

A rapid protocol for ribosome profiling of low input samples

Andreas Meindl¹, Markus Romberger¹, Gerhard Lehmann², Norbert Eichner², Leon Kleemann³, Jie Wu³, Johannes Danner², Maria Boesl¹, Mikhail Mesitov⁴, Gunter Meister², Julian König⁴, Sebastian Andreas Leidel³ and Jan Medenbach^{1,*}

¹Regensburg Center for Biochemistry, University of Regensburg, Regensburg, Germany, ²Biochemistry I, University of Regensburg, Regensburg, Germany, ³Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Bern, Switzerland and ⁴Institute of Molecular Biology (IMB), Mainz, Germany

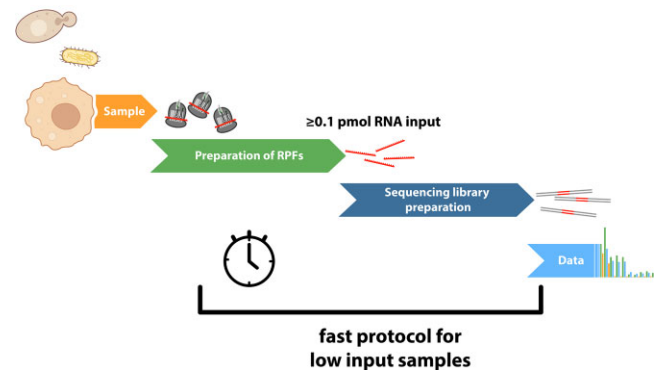
Received October 28, 2022; Revised May 04, 2023; Editorial Decision May 08, 2023; Accepted May 13, 2023

ABSTRACT

Ribosome profiling provides quantitative, comprehensive, and high-resolution snapshots of cellular translation by the high-throughput sequencing of short mRNA fragments that are protected by ribosomes from nucleolytic digestion. While the overall principle is simple, the workflow of ribosome profiling experiments is complex and challenging, and typically requires large amounts of sample, limiting its broad applicability. Here, we present a new protocol for ultra-rapid ribosome profiling from low-input samples. It features a robust strategy for sequencing library preparation within one day that employs solid phase purification of reaction intermediates, allowing to reduce the input to as little as 0.1 pmol of ~30 nt RNA fragments. Hence, it is particularly suited for the analyses of small samples or targeted ribosome profiling. Its high sensitivity and its ease of implementation will foster the generation of higher quality data from small samples, which opens new opportunities in applying ribosome profiling.

GRAPHICAL ABSTRACT

Ribosome Profiling



INTRODUCTION

Ribosome profiling generates genome-wide snapshots of cellular translation based on the analysis of short mRNA fragments that are protected by ribosomes from nucleolytic digestion. These ribosome footprints or ribosome-protected fragments (RPFs) yield the precise position of translating ribosomes on mRNAs at the moment of cell lysis (1) and their highly parallel sequencing yields comprehensive and quantitative insight into translation dynamics (2).

Ribosome profiling has significantly advanced our understanding of the plasticity and regulation of translation during development or in response to stimulation. It has provided evidence for pervasive translation of RNAs and revealed translation outside previously annotated protein-coding regions, thus spurring the discovery of novel peptides. Moreover, its high resolution has unveiled codon-dependent differences in elongation rates and ribosome collisions (3–5) and selective ribosome profiling has provided deep and novel insights into the mechanism of translation (6). For the latter, select sub-populations of ribosomes

*To whom correspondence should be addressed. Tel: +49 941 943 1721; Fax: +49 941 943 2936; Email: Jan.Medenbach@ur.de

are purified prior to the analyses of the associated footprints; these include e.g. initiating ribosomes (7–9) or ribosomes associated with specific proteins such as cotranslational chaperones (10,11).

Numerous variations have been introduced to the protocol, adapting it to a wide range of model organisms and tailoring it to address different aspects of translation and its regulation. This includes the use of different nucleases (12) or translation inhibitors to stall ribosomes on initiation codons (Harringtonine or Lactimidomycin) or at specific stages during the translation cycle (e.g. Cycloheximide, Anisomycin, Emetine, Chloramphenicol or Tigecyclin) (13–15). Furthermore, different methodological approaches have been employed for the preparation of ribosomal monosomes that contain ribosome-protected fragments. This includes enrichment of monosomes by sucrose density gradient centrifugation (16), via sucrose cushions (17,18), or by size exclusion chromatography using spin columns (ARTSeq protocol, Illumina), but also the affinity purification of elongating ribosomes via puromycin derivatives that are integrated into the nascent peptide chain (19).

Even though the original ribosome profiling protocol produces datasets of excellent quality, it is complex and challenging to perform and therefore numerous changes have been introduced over the years (see 20 for a recent version of the protocol). For the generation of high-quality datasets, a careful fine-tuning of the nucleolytic digestion, the purification of ribosomal complexes, and the precise size selection of RNA for library preparation are critical. To date, numerous ribosome profiling protocols exist that vary in experimental details, and, as many of the experimental steps have not been standardized, a comparison of the data generated by the individual workflows is difficult.

Moreover, in the case of low-input samples, the efficiency of library preparation is often becoming limiting. This is particularly the case for selective ribosome profiling where specific ribosomal sub-populations are analyzed. Similarly, small tissue samples (e.g. from clinical biopsies) have proven difficult to profile. To overcome these challenges, a standardized ribosome profiling protocol with a sensitive and efficient sequencing library preparation is highly desirable.

Here we present a novel library preparation protocol for ribosome profiling that allows the robust generation of high-quality libraries from as little as 0.1 pmol of RPFs. The simplified workflow produces high-quality libraries at little effort and within 24 h (Figure 1). This is achieved by combining the enzymatic steps of the original protocol with recently introduced innovations from different sequencing library preparation protocols into a streamlined and optimized workflow: (i) solid-phase reversible immobilization (SPRI) offers convenient and rapid purification of reaction intermediates while minimizing loss of material; (ii) an optimized strategy for PCR amplification prevents over-amplification to ensure the high quality of the final libraries; (iii) highly efficient ligation of adapters and newly designed primers enhance the sensitivity and facilitate experimental multiplexing (Figure 2); (iv) degenerate nucleotides in the ligation adapters reduce ligation bias and allow the identification of PCR duplicates during data analysis (21). In summary, the presented protocol is simple to perform,

cost efficient, more sensitive, and faster than the original protocol.

MATERIALS AND METHODS

Recombinant protein production

The open reading frame encoding TS2126 Ligase with a C-terminal His-tag was synthesized [Geneart] and cloned into pET22B using the NdeI and EcoRI restriction sites. Expression and purification occurred according to Blondal et al. (22). To produce T4 RNA ligase 2 the sequence encoding amino acids 1–294 was amplified by PCR from *Escherichia coli* and cloned into pET16B using the NdeI and BamHI restriction sites. Subsequently, the point mutations R55K and K227Q were introduced by PCR mutagenesis. Expression and purification followed the protocol by Ho et al. (23). TS2126 Ligase can be replaced by the commercially available CircLigase II (Lucigen) without affecting performance of the protocol. Similarly, T4 RNA ligase 2 (truncated, KQ mutant) can be substituted by the enzyme available from New England Biolabs.

Preparation of ribosome-protected fragments from *H. sapiens* cells

Human K562 cells were grown in DMEM supplemented with 10 % FCS at 37°C and 5% CO₂ to a maximal density of 1–1.25 million cells per ml. K562 cells were treated for 3 min with 100 µg/ml Cycloheximide and subsequently harvested by centrifugation for 2 min at 2000 × g and 4°C. The cell pellet was resuspended in 300 µl lysis buffer (5 mM Tris/Cl pH7.4, 1.5 mM KCl, 5 mM MgCl₂, 0.5% Triton X-100, 0.5% Sodiumdeoxycholate, 1 mM DTT, 100 µg/µl Cycloheximide, 10 U/ml Turbo DNase [ThermoFisher], 100 µM RNaseOUT [ThermoFisher], 1 × cOmplete protease inhibitor cocktail [Roche]) and incubated on ice for 15 min. After centrifugation at 14 000 × g for 10 min at 4°C, the cell lysate was transferred into a new reaction tube and UV absorption was determined at 260 nm in a spectrophotometer. In parallel, the RNA concentration of the lysate was determined using the Qubit RNA BR Assay-Kit and a Qubit fluorometer [ThermoFisher]. For nucleolytic digestion, 8 U of RNase I [ThermoFisher] were used per AU₂₆₀ of lysate and digestion was performed for 15 min at 20°C under slow rotation. The reaction was stopped by the addition of 100 U of SUPERase•In RNase Inhibitor [ThermoFisher].

Density gradient ultracentrifugation

After nucleolytic digestion, the samples were subjected to density gradient ultracentrifugation on 10–50% linear sucrose gradients in a buffer containing 20 mM Tris/Cl pH7.4, 150 mM NaCl, 5 mM MgCl₂, 1 mM DTT, and 100 µg/ml Cycloheximide. Centrifugation occurred in either a SW40Ti rotor (3 h at 35 000 rpm (155k × g) at 4°C), a SW55Ti rotor (1.5 h at 50 000 rpm (237k × g) at 4°C), or a SW60Ti rotor (70 min at 50 000 rpm (257k × g) at 4°C) [all rotors from Beckman]. The gradients were fractionated into 1 ml (SW40 rotor) or 400 µl fractions (SW55 and SW60

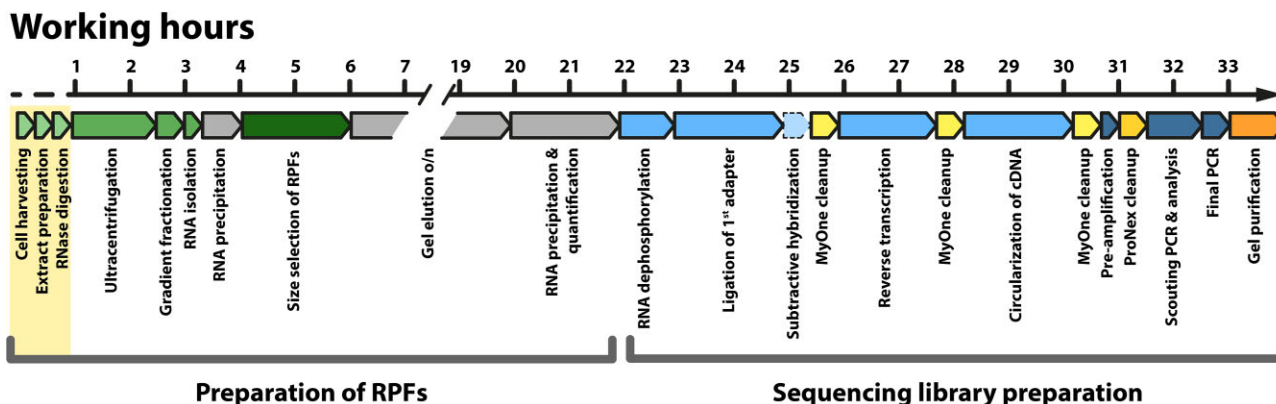


Figure 1. Overview of the workflow. As in the original protocol, ribosome protected fragments (RPFs) are produced by nucleolytic treatment of native cell extracts and purified by sucrose gradient density centrifugation and size selection. For sequencing library preparation, terminal phosphates are enzymatically removed from the RNA, followed by ligation of an adapter to the 3' end to support reverse transcription. After circularization of the cDNA, the product is amplified by PCR to generate an amplicon for Next Generation Sequencing on an Illumina platform. Purification of reaction intermediates occurs mostly via solid phase extraction with magnetic beads. Optionally, a depletion of rRNA-derived, contaminating sequences by subtractive hybridization can be integrated at a strategically optimized position into the workflow (box with dashed outline). To adapt the protocol to different samples, only the initial steps require optimization (highlighted in yellow). Working hours are given at the top.

rotors), under continuous UV monitoring using a siFractor fractionation unit [siTOOLs Biotech] connected to an ÄKTApurifier FPLC [Cytiva].

For baseline correction, a sliding window approach was employed. From a reference point, starting at the beginning of the elution profile, it searches for the lowest absorption in a defined elution volume (1.5 ml) downstream. If all absorption values within the window are higher than the reference point, the lowest value is used as the next reference point and calculation is repeated at this point. If absorption values lower than the reference point are detected within the search window, the next point for calculation is the one at the steepest angle below the reference point. Finally, all calculated reference points are connected to generate a profile that is subtracted from the data.

Purification of RPFs

The fraction containing the 80S monosomes was diluted with an equal volume of nuclease-free water and extracted with phenol:chloroform:isoamylalcohol (25:24:1). The aqueous phase was transferred to a new reaction tube and RNA was precipitated with 2.5 volumes ethanol in the presence of 3 μ l of linear acrylamide solution [ThermoFisher]. Subsequently, the RNA was separated by denaturing PAGE (15%) and visualized by SYBR Gold [ThermoFisher] staining. RNAs in a size range from 26 to 30 nts were excised from the gel and eluted in a buffer containing 300 mM NaOAc, 1 mM EDTA, 0.25% SDS, and 20 U/ml SUPERase \bullet In RNase Inhibitor. After precipitation, the RNA was resuspended in nuclease free water and the concentration was determined using a Qubit fluorometer with the microRNA Assay Kit [ThermoFisher].

Preparation of ribosome-protected fragments from *S. cerevisiae* BY4247

BY4742 yeast cells were grown at 30°C in YPD medium. Whole cell extracts were prepared from exponentially grow-

ing cultures essentially as described in Schwank et al. 2022 (24). In brief, at an OD₆₀₀ of 0.6–0.7, the cells were harvested by centrifugation and flash frozen in liquid nitrogen. Cell lysis occurred mechanically under cooling with liquid nitrogen (Cryolys Cooling System) using 0.75–1mm glass beads in a Precellys tissue homogenizer at 6000 rpm in a buffer containing 20 mM Tris/Cl pH 7.4, 150 mM NaCl, 5 mM MgCl₂, 1% Triton-X100, 1 mM DTT, 25 U/ml Turbo DNase and 100 μ g/ml Cycloheximide. After clearing the extract by centrifugation at 20 000 \times g for 10 min at 4°C, it was subjected to RNase I digestion for 5 min on ice using 30 U of enzyme per AU₂₆₀.

Subsequently, ribosomal complexes were separated by sucrose density ultracentrifugation as described in (25). In brief, 15–50% linear sucrose gradients were prepared in a buffer containing 20 mM Tris/Cl pH 7.4, 150 mM NaCl, 5 mM MgCl₂, 1 mM DTT and 100 μ g/ml Cycloheximide. Ribosomal complexes were separated by centrifugation in a SW40Ti rotor for 2.5 h at 35 000 rpm at 4°C or in a SW55 rotor for 1.5 h at 50 000 rpm at 4°C. Fractionation of the gradients and purification of the ribosomal footprints occurred as described above.

Sequencing library preparation

Optimized protocol. The indicated amounts of gel-purified RNA were dephosphorylated by treatment with 5 U of T4 PNK [New England Biolabs] in a 20 μ l reaction supplemented with 1x reaction buffer and 20 U of RNaseOUT [ThermoFisher] for 30 min at 37°C. After heat inactivation of the enzyme for 20 min at 65°C, the reaction was supplemented with 4.5 μ l DMSO, 6 μ l 50% PEG400, 1.5 μ l RNA-ligase buffer, 1 μ l T4 RNA Ligase 2 (aa 1–294, R55K, K227Q) and 20 pmol of rApp-L7 linker oligonucleotide. The reaction was incubated for 3h at 25°C, or over night at 16°C followed by an incubation for 1h at 37°C and by heat inactivation of the enzyme for 15 min at 65°C. For depletion of contaminating rRNA-derived sequences, the entire reaction was subjected to subtractive hybridization using a

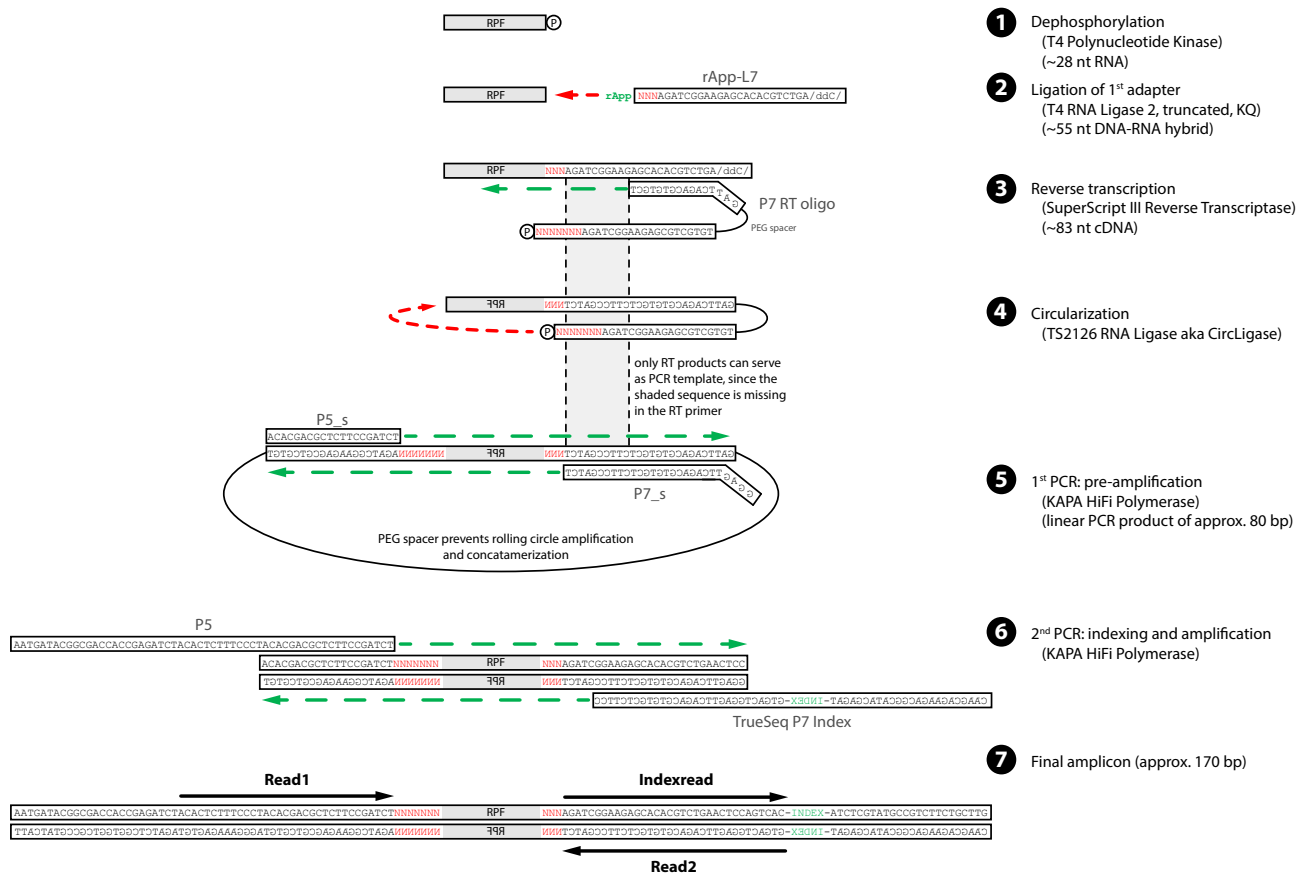


Figure 2. Schematic overview of sequencing library preparation. Individual steps in the protocol are described on the right; identity of the enzymes employed in each step is provided in brackets. Degenerate nucleotides that reduce ligation bias and that are used for the identification of PCR duplicates are highlighted in red, the position of an index for experimental multiplexing in green. Information on the approximate length of the individual products is provided for each step.

riboPOOL kit [siTOOLS Biotech] according to the manufacturer's instructions. Subsequently, the ligation product was extracted using 20 μ l MyONE Silane beads [ThermoFisher] in 650 μ l RLT Buffer [Qiagen] and 720 μ l ethanol. Next, 0.5 pmol P7 RT oligonucleotide were annealed to the ligation product (70°C for 5 min, then cooling to 25°C). Afterwards the reaction was supplemented with 0.5 mM dNTPs, 0.05 mM DTT, 1 \times first strand buffer (ThermoFisher), 100 U Superscript III (ThermoFisher), and 20 U RNaseOUT (ThermoFisher) in a total of 20 μ l. For reverse transcription the reaction was incubated for 5 min at 25°C, 20 min at 42°C and 40 min at 50°C. For hydrolysis of the RNA, the reaction was supplemented with 1.65 μ l of 1M NaOH and incubated for 20 min at 90°C, followed by neutralization with 20 μ l of 1M HEPES/KOH pH7.3. The cDNA was purified with 10 μ l of MyONE Silane beads as described above and eluted in 14 μ l of nuclease-free water. Circularization of the cDNA occurred in a 20 μ l reaction containing 0.5 mM ATP, 2.5 mM MnCl₂, 0.05 mM ATP, 50 mM MOPS (pH 7.5), 10 mM KCl, 1 mM DTT and 1 μ l TS2126 ligase for 2 h at 60°C followed by an incubation for 10 min at 80°C. The circular product was extracted with 10 μ l MyOne Silane beads in 650 μ l buffer RLT and 720 μ l ethanol. After elution it was subjected to a pre amplification by PCR using the P5_s and P7_s primers and KAPA

HiFi HotStart ReadyMix [Roche] according to the manufacturer's instructions. The PCR product was purified using the ProNex size-selective purification system [Promega] with a 1:2.95 v/v ratio of sample to beads. Scouting PCRs and final amplification of the library were performed as described in (26). In brief, for scouting PCRs, 10 μ l reactions were set up using the KAPA HiFi HotStart ReadyMix [Roche] with P5 and TrueSeq P7 indexed primers and subjected to 6–16 cycles of PCR. After analysis of the PCR products either by native PAGE or a high sensitivity screen tape assay on a TapeStation [Agilent], optimal PCR conditions were chosen according to the concentration of the PCR product and the appearance of overamplification artifacts. These 'daisy chains' consist of large assemblies of improperly annealed, partially doublestranded, heteroduplex DNA that typically appear after depletion of the PCR primers in library amplification reactions (compare 27). Final amplification of the library occurred in a 40 μ l reaction using the conditions determined in the scouting runs. Subsequently, for the removal of primers and the depletion of empty amplicons, the PCR reaction was subjected to purification using a Blue Pippin [Sage Science] with 3% agarose cassettes selecting fragments of a length of 168 bp with the setting 'tight'. Alternatively, the final library can be purified via 3% agarose gel electrophoresis or 10% native PAGE

and subsequent elution from the gel. All primer sequences are listed in Supplementary Table S3.

Comparison with the Kim et al. (2020) library preparation protocol. HEK293 cells were harvested as described previously (28), with the following adaptation: the lysis buffer was supplemented with 25 U/ml Turbo DNase [ThermoFisher] before use. Ten A260 units of lysate were digested with 1000 U RNase I [ThermoFisher]. After RNA dephosphorylation by PNK, the samples were split into two tubes to perform the Kim et al. protocol with one half and this new protocol with the other half. Ribosomal RNA was depleted using biotinylated oligonucleotides as described before (17), Supplementary Table S4.

Comparison to the McGlincy and Ingolia library preparation protocol. RPFs were prepared from K562 cells as described above. Different quantities of the same sample (ranging from 0.1 to 5 pmol) were subjected to sequencing library preparation as described in McGlincy and Ingolia (2017) with one exception: instead of the reverse transcription primer NI-802 described in the McGlincy and Ingolia protocol, we used the P7 RT Oligo (see Supplementary Table S3). In parallel, the same samples were used to prepare sequencing libraries with our novel protocol.

iCLIP2 protocol. RPFs were subjected to library preparation as described (26) with the following modification: ligation of the first adapter occurred in solution in a 20 μ l reaction containing $1 \times$ T4 RNA ligase buffer [New England Biolabs], 20 U RNaseOUT [Invitrogen], 10 U T4 RNA ligase 10 [New England Biolabs], 4 μ l PEG400 [Sigma Aldrich]. Subsequently, the ligation product was purified via MyONE Silane beads and the protocol was continued as described in (26).

Sequencing and data analyses

For sequencing, libraries were quality controlled using a high sensitivity screen tape assay on a TapeStation [Agilent] and quantified using the KAPA Library Quant Kit [Roche]. Sequencing occurred on a MiSeq (V3 kit, 150 cycles, 90 cycles single end, 6 nt index read), or on a NextSeq2000 (V3 kit, 50 cycles, 80 cycles single end, 6 nt index read).

For bioinformatic analyses, R version 3.6.3 was used (29). After de-multiplexing of the sequencing reads, adapter trimming occurred with Cutadapt (version 2.8, parameters: adapter = AGATCGGAAGAGCACACGTCT, overlap = 10, minimum-length = 0, discard-untrimmed) (30). UMI tools (version 1.0.1) (31) was used to extract the UMIs and to associate their sequence with the read identifier. Trimmed reads were mapped against hg38 (*H. sapiens*) or R64 (sacCer3, *S. cerevisiae*) using bowtie2 (version 2.4.1) (32) using the standard parameters. PCR duplicates were subsequently identified using UMI tools (parameters: –extract-umi-method read_id –method unique).

Transcript annotation was performed as described in Hasler *et al.* (33). In brief, biotypes were retrieved from ENSEMBL hsa GRCh38.92 (34), the UCSC table browser hg 38 (35, status February 2018), RepeatMasker, NCBI

GRCh38 (status July 2018), and the snoRNA atlas (36, status July 2018). The category “Pol III” was created by combining the relevant annotations from ENSEMBL, RepeatMasker, and UCSC. SnoRNA annotations from snoRNA atlas were converted from h19 to hg38 coordinates and were combined with ENSEMBL snoRNAs. In all cases, overlapping and multiple annotations from these two databases were merged. Several sub-biotypes from the ENSEMBL annotations were grouped into the following categories: ‘Protein coding’, ‘Pseudogene’, ‘Long noncoding’ and ‘RepeatMasker’. To reduce multiple counting of the annotations for small RNAs (i.e. miRNAs, Pol III transcripts, snoRNAs and Pol I rRNAs), they were cut out from the remaining annotations with an additional spacer of ten bases.

Pairwise comparison of sequencing libraries was performed on sequencing reads of 26–30 nt that can be mapped on genes with the ENSEMBL biotype ‘protein coding’. Only genes with at least one read in all experiments were used for the analysis (encompassing approx. 9500 transcripts). Plotted are read counts per gene locus (\log_2) normalized with the function ‘estimateSizeFactorsForMatrix’ of the DESeq2 package.

A-site coverage calculation around initiation and termination codons and periodicity analyses of RPFs were performed as described in (37).

RESULTS

Efficient sequencing library preparation from minute amounts of RNA

We have developed an improved and rapid strategy for the generation of sequencing libraries from minute amounts (≥ 0.1 pmol) of ribosome-protected fragments (Figure 1). The original ribosome profiling protocol employs polyacrylamide gel electrophoresis (PAGE) for the purification of reaction intermediates which is laborious and time consuming (2,16,17). Furthermore, elution of the sample from gel slices typically results in significant sample loss, particularly when working with small amounts of RNA (we typically recover $\leq 30\%$ of a sample of 30 pmol of RNA, data not shown). A more recent protocol has therefore introduced the use of spin columns for the purification of reaction intermediates (20). As an alternative to this, solid-phase reversible immobilization (SPRI) can be employed, e.g. for sequencing library preparation from samples obtained from cross-linking immunoprecipitation experiments (iCLIP2, 26). Extraction with Dynabeads MyOne silane is rapid, robust, and yields a higher sample recovery in comparison to gel purification.

Building on the iCLIP2 library preparation approach, we developed a protocol tailored specifically to the sequencing of small RNAs derived e.g. from ribosome profiling experiments. Besides employing solid-phase extraction of reaction intermediates, we re-designed the adapter sequences, allowing the generation of shorter amplicons with P5 and P7 adapters. These amplicons contain bona fide P5 and P7 sequences and differ slightly from the adapters used in recent ribosome profiling protocols (e.g. 20) (Figure 2). In contrast to iCLIP2-derived libraries and a more recent ribosome profiling protocol (20), for experimental multiplexing, in-line barcoding is replaced by a P7-side encoded

index that is introduced during the final PCR. This reduces primer costs and facilitates multiplexing with libraries derived from different experimental approaches such as e.g. RNA sequencing for expression profiling.

In agreement with previous reports, ligation of the first linker to the RNA fragments by T4 RNA ligase I has proven inefficient in our hands (Supplementary Figure S1A) (23,38). We therefore opted for T4 RNA ligase II (truncated aa 1–294 and carrying the mutations R55K and K227Q). In the presence of PEG400 and DMSO, ligation of 5 pmol of a 30 nt RNA oligonucleotide to a pre-riboadenylated adapter proceeds almost to completion (Supplementary Figure S1A). Moreover, as in the original ribosome profiling protocol, we employ circularization of the cDNA instead of ligation of a second adapter. The efficiency of circularization with TS2126 ligase (aka CirLigase) is effective and comparable in efficiency to ligation of a second adapter with T4 RNA ligase I and avoids known ligation biases (Supplementary Figure S1B and C) (21). Furthermore, the reaction proceeds much faster, requiring only a 2 h incubation instead of an overnight reaction and the enzyme TS2126 works at elevated temperatures, which reduces the impact of local secondary structures on ligation and it also exhibits no significant preference in the positions surrounding the ligation site, reducing bias during sequencing library generation (39). To further minimize biases, the ligation adapters contain degenerate positions at their 5' ends (Figure 2) (40). Simultaneously, these sequences serve as unique molecular identifiers (UMIs) for the identification of PCR duplicates during data analysis after sequencing.

To evaluate the performance of the optimized library preparation protocol, we first compared it to the recently developed iCLIP2 protocol that employs a similar strategy for purification of reaction intermediates during sequencing library preparation. We reasoned that when using small amounts of input material for library preparation, an inefficient strategy will necessitate extensive PCR amplification to yield sufficient material for sequencing. This results in an increase of the fraction of PCR duplicate reads that can be identified during bioinformatic analysis based on mapping position and unique molecular identifier (Supplementary Figure S2A). In contrast, a library generated by an efficient protocol should contain very few PCR duplicates. Legacy sequencing library preparation protocols employed for ribosome profiling typically require 10–20 pmol of RNA input for production of a high-quality library (2,16,17). We therefore subjected 5 pmol of purified RNA (26–30 nt in length, generated by RNase I digestion) to sequencing library preparation using either the iCLIP2 protocol or our optimized library preparation protocol. The libraries were amplified to yield similar amounts of product and sequenced at comparable depth with approx. 20 million reads per sample. Analysis of the sequencing data revealed that the iCLIP2-derived libraries are dominated by PCR duplicates (~70%), while the optimized protocol produces high-quality libraries that contain less than 2% PCR duplicates (Figure 3A).

We next tested the sensitivity of the optimized library preparation protocol by subjecting different amounts of purified ribosome-protected fragments (ranging from 5 to 0.1 pmol) to library preparation. In all cases we obtained se-

quencing libraries and, as expected, low amounts of input material require an increased number of PCR cycles (Figure 3B). After sequencing, analyses of the PCR duplicates in the libraries revealed that even with minute amounts of input material (0.1 pmol), duplicate reads do not exceed 2% when sequencing the libraries at a depth of approx. 5 million reads per sample (Figure 3C). This demonstrates that the protocol is well suited for the generation of sequencing libraries from as little as 0.1 pmol of ~30 nt RNA fragments (corresponding to approx. 1 ng of RNA).

With decreasing amounts of input material, we detect an exponential increase in amplicons that lack an insert. These 'empty' amplicons are generated from adapters (rApp-L7) that fail to ligate to an RNA molecule and which provide a template for the extension of the RT oligonucleotide (P7 RT oligo) during reverse transcription (Figure 2 and Supplementary Figure S2B). After circularization, non-extended RT oligonucleotides are not subject to PCR amplification as the PCR primers amplify a sequence that is generated during reverse transcription. In contrast, the partially extended RT oligos efficiently engage in PCR amplification, resulting in the generation of 'empty' amplicons. Different strategies can be employed for the removal of these unwanted sequences. (i) In the first ligation reaction, the concentration of the rApp-L7 adapter can be adjusted to the amount of input material. A decrease in the concentration of the adapter oligonucleotide, however, can significantly reduce ligation efficiency. (ii) Non-ligated adapters can be degraded enzymatically using a 5' deadenylase and the DNA-specific exonuclease RecJ. While this reaction works efficiently, we reasoned that removal of the adapters at an early stage will result in an increased loss of sample during the workflow due to unspecific adsorption onto the plasticware (the leftover adapters can block surfaces, reducing adsorption of the actual sample). For removal of the empty amplicons, we therefore introduced a final purification step using an automated agarose gel electrophoresis system (Blue Pippin, Sage Science). It allows the efficient depletion of empty amplicons prior to sequencing (Figure 3D & Supplementary Figure S3), decreasing their abundance from ~70% (0.1 pmol input) to approx. 1% ($N = 4$) in the final sequencing library.

Next, we compared library preparation to another state-of-the-art ribosome profiling protocol that employs spin columns for the purification of reaction intermediates (20), hereafter referred to as 'reference protocol'. We subjected different amounts of the same sample (ranging from 5 to 0.1 pmol) to sequencing library preparation using either protocol. After sequencing, samples processed according to the reference protocol showed a shortening of the average insert size, particularly when using low amounts of input (Figure 4A). In contrast, with our newly developed protocol, the insert size does not noticeably change, even when processing low-input samples. Since the same input sample was used for all reactions, the appearance of shortened inserts suggests a partial degradation during sequencing library preparation. For libraries prepared according to the reference protocol, this generally results in a significantly reduced number of sequence reads that can be mapped to mRNAs (Supplementary Figure S4A, $3.3 \pm 1.5\%$ versus $5.2 \pm 0.4\%$, $P = 0.016$, see Supplementary Figure S9 for statistics on PCR duplicates and mRNA reads with our

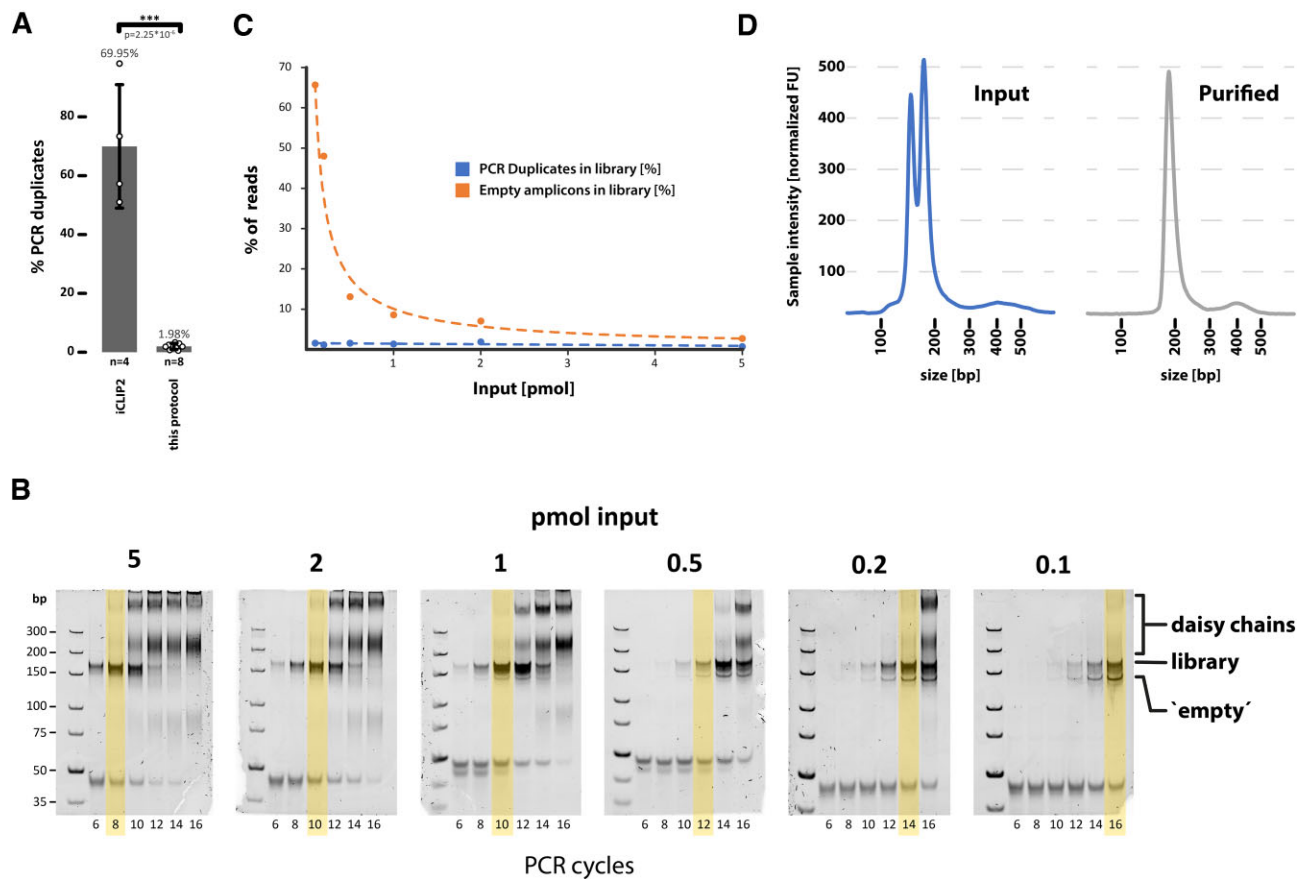


Figure 3. Sequencing library preparation from minute amounts of sample. (A) Analysis of PCR duplicates from sequencing libraries generated from 5 pmol of RNA (26–30 nt in length, generated by RNase I digestion) using the iCLIP2 library preparation protocol (left), or the protocol described here. Number of replicates are provided below each bar, depicted are mean \pm SD. The p -value was calculated using a two-tailed Student's t -test. (B) Results of scouting PCRs of libraries prepared from different amounts of input material (as noted above each gel). Provided below each lane are the number of PCR cycles that were performed. The positions of high molecular weight heteroduplex DNA ('daisy chains', derived from over-amplification of the libraries) and 'empty' amplicons are indicated on the right. Yellow boxes indicate conditions that have been used to generate the final samples for high throughput sequencing. (C) Analysis of PCR duplicates and 'empty' amplicons contained in the libraries depicted in panel B. (D) Depletion of 'empty' amplicons by automated agarose gel electrophoresis. Sample prior (left) and after purification (right).

novel protocol). Notably, the samples with the lowest input, which show signs of degradation, also exhibit the lowest number of reads that can be mapped to mRNAs (Supplementary Tables S1 and S2). Note that the libraries used for this analysis have not been subjected to subtractive hybridization for the removal of contaminating rRNA fragments. As previously described, depletion of rRNA contaminants can significantly increase the fraction of mRNA-derived sequence reads (see e.g. 17,41).

The analyses of ribosome profiling data often focus on a very narrow size range of RNA fragments that are protected by ribosomes from nucleolytic digestion – with cycloheximide-stalled ribosomes typically generating 28–30 nt fragments (42). Limiting our analyses to RNA fragments with a length of 26–30 nt, our protocol yields on average substantially more mRNA reads (Supplementary Tables S1 and S2) than the protocol used for comparison, particularly when using small amounts of input (Supplementary Figure S4B). While our protocol shows a high degree of correlation between experiments that use different amounts of input material, the average Pearson correlation coefficient of

libraries prepared with the reference protocol is significantly lower (0.918 ± 0.036 versus 0.982 ± 0.008 , $P = 4.11 \times 10^{-7}$) (Supplementary Figures S5 and S6). This can be attributed mostly to differences in protocol performance when preparing sequencing libraries from samples containing 0.5 pmol RNA or less.

To test whether our novel library preparation strategy can be easily adopted by additional labs, samples were processed in a non-developer lab. The results after library preparation were comparable between both labs (Figure 3 and Supplementary Figure S7) attesting to the simple implementation of the protocol.

The final data exhibit features of high-quality ribosome profiling datasets. In metagene analyses, data generated by the new protocol is indistinguishable from data that was obtained from the same sample following the protocol by Kim *et al.* (28). Reads are strongly enriched in open reading frames compared to untranslated regions with characteristic peaks at translation initiation and -termination codons (Figure 4B). Moreover, depending on the degree of nucleolytic digestion, a characteristic three-base periodicity for

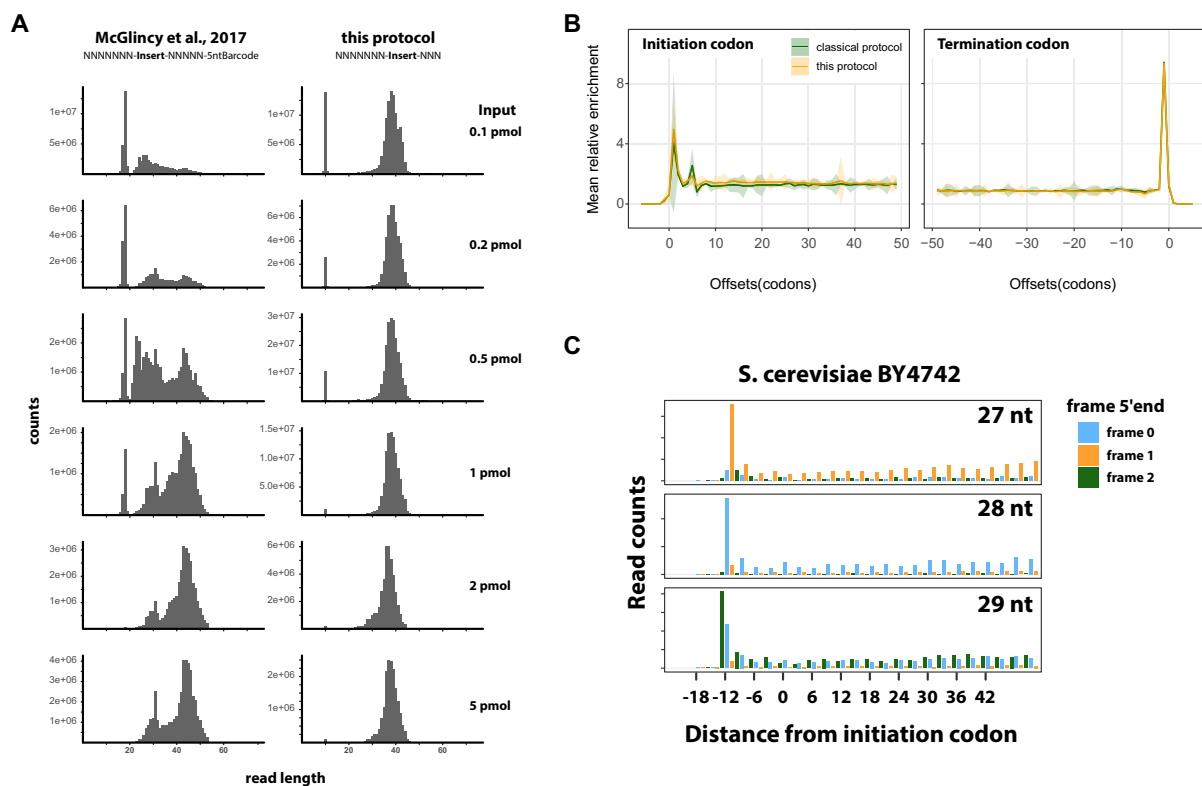


Figure 4. The novel protocol generates high quality ribosome profiling data. (A) Comparison of insert length distribution of sequencing libraries prepared with either the protocol by McGlincy et al. (20, left), or this protocol (right). The different amounts of RNA input used for library preparation are provided on the right. (B) Normalized A-site coverage observed around initiation- (left panel) and termination codons (right panel). Compared are libraries generated with the protocol by Kim et al. 2021 (28, shown in green) versus libraries generated with our novel protocol (shown in yellow). Depicted are mean values (solid line) $\pm 95\%$ confidence interval (shaded area) for three biological replicates ($n = 3$). (C) The 5' ends of ribosome footprints generated with this protocol from *S. cerevisiae* exhibit 3 base periodicity. Footprint lengths of 27, 28 or 29 nt in length (as indicated in each panel) were separately analyzed and read counts for each position relative to the initiation codon are plotted. For simplicity each frame is color coded as shown on the right.

RNA fragments of 27, 28 and 29 nt in length can be observed (Figure 4C), which is characteristic for RPFs (2).

Ribosome profiling from low input samples

The high sensitivity of the newly developed library preparation protocol allowed us to reduce the amount of input material used for the preparation of ribosome-protected fragments. For purification of ribosomal complexes after nucleolytic digestion, we routinely employ sucrose density gradient ultracentrifugation. This allows to simultaneously assess the quality of the sample, the degree of the nucleolytic digestion and the yield. Using continuous UV monitoring during fractionation of the density gradients, we could separate and harvest ribosomal complexes from small amounts of cellular extract (0.625 AU₂₆₀ containing approx. 15 μ g of total RNA as measured by fluorometry) with high accuracy and reproducibility (Figure 5A, Supplementary Figure S8).

With reduction of the input material, we observed a gradual distortion of the UV profile which is caused by differences in UV absorption through the gradient, the buffer used for cell lysis, and components of the cell lysate other than RNA. To enable a better comparability of the gradient profiles, we therefore introduced a baseline correction. Despite the inherent limitations of the mathematical correction, a better visualization of the data, such as the precise

positions of the peaks, can be achieved. It performs reasonably well as determined by the quantification of ribosomal monosomes from a serial dilution (Supplementary Figure S8).

To enable accurate size selection of ribosome protected fragments from denaturing acrylamide gels, we furthermore developed a DNA-based molecular weight marker to guide the precise selection of RNA fragments between 26 to 30 nucleotides in length for preparation of the sequencing library, even if by eye no sample is visible in the 30 nt region of the gel (Figure 5B). It is more cost effective and more stable than RNA-based molecular weight markers that are commonly employed. Moreover, protective groups have been introduced that prevent ligation of the molecular weight marker molecules to the adapter oligonucleotides used for the generation of sequencing libraries. Hence, a potential carryover of the marker does not contribute to the generation of amplicons during sequencing library preparation.

DISCUSSION

Ribosome profiling experiments are tailored to address various research questions and they are typically optimized for different model organisms and sample types. Parameters that vary between experiments typically include the application of translation inhibitors, different protocols for sample

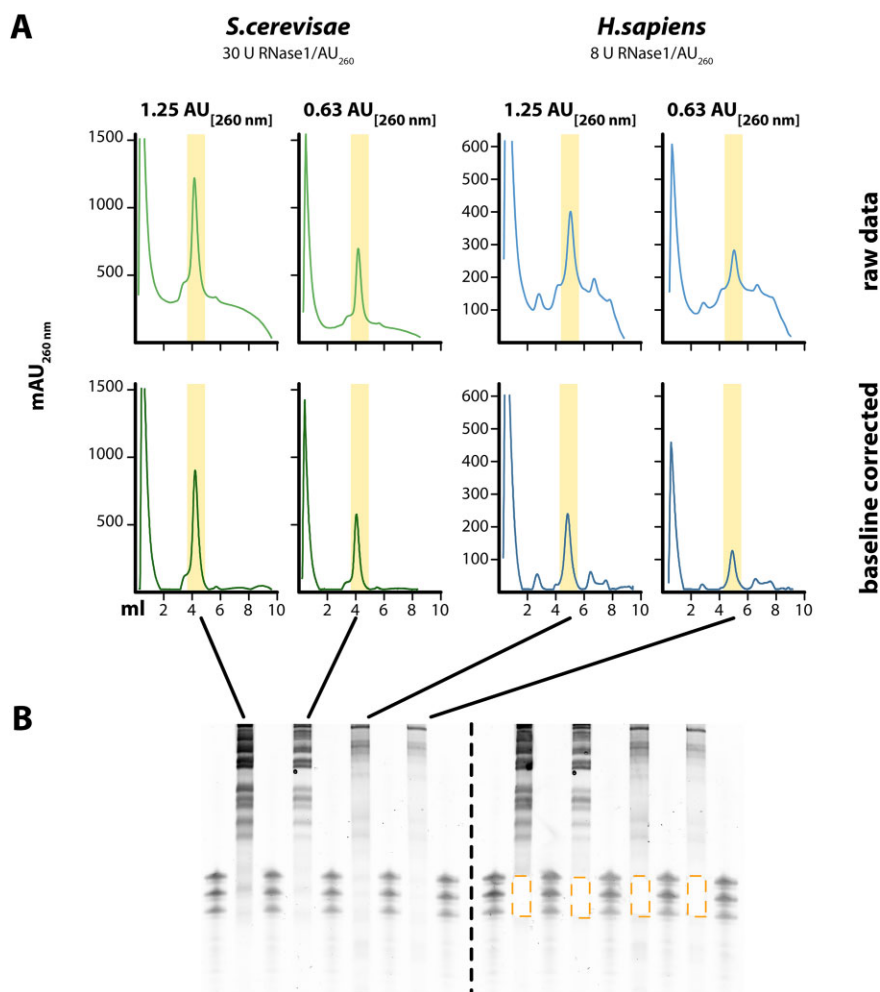


Figure 5. Preparation of RPFs from minute amounts of sample. (A) Sucrose density gradient ultracentrifugation of nuclease digested extracts from *S. cerevisiae* (BY4742) and *H. sapiens* K562 cells. Different amounts of extracts (1.25 AU and 0.63 AU corresponding to ~30 and ~15 μg of RNA) were separated on linear 10–50% sucrose gradients and fractionated while monitoring UV absorption ($\lambda = 260\text{ nm}$). Raw (top) and baseline-corrected (bottom) UV profiles are depicted. Fractions containing 80S ribosomal monosomes (highlighted in yellow) were subjected to denaturing PAGE (B) to purify RNA fragments of 26–30 nt in length. Gel prior to (left) and after excision (right) of the RNA fragments.

harvesting or preparation of native extracts, and the conditions and the choice of the enzyme for limited nucleolytic digestion to generate the ribosome footprints. All subsequent steps, however, can in principle be standardized.

Here, we provide such a standardized protocol for ribosome profiling starting from purification of ribosome protected fragments after nucleolytic digestion to the production of the final sequencing library. It is extremely rapid, cheap, broadly applicable and performs robustly even with minute amounts of input. Density gradient ultracentrifugation and scouting of PCR conditions allow quality control after critical steps to ensure processing of only high-quality samples. Due to its high reproducibility, sequencing libraries generated by this protocol are highly comparable facilitating comparisons between different experiments, which has previously been hampered by the large experimental variation between individual protocols. High sensitivity and precision during preparation of ribosomal monosomes and accurate size selection of RNA fragments contribute to the quality of the experiment and limit contami-

nations derived from ribosomal RNA. However, both steps become challenging when performing ribosome profiling with low input samples that exhibit little UV absorbance. To meet these challenges, we have developed a gradient fractionation device that can be integrated into an FPLC system, enabling detection of UV absorption with high sensitivity and precise fractionation of the sample into small volumes. Furthermore, an improved molecular size marker for gel electrophoresis guides the precise excision of low input RNA fragments with sizes of 26–30 nucleotides that are hardly visible on the gel by eye (even after staining with sensitive dyes such as SYBR-Gold). Protection groups prevent marker-derived nucleotides from contributing to the sequencing library, suppressing contaminations.

Often, ribosome profiling has been hampered by limiting amounts of starting material or low RNA yield after partial nucleolytic digestion. We have developed a high-sensitivity protocol that allows the preparation of high-quality sequencing libraries from low input samples. We observe a high correlation between libraries irrespective of the

amount of input material used. Our protocol also faithfully maintains the fragment length with low sample input and allows precise mapping of the ends of the RPFs to determine the position of translating ribosomes and e.g. their respective A-sites. This type of analysis is hampered when samples become partially degraded during sequencing library preparation, confounding data analyses.

To overcome experimental challenges such as partial degradation of low input samples during library preparation, other revised protocols for ribosome profiling have introduced barcoded adapters for ligation to the RNA fragments. This allows the pooling of samples for subsequent steps and hence increases the amount of starting material for library production, enabling ribosome profiling from samples containing as little as 30 µg of total RNA (20). However, pooling of samples with different concentrations typically results in an uneven distribution of reads during sequencing, confounding the analyses under-represented samples. Hence, for many applications, a protocol for the independent and individual preparation of low input samples can be beneficial. Our protocol provides this advantage and, moreover, it is fully compatible with the previously described barcoded adapters (20) allowing the pooling and parallel processing of multiple samples, if desired.

We typically perform ribosome profiling experiments using cell lysates that contain as little as 0.625 AU₂₆₀ (approx. 15 µg of total RNA as determined by fluorometry). Dependent on the degree of nucleolytic digestion, this yields 0.7–1 pmol of ~30 nt RNA fragments for library preparation. Since our improved library preparation protocol performs well with an input as little as 0.1 pmol of RNA fragments, future advances will enable the processing of even smaller samples. Our protocol will therefore stimulate analyses of translation from samples that previously were not amenable to this type of methodology—e.g. FACS-sorted cells or patient derived samples such as small tissue biopsies, organoids, or cells - as well as low input applications such as selective ribosome profiling that targets sub-populations of ribosomes.

DATA AVAILABILITY

All sequencing data have been deposited at the Sequence Read Archive and are accessible via the following BioProject: PRJNA883055.

SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Joachim Griesenbeck for support with the experiments using *S. cerevisiae* and the EMBL GeneCore facility for support with high throughput sequencing.

FUNDING

German Research Foundation [SFB960 TP B11 to JM and FOR2333 KO4566/5-1 to JK]; German Federal Ministry

of Education and Research (BMBF) within the framework of the e:Med research and funding scheme [junior research alliance SUPR-G, 01ZX1401D to JM]; Swiss National Science Foundation [310030_184947 to SAL]. Funding for open access charge: Institutional funding.

Conflict of interest statement. G.M. is a co-founder of and J.M. a consultant to siTOOLS Biotech GmbH, Martinsried.

REFERENCES

- Steitz, J.A. (1969) Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature*, **224**, 957–964.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R. and Weissman, J.S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, **324**, 218–223.
- Ingolia, N.T. (2014) Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.*, **15**, 205–213.
- Ingolia, N.T., Hussmann, J.A. and Weissman, J.S. (2019) Ribosome profiling: global views of translation. *Cold Spring Harb. Perspect. Biol.*, **11**, a032698.
- Nedialkova, D.D. and Leidel, S.A. (2015) Optimization of codon translation rates via tRNA modifications maintains proteome integrity. *Cell*, **161**, 1606–1618.
- Andreev, D.E., Smirnova, V.V. and Shatsky, I.N. (2021) Modifications of ribosome profiling that provide new data on the translation regulation. *Biochemistry (Mosc)*, **86**, 1095–1106.
- Archer, S.K., Shirokikh, N.E., Beilharz, T.H. and Preiss, T. (2016) Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature*, **535**, 570–574.
- Bohlen, J., Fenzl, K., Kramer, G., Bukau, B. and Teleman, A.A. (2020) Selective 40S footprinting reveals cap-tethered ribosome scanning in human cells. *Mol. Cell*, **79**, 561–574.
- Wagner, S., Herrmannova, A., Hronova, V., Gunisova, S., Sen, N.D., Hannan, R.D., Hinnebusch, A.G., Shirokikh, N.E., Preiss, T. and Valasek, L.S. (2020) Selective translation complex profiling reveals staged initiation and Co-translational assembly of initiation factor complexes. *Mol. Cell*, **79**, 546–560.
- Oh, E., Becker, A.H., Sandkci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G. *et al.* (2011) Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell*, **147**, 1295–1308.
- Doring, K., Ahmed, N., Riemer, T., Suresh, H.G., Vainshtein, Y., Habich, M., Riemer, J., Mayer, M.P., O'Brien, E.P., Kramer, G. *et al.* (2017) Profiling ssb-nascent chain interactions reveals principles of Hsp70-assisted folding. *Cell*, **170**, 298–311.
- Gerashchenko, M.V. and Gladyshev, V.N. (2017) Ribonuclease selection for ribosome profiling. *Nucleic Acids Res.*, **45**, e6.
- Ingolia, N.T., Lareau, L.F. and Weissman, J.S. (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, **147**, 789–802.
- Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B. and Qian, S.B. (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E2424–E2432.
- Wu, C.C., Zinshteyn, B., Wehner, K.A. and Green, R. (2019) High-resolution ribosome profiling defines discrete ribosome elongation states and translational regulation during cellular stress. *Mol. Cell*, **73**, 959–970.
- Ingolia, N.T. (2010) Genome-wide translational profiling by ribosome footprinting. *Methods Enzymol.*, **470**, 119–142.
- Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M. and Weissman, J.S. (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.*, **7**, 1534–1550.
- Lacsina, J.R., LaMonte, G., Nicchitta, C.V. and Chi, J.T. (2011) Polysome profiling of the malaria parasite *Plasmodium falciparum*. *Mol. Biochem. Parasitol.*, **179**, 42–46.
- Clamer, M., Tebaldi, T., Lauria, F., Bernabo, P., Gomez-Biagi, R.F., Marchiorretto, M., Kandala, D.T., Minati, L., Perenthaler, E., Gubert, D. *et al.* (2018) Active ribosome profiling with RiboLace. *Cell Rep.*, **25**, 1097–1108.

20. McGlincy, N.J. and Ingolia, N.T. (2017) Transcriptome-wide measurement of translation by ribosome profiling. *Methods*, **126**, 112–129.
21. Lecanda, A., Nilges, B.S., Sharma, P., Nedialkova, D.D., Schwarz, J., Vaquerizas, J.M. and Leidel, S.A. (2016) Dual randomization of oligonucleotides to reduce the bias in ribosome-profiling libraries. *Methods*, **107**, 89–97.
22. Blondal, T., Thorisdottir, A., Unnsteinsdottir, U., Hjorleifsdottir, S., Aevarsson, A., Ernstsson, S., Fridjonsson, O.H., Skirnisdottir, S., Wheat, J.O., Hermannsdottir, A.G. *et al.* (2005) Isolation and characterization of a thermostable RNA ligase 1 from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties. *Nucleic Acids Res.*, **33**, 135–142.
23. Ho, C.K., Wang, L.K., Lima, C.D. and Shuman, S. (2004) Structure and mechanism of RNA ligase. *Structure*, **12**, 327–339.
24. Schwank, K., Schmid, C., Fremter, T., Milkereit, P., Griesenbeck, J. and Tschochner, H. (2022) RNA polymerase I (Pol I) lobe-binding subunit Rpa1.2.2 promotes RNA cleavage and proofreading. *J. Biol. Chem.*, **298**, 101862.
25. Pospisek, M. and Valasek, L. (2013) Polysome profile analysis—yeast. *Methods Enzymol.*, **530**, 173–181.
26. Buchbender, A., Mutter, H., Sutandy, F.X.R., Kortel, N., Hanel, H., Busch, A., Ebersberger, S. and Konig, J. (2020) Improved library preparation with the new iCLIP2 protocol. *Methods*, **178**, 33–48.
27. Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., Konig, J. and Ule, J. (2014) iCLIP: protein-RNA interactions at nucleotide resolution. *Methods*, **65**, 274–287.
28. Kim, Y., Eggers, C., Shvetsova, E., Kleemann, L., Sin, O. and Leidel, S.A. (2021) Analysis of codon-specific translation by ribosome profiling. *Methods Enzymol.*, **658**, 191–223.
29. R_Core_Team. (2019) In: *R Foundation for Statistical Computing*. Vienna, Austria.
30. Kechin, A., Boyarskikh, U., Kel, A. and Filipenko, M. (2017) cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput. Biol.*, **24**, 1138–1143.
31. Smith, T., Heger, A. and Sudbery, I. (2017) UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.*, **27**, 491–499.
32. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
33. Hasler, D., Meduri, R., Bak, M., Lehmann, G., Heizinger, L., Wang, X., Li, Z.T., Sement, F.M., Bruckmann, A., Dock-Bregeon, A.C. *et al.* (2020) The Alazami Syndrome-associated protein LARP7 guides U6 small nuclear RNA modification and contributes to splicing robustness. *Mol. Cell*, **77**, 1014–1031.
34. Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C.G. *et al.* (2018) Ensembl 2018. *Nucleic Acids Res.*, **46**, D754–D761.
35. Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J. *et al.* (2003) The UCSC genome browser Database. *Nucleic Acids Res.*, **31**, 51–54.
36. Jorjani, H., Kehr, S., Jedlinski, D.J., Gumienny, R., Hertel, J., Stadler, P.F., Zavolan, M. and Gruber, A.R. (2016) An updated human snoRNAome. *Nucleic Acids Res.*, **44**, 5068–5082.
37. Sharma, P., Wu, J., Nilges, B.S. and Leidel, S.A. (2021) Humans and other commonly used model organisms are resistant to cycloheximide-mediated biases in ribosome profiling experiments. *Nat. Commun.*, **12**, 5094.
38. Pfeffer, S., Lagos-Quintana, M. and Tuschl, T. (2005) Cloning of small RNA molecules. *Curr. Protoc. Mol. Biol.*, **Chapter 26**, Unit 26.4.
39. Lama, L., Cobo, J., Buenaventura, D. and Ryan, K. (2019) Small RNA-seq: the RNA 5'-end adapter ligation problem and how to circumvent it. *J. Biol. Methods*, **6**, e108.
40. Kim, H., Kim, J., Kim, K., Chang, H., You, K. and Kim, V.N. (2019) Bias-minimized quantification of microRNA reveals widespread alternative processing and 3' end modification. *Nucleic Acids Res.*, **47**, 2630–2640.
41. Zinshteyn, B., Wangen, J.R., Hua, B. and Green, R. (2020) Nuclease-mediated depletion biases in ribosome footprint profiling libraries. *RNA*, **26**, 1481–1488.
42. Lareau, L.F., Hite, D.H., Hogan, G.J. and Brown, P.O. (2014) Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *Elife*, **3**, e01257.