareau, L.F. & Weissman, J.S., 2011. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity of Mammalian Proteomes. *Cell*, 147(4), pp.789–802.

Ledergerber, C. & Dessimoz, C., 2011. Base-calling for next-generation sequencing platforms. *Briefings in bioinformatics*, 12(5), pp.489–97.

Li, G.-W., Oh, E. & Weissman, J.S., 2012. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature*, 484(7395), pp.538–41.

LifeTechnologies, SuperScript® III Reverse Transcriptase: Evolution of the SuperScript® Reverse Transcriptases. Available at: https://www.lifetechnologies.com/uk/en/home/life-science/pcr/reverse-transcription/reverse-transcriptase-enzymes/superscript-iii-reverse-transcriptase.html [Accessed September 10, 2014].

Liu, B., Han, Y. & Qian, S.-B., 2013. Co-Translational Response to Proteotoxic Stress by Elongation Pausing of Ribosomes. *Molecular Cell*, 49(3), pp.453–463.

Lodish, H. et al., 2008. Stepwise Synthesis of Proteins on Ribosomes. In *Molecular Cell Biology*. New York: W.H. Freeman and Company, pp. 132–138.

Mattson, G. et al., 1993. A practical approach to crosslinking. *Molecular Biology Reports*, 17, pp.167–183.

Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nature reviews. Genetics*, 11(1), pp.31–46.

Michel, A.M. & Baranov, P. V, 2013. Ribosome profiling: a Hi-Def monitor for protein synthesis at the genome-wide scale. *Wiley interdisciplinary reviews. RNA*, 4(5), pp.473–90.

Miyoshi, D. & Sugimoto, N., 2008. Molecular crowding effects on structure and stability of DNA. *Biochimie*, 90(7), pp.1040–51.

Mossuto, M.F., 2013. Disulfide bonding in neurodegenerative misfolding diseases. *International journal of cell biology*, 2013(Table 1), p.318319.

Nicholson, W., 1997. Escherichia coli Ribonucleass: Paradigms for Understanding Cellular RNA Metabolism and Regulation. In G. D'Alessio & J. F. Riordan, eds. *Ribonucleases – Structures and Functions*. New York: Academic Press, pp. 1–49.

Nolan, T., Hands, R.E. & Bustin, S. a, 2006. Quantification of mRNA using real-time RT-PCR. *Nature protocols*, 1(3), pp.1559–82.

O'Connor, P.B.F. et al., 2013. rRNA:mRNA pairing alters the length and the symmetry of mRNA-protected fragments in ribosome profiling experiments. *Bioinformatics (Oxford, England)*, 29(12), pp.1488–91.

Oh, E. et al., 2011. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell*, 147(6), pp.1295–308.

Oka, O.B. V & Bulleid, N.J., 2013. Forming disulfides in the endoplasmic reticulum. *Biochimica et biophysica acta*, 1833(11), pp.2425–9.

Rasheed, S. et al., 1974. Characterization of a Newly Derived Human Sarcoma Cell Line (HT-1080). *Cancer*, 33, pp.1027–1033.

Ribrioux, S. et al., 2008. Bioinformatics prediction of overlapping frameshifted translation products in mammalian transcripts. *BMC genomics*, 9, p.122.

Richardson, C.C., 1981. *The Enzymes - Bacteriophage T4 Polynucleotide Kinase* 3rd ed. P. Boyer, ed., London: Academic Press.

Sagliocco, F. a, Moore, P. a & Brown, a J., 1996. Polysome analysis. In *Methods in molecular biology (Clifton, N.J.)*. pp. 297–311.

Schneider-Poetsch, T. et al., 2010. Inhibition of Eukaryotic Translation Elongation by Cycloheximide and Lactimidomycin. *Nature Chemical Biology*, 6(3), pp.209–217.

Sousa, R. & Mukherjee, S., 2003. T7 RNA Polymerase I . Introduction II . Structures and Structure – Function Relationships. *Progress in Nucleic Acid Research and Molecular Biology*, 73, pp.1–41.

Takanami, M., Yan, Y. & Jukes, T.H., 1965. Studies on the Site of Ribosomal Binding of f2 Bacteriophage RNA. *Journal of Molecular Biology*, 12(3), pp.761–773.

Thermo Scientific, Assessment of Nucleic Acid Purity.

Vilela, C. & McCarthy, J.E.G., 2003. Regulation of fungal gene expression via short open reading frames in the mRNA 5′untranslated region. *Molecular Microbiology*, 49(4), pp.859–867.

Viollet, S. et al., 2011. T4 RNA ligase 2 truncated active site mutants: improved tools for RNA analysis. *BMC biotechnology*, 11(1), p.72.

Zong, Q. et al., 1999. Messenger RNA translation state: the second dimension of high-throughput expression screening. *Proceedings of the National Academy of Sciences of the United States of America*, 96(19), pp.10632–6.