sCLIP—an integrated platform to study RNA–protein interactomes in biomedical research: identification of CSTF2tau in alternative processing of small nuclear RNAs

Yulia Kargapolova1,2, Michal Levin1,2, Karl Lackner1,3 and Sven Danckwardt1,2,3,*

1Institute for Clinical Chemistry and Laboratory Medicine, University Medical Center Mainz, Germany, 2Center for Thrombosis and Hemostasis (CTH), University Medical Center Mainz, Germany and 3German Center for Cardiovascular Research (DZHK), Germany

ABSTRACT

RNA-binding proteins (RBPs) are central for gene expression by controlling the RNA fate from birth to decay. Various disorders arising from perturbations of RNA–protein interactions document their critical function. However, deciphering their function is complex, limiting the general functional elucidation of this growing class of proteins and their contribution to (patho)physiology. Here, we present sCLIP, a simplified and robust platform for genome-wide interrogation of RNA–protein interactomes based on crosslinking-immunoprecipitation and high-throughput sequencing. sCLIP exploits linear amplification of the immunoprecipitated RNA improving the complexity of the sequencing-library despite significantly reducing the amount of input material and omitting several purification steps. Additionally, it permits a radiolabel-free visualization of immunoprecipitated RNA. In a proof of concept, we identify that CSTF2tau binds many previously not recognized RNAs including histone, snoRNA and snRNAs. CSTF2tau-binding is associated with internal oligoadenylation resulting in shortened snRNA isoforms subjected to rapid degradation. We provide evidence for a new mechanism whereby CSTF2tau controls the abundance of snRNAs resulting in alternative splicing of several RNAs including ANK2 with critical roles in tumorigenesis and cardiac function. Combined with a bioinformatic pipeline sCLIP thus uncovers new functions for established RBPs and fosters the illumination of RBP–protein interaction landscapes in health and disease.

INTRODUCTION

RNA-binding proteins (RBPs) are key factors regulating the fate of virtually all classes of RNA molecules throughout their lifespan by controlling their temporal, spatial and functional dynamics. Proper RBP functions are critical for a plethora of cellular programs controlling for instance normal cell growth and development by modulating gene expression posttranscriptionally. Accumulating evidence suggests that many disorders are linked to alterations in the abundance or functionality of RBPs (1,2). Thus exploring their role in posttranscriptional gene regulation has great potential to unravel underlying disease mechanisms and is key to identify novel diagnostic and therapeutic avenues.

Recently, developed methods such as ‘interactome capture’ identified numerous novel RBPs (3–7). Many RBPs are multifunctional; apart from binding to RNA they exhibit moonlighting functions, they are involved in metabolic processes and possess enzymatic activity (so-called ‘enigmatic RBPs’ (6,8,9)). RBPs can also integrate external biological signals leading to rapid structural remodeling of RBP–RNA complexes with important roles in health and disease (i.e.(10) and refs. therein). However, the function of the vast majority of the growing class of RBPs is still enigmatic.

In spite of numerous attempts to characterize the role of RBPs in health and disease, only a small proportion of them has been addressed. This reflects the enormous number and diversity of RBPs identified within recent years establishing complex, sheer endless landscapes of RBP–RNA interactions differing from cell-to-cell and context-to-context most probably also in a species-dependent manner (3–7). Nevertheless, the first attempts to unravel the role of RBPs in health and diseases already provided intriguing insights.

Crosslinking and immunoprecipitation (‘CLIP’ and variants) have been applied successfully to identify specific RNA–protein interactions in high resolution both in cell...
culture as well as in living organisms or tissues (11–18 and references therein). Here, RNA–protein interactions are preserved in living cells by UV-irradiation forming covalent bonds between amino acid residues bound directly to the RNA molecule. After irradiation, the cells are solubilized and the RNA is partially digested. Next the RNA–protein complexes are purified under stringent conditions by immunoprecipitation (IP) using an antibody directed against the protein of interest, and the bound RNA is ultimately released from the RBP. Finally, the RNA is used to generate a cDNA library for high-throughput sequencing. By mapping the reads to the transcriptome, the whole repertoire of RBP targets can be ultimately determined.

Current CLIP protocols include several critical steps such as the specificity and efficiency of the IP of the protein of interest. CLIP is thus often applied on exogenously expressed protein variants allowing to exploit standardized IP protocols for example based on FLAG-tag (CLIP, HITS-CLIP, iCLIP), streptavidine and polyhistidine epitopes or other protein tags (19). These modifications overcome the issue of inefficient antibodies, but deprive the method of its benefits to study endogenously proteins in native and/or dynamic conditions.

There are however further (critical) limitations: Most techniques require radioactive labeling of the IPed RNA (as an important step of quality control) and high amounts of input material. Further many rely on complex protocols for library synthesis employing RNA ligation with an inherently poor efficiency (20) on low input material and including many steps with a high risk of material loss (such as purification/washing steps and/or gel purifications). Finally, exploring RBP–RNA interactomes requires profound bioinformatical know-how and strategies to process big sequencing datasets. These limitations restrict the broad usage of the technique in biomedical research and make it mostly inapplicable to non-expert laboratories.

Here, we set out to overcome these restrictions. sCLIP omits the use of radioactivity for the visualization of the IPed RNA by applying a highly sensitive HRP-based detection strategy. Further, by using a linear in vitro transcription amplification reaction, sCLIP avoids using RNA-ligation on low input material and omits size-selection of cDNAs while reducing the number of clean-up steps hence minimizing material loss. This improves the complexity of the sequencing-library, resulting in a simplified and efficient procedure with an extremely high proportion of informative reads. Eventually this allows working with minimal starting material, which is often limiting in clinical samples.

Applying sCLIP along with a newly designed bioinformatics pipeline we confirm previous findings and provide important new insights into the role of cleavage and stimulation factor 2 tau (CSTF2tau) an essential component involved in 3′ end maturation of mRNAs. CSTF2tau and its paralog CSTF2 are expressed at different levels in various tissues and represent important regulators of alternative cleavage and polyadenylation of mRNAs with redundant functions (21). Here we show that the depletion of CSTF2tau affects the expression of various protein-coding and, interestingly, non-coding genes. We demonstrate that CSTF2tau binding to snRNAs is associated with their oligoadenylation and (rapid) degradation, which in turn is coupled to alternative splicing of several transcripts with varying exon organizations including the cancer-associated ANK2, CNRIP1 and RAP1GDS1 RNAs. Combined with the bioinformatics pipeline sCLIP thus facilitates deep insights into RNA interactome(s) and helps illuminating the downstream mechanisms of RBP functions in biomedical research.

MATERIALS AND METHODS
Methods

sCLIP-seq library preparation. A detailed protocol with a step-by-step instruction is provided in the supplementary information.

sCLIP-seq data processing and quality control

A general overview of the sCLIP-seq data processing steps is provided in the supplementary information.

Poly(A)-tail length assay (e-PAT)

To tag the polyadenylated RNAs and measure the length of poly(A) tail an ePAT method was applied as described recently (22) (further experimental details along with a table of primers can be found in supplementary materials).

Cell culture and knockdown experiments

BE(2)-C cells were cultured and transfected with siRNAs under standard conditions. Protein expression analysis, RNA extraction and PCR have been carried out as described earlier (10). Further information on the RNA decay experiments, splicing analysis and RNA seq data analysis is provided in the supplementary data.

RESULTS

A radio-label free crosslinking and immunoprecipitation protocol with optimized CLIP biochemistry and an automated bioinformatics pipeline (sCLIP)

Typically CLIP-seq is carried out after crosslinking of the RBPs to the target RNA by applying UV-light to intact cells (Figure 1A). This preserves in vivo relevant RBP–RNA interactions after disintegration of the cell and allows to IP the RNP complexes (after partial RNA digestion) under conditions, which prevent at the same time the formation of physiologically non-existing ‘artificial’ RNA–RBP-interactions after disrupting the cell compartmentalization. Thus, when carefully executed, only RNA molecules loaded onto RBPs under physiological conditions are studied. Next, the RBP–RNA complexes are typically visualized by labeling of the RNA with radioactivity. Upon gel purification of the complexes, bound RNAs are eluted and linkers are ligated to the 3′ (and 5′) end, which eventually allows the cDNA synthesis for the generation of the sequencing library.

RNA-ligases suffer from very poor efficiency (~10% (20)). Thus the sequencing linker ligation reaction on RNA released from RNP complexes is the most crucial and limiting step of many CLIP protocols. Inefficient RNA-ligation
lowers the library complexity thus making low-abundant target RNAs difficult to be detected in CLIP-seq data. This also impacts on the amount of starting material needed, which is often limited in case of precious clinical samples. Based on these considerations we set out to develop a simplified CLIP protocol, which (I.) avoids using RNA-ligation on low input material and (II.) allows reducing the number of purification steps to improve the complexity of the CLIP-seq library despite using minimal input material. Finally, we aimed (III.) to overcome the use of radioactivity for the visualization of the IPed RNA.

To this end, we set out to advance CLIP by testing three alternatives of the library synthesis (Supplementary Figure S1A), which allow omitting RNA-ligation in the first step of the protocol. As non-templated 3′ poly(A) tailing is the basis of all CLIP variants tested here, we carefully examined whether the poly(A) polymerase may have substrate preferences. Although substrate preference has been reported (23), we observe only marginal preferences of the poly(A) polymerase when using synthetic RNA oligonucleotides with differing 3′ end terminal nucleotides (Supplementary Figure S1B). For instance, oligonucleotides terminated with A and G are most efficiently tailed, whereas C and G terminated oligonucleotides appeared to be slightly less efficiently tailed. Yet this subtle tendency can be overcome by using different amounts of ATP (Supplementary Figure S1B). Further, prolonging the reaction time allows to fully polyadenylate even 5 μg of highly structured RNA oligonucleotides (Supplementary Figure S1C). The amount of RNA tested in this experiment however is much higher than that processed in conventional CLIP protocols (in our experience 50–100 pg RNA eluted from the RBP after IP). Thus, although a substrate preference cannot be completely excluded, there is no evidence that the poly(A) polymerase introduces a significant bias under the conditions applied here.

Next, the polyadenylated RNA was reversely transcribed using anchored oligo (dT) primers with Illumina adaptors followed by three different approaches to generate the sequencing library (Supplementary Figure S1A). While approach A (which is based on G-tailing of the cDNA 3′ end by using terminal deoxynucleotidyl transferase and subsequent PCR amplification of the product) and approach B (employing a circular ligation activity, an enzyme which was successfully applied in a recently published iCLIP protocol (24)), gave reasonable results (not shown), approach C showed the best performance (a simplified experimental outline is shown in Figure 1A; a detailed step-by-step protocol complements this outline in the supplementary material). After RNA purification, poly(A) tailing and reverse transcription, this approach utilizes an in vitro transcription reaction (IVT), which serves to boost the amount of original RNA species linearly and in an unbiased manner (25–27) and thereby allows to work with higher amounts of material in the following steps. Most importantly however, this protocol does not require an additional step of cDNA size selection (for instance to remove concatenated sequencing adaptor primers) and therefore avoids loosing valuable material, which overall helps improving the library complexity. Accordingly, we were able to reduce the amplification of the library for sequencing to a minimum of 9–11 PCR cycles required to complete the library preparation. In contrast, most of the currently available protocols use 25 cycles of PCR amplification (Supplementary Table S1) in order to generate sufficient material for sequencing. This is critical as a high number of PCR cycles bears the risk of a disproportionate amplification of sequences with different nucleotide composition, which can ultimately lead to a false interpretation of the RBPs’ binding preference.

Although the number of PCR cycles used in the sCLIP protocol is very low (≤11 cycles), an amplification bias cannot be excluded a priori. To correct for a potential bias, unique molecular identifiers were used to label each single molecule as described previously (28) allowing for removal of PCR duplicates in the data processing step. Further for multiplexing purposes, the first oligo (dT) primer contains a 6 digit barcode sequence, which allows to distinguish different experimental samples pooled together for sequencing.

To simplify the sCLIP sequencing data analysis, we developed a bioinformatics pipeline, which allows an automated and standardized analysis for all samples. The pipeline integrates all crucial steps of bioinformatic processing of CLIP-seq data (such as adaptor removal, reads quality assessment, alignment, collapsing duplicated reads, peak calling, merging replicates, annotation to genome features, differential binding analysis and motif search; Supplementary Figure S2A and B).

The performance of sCLIP as an integrated platform with improved CLIP biochemistry and automated bioinformatic processing for comprehensive RBP–RNA interaction analysis was next tested for CSTF2tau, an established RBP. CSTF2tau is involved in RNA processing by guiding the core cleavage and polyadenylation machinery to 3′end processing sites (21,29). The RNA binding specificities of CSTF2tau have recently been studied with HITS-CLIP and iCLIP approaches independently by two research groups (21,30,31), making this protein an ideal candidate for evaluating the performance of the sCLIP method developed here.

Endogenous CSTF2tau was specifically IPed from BE(2)-C cells under optimized conditions (Supplementary Figure S1D–G). CLIP-type approaches are very sensitive to IP efficiency and purity. Accordingly co-IP of interacting partners of the protein of interest can lead to a false discovery of RNA targets. Apart from exploring the specificity of the IP antibody (Supplementary Figure S1D), we therefore assessed the purity of the IP by staining for CSTF77, a known interaction partner of CSTF2tau (32), by western blotting (Supplementary Figure S1E). After IP under most optimal conditions, an aliquot of the CSTF2tau protein–RNA complexes was visualized by using a newly established non-radioactive labeling procedure (Supplementary Figure S1F). Briefly, this method is based on biotin labeling and a highly sensitive HRP-coupled detection (see detailed protocol in the supplementary materials and methods section). Thus, unlike most available protocols (Supplementary Table S1), we were able to implement a strategy that acknowledges the increasingly widespread banning of radioactivity in research facilities. Finally, the integrity of the eluted RNA was verified (Supplementary Figure S1G) and libraries were generated as detailed above (Figure 1A and Supplementary Figure S1H), and ultimately subjected to sequencing.
To evaluate the performance of the newly established protocol we next assessed the reproducibility of biological replicates. We observed a high consistency between two independent replicates of sCLIP for CSTF2tau (Pearson correlation 0.77; Figure 1B). As reported previously, correlations between replicates produced by CLIP variants usually vary between 0.3 and 0.9 (Supplementary Table S1). Thus, a correlation of above 0.77 observed in this study suggests that sCLIP delivers highly reproducible results.

We next determined the relative proportion of ‘usable reads’, compared to the fraction of reads, which were either classified as reads that were ‘too short’ (and thus discarded), ‘PCR duplicates’, or reads that did not uniquely map to the genome (Figure 1C). Importantly, despite omitting gel purification of the cDNA library before sequencing, we identify less than 8.5% of the reads being too short (applying a threshold of 18 nt) and thus discarded. Compared to other CLIP protocols the fraction of reads being discarded when applying sCLIP is thus at the very low end (Supplementary Table S1). Most remarkably, however, ~60% of the input reads (corresponding to >95% of the uniquely mapped reads) were found to be usable to determine the respective RBP interactome, which is significantly higher compared to most of the other CLIP protocols (Figure 1D and Supplementary Table S1). In contrast, the number of undesired PCR duplicates is extremely low (<5%) likely due to the
limited numbers of PCR cycles during the library generation (Supplementary Table S1).

We thus conclude that omitting gel purification of the cDNA library in sCLIP (together with other tedious purification steps, which are typically a source of significant loss of material) does not result in an undesired high proportion of unusable reads. Instead, together with the IVT amplification reaction it permits working with significantly lower amounts of input material and simplifies the procedure. Thus sCLIP is suited to produce highly complex sequencing libraries in a robust and reproducible manner providing deep insights into a RBPs’ interactome with high resolution in a radio-label free manner.

**sCLIP recapitulates previous CLIP data and identifies novel CSTF2tau binding to histone and non-coding RNAs**

We next examined CSTF2tau sCLIPs further by applying the bioinformatic sCLIP pipeline to extract the binding preferences of this protein (Supplementary Figure S2A and B). Analyzing the genome-wide distribution of CSTF2tau sites we identified >16 000 binding sites of CSTF2tau and revealed that 31% are localized in introns, 30% align to 3’UTRs, 7% belong to 5’UTRs, 8% were detected on non-coding genes and 2% on exons (Figure 2A). After correction for the length of each feature, 3’UTRs and 5’UTRs are most prevalently covered by the protein (Figure 2A, bar diagram). The observed binding preference to 3’UTRs recapitulates previously described properties of CSTF2tau (21,30–32). In accordance to its function in 3’end processing, CSTF2tau sCLIP reads are focused at the 3’end of polyadenylated pre-RNAs and co-localize with annotated cleavage and polyadenylation sites (Figure 2B). Extracting the exact location of the sCLIP reads transcriptome-wide, we identify that they typically localize 40–90 nucleotides downstream of the cleavage and polyadenylation sites (Figure 2C). Further, CSTF2tau binding sites are characterized by high frequencies of A and T (U) residues and, among dinucleotides, AA, TG (UG) and TT (UU) residues are enriched (Supplementary Figure S2C and D). The over-representation of A and T rich regions within the CSTF2tau binding sites (and in close vicinity) can also be observed when applying a de novo motif search algorithm to all determined sites. This analysis reveals enrichment of several motifs (Figure 2D and Supplementary Figure S2E). The first resembles the AUAUA hexamer sequence (reflecting one of the most prevalent poly(A) signal sequence motifs), while the other two motifs are U/GU rich and represent the so called ‘downstream U/GU-rich sequence elements’ (33) required for efficient 3’end processing (which are typically recognized by the CSTF2/2tau proteins (21)). This suggests that sCLIP reliably identifies binding sites of endogenously expressed CSTF2tau, in accordance to previously published biochemical and CLIP-seq data (21,30–32).

In this context, it is important to point out that we did not observe an enrichment of CLIP-reads 20–30 nt downstream of the poly(A) signal, where endogenous RNAs are typically cleaved and polyadenylated (Figure 2C, and Supplementary Figure S2F). Thus although the sCLIP library preparation is based on in vitro poly(A) tailing of the IPed RNAs and subsequent oligo d(T)-based reverse transcription, we do not observe an increased occurrence of 3’ends of the CLIP RNAs that coincide with cleavage sites (Supplementary Figure S2F). This indicates that sCLIP does not introduce a bias towards mRNA 3’ends (reflecting intracellularly polyadenylated RNA species), and highlights that sCLIP represents a reliable alternative compared to other CLIP strategies.

CSTF2tau binding sites are most prevalently located in the 3’UTRs of coding RNAs (Figure 2B). In accordance to its role in 3’end processing of polyadenylated RNAs binding of CSTF2tau to the 3’ends is easily explained and even expected. Surprisingly, however, we also observe the binding of CSTF2tau on replication-dependent histones (Supplementary Figure S2H) and on U7 snRNA (not shown), a component of the small nuclear ribonucleoprotein complex (U7 snRNP), involved in histone processing.

The position of CSTF2tau binding on replication-dependent histone RNAs deserves special attention as it occurs at the 5’end of the open reading frame (Supplementary Figure S2H). Interestingly, while replication-dependent histones, which are processed by a unique 3’end histone-processing mechanism (HIST1H4E and HIST1H3D) show this ‘atypical’ location, replication-independent histones, which are processed by the conventional cleavage and polyadenylation machinery (34) (for example H3F3A and H1FX) show a ‘typical’ binding at the mRNA 3’end (Supplementary Figure S2H; further examples showing histones with binding sites of CSTF2tau protein according to the same logic are listed in Supplementary Tables S3 and S4).

Most interestingly however, in addition to previously reported interactions of CSTF2tau with protein-coding RNAs, we also observe numerous binding sites on non-coding RNAs (Figure 2E). Approximately 10% of RNAs bound by the protein originate from ‘non-coding genes’, the majority of which belong to antisense (63), lincRNA (67) and pseudogene (36) type of transcripts (Supplementary Figure S2G).

We next carried out hypergeometric analyses to discover RNA types, which are over- or under-represented in the cohort of CSTF2tau targets (Figure 2E, bar diagram). Strikingly, transcripts of protein-coding genes were under-represented, whereas some categories of non-coding genes, for example sense intronic, linc RNAs and antisense RNAs were over-represented. This suggests that CSTF2tau binds protein-coding genes selectively. Over-represented binding of CSTF2tau to non-coding RNAs was not reported before and suggests that the protein possibly possesses novel, yet not identified functions.

Taking a closer look at non-coding RNAs, we identify that CSTF2tau binds Small Cajal body-specific RNAs (SCARNAs), a class of small nuclear RNAs (snRNAs) localized in the Cajal body (an organelle involved in biogenesis of snRNPs). Among those CSTF2tau recognizes SCARN A7 and SCARN A9, which guide the modification of RNA polymerase II transcribed spliceosomal RNAs U1 and U2 (35). Interestingly, we also observe several U-type snRNAs bound by CSTF2tau (further detailed below).

Altogether sCLIP delivers several novel so far not reported RBPs-protein interactions (Supplementary Figure S2G). Intriguingly, some of the CSTF2tau targets (i.e.
Figure 2. sCLIP recapitulates previously reported RBP–RNA interactions data and identifies CSTF2tau binding to non-coding RNAs. (A) CSTF2tau preferentially binds to 3′ and 5′ untranslated regions (UTRs). The pie diagram (left) shows the distribution of binding sites of CSTF2tau on the genomic features in percent; the bar diagram (right) shows enrichment in coverage of sites over the total length of the genomic feature of the gene. (B) CSTF2tau binding sites are focused in the 3′ UTR and are located in close proximity to polyadenylation sites, illustrated for YY1AP1, FUBP1, GDI2 and BUB3 transcripts for two independent experiments; samples, which were generated by immunoprecipitation with an IgG antibody ‘IgG only’, serve as control (overall estimated proportion of CSTF2tau cCLIP Rep 1 and Rep 2 libraries, see also Figure 1 having the same reads as the IgG control sCLIP is 1.05% and 1.15%). A similar binding pattern can also be found for replication-dependent histones (see Supplementary Figure S2H). (C) CSTF2tau CLIPs are predominantly located downstream of the endonucleolytic cleavage site (where 3′ end processing of polyadenylated mRNAs occurs) with a peak centred 50–70 nucleotides downstream of the poly(A) signal AAUAAA hexamer (further binding preferences are shown in Supplementary Figure S2E). Of note, in replication dependent histones (in which processing at the mRNA 3′ end is executed by other mechanisms) the CTF2tau CLIPS are found to predominate at the ORF 5′ end (Supplementary Figure S2H) thereby confirming the specificity of the CSTF2tau CLIPs studied here. (D) Three most prevalent motifs found within a 100nt region around the center of the peak of CSTF2tau sCLIP tags. (E) In addition to protein-coding RNAs, CSTF2tau also binds non-coding RNAs (pie diagram; for further details on the distribution among the different classes of ncRNAs, see Supplementary Figure S2G). The hypergeometric distribution analysis (bar diagram) shows that some classes are significantly overrepresented among the RNAs bound by CSTF2tau, while protein coding RNAs and pseudogenes are underrepresented (dashed line labels probability below 5%; red bars indicate over-representation; blue bars show under-representation; * indicates a statistically significant change, P-value <0.05; absolute numbers of RNAs and respective P-values see Supplementary Table S11).
SCARNAs and U7 snRNAs) play a role in the biogenesis of other RNAs (U-type snRNAs and replication-dependent histones, respectively), which in turn are bound by CSTF2tau. This may reflect a functional RNA regulon (36) ensuring a tight coupling of both, biogenesis and post-transcriptional regulation of specific RNA molecules.

CSTF2tau regulates the steady state abundance of snRNAs and snoRNAs

The strength of crosslinking and immunoprecipitation protocols lies in deep and multi-layer target identification as illustrated here. To further corroborate the functional importance of CSTF2tau for the respective target RNAs, we next performed RNA-sequencing in the presence and absence of the protein. To that end, BE(2)-C cells were transiently transfected with a pool of 4 siRNAs targeting CSTF2tau and the depletion efficiency was confirmed by western blot analysis (Figure 3A).

Upon depletion we retrieved 80 million reads on average per sample by RNA sequencing, which did not significantly differ between the samples (Supplementary Figure S3A). The relative similarity between replicates was further assessed by multidimensional scaling (MDS) plot confirming that the biological replicates are highly consistent between each other, while the different treatments are well separated from each other (Supplementary Figure S3B). Applying a threshold of a false discovery rate (FDR) of <0.05 we identify ∼2000 genes to be differentially expressed (Figure 3B and C, and Supplementary Figure S3C). Interestingly, there was no obvious directionality of the change of the steady state RNA level; RNAs were both up- and downregulated (FDR < 0.05). However, only 215 genes were highly regulated (>1.5-fold). Among those the majority of genes (88%) were up-regulated in CSTF2tau depleted cells (Figure 3C).

We next performed hypergeometric analyses, which revealed that only a minority of protein-coding genes were highly regulated upon CSTF2tau depletion (Figure 3D). In contrast, non-coding RNAs, especially snRNAs and snoRNAs were significantly over-represented within the cohort of genes differentially expressed upon CSTF2tau depletion. Strikingly, the expression of genes belonging to the snRNA and snoRNA group was exclusively up-regulated even when a less stringent threshold (>1.25) of regulation was applied (Supplementary Figure S3D-G). If the regulation were unbiased, the expected number of downregulated sn/snoRNA genes should be ∼24 (instead no gene is found; Supplementary Figure S3E). The probability of the bias observed here is very low (1.4 × 10^{-14}), suggesting that the directional regulation of snRNA and snoRNA group upon CSTF2tau RNAi is caused by the depletion of the protein and does not occur by chance.

To identify transcripts, which are directly bound and functionally affected by CSTF2tau, we merged the list of bound RNA targets with the list of differentially expressed RNAs. Strikingly, protein-coding genes are under-represented in the overlapping cohort (Figure 3E). This suggests that there is no direct effect of CSTF2tau depletion on the abundance of this class of mRNA targets. This is in line with a CSTF2tau knockout in mouse testis cells, which did not reveal a direct effect of CSTF2tau on the expression of its targets (32). However, our analysis reveals that the class of snRNA genes are over-represented among the regulated CSTF2tau targets (Figure 3E). This finding invites speculations that CSTF2tau binding might have a direct effect on the abundance of the snRNAs.

CSTF2tau binding to the 3′ end of snRNAs mediates internal oligoadenylation leading to snRNA degradation

CSTF2tau plays an important role in mRNA 3′ end cleavage and polyadenylation (21,30,31). Based on this function we wondered whether binding of CSTF2tau to snRNAs might explain the change of steady state mRNA expression via differential 3′ end cleavage and polyadenylation.

Interestingly, CSTF2tau specifically recognizes the 3′ end of these molecules (Figure 4A) - reminiscent to the binding position of this protein to the precursors of polyadenylated mRNAs (Figure 2B). To identify whether human U-type snRNAs are potentially polyadenylated we carried out an ePAT assay, which allows to detect polyadenylated molecules and to measure the length of their poly(A) tails (26). Upon ePAT PCR (Supplementary Figure S4A), the products were cloned and sequenced using Sanger sequencing. ePAT confirmed that fractions of U1, U4, U5 and U11 snRNAs are oligoadenylylated (Figure 4B). In contrast, for U4 ATAC the length of the poly(A) tail was short (6 nucleotides), but the occurrence of the tail could not be explained by internal priming. Most interestingly, however with the exception of U4 ATAC, the ‘body’ of all oligoadenylylated U-type snRNA studied here (U1, U4, U5, U11) is significantly shorter than the annotated cDNA. Thus these molecules appear to be trimmed or endonucleolytically cleaved before oligoadenylation. Further, binding of CSTF2tau occurs downstream of the oligoadenylylated region (Figure 4A, site of polyadenylation highlighted with red arrow head) and thus apparently takes place before trimming (or endonucleolytic cleavage). This functional architecture is highly reminiscent to 3′end processing of polyadenylated RNAs.

The RNA-abundance of snRNAs such as U4 and U5 changed >1.5-fold upon CSTF2tau depletion (as identified by RNA-seq). Of note, the RNA-seq analysis is based on random priming and thus both, polyadenylated as well as non-polyadenylated RNA fractions are determined. Based on these findings we asked whether the amount of the polyadenylated fraction of snRNAs changes upon CSTF2tau depletion (Supplementary Figure S4B). After RNA extraction, the cDNA was generated with two different protocols that allow the quantification of both, the total and the polyadenylated fraction of the snRNAs individually. In qRT-PCR experiments we observe that depletion of CSTF2tau leads to significant down-regulation of the polyadenylated snRNA isoforms of U1, U4, U5 and U11 relative to the total amount (Figure 4C). Thus, in concordance with the expectation for conventional 3′end processing of polyadenylated mRNAs, depletion of CSTF2tau reduces the fraction of oligoadenylylated U-type snRNA isoforms (U1, U4, U5 and U11). Strikingly, the abundance of U-type snRNAs increases upon depletion of CSTF2tau, which we would not expect to occur for RNAs undergoing canonical 3′end formation. In contrast, the abundance
Figure 3. CSTF2tau binds to and regulates the abundance of snRNAs. (A) Transfection of BE2C cells with target-specific siRNAs down-regulates CSTF2tau protein expression to <10%. (B) Depletion of CSTF2tau by siRNAs leads to significant regulation of the steady-state RNA abundance of over 2000 RNAs (determined by RNA seq; FDR < 0.05; for further information on the quality of reads retrieved per sample by RNA sequencing and the relative similarity between replicates and differences between the treatments see Supplementary Figure S3A and B). (C) The steady state expression of the majority of significantly regulated RNAs are only mildly changed. A small proportion of genes is regulated at over 1.5-fold. Out of these genes 80% are up-regulated (P-value < 0.05). (D) Hypergeometric distribution analysis revealing that snRNAs and snoRNA are significantly overrepresented among the RNAs regulated by CSTF2tau, while protein coding RNAs are underrepresented (dashed line labels probability <5%; red bars indicate over-representation; blue bars show under-representation; * indicates a statistically significant change, P-value <0.05; further information illustrating and confirming the abundance changes of the snRNAs and snoRNAs is provided in Supplementary Figure S3F and Figure 4C; absolute numbers of RNAs and respective P-values see Supplementary Table S11). (E) Hypergeometric distribution analysis showing that snRNAs are significantly overrepresented among the RNAs bound and regulated by CSTF2tau (dashed line labels probability below 5%; * indicates a statistically significant change, P-value <0.05; absolute numbers of RNAs and respective P-values see Supplementary Table S11).

Depletion of CSTF2tau increases the stability of U-type snRNAs and causes alternative splicing of several transcripts including the tumor-associated ANK2, CNRIP1 and RAP1RDS1 RNAs

Previously oligoadenylation of snoRNAs has been shown to initiate their degradation (42). We therefore wondered whether the inhibition of polyadenylation after CSTF2tau depletion might affect the RNA decay rate of the U-type snRNAs studied here. To address this we carried out actinomycin D (ActD)—RNA decay experiments with cells that were first treated with siRNAs against CSTF2tau for 48 hours followed by addition of ActD for 0, 2, 4 and 6 h, respectively (Figure 5A). For each time point total RNA was assessed by qRT-PCR. Interestingly, U1, U4 and U5 snRNAs are faster degraded in the control sample compared to the CSTF2tau depleted sample (Figure 5A), whereas the relative stability of a control mRNA (the proto-oncogene Jun mRNA) remained unchanged upon CSTF2tau deple-
Figure 4. snRNAs are internally poly/oligoadenylated in a CSTF2tau dependent manner. (A) Illustration of the CSTF2tau binding on the snRNAs’ 3’ end (highlighted in red). The site of oligoadenylation is marked by a red arrow head and asterisk (see also Figure 4B). (B) ePAT assay based identification of polyadenylated snRNA isoforms after cloning and conventional sanger sequencing. The canonical (matured) 3’ end of the respective snRNA is shown in bold, (internally) oligoadenylated snRNA isoforms are shown underneath (for further details see Supplementary Figure S4; some oligoadenylated isoforms are found more than once as indicated i.e. 4×, 2x etc.). (C) Quantitative RT-PCR analysis to determine the fold change of the steady state RNA abundance of the total (blue) and poly/oligoadenylated (red) snRNA isoforms upon depletion of CSTF2tau (log2 scale; for details on normalization controls see Supplementary Figure S4, * indicates a statistically significant change, P-value < 0.05).

In contrast, the long-lived U11 snRNA and the U4 snRNAs are represented by a short 3’ terminal oligo(A) tail, does not show significantly reduced decay upon CSTF2tau depletion. This suggests that oligoadenylation of snRNAs studied here specifically down-regulates the stability of U1, U4 and U5. U1, U4 and U5 belong to the SM class of snRNAs resembling components of the major and the minor spliceosome (43,44). We thus finally asked whether the CSTF2tau effect on the abundance of snRNAs may result in a splicing phenotype. To this end, we analysed global splicing changes upon CSTF2tau depletion based on RNAseq by applying DEXseq (37) (Figure 5B). We detect 126 events of significant differential exon usage (FDR < 0.05) comprising 61 skipped and 65 retained exons (affecting 86 genes in total; Supplementary Table S5). This number is well within the range of regulated events upon depletion of (core) splicing components known to affect alternative splicing (14,38). We next confirmed the splicing pattern of three genes randomly chosen from the alternatively spliced candidates, ankyrin-2 (ANK2), SmgGDS (RAP1GDS1) and the cannabinoid receptor interacting protein 1 (CNRIP1; Figure 5C) by quantitative RT-PCR. Indeed we observed significant changes for all three transcripts tested here, resembling either phenotypes of exon skipping (ANK2), exon retention (RAP1GDS1) or intron skipping (CNRIP1). Intriguingly, all three genes have recently been linked to tumor formation ((39–41) and refs. therein). This may reflect the overall surprisingly high prevalence of cancer-related genes being affected by alternative splicing upon depletion of CSTF2tau (19 out of 86 genes).

Altogether, these data strongly suggest that the modulation of CSTF2tau regulates splicing of target genes by controlling the abundance of spliceosomal RNA components of the major and the minor spliceosome via internal oligoadenylation and subsequent RNA degradation. Further, these functional findings substantiate and exemplify the power of the integrated sCLIP platform, and reveal novel insights into the role of CTSF2tau in specifying the fate of snRNAs with potentially important clinical implications.

DISCUSSION

RNA binding proteins are crucial mediators of gene regulation. Here we present a CLIP approach, which allows studying RBP–RNA interactomes in a simplified and efficient manner resulting in an extremely high proportion of informative reads. Our approach addresses and overcomes several limiting steps of previously described CLIP versions (Supplementary Table S1 and specified below) and integrates the improvements previously introduced by recent CLIP approaches. sCLIP avoids using radioactive labeling of RNA (akin to recently published irCLIP, where the radioactivity is substituted by fluorescent labeling (17)).
Figure 5. CSTF2tau-dependent oligoadenylation of U-type snRNAs enhances their decay and controls alternative splicing of tumor-associated RNAs. (A) RNA decay determined by qRT-PCR analysis 2, 4 and 6 h after transcriptional halt via addition of actinomycin D, with (blue line) and without depletion (red line) of CSTF2tau. The decay curve of the Jun RNA served as a positive and specificity control for the decay analysis upon actinomycin D treatment and depletion of CSTF2tau (* indicates a statistically significant change, P-value < 0.05). (B) Global splicing analysis to monitor the functional consequences of snRNA abundance alterations upon CSTF2tau depletion by applying DEXseq. Each spot in red denotes a gene for which a significant differential exon usage can be detected (FDR < 0.05; with the y-axis indicating the fold change of exon usage (log2) and the x-axis showing the average normalized count value (log2)). Detailed information concerning the identity of differentially spliced genes is provided in Supplementary Table S5. (C) qRT-PCR-based splicing analysis to validate the functional consequences of snRNA abundance alterations upon CSTF2tau depletion for three randomly selected RNAs (ANK2, RAP1GDS1 and CNRIP1, * denotes a statistically significant change, P-value < 0.05; a more detailed information on the validation of alternative splicing of the transcripts tested by RT-PCR is provided in Supplementary Figure S5). (D) Proposed model for a CSTF2tau-dependent buffering of snRNA levels for the modulation of (alternative) splicing (further details see Discussion section).

sCLIP exploits labeling of RNA with biotin coupled to a HRP-based detection method with highest sensitivities and thus allows to keep this important step of quality control even in the absence of radioactivity. This accounts for the increasingly widespread banning of radioactivity due to safety measures in laboratories. sCLIP adopts a linker-cDNA synthesis strategy employed for single cell sequencing (27). The optimized protocol allows working with input material as little as 700 μg of protein (1–2 million cells). Despite this, it generates highly reproducible results with a consistency between replicates above 0.77. Additionally, sCLIP exploits the benefits of an experimental barcoding approach (14,42) and thus allows multiplexing of several samples. In the context of dynamic CLIP (comparing two or more conditions), this has the advantage that all subsequent steps of library preparation are carried out in one batch, allowing to reduce technical variabilities between the experiments. Further it reduces material and sequencing costs. In addition, sCLIP makes use of a second type of barcode (in analogy of iCLIP), which allows distinguishing and discarding amplification artifacts. Most importantly, however, libraries prepared by sCLIP require only 9–11 cycles of PCR amplification, which are 10–15 cycles less than in most other CLIP protocols (Supplementary Table S1). We refer this to three major advancements: First we employ a linear in vitro transcription amplification strategy of the input RNA. Second, we omit an inherently inefficient RNA ligase reaction on low input material. Third, the sCLIP protocol does not require size selection of the cDNA, as adaptor contamination of sCLIP libraries is virtually impossible. Therefore loss of material at the step of size selection is excluded, preventing low complexity of the library due to artificial multiplication of single reads. Accordingly, we obtain a very high proportion of useable reads with a high reproducibility, which permit deep insights into protein-RNA-interactomes despite significantly reducing the amount of input material and omitting several purification steps (Supplementary Table S1).
As a proof of concept, we applied the newly developed sCLIP pipeline for studying interacting RNA of CSTF2tau. The method successfully recapitulates previously described binding preferences of the protein and confirms that CSTF2tau binds 3′ ends most prevalently downstream of the cleavage and polyadenylation sites of pre-mRNAs (30). In line with previously published data (31), we also observed that binding of CSTF2tau downstream of the polyadenylation site predicts the usage of the site by cleavage and polyadenylation machinery (not shown).

Analyzing the sCLIP data in more detail, we further identify that the protein recognizes not only replication-independent histones, which are polyadenylated by the canonical cleavage and polyadenylation machinery, but also replication-dependent (RD) histones. In contrast to replication-independent histones, the processing of RD histones is uncoupled from polyadenylation and is carried out by a unique processing complex (32). This complex recovers the components of the canonical polyadenylation machinery, for example CPSF73, which executes the endonucleolytic cleavage (43). It has been reported that this complex also contains CSTF2, the paralog of CSTF2tau (44). In the context of this study, we reveal that CSTF2tau is bound to RD histones and can also be part of the cleavage complex. Of particular interest is the location of peaks on RD histones; they are located at the 5′ end of the molecule together. Furthermore, sCLIP reveals many non-coding RNAs as targets of CSTF2tau (Figure 2E).

CSTF2tau is known to contribute to 3′ end processing of pre-mRNAs. As revealed by sCLIP, the binding of CSTF2tau on snRNAs occurs at the 3′ end of canonically matured isoforn (or further downstream; Figure 4A). We report for the first time that CSTF2tau controls the abundance of snRNAs. Moreover, their steady-state RNA level was uniformly up-regulated upon depletion of CSTF2tau (except for U4 ATAC). Further we demonstrate that a fraction of snRNAs is oligoadenylated—presumably by recruiting the processing complex independently from a functional AAUAAA poly(A) signal (45). Most remarkably, the relative expression of oligoadenylated snRNAs (compared to total snRNAs) decreased upon depletion of CSTF2tau, which results in prolonged half-life (except for the long-lived U4 ATAC and U11 RNA).

The intriguing property of the poly(A) tail to determine RNA longevity on the one hand and a rapid decay on the other, has been discovered earlier (46,47). Canonical polyadenylation of mRNA precursors typically results in stabilization of transcripts as this is intricately coupled to the polyadenylation machinery that we identified here may—when oligoadenylated ‘internally’ (with exception of the U4 ATAC). Truncated U1 snRNA (U1-tfs) molecules have previously been described (50). U1-tfs lacking the Sm site are unable to form the Sm heptamer. Intriguingly, the last nucleotide of U1-tfs is exactly the same, at which we detect the oligo(A) attached. Moreover U1-tfs are more rapidly degraded and localized primarily to P-bodies (50).

The binding of CSTF2tau to snRNAs downstream of the endonucleolytic cleavage site is reminiscent to the processing of conventionally polyadenylated mRNAs. Accordingly, the depletion of CSTF2tau leads to a decrement of the oligoadenylated fraction suggesting that the binding of the protein at the 3′ ends of snRNAs contributes to their oligoadenylation (most likely together with other factors not yet identified here). In contrast to conventional mRNAs, the stability of total snRNAs increases upon depletion of CSTF2tau. This is in line with previous reports that oligoadenylated RNAs are substrates for fast decay (46,47). We are thus proposing a model, according to which CSTF2tau regulates the abundance of snRNAs. High-level CSTF2tau expression promotes the oligoadenylation of snRNAs resulting in a faster degradation of the affected molecules and lowering the steady state level of total snRNAs. Under conditions of low-level CSTF2tau, the proportion of oligoadenylated molecules decreases, leading to higher stability of snRNAs and elevated steady state mRNA levels. This mechanism eventually appears to control alternative splicing of several target genes (Figure 5B, and Model, Figure 5D). In this context it is interesting to note that all three genes further tested to be affected by alternative splicing upon CSTF2tau depletion (ANK2, CNRIP1 and RAP1GDS) show complex exon organizations, and are linked to tumorigenesis (39–41,51) and the regulation of neuronal signaling (52,53). For ANK2 extensive alternative splicing has been connected to (ab)normal cardiac function (54) including the ankyrin-B syndrome, characterized by sinus node dysfunction, susceptibility to ventricular arrhythmias, and sudden death (55).

Apart from this specific role, the insights provided by applying the sCLIP protocol developed here, might have broader medical implications. SnRNPs play crucial roles in many physiologically important processes. Accordingly, deregulation of their complex biogenesis typically results in most devastating consequences. This can globally be exemplified by Spinal muscular atrophy in which mutations in the survival motor neuron-1 gene result in the degeneration of spinal motor neurons and severe muscle wasting. Normally, the SMN protein acts as a key factor in the assembly of Sm-snRNPs, and probably also snoRNPs as well as other RNP s (56,57). More specifically, defects in snRNAs have recently been shown to result in serious disorders such as microcephalic osteodysplastic primordial dwarfism type I (MOPD I), a severe developmental disorder characterized by extreme intrauterine growth retardation and multiple organ abnormalities (58), or neurodegeneration (59). Thus, a dysfunction of essential factors in messenger RNA splicing is not only relevant for proper RNA processing; it can also result in most significant pathophysiologicals.

Based on these data it is tempting to speculate that the mechanism that we identified here may—when dysregulated—also lead to clinically relevant phenotypes.
In this context, it is noteworthy that CSTF2tau often shows high expression in numerous tumor entities across various organ systems (www.proteinatlas.org). This invites speculations that aberrant CSTF2tau may influence the turnover of snRNAs in these conditions, which in turn may influence critical downstream cellular functions in the course of serious pathophysiologicals.

Illuminating the RNA–protein interactomes is relevant for the understanding of a variety of biological processes to which the growing number of RBPs contribute to (1). Here we provide the broad biomedical scientific community a simplified and robust tool to explore RNA–protein interactomes with high resolution on limited input material while omitting some of the inherently critical and tedious experimental steps that are normally prone to result in a significant loss of material (such as purifications). The potential of this technique lies in the elucidation of novel mechanisms regulating the RNA fate from birth to decay thereby helping to illuminate the basis for potentially most devastating dysregulated processes even in systems with limited material. Apart from this, the definition of RBP target sites is key for possible novel targeted therapeutics in the future.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We thank members of the Danckwardt lab for helpful discussions. We thank EMBL Genomics Core Facility and Dr Vladimir Benes for helpful support. This work was completed in part with resources provided by the High Performance Computing Center of Mainz University.

FUNDING
DFG [DA 1189/2-1 to S.D., GRK 1591]; Federal Ministry of Education and Research [BMBF 01EO1003]; Hella Bühler Prize for Cancer Research; DGKL and the Institute of Clinical Chemistry, University Medical Center Mainz; Humboldt Research Fellowship (to M.L.). Funding for open access charge: DFG, BMBF.

Conflict of interest statement: None declared.

REFERENCES
1. Lukong, K.E., Chang, K.-w., Khandjian, E.W. and Richard, S. (2008) RNA-binding proteins in human genetic disease. Trend Genet., 24, 416–425.
2. Castello, A., Fischer, B., Hentze, M.W. and Preiss, T. (2013) RNA-binding proteins in Mendelian disease. Trend Genet., 29, 318–327.
3. Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M. et al. (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell, 149, 1393–1406.
4. Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasilieva, A., Murakawa, Y., Schueller, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M. et al. (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell, 46, 674–690.
5. Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijsveld, J., Hentze, M.W. and Kim, V.N. (2013) The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol., 20, 1122–1130.
6. Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.-M., Schwarzl, T., Curk, T., Foehr, S., Huber, W. et al. (2015) The RNA-binding proteomes from yeast to man harbor conserved endomRBPs. Nat. Commun., 6, 10127.
7. Liao, Y., Castello, A., Fischer, B., Leicht, S., Foehr, S., Frese, C., Ragan, C., Kurscheid, S., Pagler, E., Yang, H. et al. (2016) The cardiomyocyte RNA-binding proteome: links to intermediary metabolism and heart disease. Cell Rep., 16, 1456.
8. Huberta, D. and van der Klei, I. (2010) Moonlighting proteins: an intriguing mode of multitasking. Biochim. Biophys. Acta, 1803, 520.
9. Castello, A., Hentze, M.W. and Preiss, T. (2015) Metabolic enzymes enjoying new partnerships as RNA-binding proteins. Trends Endocrinol. Metab., 25, 746–757.
10. Danckwardt, S., Gantzert, A.S., Macher-Goeppinger, S., Probst, H.C., Gentzel, M., Wilm, M., Grone, H.J., Schirrmacher, P., Hentze, M.W. and Kulozik, A.E. (2011) p38 MAPK controls prothrombin expression by regulated RNA 3′ end processing. Mol. Cell, 41, 298–310.
11. Ule, J., Jensen, K., Mele, A. and Darnell, R. (2005) CLIP: a method for identifying protein–RNA interaction sites in living cells. Methods, 37, 376.
12. Hafner, M., Landthaler, M., Burger, L., Khoshshirid, M., Hauser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M. et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell, 141, 129–141.
13. Huppertz, I., Attig, J., D’Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J. and Ule, J. (2014) iCLIP: protein–RNA interactions at nucleotide resolution. Methods, 65, 274–287.
14. Licatalosi, D.D., Mele, A., Fak, J.I., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X. et al. (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature, 456, 464–469.
15. Flynn, R.A., Martin, L., Spitale, R.C., Do, B.T., Sagan, S.M., Zarnegar, B., Qu, K., Khavari, P.A., Quake, S.R., Sarnow, P. et al. (2015) Dissecting noncoding and pathogen RNA–protein interactomes. RNA, 21, 135–143.
16. Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundaramaran, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K. et al. (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods, 13, 508–514.
17. Zarnegar, B.J., Flynn, R.A., Shen, Y., Do, B.T., Chang, H.Y. and Khavari, P.A. (2016) iuCLIP platform for efficient characterization of protein–RNA interactions. Nat. Methods, 13, 489–492.
18. Saulière, J. and Le Hir, H. (2015) In: Rederstorff, M (ed). Non-Coding RNAs: Methods and Protocols. Springer, NY, pp. 151–160.
19. Bohnsack, M.T., Martin, R., Granneman, S., Ruprecht, M., Schleiff, E. and Tollervey, D. (2009) Pпр3 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. Mol. Cell, 36, 583–592.
20. Zhuang, F., Fuchs, R.T., Sun, Z.Z., Zheng, Y. and Robb, G.B. (2012) Structural bias in T4 RNA ligase-mediated 3′-adapter ligation. Nucleic Acids Res., 40, 654.
21. Yao, C., Choi, E.-A., Weng, L., Xie, X., Wan, J., Xing, Y., Moresco, J.J., Tu, P.G., Yates, J.R. and Shi, Y. (2013) Overlapping and distinct functions of CstF64 and CstF64′ in mammalian mRNA 3′ processing. RNA, 19, 1781–1790.
22. Jänicke, A., Vancuyenberg, J., Boag, P.R., Traven, A. and Beilharz, T.H. (2012) ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3′ RACE applications. RNA, 18, 1289–1295.
23. Yehuda-Reshkev, S. and Schuster, G. (2000) Characterization of the E.coli poly(A) polymerase: nucleotide specificity, RNA-binding affinities and RNA structure dependence. Nucleic Acids Res., 28, 1139–1144.
24. Huppertz, I., Attig, J., D’Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J. and Ule, J. (2014) iCLIP: protein–RNA interactions at nucleotide resolution. Methods, 65, 274–287.
