



## OPEN Quantitative detection of pseudouridine in RNA by mass spectrometry

Shanice Jessica Hermon<sup>1,2</sup>, Anastasia Sennikova<sup>1,2</sup> & Sidney Becker<sup>1,2</sup>✉

Pseudouridine ( $\Psi$ ) is one of the most prevalent and dynamic modification in RNA, and was shown to evade the host immune response in mRNA vaccines. Despite its significance, the biological role of  $\Psi$  remains poorly understood as certain key limitations and challenges in the detection of  $\Psi$  are yet to be overcome. In account of this, we report the usage of a chemical labelling strategy for the first quantitative detection of  $\Psi$  by mass spectrometry. We demonstrate a labelling efficiency exceeding 99% in isolated yeast tRNAs hosting multiple  $\Psi$ s. LC-MS/MS analysis enables precise mapping of  $\Psi$  at single-base resolution, while simultaneously capturing a wide array of additional post-transcriptional modifications, which is not achieved with current sequencing technologies. This advancement may help unravel the dynamics and biological implications of  $\Psi$ , shedding light on its interplay with other modifications and deepening our understanding of its functional role.

**Keywords** RNA modifications, Mass spectrometry, Pseudouridine, tRNA

Post-transcriptional modifications are involved in fundamental cellular processes. They modulate gene expression by fine-tuning transcription, RNA processing and protein synthesis, thereby establishing multiple regulatory layers. Therefore, their spatiotemporal deposition is vital for normal human development and dysregulation is connected to pathology. So far, however, the functional implications of most RNA modifications are largely elusive. More than 150 different modifications have been discovered in both coding and noncoding RNAs<sup>1</sup>. They alter the physicochemical properties of nucleic acids to influence RNA structure, function and stability<sup>1,2</sup>.

Pseudouridine ( $\Psi$ ) is one of the most prevalent and dynamic modification, found in many functional RNA species including tRNAs, rRNAs and snRNAs<sup>3–5</sup>.  $\Psi$  is installed by the pseudouridine synthases (PUS), from isomerization of uridine, which creates an additional hydrogen bond donor at the N1-position<sup>3</sup>. These structural changes improve base stacking and stabilize the folded RNA structure<sup>6</sup>.  $\Psi$  is believed to be involved in regulating translation<sup>5,7</sup>, splicing<sup>5,8</sup>, RNA–protein interactions<sup>5</sup> and in modulating environmental stress response<sup>3,9</sup>. Investigating  $\Psi$ 's distribution, dynamics and interdependency with other epitranscriptional modifications will help to provide a better mechanistic understanding to uncover its exact biological role. This, however, requires reliable detection methods at single base resolution.

Current approaches to map  $\Psi$  are mainly based on mass spectrometry (LC–MS) or next generation sequencing (NGS) technologies, both of which are unable to directly detect  $\Psi$ <sup>10</sup>. Consequently, NGS methods mostly utilize chemical labelling of  $\Psi$  followed by its indirect detection from deletion signatures from reverse transcription.  $\Psi$  detection, however, varies significantly between the different NGS methods and shows strong sequence context dependency, causing limited overlap of the identified  $\Psi$ -sites<sup>11</sup>. Moreover, as NGS methods follow an indirect mode of detection, the extend of false positives/negatives are high, especially if  $\Psi$ -sites are near uridines<sup>12–14</sup>. These methods cannot provide quantitative estimates of  $\Psi$  at a given site. Additionally, as NGS methods can only analyze a limited number of modifications in parallel, the effect of these methods on other RNA modifications is therefore not well established. These concerns have led to the recent developments of nanopore sequencing which identifies  $\Psi$  sites as high confidence U to C base calling errors<sup>15–19</sup>. Many of these methods, however, cannot truly discriminate  $\Psi$  from other U-derived modifications such as Dihydrouridine (D) and 2'-O-methyluridine (Um)<sup>15</sup>. To overcome these limitations, subsequent approaches utilize chemically assisted nanopore sequencing for the detection of  $\Psi$  as in-del signatures<sup>20</sup>. Nevertheless, the sequence context dependency of  $\Psi$  in base calling and the occurrence of false positives inevitably lead to high error rates in nanopore sequencing methods<sup>16–20</sup>.

<sup>1</sup>Max-Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. <sup>2</sup>Department of Chemistry and Chemical Biology, Technical University Dortmund, 44227 Dortmund, Germany. ✉email: sidney.becker@tu-dortmund.de; sidney.becker@mpi-dortmund.mpg.de

As an orthogonal strategy, mass spectrometry-based analysis is a powerful tool that allows for the direct, sequence independent detection of multiple modifications in parallel, thereby significantly reducing potential biases. As  $\Psi$  exhibit the same mass as uridine, most mass spectrometry-based methods rely on chemical derivatization strategies to generate a mass tag on  $\Psi$  for its detection. A widely used method is the chemical derivatization with CMC (N-cyclohexyl-N'- $\beta$ -(4-methylmorpholinium) ethylcarbodiimide), which generates a mass tag of 252 Da<sup>21</sup>, while acrylonitrile provides a mass difference of 53 Da<sup>22–24</sup>. However, the labelling efficiency of these reported methods is significantly low, leading to the loss of valuable quantitative information. In addition, false positives are observed due to uridine labelling. CMC is also reported to label 2-methylthio-N6-isopentenyladenosine<sup>25</sup>, complicating analysis. Alternatively, stable isotope labelling for  $\Psi$  quantification<sup>26</sup> and characteristic  $\Psi$ -specific collision induced dissociation (CID) pattern for  $\Psi$ -detection are also reported<sup>27,28</sup>. These methods, however, can reliably detect  $\Psi$  only in short RNA sequences in the absence of neighboring uridines.

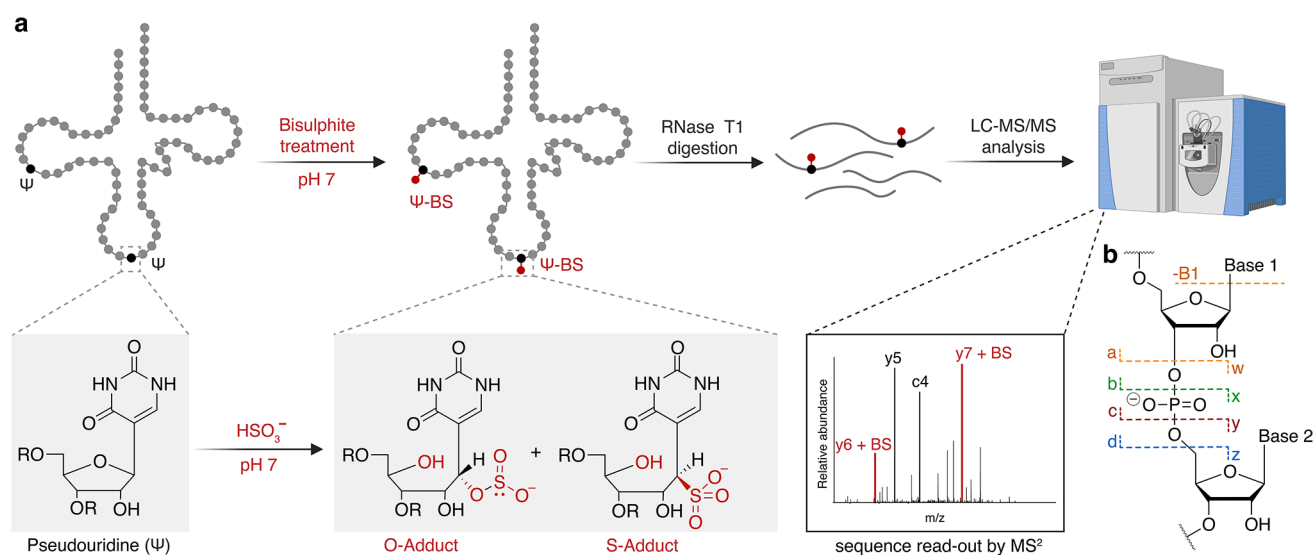
The key limitations and challenges that are yet to be overcome calls for the development of efficient methods for both qualitative and quantitative  $\Psi$ -detection. The parallel read-out of various modifications concurrently can shed light on  $\Psi$ s interplay with other post-transcriptional RNA modifications. Therefore, we considered potential  $\Psi$  labelling reagents that could provide quantitative sequence read-out by mass spectrometry. A promising candidate is bisulphite (BS), which is known to convert cytosine to uridine (C-to-U) at acidic pH. Bisulphite reacts covalently with  $\Psi$  as well, attaching to the C1' of the ring opened ribose via the sulphur (S adduct) or oxygen (O adduct) atom of the bisulphite nucleophile (Fig. 1a). This reaction has been applied in recent NGS methods for the indirect detection of  $\Psi$ <sup>29–32</sup>. If the reaction is performed at neutral pH, labelling of  $\Psi$  occurs without C-to-U conversion. Slightly basic pH treatment reverses any unspecific bisulphite label on uridine or cytosine nucleosides<sup>30,33</sup>.

Thus, we explored the specificity and efficiency of bisulphite labelling on  $\Psi$ -containing RNA oligonucleotides. We further evaluated if the labelling could be exploited for the efficient and quantitative mass spectrometry-based detection of  $\Psi$ . For this, we first optimized labelling and sample preparation procedures as well as LC-MS/MS methodologies on RNA oligonucleotides. Upon comparison with existing LC-MS based  $\Psi$ -labelling methods, we found that bisulphite labelling shows superior selectivity and efficiency, which allows for the first quantitative detection of  $\Psi$  at single base resolution. We further show the compatibility of bisulphite labelling within biological tRNA samples. Isolating tRNA<sup>His</sup>, tRNA<sup>Leu</sup> and tRNA<sup>Arg</sup> from yeast tRNA extract allowed for LC-MS analysis of the intact tRNAs. RNase T1 digestion and subsequent LC-MS/MS analysis of tRNA<sup>His</sup> and tRNA<sup>Leu</sup>, enabled the identification of the exact sequence position of bisulphite labelled  $\Psi$  in parallel with all other RNA modifications that are present in the mentioned tRNAs.

## Results

### Detection and quantification of $\Psi$ in synthetic RNA oligonucleotides

For initial trials, we tested the efficiency of bisulphite labelling (pH 7) on synthetic 30-mer RNA oligonucleotides with one or more  $\Psi$ s in various sequence contexts. As negative control, we performed labelling on the exact same RNA oligonucleotide sequence with uridine (U) instead of  $\Psi$ . Following labelling, the samples were subjected



**Fig. 1.** Overview of the bisulphite labelling strategy used for the LC-MS/MS detection of pseudouridine ( $\Psi$ ) in tRNA. **(a)** Bisulphite labelling of  $\Psi$  in tRNA generates either the S adduct or the O adduct, both of which are indistinguishable in LC-MS/MS analysis. Following labelling, the tRNAs are treated with RNase T1 to generate smaller fragments for analysis. In the MS/MS analysis, all CID fragments which hosts the  $\Psi$ -BS adducts show a mass shift of 82 Da, enabling the identification of  $\Psi$  in the sequence. **(b)** CID fragmentation pattern of RNA producing 9 major fragment ion series (*a*, *a-B*, *b*, *c*, *d*, *w*, *x*, *y*, and *z*). Part of the figure is created with BioRender.com.

to slightly basic pH conditions (pH 9) for desulphonation to remove any non-specific label. Omitting this step results in the detection of false positives (Supplementary Fig. S1). We further confirmed that no degradation of the RNA oligonucleotides is observed after bisulphite treatment (Supplementary Fig. S2).

In all tested  $\Psi$ -containing RNAs, we observed quantitative labelling regardless of the sequence context or number of  $\Psi$ s (Fig. 2, Supplementary Figs. S3–S5). The bisulphite label introduces a mass tag of 82 Da onto  $\Psi$ , which is stable under ESI–MS conditions and can be detected by intact mass analysis. The use of an RNA oligonucleotide containing three  $\Psi$ s shows complete labelling with a mass shift of 246 Da ( $3 \times 82$  Da), confirming the presence of the expected number of  $\Psi$ s. In contrast, the mass of the control oligonucleotide (where  $\Psi$  is replaced by U) remained unchanged following bisulphite treatment (Fig. 2a,b). This confirmed the high specificity of the label. We further tested for C-BS and U-BS adduct formation in the currently optimized conditions. For this, the bisulphite labelling procedure was tested on multiple RNA oligonucleotides with varying sequence contexts of U and C. Nonspecific labelling was not observed (Supplementary Fig. S6). In addition, no cytosine deamination was detected under neutral bisulphite treatment (pH 7). It is known, however, that C-to-U conversion takes place at lower pH<sup>34</sup>. Therefore, as a positive control we confirmed cytosine deamination with a commercially available RNA bisulphite kit that uses acidic pH conditions (Supplementary Fig. S6).

To verify if the bisulphite labelling is indeed quantitative, we tested the labelling efficiency in mixed oligonucleotide samples containing different ratios of the  $\Psi$ -containing RNA vs. control sequence. The samples were treated with bisulphite and the intact mass of the labelled and unlabeled sequences were quantified by signal integration. The expected fraction was plotted against the measured fraction (Supplementary Table S1). Fitting a linear regression provided a slope of 1.00, confirming highly specific and quantitative  $\Psi$  labelling. Indeed, the experimental values showed a standard deviation  $< 6\%$  (from three replicates) for all tested fractions (Fig. 2c). This confirmed that bisulphite labelling at neutral pH enables robust and quantitative detection of  $\Psi$  without effecting other canonical bases in the RNA.

Next, we confirmed the exact sequence position of  $\Psi$ . For this, we first optimized MS parameters for collision induced dissociation (CID) of RNA oligonucleotides, to obtain full sequence coverage (Supplementary Fig. S7). CID generates *a/a-B*, *b*, *c*, *d*, *w*, *x*, *y* and *z* type ions in a ladder-like fashion, with the *c* and *y* ions most common for RNA<sup>35,36</sup> (Fig. 1b). As  $\Psi$  and U containing RNA oligonucleotides cannot be discriminated by intact mass, the fragment ions generated upon CID are also identical.  $\Psi$ -labelling with bisulphite will not only change the intact mass but also create altered fragment ions upon CID fragmentation. After optimizing MS-parameters, we confirmed that the bisulphite label is stable under these MS/MS conditions to identify its exact sequence position (Fig. 2d). The control oligo (without  $\Psi$ ) does not have an identifiable label in any of the identified fragments (Fig. 2e). In the labeled oligo (with  $\Psi$ ), the fragments *c3*, *y11* and *x6* differs from *b2*, *y10* and *w5*, by an additional mass of the  $\Psi$ -BS adduct (+ 82 Da), perfectly demarcating the position of  $\Psi$  in the sequence (Fig. 2d,e). Overall, all 30-mer RNA oligonucleotides showed 100% sequence coverage and stable bisulphite labelling upon optimized CID fragmentation, allowing to obtain the exact  $\Psi$  position within the sequence (Fig. 2f,g, Supplementary Figs. S3–S5).

### Comparison with reported labelling strategies

Next, we tested how the bisulphite labelling compares to previously reported  $\Psi$ -labelling methods. Of these, the CMC labelling strategy and cyanoethylation are the most commonly applied methods<sup>21,24,25</sup>. Both methods have contributed significantly in the detection of novel  $\Psi$  sites<sup>37</sup>. In account of this, we compared the efficiency and specificity of bisulphite, CMC and acrylonitrile derivatization for LC–MS detection of  $\Psi$  (Fig. 3).

We first treated an oligonucleotide carrying two  $\Psi$ s with bisulphite and observed a single mass peak. The peak corresponds to a mass shift of 164 Da from the unlabeled intact mass, confirming a complete conversion of  $\Psi$  to  $\Psi$ -BS adducts. No false positives or partially labelled species were observed upon bisulphite treatment (Fig. 3b).

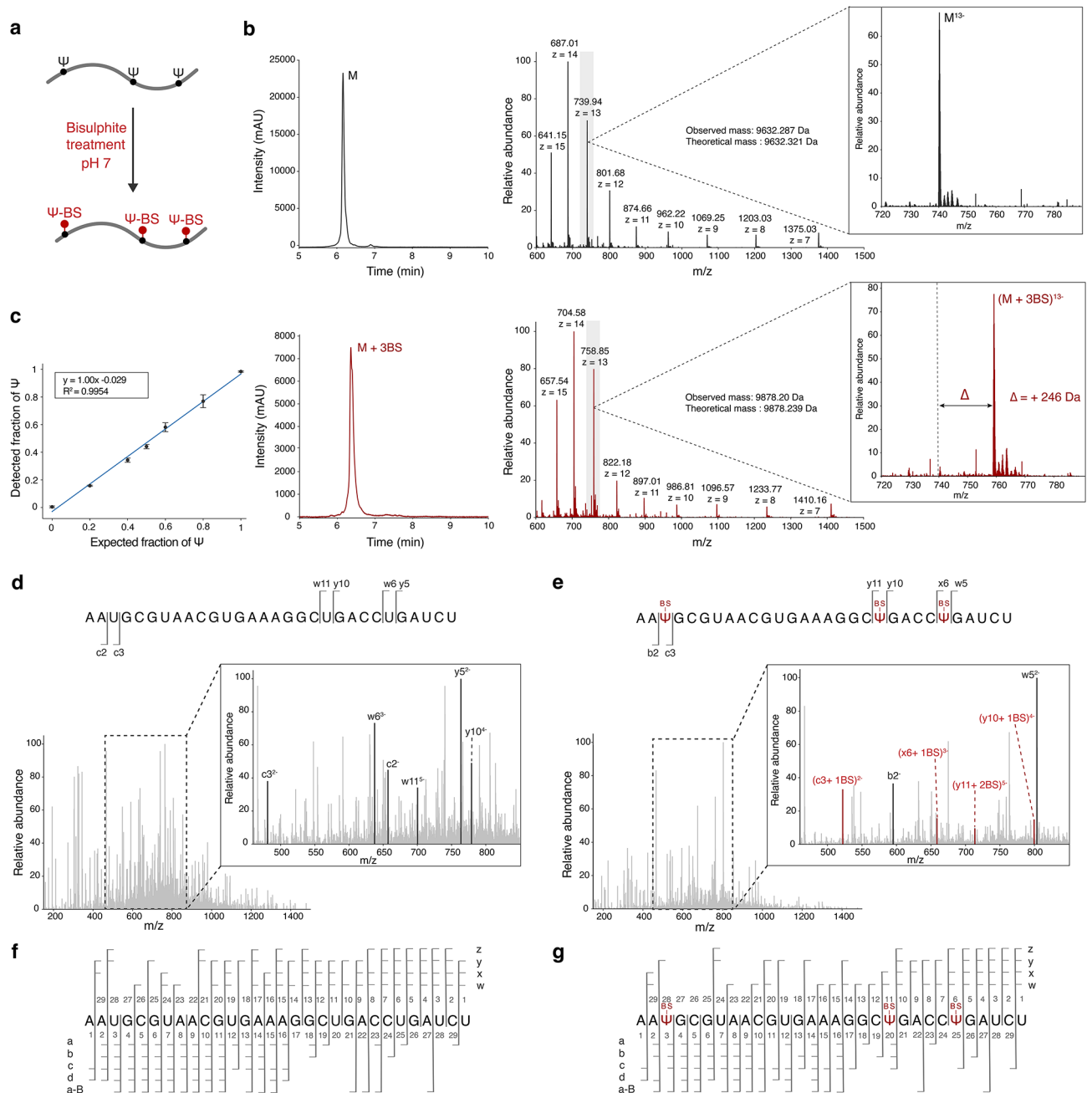
Next, we performed CMC labelling on the same oligonucleotide (containing two  $\Psi$ s), using reported experimental conditions<sup>37–39</sup>. Derivatization with CMC modifies the N1 or N3 of  $\Psi$ , N1 of guanine and N3 of uridine. Further treatment at a pH of 10.4 for 4 h removes the CMC from guanine and uridine, maintaining the specificity to  $\Psi$ <sup>25</sup>. This generates a  $\Psi$ -mass tag of 252 Da, enabling its discrimination from uridine. Following CMC treatment, we observe the labelling to be incomplete as multiple LC–MS peaks are identified corresponding to unlabeled, single labelled and double labelled oligonucleotide. A triple labelled specie suggests formation of false positives (Fig. 3c). Even though CMC is widely used, this method suffers from several limitations including potential false negatives/positives due to low labelling efficiency and specificity<sup>25</sup>. The high pH conditions for reversing unspecific CMC labelling create multiple degradation products, thus reducing overall data quality (Fig. 3c).

Lastly, we tested acrylonitrile labelling of  $\Psi$ , which generates 1-cyanoethyl- $\Psi$  with a mass shift of 53 Da from unlabeled  $\Psi$ <sup>24</sup>. Upon treatment with acrylonitrile following a reported procedure<sup>24</sup>, multiple species with two, three, four and five labels, differing by a mass of 53 Da were observed (Fig. 3d). This suggests that the method generates significant false positives. A reason could be the nonselective derivatization of uridine<sup>23</sup>. In addition, non-labelled as well as single labelled oligonucleotides are observed, which demonstrates low labelling efficiency.

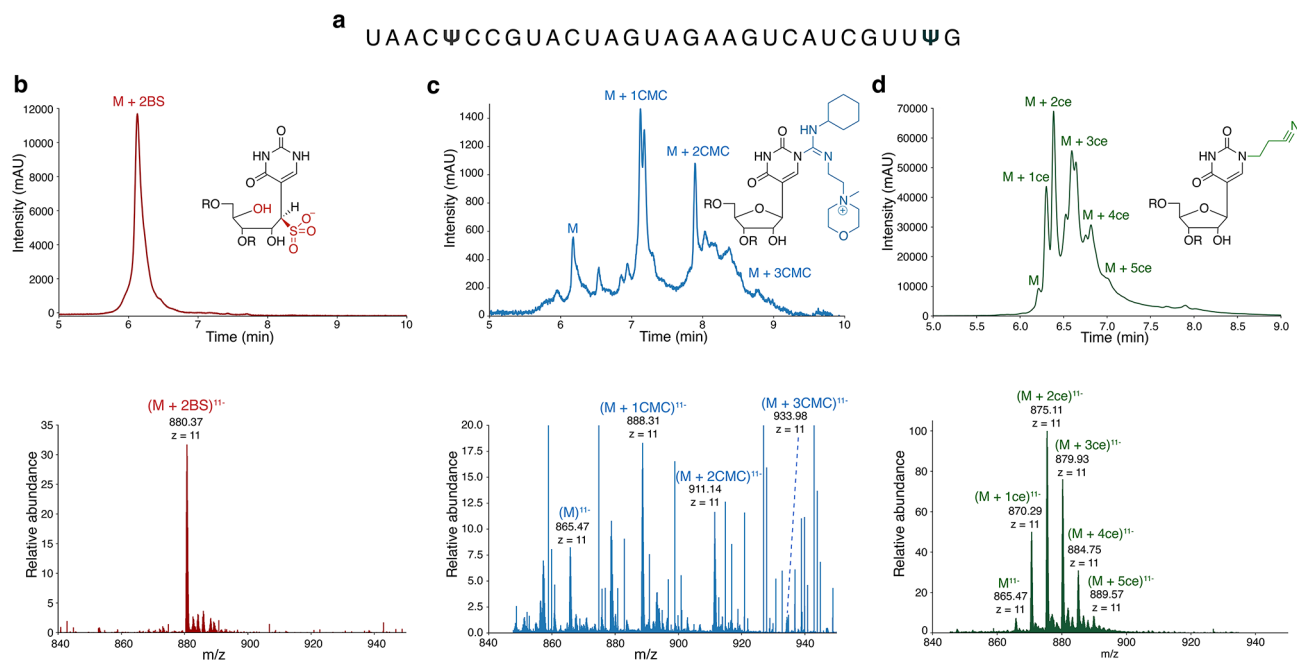
We repeated the three labelling reactions (bisulphite, CMC and acrylonitrile) on another oligonucleotide to confirm that the results are reproducible (Supplementary Fig. S8). Overall, our experiments suggest that bisulphite labelling surpasses existing mass spectrometry based pseudouridine detection strategies in terms of efficiency and specificity. This allows for the first direct and quantitative read-out of  $\Psi$  in RNA.

### Detection of $\Psi$ in isolated yeast tRNAs

Finally, we performed labelling on biologically relevant tRNA samples, which are heavily modified. Therefore, the effect of bisulphite labelling on other RNA modifications can be assessed. In addition, it allows for the base



**Fig. 2.** LC-MS/MS analysis of bisulphite treated synthetic RNA oligonucleotides. **(a)** Schematic representation of the bisulphite treatment of synthetic RNA oligonucleotides RNA1U and RNA1P. **(b)** top: HPLC profile (left) and MS1 spectra (right) of the control RNA oligonucleotide (RNA1U) devoid of Ψ, following treatment with bisulphite. The multiple charged negative ion series of the intact RNA1U is shown. The mass peak of 739.94, with a charge state of -13 ( $M^{13-}$ ) shown in the inset. bottom: HPLC profile (left) and MS1 spectra (right) of the RNA oligonucleotide hosting three Ψs in parallel (RNA1P), following treatment with bisulphite. The multiple charged negative ion series of the bisulphite labelled intact RNA1P is shown. The mass peak of 758.85, with a charge state of -13 ( $(M + 3BS)^{13-}$ ) is shown in the inset. The mass peak of the labelled RNA1P shifts from the control RNA1U (as indicated by the dotted line) by 246 Da. **(c)** Quantification of Ψ from sample mixtures of control RNA2U and RNA oligonucleotide with a single Ψ (RNA2P) in variable ratios. The sequences of the control RNA-RNA1U and the RNA with Ψ (RNA2P) is GUUCCCAUCCGUAGCUAUGAAGACGUGCGU and GUUCCCAUCCGUAGCUAΨGAAGACGUGCGU respectively. A fitting curve is plotted based on the observed Ψ fraction and the expected Ψ fraction. **(d, e)** CID spectra of the control (RNA1U) and the RNA with Ψ (RNA1P) treated with bisulphite (sequences mentioned at the top). The precursor ions for fragmentation are  $M^{13-}$  and  $(M + 2BS)^{13-}$  respectively. Assignment of the CID fragment ions reveal the positions of the Ψ at 3, 20 and 25 as expected. **(f, g)** Representation of all fragment ions identified in both RNA1U and RNA1P following BS treatment.

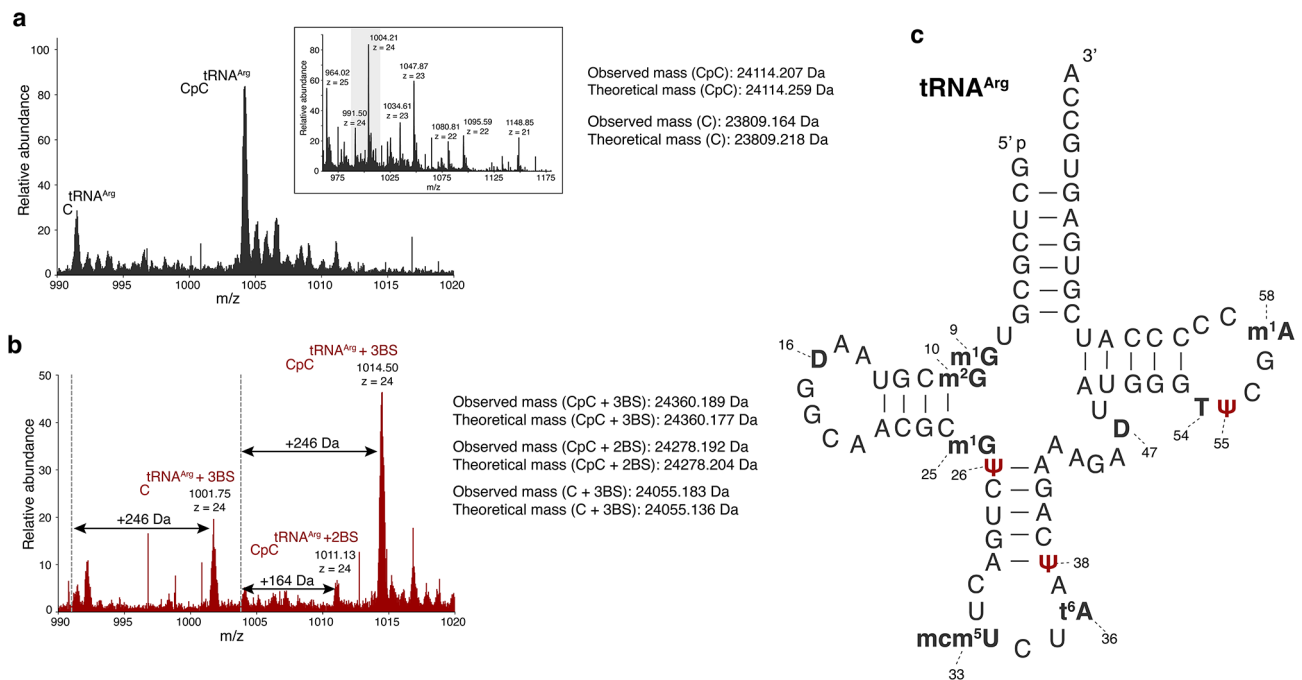


**Fig. 3.** Comparison of the specificity and efficiency of labelling  $\Psi$  by CMC, acrylonitrile and bisulphite. **(a)** RNA sequence with two  $\Psi$ s in parallel (RNA3P) used for all three labelling studies. **(b)** HPLC profile of the bisulphite labelled RNA3P showing a single clean peak and the chemical structure of  $\Psi$ -BS adduct (top). MS1 spectra of the BS labelled RNA3P showing a single clean peak with a mass of 880.37, with a charge state of  $-11$  (bottom). **(c)** HPLC profile of the CMC labelled RNA3P, and the chemical structure of the CMC labelled  $\Psi$  (top). MS1 spectra of the CMC labelled RNA3P, the mass peak of the intact mass of the RNA3P is labelled as  $M^{11-}$ , CMC labelled peaks are represented as  $(M + n \text{ CMC})^{11-}$ , where  $n$  = number of CMC labels on the RNA. **(d)** HPLC profile of the acrylonitrile (ce) RNA3P, and the chemical structure of the acrylonitrile labelled  $\Psi$  (top). MS1 spectra of the acrylonitrile labelled RNA3P, the mass peak of the intact mass of the RNA is labelled as  $M^{11-}$ , cyanoethylated peaks are represented as  $(M + n \text{ ce})^{11-}$ , where  $n$  = number of acrylonitrile labels on the RNA.

resolution read-out of  $\Psi$  in parallel with all other post-transcriptional tRNA modifications. We used yeast tRNA extract and analyzed its composition by LC-MS to identify the most abundant tRNA species for isolation. Using biotin labelled DNA oligonucleotide probes targeting the respective tRNAs, we performed streptavidin-based pull down to isolate tRNA<sup>His</sup>, tRNA<sup>Leu</sup> and tRNA<sup>Arg</sup> from a yeast tRNA mixture. The identity of the isolated tRNAs were confirmed by LC-MS analysis by comparing the respective masses of the tRNA variants (3'-end: CCA, CC or C) to the masses of the sequences reported in the database<sup>40,41</sup>. All tRNA variants were identified with an error of less than 3 ppm (Fig. 4, Supplementary Figs. S9,S10).

Most tRNA modifications are methylations such as  $m^2C$ ,  $m^1G$ ,  $m^2G$ ,  $m^{22}G$ , Am, which are not expected to be altered by bisulphite (Fig. 4, Supplementary Figs. S9, S10). The three isolated tRNAs (tRNA<sup>His</sup>, tRNA<sup>Leu</sup> and tRNA<sup>Arg</sup>) also host dihydrouridines (D) at multiple sequence positions and thymine in the T loop (Fig. 4, Supplementary Figs. S9, S10). In addition, tRNA<sup>Arg</sup> contains mcm5U (5-methoxycarbonylmethyluridine) and t6A (N6-threonylcarbamoyl-adenosine) (Fig. 4)<sup>40,41</sup>. Upon bisulphite labelling, the intact mass analysis of all three tRNAs suggests the presence of three  $\Psi$ s in each of the sequences, as reported in the tRNA database (Fig. 4, Supplementary Figs. S9,S10). For tRNA<sup>His</sup>, we observed a complete conversion from the unlabeled to the triple labelled specie (Supplementary Fig. S9). In contrast, tRNA<sup>Arg</sup> provides two differently labelled species, hosting either two  $\Psi$ s (7.82%) or three  $\Psi$ s (92.18%). In tRNA<sup>Leu</sup>, we assumed the loss of an acetyl group (from ac<sup>4</sup>C) since the mass difference was only 204 Da (3BS - ac) instead of the expected 246 Da (Supplementary Fig. S10). However, all other modifications present in tRNA<sup>His</sup>, tRNA<sup>Leu</sup> and tRNA<sup>Arg</sup> were not affected by the labelling.

We finally verified that our bisulphite label is found in the expected  $\Psi$ -positions and further profiled all other post-transcriptional modifications within tRNA<sup>His</sup> and tRNA<sup>Leu</sup>. The LC-MS/MS analysis of full length tRNA is currently problematic due to technical limitations such as sensitivity. Therefore, we performed a partial digestion by RNase T1 to obtain shorter RNA sequences. Chromatographic separation and CID-fragmentation allowed us to accurately read the full tRNA sequence including its modifications (Fig. 5a,b). For analysis, we only considered the MS/MS data of RNase digested fragments with an error of less than 10 ppm. For unlabeled and labelled tRNA<sup>His</sup> we found in each case 100% sequence coverage (Fig. 5c,d, Supplementary Fig. S11). Moreover, all tRNA modifications were identified, which are covered by multiple digestion products. The data confirms that the three bisulphite labels are mapped to position 14, 33 and 55, the expected  $\Psi$ -positions provided in the tRNA modification database (Fig. 5e)<sup>40,41</sup>. Finally, we performed the same experiment with tRNA<sup>Leu</sup>, not only to confirm the  $\Psi$ -positions but also the loss of the acetyl group of ac<sup>4</sup>C in position 12. Again, the sequence was fully mapped, confirming the correct positions of tRNA alteration (Supplementary Fig. S12).



**Fig. 4.** LC-MS analysis of bisulphite labelled tRNA<sup>Arg</sup> (a) LC-MS analysis of the intact tRNA<sup>Arg</sup>. The mass peaks of 1004.21, and 991.50 with charge states of -24, of the unlabeled intact tRNA is shown. The series of multiply charged negative ions of the intact tRNA<sup>Arg</sup> is shown in the inset. (b) LC-MS analysis of the bisulphite labelled intact tRNA<sup>Arg</sup>, the labelled mass peaks, with charge states of -24 is shown. The mass peak of the labelled tRNA<sup>Arg</sup> shifts from the unlabeled one (as indicated by the dotted line) by 246 Da or 164 Da. (c) Secondary structure of yeast tRNA<sup>Arg</sup> (database ID: tdbR00000370)<sup>40,41</sup>. The positions of all modified residues are marked. The symbols for the modified nucleosides are as follows: m<sup>1</sup>G, 1-methylguanosine; m<sup>2</sup>G, 2-methylguanosine; Ψ, Pseudouridine; mcm<sup>5</sup>U, 5-methoxycarbonylmethyluridine; t<sup>6</sup>A, N6-threonylcarbamoyladenine; D, Dihydrouridine; T, Thymidine; m<sup>1</sup>A, 1-methyladenine.

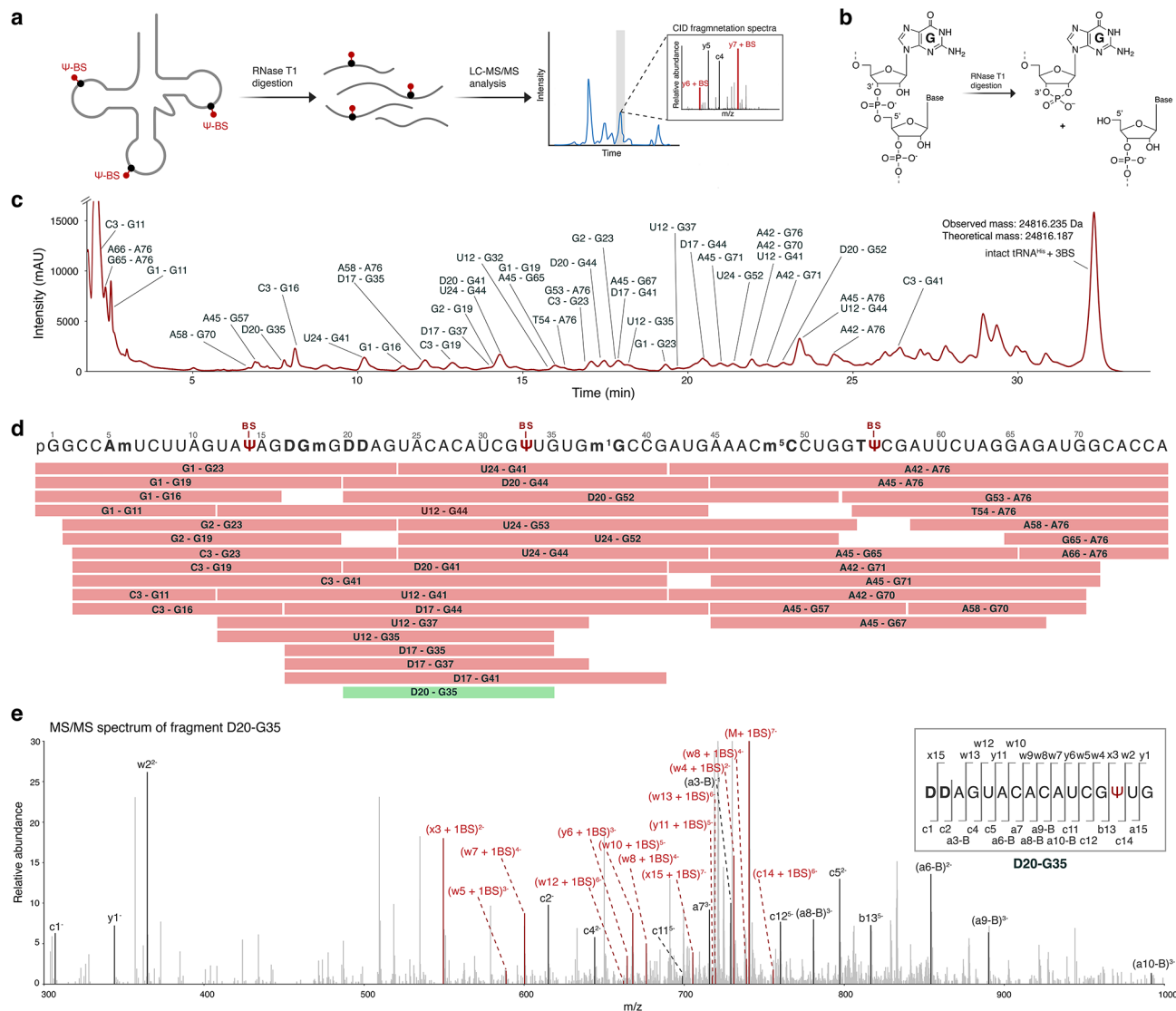
In summary, we show that bisulphite labelling allows for the quantitative and specific detection of Ψ in synthetic and biological RNA samples by LC-MS/MS at single base resolution. Additionally, numerous other post-transcriptional modifications can be assessed. Even the loss of a modification upon bisulphite labelling was mapped to the correct position, illustrating the profiling of modification dynamics in tRNA.

## Discussion

Ψ is one of the most abundant RNA modifications and has been linked to cellular stress responses (by increase of the RNA half-life), stem cell differentiation and cancer progression<sup>42,43</sup>. Several methods exist for the labelling and detection of Ψ by LC-MS. These methods, however, fail to provide quantitative data and show significant rates of false positives. This provides a major challenge to better understand the dynamics and abundance of Ψ in biological RNAs, which may host multiple Ψs in parallel. With an increasing number of Ψs per sequence, the efficiency of detection is significantly reduced while data analysis becomes problematic.

Here, we report a Ψ-labelling method that allows for the robust and quantitative detection of Ψ and its positional mapping by mass spectrometry in parallel with many other post-transcriptional RNA modifications. We show a near quantitative labelling efficiency, in both synthetic oligonucleotides and isolated yeast tRNAs hosting multiple Ψs simultaneously. In contrast to other LC-MS based labelling detection methods, our analysis did not detect false positives. The workflow requires only low amounts of RNA (a few picomoles), representing a major advantage in contrast to previous labelling strategies. The labelling efficiency in RNAs that contain four or more Ψs is yet to be explored. Given our data we don't expect a significant drop in labelling efficiency. With the exception of ac<sup>4</sup>C, the labelling method is compatible with all of the tested RNA modifications, but we are currently unable to test all possible modifications for their compatibility with bisulphite labelling. This will be further investigated in the future.

Unlike global quantification of RNA modifications at the nucleoside level from fully digested samples, high-resolution LC-MS/MS offers direct sequence readout of RNA oligonucleotides, providing both sequence context and relative quantification. This approach can distinguish chemical modifications based on their mass, enabling the investigation of modification patterns and interdependencies among multiple RNA modifications. It holds a significant advantage over current sequencing technologies, which are limited to detecting only a few modifications at a time. NGS methods rely on deletion or mutation signatures during reverse transcription for indirect detection of RNA modifications, which can introduce sequence context-dependent biases during library preparation. In contrast, LC-MS/MS allows for a direct mode of detection, minimizing such biases.



**Fig. 5.** LC-MS/MS analysis of bisulphite labelled tRNA<sup>His</sup> following digestion with RNase T1. **(a, b)** Schematic representation of RNase T1 digestion of tRNA and LC-MS/MS analysis. RNase T1 cleaves after guanosine, generating 5'-OH at the adjacent nucleoside residue and 2'3'-cyclic phosphate termini at the guanosine<sup>45</sup>. **(c)** HPLC profile of the RNase T1 digested fragments of the bisulphite labelled tRNA<sup>His</sup>, all identified fragments are labelled. **(d)** Sequence map showing all MS/MS identified RNase T1 digestion fragments from the bisulphite labelled tRNA<sup>His</sup>. Only fragments with a 2'3'-cyclic phosphate, full structural resolution and error < 5 ppm are considered for sequence mapping. The sequence map covers 100% of the tRNA<sup>His</sup> sequence identifying all expected modifications including labelled Ψ. **(e)** CID spectra of the fragment D20 - G35 (green fragment in **d**). The precursor ion for fragmentation is (M + 1BS)<sup>7-</sup>.

However, LC-MS sensitivity decreases for large molecules, limiting the practical oligonucleotide analysis size to approximately 50 nucleotides. This allows short RNAs, such as microRNAs or tRNA fragments (TRFs), to be analyzed without RNase digestion, while longer RNAs like tRNAs or mRNAs require enzymatic digestion to achieve adequate sequence coverage in MS/MS analysis. Notably, LC-MS analysis of mRNA vaccines has highlighted the challenges in achieving full sequence coverage<sup>44</sup>, indicating that transcriptomic analysis with current LC-MS technologies is unlikely to compete with sequencing methods. Nevertheless, LC-MS's ability to detect multiple modifications simultaneously offers valuable insights into modification patterns and even the discovery of new RNA modifications in short RNA sequences. Realizing this potential will depend on significant advancements in software development to improve analysis and data interpretation.

Although software for the MS/MS analysis of oligonucleotides exists nowadays, the development lags far behind in comparison with proteomics software solutions. Sequence-dependent analysis of nucleic acid modifications still requires substantial manual effort and is limited to known sequences. The presence of adducts, such as Na<sup>+</sup> or K<sup>+</sup>, further complicates this process. However, with continued advancements in software

solutions, LC–MS-based analysis of short nucleic acid sequences has the potential to become a cost-effective alternative to NGS and third-generation sequencing technologies.

## Methods

### Bisulphite treatment and LC–MS/MS analysis of synthetic RNA oligonucleotides

20 pmols of the RNA oligonucleotide (2  $\mu$ l of 10  $\mu$ M stock solution) was mixed with 18  $\mu$ l of freshly prepared bisulphite reagent (270 mg  $\text{Na}_2\text{SO}_3$  and 34 mg  $\text{NaHSO}_3$  dissolved in  $\text{H}_2\text{O}$  (total volume: 1 mL, pH 7) and incubated at 70 °C for 4 h. Further, 20  $\mu$ l (equal volume) of tris–HCl buffer (1.5 M, pH 9) was added to the sample mixture and incubated for 1 h at 37 °C for desulphonation. The oligonucleotides are then desalted from the sample mixture with micro bio spin P6 columns (Bio-Rad) and stored at -20 °C prior to LC–MS analysis.

#### LC–MS/MS analysis

20  $\mu$ l of the RNA sample is analyzed per injection in a Vanquish HPLC system connected to an Orbitrap Exploris 120 ESI–MS (Thermo Fisher Scientific). The sample components are separated with a flow rate of 0.3 ml/min in a solvent system of 50 mM HFIP (1,1,1,3,3,3-hexafluoro-2-propanol), 15 mM TEA (triethylamine), pH 9.00 (solvent A) and methanol (solvent B) at 80 °C on DNAPac RP 4  $\mu$ m column (2.1  $\times$  100 mm, Thermo Scientific). A multistep gradient (3% of solvent B in 0–2 min, followed by an increase from 3 to 10% until 5 min and 10 to 20% until 20 min) was used for separation. MS analysis is performed in the negative mode with a spray voltage of 2500 V, flow of sheath gas, aux gas and sweep gas maintained at 35, 7 and 0 respectively, and ion transfer tube temperature and vaporizer temperature set at 300 °C and 275 °C respectively. For MS1 analysis, the orbitrap resolution of 120,000 was used, the scan range was set from 550 to 3000 m/z, RF lens at 70%, normalized AGC target of 200% and data type set to profile mode. Intensity threshold was set to  $1 \times e^4$ .

For MS/MS analysis, masses of the peaks to be targeted for MS/MS are defined for every run, with a mass tolerance of 3 m/z. The normalized collision energies for CID are set at 15, 17 and 19. An orbitrap resolution of 120,000, normalized AGC of 200 and maximum injection time of 400 ms are used. MS/MS data processing is performed in BioPharma Finder v.05.1.

### Quantification of bisulphite labelling

A 30-mer  $\Psi$ -containing RNA oligonucleotides and the corresponding 30-mer control RNA ( $\Psi$  replaced by U) are mixed in varying ratios ( $\Psi$ : 0%, 20%, 40%, 50%, 60%, 80%, 100%). The 100%  $\Psi$  sample contains only the RNA with  $\Psi$ , while the 0%  $\Psi$  sample contains only the control RNA. The final concentration of RNA was maintained at 20 pmols in 20  $\mu$ l of bisulphite reagent. All samples were treated with bisulphite and analyzed in the LC–MS as described above, with 3 technical replicates. A fitting curve is plotted based on the observed  $\Psi$  fraction and the expected  $\Psi$  fraction.

### CMC treatment of synthetic RNA oligonucleotides

The experimental method for CMC treatment of oligonucleotides was followed as reported before<sup>37–39</sup>. 1  $\mu$ g of oligonucleotide is treated with 30  $\mu$ l of freshly prepared CMC solution (0.17 M CMC, 7 M urea, 4 mM EDTA, 50 mM bicine, pH 8.3) at 37 °C for 20 min. The treatment is stopped with the addition of 100  $\mu$ l of 0.3 M NaOAc (pH 5.6) and 0.1 M EDTA. 700  $\mu$ l of chilled ethanol was added to the mixture to precipitate the RNA oligonucleotides, the pellets are further washed twice with 70% ethanol and air dried. The pellets are resuspended in 40  $\mu$ l of 50 mM  $\text{Na}_2\text{CO}_3$  (pH 10.4) and incubated at 37 °C for 4 h (to remove nonspecific labelling), ethanol precipitated and air dried as before and stored in -20 °C prior to LC–MS analysis. The pellets are resuspended in 80  $\mu$ l of water, 20  $\mu$ l of the samples are injected per analysis in the LC–MS. The LC–MS parameters used are described as above with a modified LC gradient. A multistep gradient with 3% of solvent B from 0–2 min, followed by an increase from 3 – 10% until 5 min and 10 to 50% until 20 min was used for separating the different sample components.

### Cyanoethylation of synthetic RNA oligonucleotides

Cyanoethylation of pseudouridine in RNA oligonucleotides was performed as reported before<sup>24</sup>. 2  $\mu$ g of synthetic  $\Psi$  containing RNA is treated with 30  $\mu$ l of 41% ethanol/1.1 M triethylammonium acetate (pH 8.6) and 4  $\mu$ l of acrylonitrile and incubated at 70 °C for 2 h. After cyanoethylation, the sample mixture is lyophilized immediately, redissolved in water and the oligonucleotide is recovered by precipitation with 2 M ammonium acetate and 3 vol chilled ethanol. The recovered oligo is dried well, resuspended in 100  $\mu$ l of water and stored in -20 °C prior to LC–MS analysis. For LC–MS analysis, 20  $\mu$ l of the treated oligonucleotide is used per injection. The LC–MS parameters used for both CMC treatment and cyanoethylation experiments are the same.

### Bisulphite treatment at acidic pH

EZ RNA methylation kit (Zymo research) was used for the bisulphite treatment of RNA oligonucleotides at acidic pH. 1  $\mu$ g of the RNA was used for the experiment, which was followed according to manufactures instructions. 1/5<sup>th</sup> of the final eluted sample was submitted per injection to the LC–MS for analysis. The LC–MS parameters used are the same as used for all other bisulphite treated RNA oligonucleotides.

### Isolation of individual tRNAs

Commercially available yeast tRNA extract (Invitrogen) were used for isolations. Individual tRNAs were isolated from the tRNA extract with biotinylated DNA oligonucleotide probes. 5–10  $\mu$ l of yeast tRNA extract is mixed with 2  $\mu$ l of 100  $\mu$ M biotinylated probe oligonucleotide and 12  $\mu$ l PEG-400 and 50  $\mu$ l 1X annealing buffer (20 mM Tris–HCl, pH 7.5, 5 mM  $\text{MgCl}_2$ , 0.1 mM DTT, 0.4% Triton X-100), incubated at 90 °C for 5 min for denaturation, followed by a stepwise descend to 25 °C in 45 min for hybridization. To the sample mixture, 50  $\mu$ l of 5X binding/

wash buffer (50 mM Tris HCl (pH 7.5), 5 mM EDTA, 2.5 M NaCl), 100  $\mu$ l of streptavidin beads (Miltenyi Biotec) and 258  $\mu$ l of water was added and incubated at 45 °C for 1 h for binding to the microbeads. The magnetically labelled DNA-tRNA complex is passed through the  $\mu$ MACS column following manufacturer's instructions, followed by washes (3 ml). The tRNA-probe complexes are eluted from the column with preheated (80 °C) elution buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA) in 3 stages of 150  $\mu$ l each, then pooled together and vacuum dried. To generate a larger quantity of tRNAs to aid in optimizations, tRNAs were isolated in multiple rounds. The elutions are pooled together and further HPLC purified to separate the intact tRNA from smaller fragments. The purified tRNAs are lyophilized and resuspended in water, stored at -20 °C for long term storage. The concentration of the resuspended tRNAs is measured with Qubit.

The biotinylated DNA probes used for the isolation of the respective tRNAs are as follows:

tRNA<sup>His</sup> (tdbR00000145): [BtN]GGCCACAACGATGTGTACTAACCCTATACTAAGATGGCC.  
 tRNA<sup>Arg</sup> (tdbR00000370): [BtN]TGATTAGAAGTCAGACGCGTTGCCATTACGCCACGCGAGC.  
 tRNA<sup>Leu</sup> (tdbR00000249): [BtN]AGCTTGAATCAGGCGCCTTAGACCGCTCGGCCAAACAACC.

### Bisulphite treatment of isolated tRNAs

50–150 ng of the isolated tRNA is mixed with 20  $\mu$ l of bisulphite reagent (see above) and incubated at 70 °C for 4 h. Further, 20  $\mu$ l (equal volume) of tris-HCl buffer (1.5 M, pH 9) was added to the sample mixture and incubated for 1 h at 37 °C for desulphonation. The tRNAs are then desalted from the sample mixture with micro bio spin P6 columns (Bio-Rad) and vacuum dried. The samples are further resuspended in 30  $\mu$ l of RNase free water and submitted in the LC-MS as a single injection. The LC-MS parameters used are described as above with a modified LC gradient. A multistep gradient starting with 0% of solvent B for 2 min, followed by an increase from 0 – 6% until 20 min, 6 to 10% until 30 min and, 10 to 80% until 35 min was used.

### Bisulphite treatment and RNase T1 digestion of tRNA<sup>His</sup> and tRNA<sup>Leu</sup>

700–1000 ng of the isolated tRNA is mixed with 20  $\mu$ l of bisulphite reagent and incubated at 70 °C for 4 h. Further, 20  $\mu$ l (equal volume) of tris-HCl buffer (1.5 M, pH 9) was added to the sample mixture and incubated for 1 h at 37 °C for desulphonation. The tRNA is then desalted from the sample mixture with micro bio spin P6 columns (Bio-Rad) and vacuum dried prior to RNase digestion. Further, the tRNA is resuspended in 40  $\mu$ l of digestion buffer and 1.15  $\mu$ l of 6X diluted RNase T1 beads (Thermo Scientific™ SMART Digest™ RNase Kits) is added to the sample. Mixing and incubation at 37 °C for 15 min allows for partial digestion. The RNase T1 beads are removed from the sample mixture via magnetic separation and the sample mixture is analyzed by LC-MS/MS. As a control, 500 ng of the intact tRNA<sup>His</sup>, without any bisulphite treatment is subjected to RNase T1 digestion and analyzed by the LC-MS/MS.

The RNase digested tRNA fragments were analyzed in the LC-MS as described above with a modified LC gradient. A multistep gradient starting with 0% of solvent B for 2 min, followed by an increase from 0 to 3% until 15 min, 3 to 10% until 40 min and, 10 to 80% until 45 min was used. MS/MS analysis was performed using DDA with a scan range of 200–2000 m/z, with an MS1 resolution of 120,000, a normalized AGC of 200%, an isolation window of 3 m/z, and normalized stepped collision energy of 15,17, and 19.

### LC-MS/MS Data analysis

All MS/MS data analysis was performed with Biopharma Finder v.05.1 Automatic parameter values were set for component detection. For fragment identification, all MS/MS data was used, with a maximum mass set to 20 kDa, mass accuracy to 10 ppm and minimum confidence to 0.8. Mass search for unspecified modifications were allowed in a mass range of -151 to +96 Da, with residue deletion allowed. Search for RNase T1 digestion fragments were allowed with a high specificity, limiting the specificity of cleavage at guanosine residues, generating 3'-cyclic phosphate (as most of the fragments generated upon RNase T1 digestion have 3' cyclic phosphate). Of the identified fragments, only those with a mass error less than 10 ppm, average spectral resolution of 1, and a confidence score of 100% were considered for further analysis.

### Data availability

All LC-MS raw data generated or analyzed during this study is available at: <https://owncloud.gwdg.de/index.php/s/n20ombQ3HK3hlsk>.

Received: 15 July 2024; Accepted: 4 November 2024

Published online: 11 November 2024

### References

1. Suzuki, T. The expanding world of tRNA modifications and their disease relevance. *Nat. Rev. Mol. Cell Biol.* **22**, 375–392 (2021).
2. Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187–1200 (2017).
3. Ge, J. & Yu, Y.-T. RNA pseudouridylation: New insights into an old modification. *Trends Biochem. Sci.* **38**, 210–218 (2013).
4. Li, X., Ma, S. & Yi, C. Pseudouridine: The fifth RNA nucleotide with renewed interests. *Curr. Opin. Chem. Biol.* **33**, 108–116 (2016).
5. Cerneckis, J., Cui, Q., He, C., Yi, C. & Shi, Y. Decoding pseudouridine: An emerging target for therapeutic development. *Trends Pharmacol. Sci.* **43**, 522–535 (2022).
6. Arnez, J. G. & Steitz, T. A. Crystal structure of unmodified tRNA<sup>Gln</sup> complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* **33**, 7560–7567 (1994).
7. Guzzi, N. et al. Pseudouridylation of tRNA-derived fragments steers translational control in stem cells. *Cell* **173**, 1204–1216.e26 (2018).

8. Basak, A. & Query, C. C. A Pseudouridine residue in the spliceosome core is part of the filamentous growth program in yeast. *Cell Rep.* **8**, 966–973 (2014).
9. Wu, G., Xiao, M., Yang, C. & Yu, Y.-T. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J.* **30**, 79–89 (2011).
10. Yoluç, Y. et al. Instrumental analysis of RNA modifications. *Crit. Rev. Biochem. Mol. Biol.* **56**, 178–204 (2021).
11. Zaringhalam, M. & Papavasiliou, F. N. Pseudouridylation meets next-generation sequencing. *Methods* **107**, 63–72 (2016).
12. Marchand, V. et al. HydraPsiSeq: A method for systematic and quantitative mapping of pseudouridines in RNA. *Nucl. Acids Res.* **48**, e110–e110 (2020).
13. Schwartz, S. et al. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162 (2014).
14. Zhang, M. et al. Quantitative profiling of pseudouridylation landscape in the human transcriptome. *Nat. Chem. Biol.* **19**, 1185–1195 (2023).
15. Tavakoli, S. et al. Semi-quantitative detection of pseudouridine modifications and type I/II hypermodifications in human mRNAs using direct long-read sequencing. *Nat. Commun.* **14**, 334 (2023).
16. Begik, O. et al. Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. *Nat. Biotechnol.* **39**, 1278–1291 (2021).
17. Huang, S. et al. Interferon inducible pseudouridine modification in human mRNA by quantitative nanopore profiling. *Genome Biol.* **22**, 330 (2021).
18. Huang, S., Wylder, A. C. & Pan, T. Simultaneous nanopore profiling of mRNA m6A and pseudouridine reveals translation coordination. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-024-02135-0> (2024).
19. Fleming, A. M., Mathewson, N. J., Howpay Manage, S. A. & Burrows, C. J. Nanopore dwell time analysis permits sequencing and conformational assignment of pseudouridine in SARS-CoV-2. *ACS Cent Sci.* **7**, 1707–1717 (2021).
20. Fleming, A. M., Zhu, J., Done, V. K. & Burrows, C. J. Advantages and challenges associated with bisulfite-assisted nanopore direct RNA sequencing for modifications. *RSC Chem. Biol.* **4**, 952–964 (2023).
21. Patteson, K. G. Identification of the mass-silent post-transcriptionally modified nucleoside pseudouridine in RNA by matrix-assisted laser desorption/ionization mass spectrometry. *Nucl. Acids Res.* **29**, 49e–449 (2001).
22. Suzuki, T. & Suzuki, T. A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. *Nucl. Acids Res.* **42**, 7346–7357 (2014).
23. Taoka, M. et al. A mass spectrometry-based method for comprehensive quantitative determination of post-transcriptional RNA modifications: the complete chemical structure of *Schizosaccharomyces pombe* ribosomal RNAs. *Nucl. Acids Res.* **43**, e115–e115 (2015).
24. Mengel-Jorgensen, J. Detection of pseudouridine and other modifications in tRNA by cyanoethylation and MALDI mass spectrometry. *Nucl. Acids Res.* **30**, 135e–1135 (2002).
25. Durairaj, A. & Limbach, P. A. Improving CMC-derivatization of pseudouridine in RNA for mass spectrometric detection. *Anal. Chim. Acta* **612**, 173–181 (2008).
26. Popova, A. M. & Williamson, J. R. Quantitative analysis of rRNA modifications using stable isotope labeling and mass spectrometry. *J. Am. Chem. Soc.* **136**, 2058–2069 (2014).
27. Yamauchi, Y. et al. A mass spectrometry-based method for direct determination of pseudouridine in RNA. *Nucl. Acids Res.* **44**, e59–e59 (2016).
28. Pomerantz, S. C. & McCloskey, J. A. Detection of the common RNA nucleoside pseudouridine in mixtures of oligonucleotides by mass spectrometry. *Anal. Chem.* **77**, 4687–4697 (2005).
29. Khoddami, V. et al. Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. *Proc. Natl. Acad. Sci.* **116**, 6784–6789 (2019).
30. Fleming, A. M. et al. Structural elucidation of bisulfite adducts to pseudouridine that result in deletion signatures during reverse transcription of RNA. *J. Am. Chem. Soc.* **141**, 16450–16460 (2019).
31. Dai, Q. et al. Quantitative sequencing using BID-seq uncovers abundant pseudouridines in mammalian mRNA at base resolution. *Nat. Biotechnol.* **41**, 344–354 (2023).
32. Zhao, Y. et al. Pseudouridine detection and quantification using bisulfite incorporation hindered ligation. *ACS Chem. Biol.* **19**, 1813–1819 (2024).
33. Fleming, A. M., Xiao, S. & Burrows, C. J. Pseudouridine and N1-methylpseudouridine display pH-independent reaction rates with Bisulfite yielding ribose adducts. *Org. Lett.* **24**, 6182–6185 (2022).
34. Hayatsu, H., Wataya, Y., Kai, K. & Iida, S. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. *Biochemistry* **9**, 2858–2865 (1970).
35. McLuckey, S. A. & Habibi-Goudarzi, S. Decompositions of multiply charged oligonucleotide anions. *J. Am. Chem. Soc.* **115**, 12085–12095 (1993).
36. McLuckey, S. A., Van Berkel, G. J. & Glish, G. L. Tandem mass spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.* **3**, 60–70 (1992).
37. Suzuki, T. et al. Complete chemical structures of human mitochondrial tRNAs. *Nat. Commun.* **11**, 4269 (2020).
38. Rajan, K. S. et al. Pseudouridines on Trypanosoma brucei spliceosomal small nuclear RNAs and their implication for RNA and protein interactions. *Nucl. Acids Res.* **47**, 7633–7647 (2019).
39. Bakin, A. & Ofengand, J. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: Analysis by the application of a new sequencing technique. *Biochemistry* **32**, 9754–9762 (1993).
40. Sajek, M. P., Woźniak, T., Sprinzl, M., Jaruzelska, J. & Barciszewski, J. T-psi-C: User friendly database of tRNA sequences and structures. *Nucl. Acids Res.* <https://doi.org/10.1093/nar/gkz922> (2019).
41. Dunin-Horkawicz, S. MODOMICS: A database of RNA modification pathways. *Nucl. Acids Res.* **34**, D145–D149 (2006).
42. Zhang, Z. et al. Global analysis of tRNA and translation factor expression reveals a dynamic landscape of translational regulation in human cancers. *Commun. Biol.* **1**, 234 (2018).
43. Cui, Q. et al. Targeting PUS7 suppresses tRNA pseudouridylation and glioblastoma tumorigenesis. *Nat. Cancer* **2**, 932–949 (2021).
44. Vanhinsbergh, C. J. et al. Characterization and sequence mapping of large RNA and mRNA therapeutics using mass spectrometry. *Anal. Chem.* **94**, 7339–7349 (2022).
45. Pace, C. N., Heinemann, U., Hahn, U. & Saenger, W. Ribonuclease T1: Structure, function, and stability. *Angew. Chem. Int. Ed. Engl.* **30**, 343–360 (1991).

## Acknowledgements

We are grateful to H. Waldmann for sharing resources and infrastructure, P. Janning for discussions and advice, as well as J. Warmers for technical support. S.H. received training by the International Max-Planck Research School for Living Matter (IMPRS-LM). The work was supported by the Max-Planck Society and the Ministry of Culture and Science of the State of Northrhine Westphalia (NRW returning scholars program and CANcer Targeting Network initiative; NW21-062 C).

### Author contributions

S.B conceptualized the project. A.S. and S.H. conducted the experiments and data analysis. S.B and S.H wrote the main manuscript text and prepared the figures.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-78734-3>.

**Correspondence** and requests for materials should be addressed to S.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024