Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2020, **11**, 1878

All publication charges for this article have been paid for by the Royal Society of Chemistry

Chemical tagging for sensitive determination of uridine modifications in RNA†

Qing-Yun Cheng, Jun Xiong, Cheng-Jie Ma, Yi Dai, Jiang-Hui Ding, Fei-Long Liu, Bi-Feng Yuan * and Yu-Qi Feng

The discovery of dynamic and reversible modifications in messenger RNA (mRNA) is opening new directions in RNA modification-mediated regulation of biological processes. Methylation is the most prevalent modification occurring in mRNA and the methyl group is mainly decorated in the adenine, cytosine, and guanine base or in the 2′-hydroxyl group of ribose. However, methylation of the uracil base (5-methyluridine, m^5U) has not been discovered in mRNA of eukaryotes. In the current study, we established a method of *N*-cyclohexyl-*N*′-*β*-(4-methylmorpholinium) ethylcarbodiimide *p*-toluenesulfonate (CMCT) labelling coupled with liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS/MS) analysis for the sensitive determination of uridine modifications in RNA. Our results demonstrated that the detection sensitivities of uridine modifications in RNA increased up to 1408 fold upon CMCT labelling. Using the developed method, we identified the distinct existence of m^5U in mRNA of various mammalian cells and tissues. In addition, the stable isotope tracing monitored by mass spectrometry revealed that the methyl group of m^5U originated from *S*-adenosyl-L-methionine (SAM). Our study expanded the list of modifications occurring in mRNA of mammals. Future work on transcriptome-wide mapping of m^5U will further uncover the functional roles of m^5U in mRNA of mammals.

Received 9th October 2019
Accepted 6th January 2020

DOI: 10.1039/c9sc05094a

rsc.li/chemical-science

Introduction

Chemical modification in nucleic acids can delicately regulate gene expression.[1] The functional roles of DNA modifications in regulation of gene expression have been well established.[2] However, how RNA modifications affect translation and eventually tune gene expression has just begun to be understood.[3] The discovery of a variety of modifications in RNA broadens our views on the structures and functions of RNA.[4] To date, over 150 chemically distinct modifications have been identified in various species of RNA in all three kingdoms of life.[5,6] Most of these modifications are characterized in the abundant ribosomal RNA (rRNA) and transfer RNA (tRNA), where they can alter the RNA secondary structure by affecting base stacking or changing hydrogen bonding patterns.[7,8]

RNA modifications are found to be involved in various RNA processing and metabolism procedures, including alternative splicing, stability, translation, folding and molecular recognition.[7–10] The discovery of dynamic and reversible modifications in messenger RNA (mRNA) is opening new directions in RNA modification-mediated regulation of biological processes.[11] The technical breakthroughs led to the discovery of a series of modifications in eukaryotic mRNA, including N^6-methyladenosine (m^6A), *N*1-methyladenosine (m^1A), N^6,2′-*O*-dimethyladenosine (m^6Am), 5-methylcytidine (m^5C), *N*3-methylcytidine (m^3C), *N*7-methylguanosine (m^7G), 2′-*O*-methylation, inosine (I), and pseudouridine (Ψ).[11–13] Very recently, N^4-acetylcytidine was newly discovered to be present in mRNA of mammals.[14] These studies demonstrated that the modifications play versatile roles in mRNA processing and eventually affect the epitranscriptomic state of cells.[15]

Methylation is the most prevalent modification occurring in mRNA, such as m^6A, m^1A, m^6Am, m^5C, m^3C, and m^7G, and 2′-*O*-methylation.[11] The methyl group is mainly decorated in adenine, cytosine, and guanine base or in ribose. However, methylation of the uracil base (5-methyluridine, m^5U) has not been discovered in mRNA of eukaryotes. m^5U is a common modification found in rRNA and tRNA.[16] Two m^5U sites were previously identified at positions 747 and 1939 in *E. coli* 23S rRNA, and one m^5U site was found at position 54 in *E. coli* tRNA.[17,18] One early study showed that m^5U was also present in the tRNA-like domain of tmRNA in *E.coli* cells.[19] It was proposed that TrmA protein can methylate U54 in the T-arm of most *E. coli* tRNAs.[18,20] And the RumA protein was responsible for the methylation of U1939 in *E. coli* 23S rRNA to form m^5U that may

Key Laboratory of Analytical Chemistry for Biology and Medicine (Ministry of Education), Sauvage Center for Molecular Sciences, Department of Chemistry, Wuhan University, Wuhan 430072, P.R. China. E-mail: bfyuan@whu.edu.cn; Fax: +86-27-68755595; Tel: +86-27-68755595

† Electronic supplementary information (ESI) available: Quantification of modifications in RNA; CeO_2 extraction of CMCT labelled nucleosides; evaluation of the purity of isolated mRNA by real-time quantitative PCR; analysis of mRNA by high-throughput sequencing; Table S1–S7; and Fig. S1–S14. See DOI: 10.1039/c9sc05094a

play an important role in translation.[17] But the existence of m^5U in mRNA of eukaryotes remains elusive.

In eukaryotic cells, the RNA pool predominantly consists of rRNA (~80%) and tRNA (~15%), whereas mRNA only constitutes a small fraction (<5%).[21] The modifications in mRNA generally exist in low abundance.[13] For example, m^3C is present at a frequency around ~1 to 5 modifications per 10^5 cytidine.[22] Therefore, the low *in vivo* contents can lead to the detection of modifications in mRNA being a challenging task. The mass spectrometry (MS)-based analytical platform has been widely employed for determining modifications due to its superior detection sensitivity and powerful capability in the identification of the chemical structures of modifications.[23–28] However, the ionization efficiencies of uridine and uridine modifications in electrospray ionization (ESI)-MS are much poorer than those of other canonical nucleosides and their corresponding modifications,[29,30] which makes the detection of uridine modifications more difficult.

The chemical labelling strategy that can introduce an easily ionizable functional group into analytes has been developed to improve the detection sensitivities of compounds while in combination with mass spectrometry analysis.[31–33] In this respect, the reagent *N*-cyclohexyl-*N*′-*β*-(4-methylmorpholinium) ethylcarbodiimide *p*-toluenesulfonate (CMCT) was previously used to label Ψ and U followed by treatment under alkaline conditions to differentiate Ψ from U.[34–36] The CMCT group can stably link to uridine and uridine modifications if omitting the alkaline treatment, which then can be utilized to detect the low-abundant uridine modifications since the introduction of the permanently positively charged quaternary ammonium group from CMCT can enhance the ionization efficiency of uridine modifications. In the current study, we established a method of CMCT labelling coupled with liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) analysis for sensitive determination of the uridine modifications in RNA. Our results demonstrated that the detection sensitivities of uridine modifications in RNA increased by 6–1408 fold upon CMCT labelling. Using the developed method, we identified, for the first time, the existence of m^5U in mRNA of mammalian cells.

Materials and methods

Chemicals and reagents

Adenosine (A), uridine (U), cytidine (C), guanosine (G), N^6-methyladenosine (m^6A) and venom phosphodiesterase I were purchased from Sigma-Aldrich (Beijing, China). N^6-isopentenyladenosine (i^6A), N^6-threonylcarbamoyladenosine (t^6A), 1-methylguanosine (m^1G), N^2-methylguanosine (m^2G), pseudouridine (Ψ), 5-methyluridine (m^5U), 1-methylpseudouridine ($m^1\Psi$), 5-hydroxymethyluridine (hm^5U), 5-methoxyuridine (mo^5U), 5-methoxycarbonylmethyluridine (mcm^5U), and 1-cyclohexyl-3-(2-(4-morpholinyl)ethyl) carbodiimide (CME) were purchased from Carbosynth China Ltd. (Jiangsu, China). 6-Methyluridine (m^6U) was purchased from Granlen Inc. (Henan, China). *N*-cyclohexyl-*N*′-*β*-(4-methylmorpholinium) ethylcarbodiimide *p*-toluenesulfonate (CMCT), D_3-iodomethane (CD_3I), cerium oxide (CeO_2), and L-methionine-(methyl-D_3) (D_3-Met) were purchased from Aladdin Industrial Inc. (Shanghai, China). S1 nuclease and calf intestinal alkaline phosphatase were obtained from Takara Biotechnology (Dalian, China). The E.Z.N.A. miRNA kit was obtained from Omega Bio-tek Inc. (Norcross, GA, USA). Dulbecco's modified Eagle medium (DMEM), RPMI-1640 medium and fetal bovine serum were obtained from Thermo-Fisher Scientific (Beijing, China). Chromatographic grade acetonitrile (ACN) and methanol (MeOH) were purchased from Merck (Darmstadt, Germany). Analytical grade formic acid, ammonium bicarbonate, diethyl ether and ammonia hydroxide were purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China).

Biological samples and stable isotope labelling tracing

Human embryonic kidney 293T cell line (HEK293T), human cervical carcinoma cell line (HeLa), human hepatic cell line (HL-7702), and human liver carcinoma cell lines (HepG2) were obtained from the China Center for Type Culture Collection. HEK293T, HeLa, and HepG2 cells were grown in DMEM medium supplemented with 10% fetal bovine serum, 100 U mL^{-1} penicillin, and 100 μg mL^{-1} streptomycin at 37 °C in a 5% CO_2 atmosphere. HL-7702 cells were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum, 100 U mL^{-1} penicillin, and 100 μg mL^{-1} streptomycin at 37 °C in a 5% CO_2 atmosphere. Male wild-type C57BL/6J mice (2 months old) were obtained from the Center for Animal Experiment/ABSL-3 Laboratory of Wuhan University and sacrificed to collect tissues. This study was performed in strict accordance with the NIH guidelines for the care and use of laboratory animals (NIH Publication No. 85-23 Rev. 1985) and was approved by the Animal Care and Use Committee of Wuhan University (Wuhan, China).

As for stable isotope labelling tracing monitored by mass spectrometry, HEK293T cells were cultured in DMEM medium supplemented with 0.3 mM of D_3-Met to metabolically label the RNA with the CD_3 group. The cells were collected after culturing for 72 h and then mRNA was extracted.

Purification and enzymatic digestion of different species of RNA

Total RNA was extracted using an E.Z.N.A. HP Total RNA kit (Omega Bio-Tek Inc., Norcross, GA, USA) according to the manufacturer's recommended procedure. Small RNA (<200 nt) was purified using the E.Z.N.A. miRNA kit (Omega Bio-Tek Inc., Norcross, GA, USA). For mRNA isolation, the obtained large RNA (>200 nt) after small RNA depletion using the E.Z.N.A. miRNA kit was further extracted twice using a PolyATtract mRNA isolation system (Promega, Madison, WI, USA). The resulting mRNA was further processed with agarose gel electrophoresis-based purification, which can remove the potential trace level of contamination of small RNA (<200 nt). The detailed procedure of agarose gel electrophoresis-based purification can be found in the ESI.† The rRNA was purified from the obtained large RNA (>200 nt) by means of size-exclusion chromatography (SEC) on a Bio SEC-5 column (5

μm, 1000 Å, 4.6 × 150 mm; Agilent Technologies, Foster City, CA, USA) according to a previous study.[37] Isocratic separation was performed at a flow rate of 0.35 mL min^{-1} at 35 °C with 100 mM NH_4Ac (pH 7.0) as the mobile phase. The fraction was collected and further processed with the E.Z.N.A. miRNA kit to remove NH_4Ac. The purification was repeated twice to obtain high purity of rRNA.

Enzymatic digestion of RNA was performed according to a previously described method under neutral conditions.[37] Briefly, 3 μL of 10 × buffer (500 mM Tris–HCl, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM $ZnSO_4$, pH 7.0), 1 μL of S1 nuclease (180 U μL^{-1}), 1 μL of venom phosphodiesterase I (0.001 U μL^{-1}), and 1 μL of calf intestinal alkaline phosphatase (30 U μL^{-1}) were added to the RNA (0.2–5 μg) dissolved in 24 μL H_2O. The mixture was incubated at 37 °C for 4 h. 270 μL of H_2O and 300 μL of $CHCl_3$ were added to the resulting solution followed by vortexing for 3 min and centrifugation at 12 000 rpm for 10 min. The upper aqueous phase was collected. The extraction process was repeated 3 times to fully remove the proteins. The aqueous phase was collected and lyophilized and then subjected to subsequent analysis.

Synthesis of *N*-cyclohexyl-*N′*-*β*-(4-methylmorpholinium) ethylcarbodiimide *p*-toluenesulfonate-(methyl-D_3) (D_3-CMCT)

D_3-CMCT was synthesized from 1-cyclohexyl-3-(2-(4-morpholinyl)ethyl) carbodiimide (CME) and D_3-iodomethane (CD_3I) according to a previous report[38] with slight modification. The synthesis route is shown in Fig. S1 in the ESI.† Briefly, 40 μL of CME (50 mM in ACN) was incubated with 10 μL of CD_3I at 30 °C for 15 h. Then 100 μL of borate buffer (50 mM, pH 8.5) and 200 μL of diethyl ether were added to the mixture. The mixture was separated into a lower aqueous phase and an upper diethyl ether. The unreacted CME and CD_3I remained in the diethyl ether phase while the newly synthesized D_3-CMCT was in the aqueous phase. After removing the diethyl ether phase, the aqueous phase was collected and used for the labelling reaction.

Chemical labelling

CMCT and D_3-CMCT were used to label uridine and uridine modifications. The labelling reaction was carried out in borate buffer (50 mM, pH 8.5). The reaction conditions, including reaction time and temperature and the concentration of CMCT, were optimized to achieve the best labelling efficiency. After the labelling reaction, the mixture was subjected to CeO_2 solid phase extraction (SPE) to remove excess CMCT according to a previously described procedure,[39] and the resulting derivatives were then subjected to LC-ESI-MS/MS analysis. The detailed procedure for the extraction of CMCT labelled nucleosides by CeO_2 SPE is shown in the ESI.†

Analysis of CMCT\D_3-CMCT labelled uridine modifications by LC-ESI-MS/MS

Analysis of CMCT\D_3-CMCT labelled uridine modifications was performed on an AB 3200 QTRAP mass spectrometer (Applied Biosystems, Foster City, CA, USA) equipped with an electrospray ionization (ESI) source (Turbo Ionspray) and a Shimadzu LC-20AD HPLC system (Tokyo, Japan). Data acquisition and processing were performed with AB SCIEX Analyst 1.5 software (Applied Biosystems). ESI-MS was conducted in positive-ion mode and multiple-reaction monitoring (MRM) was employed for the detection. To achieve the best performance, the mass spectrometric parameters were optimized, and the optimized MS conditions are listed in Table S1 in the ESI.† The separation of CMCT labelled uridine modifications was performed on an Inersil ODS-3 reversed-phase column (5 μm, 2.1 × 250 mm; GL Sciences Lnc., Tokyo, Japan). NH_4HCO_3 (2 mM, pH 4.5, solvent A) and ACN (solvent B) were used as the mobile phases. Gradients of 0–3 min, 5% B; 3–20 min, 5–40% B; 20–29 min, 40% B; 29–31 min, 40–5% B; and 31–45 min, 5% B were used. The flow rate was set at 200 μL min^{-1}. Quantification of uridine modifications is shown in the ESI.†

High-resolution mass spectrometry analysis

The CMCT\D_3-CMCT labelled uridine modifications were also examined with a high-resolution LTQ-Orbitrap Elite mass spectrometer (Thermo-Fisher Scientific, Waltham, MA, USA) equipped with an electrospray ionization (ESI) source and a Dionex Ultimate 3000 UPLC system (Thermo-Fisher Scientific, Waltham, MA, USA). The LC separation conditions were the same as those for the AB 3200 QTRAP mass spectrometer system. Full MS scans were acquired in positive-ion mode at a resolution of 60 000. Collision induced dissociation (CID) with a collision energy of 25–35 eV was used. The fragments were acquired with a mass range of *m*/*z* 200–700 at a resolution of 60 000. The source and ion transfer parameters applied were as follows: heater temperature, 300 °C; capillary temperature,

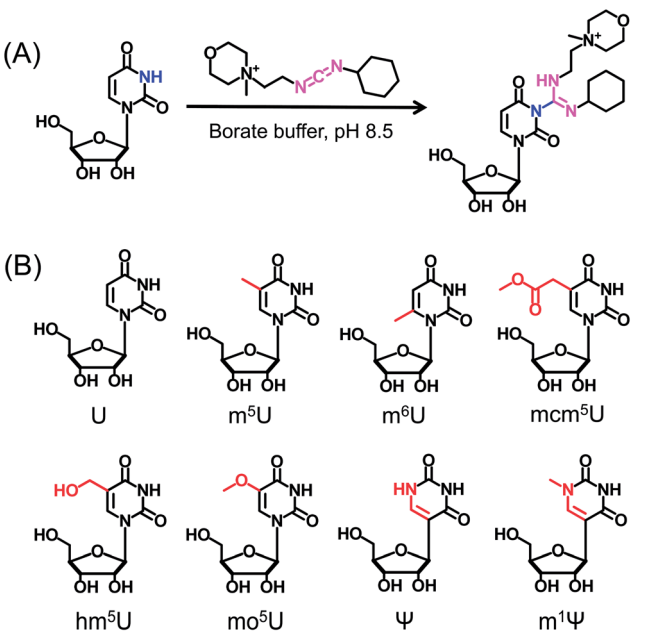

Fig. 1 Chemical reaction for labelling of uridine with CMCT. (A) CMCT carries a carbodiimide group that can selectively react with the NH group at the *N*3 position of the uracil base. (B) The chemical structures of uridine and uridine modifications.

350 °C; sheath gas flow, 35 arbitrary units; auxiliary gas flow, 15 arbitrary units; spray voltage, 3.5 kV; capillary voltage, 35 V; and the S-lens RF level, 60%. Data analysis was carried out using Xcalibur v3.0.63 (Thermo-Fisher Scientific, Waltham, MA, USA).

siRNA knockdown and overexpression of *TRMT2A*

Knockdown of *TRMT2A* was performed using siRNA (Sangon Biotech, Shanghai, China) targeting human *TRMT2A* mRNA. Non-targeting siRNA was used as a negative control. The sequences of *TRMT2A* siRNA are 5′-GCUACAUCAGGGAUGACUUTT-3′/3′-TTCGAUGUAGUCCCUACUGAA-5′, 5′-GCAGACUGAGUAUCGUAAUTT-3′/3′-TTCGUCUGACUCAUAGCAUUA-5′, and 5′-GGACAAUGAGUUGAGUAAUTT-3′/3′-TTCCUGUUACUCAACUCAUUA-5′ and the sequence of the control siRNA is 5′-UUCUCCGAACGUGUCACGUTT-3′/3′-TTAAGAGGCUUGCACAGUGCA-5′. The plasmid pcDNA3.1(+)-TRMT2A was purchased from Sangon Biotech (Shanghai, China). Human HEK293T cells were transfected with siRNA or pcDNA3.1(+)-TRMT2A plasmid using Lipofectamine 2000 (Invitrogen, USA) according to the manufacturer's instruction. The culture medium was exchanged at 10 h after transfection and cells were harvested at 48 h after transfection. As for the rescue experiment, HEK293T cells were transfected with siRNA for 48 h and then further

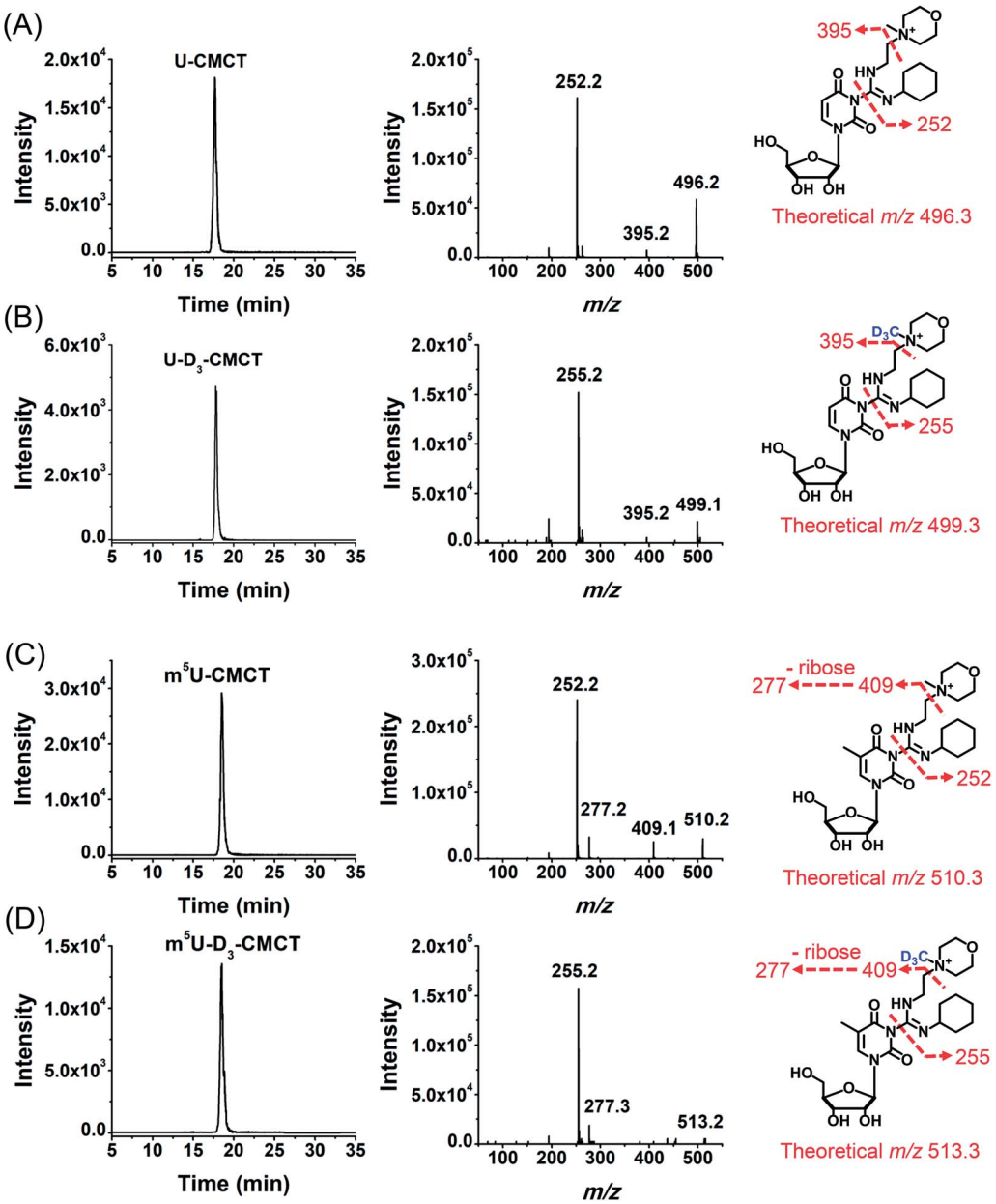

Fig. 2 Examination of the CMCT/D_3-CMCT labelling reaction. Extracted-ion chromatogram and fragment ions of (A) CMCT labelled U, (B) D_3-CMCT labelled U, (C) CMCT labelled m^5U, and (D) D_3-CMCT labelled m^5U. The mass transitions 496.3 → 252.2 for CMCT labelled U, 499.3 → 255.2 for D_3-CMCT labelled U, 510.3 → 252.2 for CMCT labelled m^5U, and 513.3 → 255.2 for D_3-CMCT labelled m^5U were used for the detection in MRM detection mode.

transfected with pcDNA3.1(+)-TRMT2A plasmid for additional 48 h. For the relative quantification of the mRNA of *TRMT2A*, 1 μg of isolated total RNA was used to generate cDNA using PrimeScript™ RT reagent with gDNA Eraser (Takara Biotechnology). Real-time quantitative PCR was performed using a CFX Connect real-time PCR detection system (Bio-Rad Laboratories) and a SYBR® Premix Ex Taq II (Tli RNaseH Plus) (Takara Biotechnology) according to the manufacturer's instructions. The levels of gene expression were normalized to glyceraldehyde 6-phosphate dehydrogenase gene (*GAPDH*). PCR primers sequences are as follows: *TRMT2A* forward primer 5′-CGGAAGGTAAAGAGGGTCATTGG-3′, *TRMT2A* reverse primer 5′-GAATGCAAGCCAGCACGGGGT-3′, *GAPDH* forward primer 5′-CTGGGCTACACTGAGCACC-3′, and *GAPDH* reverse primer 5′-AAGTGGTCGTTGAGGGCAATG-3′. Western blotting analysis of TRMT2A was carried out to evaluate the siRNA knockdown and overexpression of *TRMT2A* according to a previous study,[40] and the detailed procedure can be found in the ESI.†

Analysis of mRNA by high-throughput sequencing

The isolated mRNA was processed using a KAPA Stranded RNA-Seq Library Prep Kit (Illumina) for sequencing library construction following the manufacturer's protocol. RNA-seq was carried out on an Illumina HiSeq 4000 (Illumina) platform by Aksomics Inc (Shanghai, China). Detailed information can be found in the ESI.†

Results and discussion

Chemical labelling

CMCT carries a carbodiimide group that can selectively react with the NH group at the *N*3 position of the uracil base (Fig. 1A). The introduction of a constant positively charged quaternary ammonium group from CMCT can efficiently enhance the ionization efficiency of the derivatives in electrospray ionization (ESI)-MS. To achieve good accuracy for quantification and identification, we also synthesized D_3-CMCT (Fig. S1 in the ESI†). The high-resolution mass spectrometry analysis demonstrated that D_3-CMCT was successfully synthesized (Fig. S2 in the ESI†).

We then used 8 standards of uridine and uridine modifications to evaluate the CMCT/D_3-CMCT labelling reaction (Fig. 1B). Shown in Fig. 2 are the extracted-ion chromatograms and the fragment ions of the CMCT/D_3-CMCT labelled uridine and m^5U. The measured precursor ions and product ions of CMCT/D_3-CMCT labelled uridine and m^5U were identical to the theoretical values, suggesting that uridine and m^5U were successfully labelled and the desired derivatives were obtained. The fragmentation analysis demonstrated that the other 6 uridine modifications (Ψ, $m^1Ψ$, m^6U, hm^5U, mo^5U, and mcm^5U) were also successfully labelled with CMCT/D_3-CMCT (Fig. S3 in the ESI†). In addition, the CMCT/D_3-CMCT labelled uridine modifications were further confirmed by high-resolution mass spectrometry analysis (Fig. S4 in the ESI†).

It is worth noting that protonated molecules generally undergo fragmentation based on a charge-directed

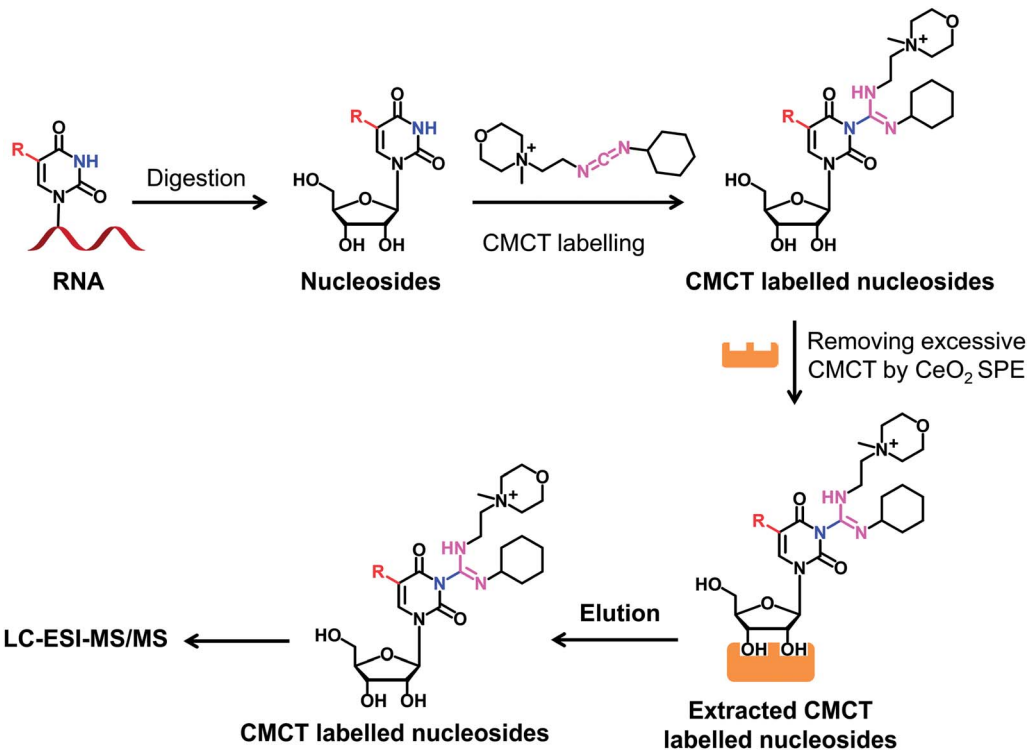

Fig. 3 Schematic illustration of the procedure for determination of uridine modifications by CMCT labelling coupled with LC-ESI-MS/MS analysis. Isolated RNA was enzymatically digested followed by CMCT labelling. Then the CMCT labelled uridine modifications were extracted and enriched by CeO_2 solid phase extraction (SPE). The resulting CMCT labelled uridine modifications were subjected to LC-ESI-MS/MS analysis.

mechanism. But when the charge is fixed at one specific atom, charge-remote fragmentation may occur,[41,42] which is caused by hydrogen transfer instead of proton transfer among different atoms of the precursor ion.[43,44] For example, CMCT labelled m^5U is constantly and positively charged due to carrying a quaternary ammonium group and CMCT labelled m^5U undergoes charge-remote fragmentation to generate the fragment ion of *m*/*z* 252.2 (Fig. 2C and S5 in the ESI†). CMCT labelled m^5U can also lose a neutral fragment of 4-methylmorpholine and generate the fragment ion of *m*/*z* 409.2, which can further lose a ribose to form the fragment ion of *m*/*z* 277.2 (Fig. 2C and S5 in the ESI†).

It can be seen that the fragment ion of *m*/*z* 252.2 is the predominant fragment ion of CMCT labelled uridine modifications (Fig. 2 and S3 in the ESI†). We examined the fragmentation pathway of the CMCT labelled m^5U with different collision energies ranging from 1 eV to 30 eV. The result showed that the ion of *m*/*z* 252.2 was the predominant fragment ion under different collision energies (Fig. S6 in the ESI†). However, we did not observe the cleavage of the *N*-glycosidic bond that would produce the fragment ion of *m*/*z* 378.2 (Fig. S6 in the ESI†). These results suggested that the N–C bond between nucleosides and CMCT was more prone to break than the *N*-glycosidic bond.

To obtain good labelling efficiency, we optimized the reaction conditions, including reaction temperature and time and the concentrations of CMCT. The result showed that the best labelling efficiencies were obtained at 40 °C for most of the uridine modifications (Fig. S7A in the ESI†). Therefore, 40 °C was used as the reaction temperature. We next optimized the reaction time. The results demonstrated that 14 h of reaction was sufficient for the CMCT labelling (Fig. S7B in the ESI†). As for the optimization of the concentration of CMCT, the results showed that the signals of CMCT labelled uridine modifications reached a plateau when the concentration of CMCT was 50 mM (Fig. S7C in the ESI†). Taken together, the optimized CMCT labelling conditions for uridine modifications were 40 °C for 14 h with 50 mM CMCT. Under the optimized reaction conditions, more than 91% of uridine modifications could react with CMCT to form their corresponding derivatives (Table S2 in the ESI†). And the formed derivatives also exhibited good stability, which was suitable for the subsequent analysis by LC-ESI-MS/MS (Fig. S7D in the ESI†).

CMCT labelling improves the detection sensitivities of uridine modifications

We next evaluated the detection sensitivities of the uridine modifications upon CMCT labelling. The CMCT labelled uridine modifications were extracted with CeO_2 to remove the excess CMCT followed by LC-ESI-MS/MS analysis (Fig. 3). The results showed that the retention of uridine modifications without CMCT labelling on the reversed-phase column was relatively weak (Fig. 4A). However, the retention greatly increased after CMCT labelling (Fig. 4B), which is due to the enhanced hydrophobicity of the derivatives by the introduction of cyclohexyl group from CMCT.

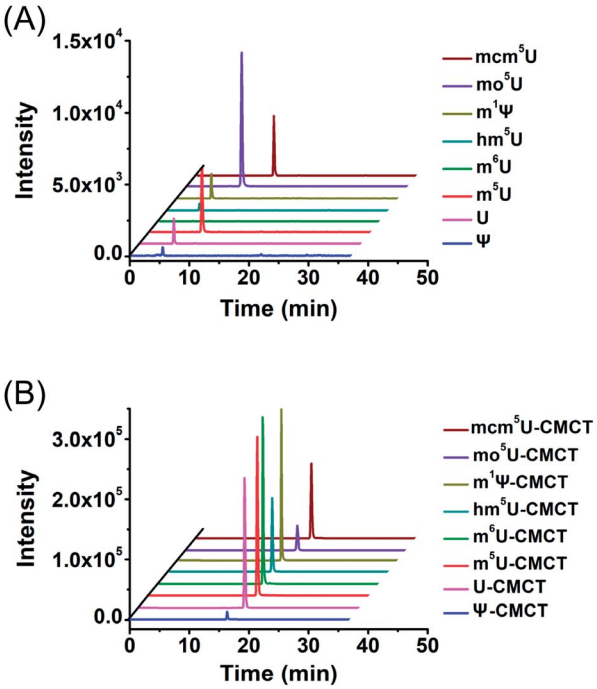

Fig. 4 Extracted-ion chromatograms of Ψ, U, m^5U, m^6U, mcm^5U, hm^5U, $m^1\Psi$ and mo^5U without (A) and with (B) CMCT labelling. The optimized MS conditions are listed in Table S1 in the ESI.†

We then examined the detection sensitivities of these uridine modifications upon CMCT labelling. The limits of detection (LODs) defined as the amounts of the analytes at a signal-to-noise ratio (S/N) of 3 were employed to evaluate the detection sensitivities of uridine modifications. The results demonstrated that the detection sensitivities of uridine modifications improved by 6 to 1408 fold upon CMCT labelling (Table 1). We reason that the introduction of positively charged quaternary ammonium group into uridine modifications could greatly improve the ionization efficiency during ESI-MS analysis, which eventually leads to enhanced detection sensitivities. Upon CMCT labelling, the LODs of the CMCT labelled uridine

Table 1 Evaluation of the detection sensitivities of uridine, uridine modifications and guanosine upon CMCT labelling. Independent measurements of three samples were performed

Analytes	Without labelling (fmol, LODs)	With CMCT labelling (fmol, LODs)	Detection sensitivities improved fold
U	160.0 ± 17.0	0.44 ± 0.04	364
Ψ	276.9 ± 22.1	2.20 ± 0.13	126
m^5U	26.4 ± 3.2	0.29 ± 0.02	91
m^6U	408.2 ± 28.1	0.29 ± 0.03	1408
$m^1\Psi$	99.0 ± 10.3	0.31 ± 0.02	319
hm^5U	145.8 ± 13.3	0.39 ± 0.03	374
mo^5U	7.8 ± 0.8	1.28 ± 0.10	6
mcm^5U	20.3 ± 4.2	0.54 ± 0.02	38
G	6.6 ± 0.6	0.90 ± 0.08	7

modifications are at similar levels (Table 1). However, the LODs of these native uridine modifications without CMCT labelling have relatively large differences (LODs ranging from 6.6 to 408.2 fmol, Table 1). Due to the large differences of the LODs among native uridine modifications, the improved folds of detection sensitivities (LODs of native uridine modifications/LODs of CMCT labelled uridine modifications) were different among these uridine modifications.

In addition to uridine modifications, guanosine can also react with CMCT as in previous studies.[45,46] CMCT labelling can increase the detection sensitivity of guanosine by 7-fold, which is smaller than that for most of the uridine modifications (Table 1). It can be seen that the ionization efficiency of guanosine is better than those of uridine modifications and the increased detection sensitivity of guanosine by chemical labelling is therefore not as dramatic as that of uridine modifications (Table 1). Although

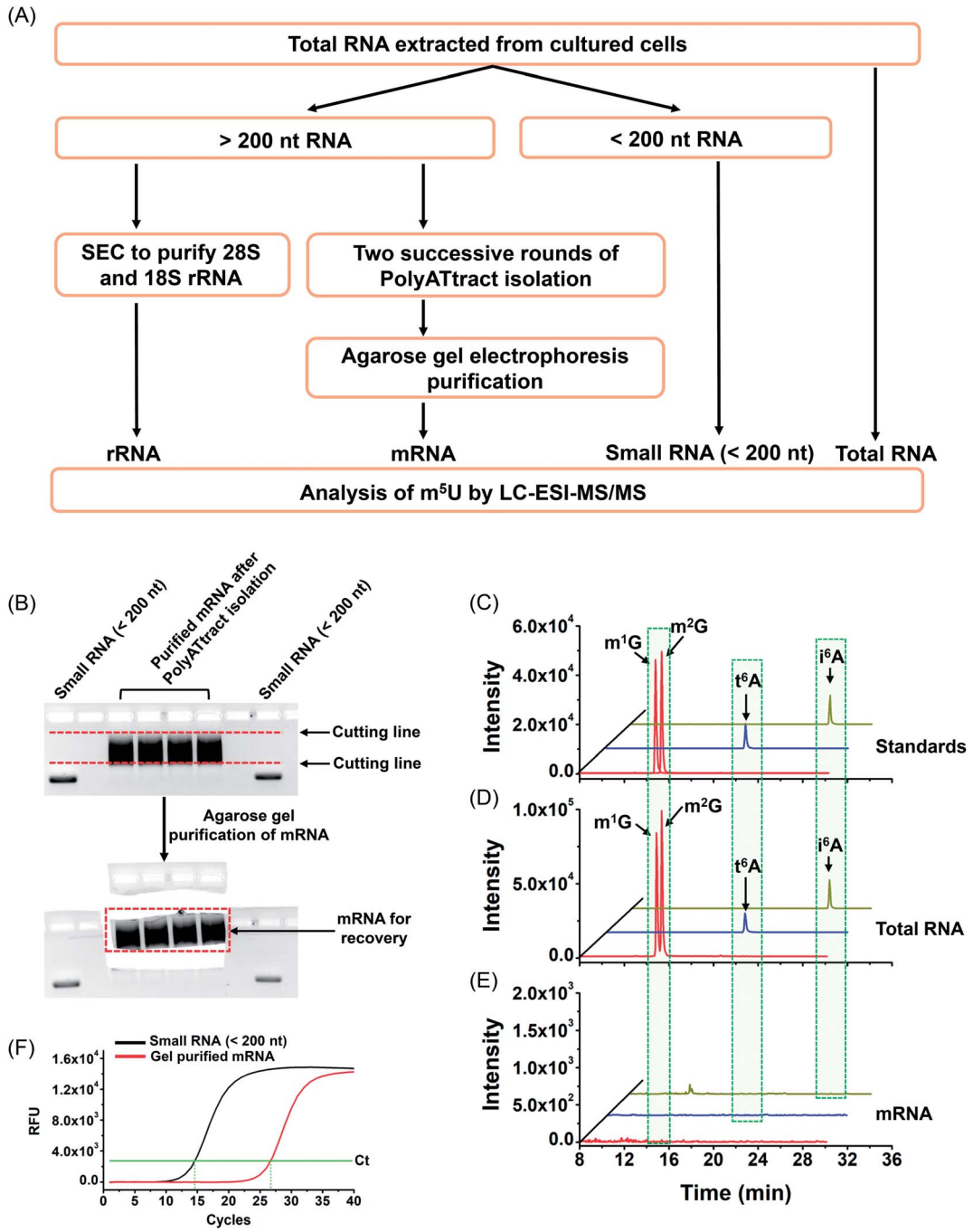

Fig. 5 Schematic illustration of the extraction of different species of RNA and evaluation of the purity of the isolated mRNA. (A) Schematic illustration of the extraction of different species of RNA. (B) Agarose gel electrophoresis-based purification. (C) Extracted-ion chromatograms of the standards of 4 modifications (m^1G, m^2G, t^6A, and i^6A). (D) Extracted-ion chromatograms of the 4 modifications detected in total RNA of HEK293T cells. (E) Extracted-ion chromatograms of the 4 modifications detected in mRNA of HEK293T cells. (F) Evaluation of the purity of the isolated mRNA by real-time quantitative PCR.

guanosine reacts with CMCT, this will not affect the sensitive detection of m^5U and other uridine modifications since they have distinctly different molecular weights (Table S1 in the ESI†).

Method validation

The calibration curves were constructed to quantitatively measure the modifications in RNA. A mixture of uridine modification standards with a series of amounts ranging from 0.016 pmol to 16 pmol and fixed amounts of isotope internal standards (D_3-CMCT labelled derivatives) was prepared. The calibration curves were constructed by plotting the peak area ratios (analytes/IS) against the amounts of uridine modifications with triplicate measurements. The results showed that good linearities were obtained with the coefficient of determination (R^2) being great than 0.9994 (Table S3 in the ESI†).

In addition, the accuracy of the established method was evaluated by comparing the measured contents of uridine modifications to the theoretical contents of uridine modifications. The intra- and inter-day relative standard deviations (RSDs) were calculated with different amounts of uridine modification standards for evaluation of the precision of the method. Three measurements over a day gave the intra-day RSDs, and the inter-day RSDs were determined by measuring samples for three consecutive days. The results showed that the intra- and inter-day RSDs for uridine modifications were less than 15.3% and relative errors (REs) were less than 7.8% (Table S4 in the ESI†), indicating that good precision and accuracy were achieved. Taken together, the results demonstrated that the developed method was reliable for the accurate quantification of uridine modifications.

Determination of m^5U in mRNA in mammalian cells

With the established method, we explored the potential existence of endogenous uridine modifications in mRNA from cultured human cell lines. Obtaining highly pure mRNA is very important for the determination of the existence of new modifications in mRNA. In this respect, here we developed a method to purify mRNA by successive three different methods, including (1) small RNA (<200 nt) depletion, (2) PolyATtract mRNA isolation (twice), and (3) agarose gel electrophoresis-based purification (Fig. 5A and B).

As tRNA contains abundant m^5U modification,[47] tRNA contamination could lead to a false positive result for the determination of m^5U in mRNA. In this respect, small RNA (<200 nt) was first depleted from total RNA using the E.Z.N.A. miRNA kit. The resulting RNA was further purified twice with the PolyATtract mRNA isolation system followed by agarose gel electrophoresis-based purification (Fig. 5A and B). Four modifications (m^1G, m^2G, i^6A, and t^6A) that were reported to be only present in tRNA or rRNA were used to evaluate the purity of the isolated mRNA.[9,48,49] Shown in Fig. 5C, D and 5E are the extracted-ion chromatograms of the standards of these 4

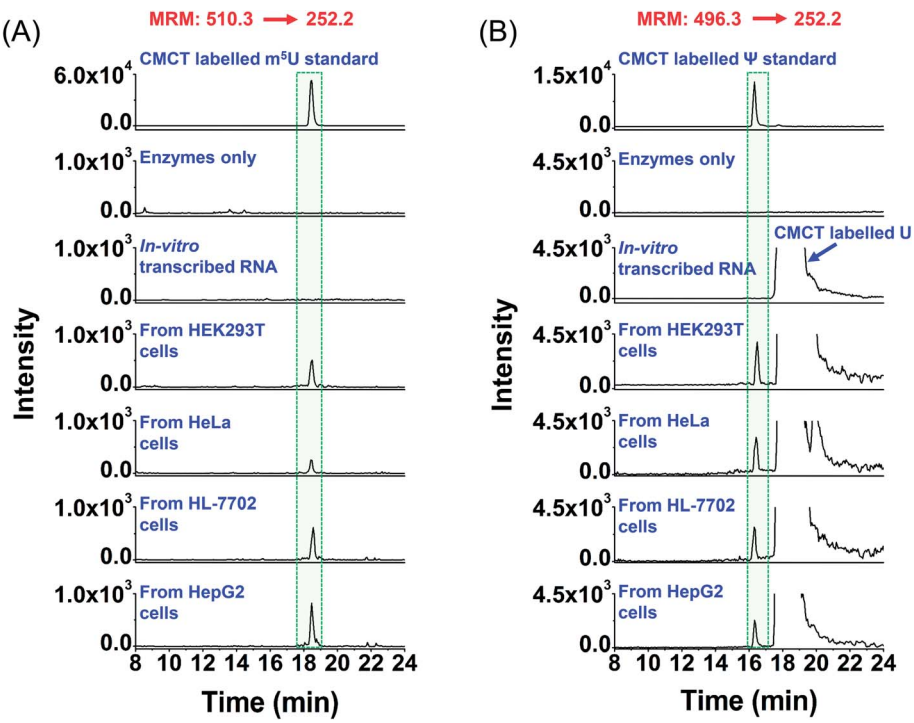

Fig. 6 Extracted-ion chromatograms of the detected m^5U and Ψ in mRNA of various human cells. (A) Extracted-ion chromatograms of m^5U standards and the m^5U detected in mRNA from human HEK293T, HeLa, HL-7702, and HepG2 cells. (B) Extracted-ion chromatograms of Ψ standards and the Ψ detected in mRNA from human HEK293T, HeLa, HL-7702, and HepG2 cells. "Enzyme only" represents the sample only containing the enzymes used for digestion (S1 nuclease, venom phosphodiesterase I, and calf intestinal alkaline phosphatase) and omitting the mRNA. The sequence of the *in vitro* transcribed RNA can be found in Table S7 in the ESI.† The mass transitions 510.3 → 252.2 for CMCT labelled m^5U and 496.3 → 252.2 for CMCT labelled Ψ were used for the detection in MRM detection mode.

modifications, the detected modifications from total RNA and from mRNA of HEK293T cells, respectively. The results demonstrated that all these 4 modifications were clearly detected in total RNA (Fig. 5D), but were undetectable in mRNA (Fig. 5E), indicating the high purity of mRNA.

In addition, we used real-time quantitative PCR to evaluate the purity of the isolated mRNA. The detailed evaluation can be found in the ESI.† The real-time quantitative PCR analysis demonstrated that the potential contamination of small RNA was less than 0.023% in mRNA (Fig. S8 in the ESI†), *i.e.* less than 0.12 ng of small RNA in 500 ng of mRNA. We then analyzed m^5U from 500 ng of mRNA. The peak of m^5U from 500 ng of mRNA was clear and distinct (Fig. S9 in the ESI†). However, no signal of m^5U was observed from 0.12 ng of small RNA (Fig. S9 in the ESI†), indicating that the m^5U signal from 500 ng of mRNA indeed comes from mRNA, but not the potential trace level of small RNA. Similarly, the potential trace levels of 18S and 28S rRNA in mRNA should not compromise the detection of m^5U in mRNA (details can be found in the ESI†). Moreover, the RNA-seq analysis demonstrated that 12 299 genes were detected in the isolated mRNA (Table S5 and S6 in the ESI†), indicating the good quality of the isolated mRNA.

We then analysed the potential uridine modifications in mRNA. It turned out that among the 7 uridine modifications (Ψ, m^5U, m^6U, $m^1\Psi$, hm^5U, mo^5U, and mcm^5U), only m^5U and Ψ were detectable in mRNA. Shown in Fig. 6A and B are the extracted-ion chromatograms of CMCT labelled m^5U and Ψ in mRNA, respectively, from various cultured human cells. Although CMCT labelled m^5U, m^6U, and $m^1\Psi$ have the same molecular weight and the same MRM ion transition was used to monitor CMCT labelled m^5U, m^6U, and $m^1\Psi$ (Table S1 in the ESI†), they could be well separated in HPLC (the retention times of m^5U, m^6U, and $m^1\Psi$ were 21.06 min, 19.18 min and 18.55 min, respectively; Fig. S10A in the ESI†). Under these separation conditions, we confirmed that only m^5U was detected in mRNA; no m^6U or $m^1\Psi$ was detected (Fig. S10B in the ESI†). Since there is no detectable m^6U or $m^1\Psi$ in mRNA of cultured human cells, we used the shorter gradient (Fig. 6) instead of the longer gradient (Fig. S10 in the ESI†) to detect m^5U from mRNA.

It can be seen that m^5U and Ψ were undetectable in the control samples (only adding enzymes and omitting mRNA or *in vitro* transcribed RNA) (Fig. 6), which excludes the possibility that the detected m^5U and Ψ were from the contamination of

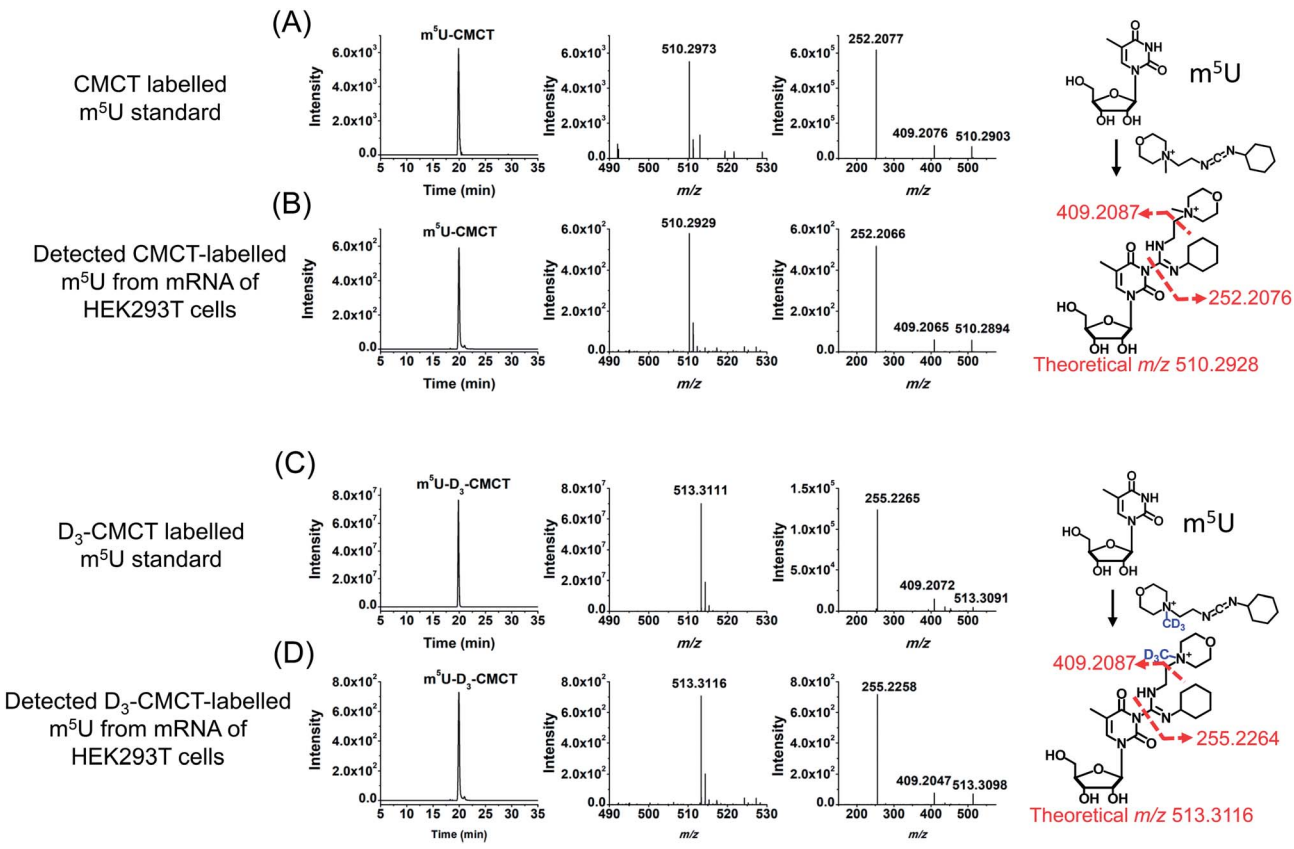

Fig. 7 Identification of m^5U in mRNA of HEK293T cells by high-resolution mass spectrometry analysis. (A) The extracted-ion chromatogram (left panel), MS spectrum (middle panel) and MS/MS spectrum (right panel) of the CMCT labelled m^5U standard. (B) The extracted-ion chromatogram (left panel), MS spectrum (middle panel) and MS/MS spectrum (right panel) of CMCT labelled m^5U from mRNA of HEK293T cells. (C) The extracted-ion chromatogram (left panel), MS spectrum (middle panel) and MS/MS spectrum (right panel) of the D_3-CMCT labelled m^5U standard. (D) The extracted-ion chromatogram (left panel), MS spectrum (middle panel) and MS/MS spectrum (right panel) of D_3-CMCT labelled m^5U from mRNA of HEK293T cells.

enzymes used in mRNA digestion. To further confirm the detected m^5U and Ψ in mRNA, we also employed high-resolution mass spectrometry to examine the CMCT labelled m^5U and Ψ. The results showed that the extracted-ion chromatograms, MS spectra and MS/MS spectra of CMCT/D_3-CMCT labelled m^5U and Ψ from the cultured HEK293T cells are identical to those of the standards (Fig. 7 and S11 in the ESI†), further confirming the detected m^5U and Ψ in mRNA. These results collectively suggested that both m^5U and Ψ were distinctly present in mRNA of mammals.

Ψ modification has been previously reported to be present in mRNA of mammals with a potential role in enhancing transcript stability.[35,36,50] It is worth noting that the universal existence of m^5U in mRNA of various cell lines is first reported in the current study. On the other hand, detection of m^5U by LC-ESI-MS/MS without CMCT labelling showed that no m^5U was observed (Fig. S12 in the ESI†), indicating that the direct analysis of m^5U is challenging and our developed chemical labelling method is capable of detecting low abundance of m^5U in mRNA.

Metabolic labelling of cells

To further confirm the occurrence of m^5U in mRNA of mammals, we used stable isotope tracing monitored by mass spectrometry (Fig. 8A). It has been known that ATP and *L*-methionine could be converted to *S*-adenosyl-*L*-methionine (SAM) by methionine adenosyltransferase, and SAM is a universal methylating agent for the methylation of nucleic acids.[51] We cultured human HEK293T cells in DMEM medium supplied with 0.3 mM stable isotope labelled *L*-methionine (D_3-Met). Then mRNA was isolated and analysed by our developed method. The result showed D_3-m^5U was distinctly detected besides m^5U (Fig. 8B). The concentration of unlabelled methionine in the DMEM medium is 0.2 mM. Thus, the percentage of D_3-Met in medium is approximately 60% of the total methionine. And approximately half of the measured m^5U carried the CD_3 group, which is comparable to the percentage of D_3-Met in total methionine of the media. Extended culturing time of cells in the DMEM medium with added D_3-Met will not increase the percentage of D_3-m^5U, further indicating that the methyl group of m^5U is from SAM. Collectively, this result indicated that SAM was the methyl donor for the methylation of uridine in mRNA.

Content of m^5U in different species of RNA

We then further quantitatively measured the content of m^5U in different RNA species. In this respect, we isolated different types of RNA, including mRNA, 28S rRNA, 18S rRNA and small RNA (<200 nt) from HEK293T and HeLa cells. The purified 28S rRNA, 18S rRNA, and small RNA (<200 nt) were confirmed by agarose and polyacrylamide gel electrophoresis analysis (Fig. S13 in the ESI†).

Notably, m^5U was detectable in all species of RNA in both HEK293T and HeLa cells. The measured contents of m^5U in mRNA of various human cells and mouse tissues ranged from 0.001% to 0.0059% (m^5U/U) (Fig. 9A and B). In addition, the measured contents of Ψ in mRNA ranged from 0.29% to 1.04% (Ψ/U) (Fig. 9C and D), which are comparable to previously

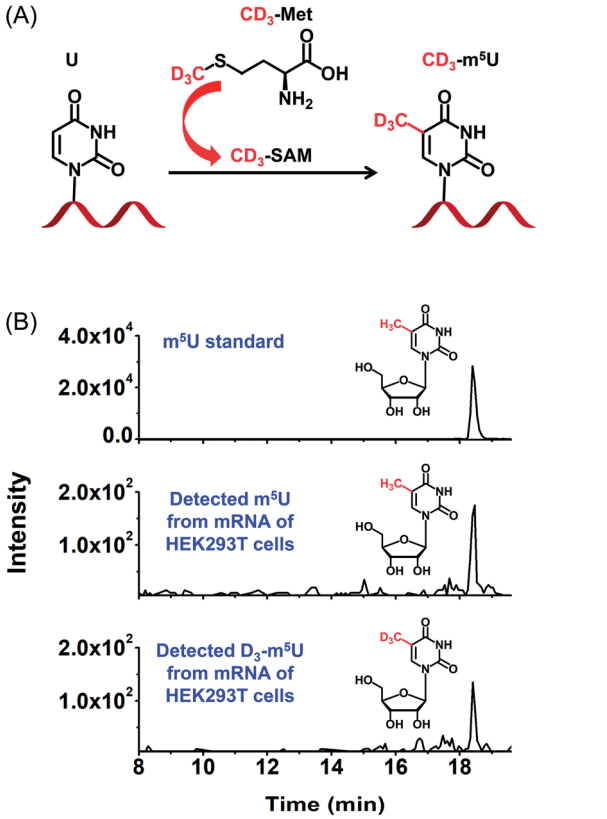

Fig. 8 Identification of m^5U in mRNA of HEK293T cells using stable isotope tracing monitored by mass spectrometry. (A) DMEM medium supplied with D_3-Met was used for the cell culturing. D_3-Met could be converted to D_3-SAM that is a methylating reagent for methylation of nucleic acids. (B) The extracted-ion chromatograms of the CMCT labelled m^5U standard (upper panel), detected CMCT labelled m^5U from mRNA of HEK293T cells (middle panel), and detected CMCT labelled D_3-m^5U from mRNA of HEK293T cells (bottom panel).

reported levels of Ψ in mRNA of mammals,[13,36] indicating the good accuracy of the developed method.

The contents of m^5U in small RNA (<200 nt) of HEK293T and HeLa cells were 1.9% and 1.6% (m^5U/U), respectively (Fig. 9E). The contents of m^5U in 28S and 18S rRNA of HEK293T and HeLa cells ranged from 0.0039% to 0.0057% (m^5U/U) (Fig. 9E). These results showed that small RNA (<200 nt) contained the most abundant m^5U, while mRNA and rRNA had similar levels of m^5U. It was reported that m^5U was an abundant modification in tRNA.[16] And a pervious study showed tRNA presents ~90% of small RNA (<200 nt).[52] Thus, the detected m^5U in small RNA should mainly come from tRNA.

It has been known that over 150 types of modifications were identified in RNA,[53] but only a small number of modifications were detected in mRNA of eukaryotes.[9,11] Here, we demonstrated the existence of appreciable levels of m^5U (0.001% to 0.0059%, m^5U/U) in mRNA of mammalian cells. The measured content of m^5U in mRNA is comparable to that of recently identified m^6Am (~0.003%, m^6Am/A) and m^3C (~0.001–0.005%, m^3C/C) in mRNA of mammals.[22,54] It has been proposed that the modifications of m^6Am and m^3C in mRNA may affect

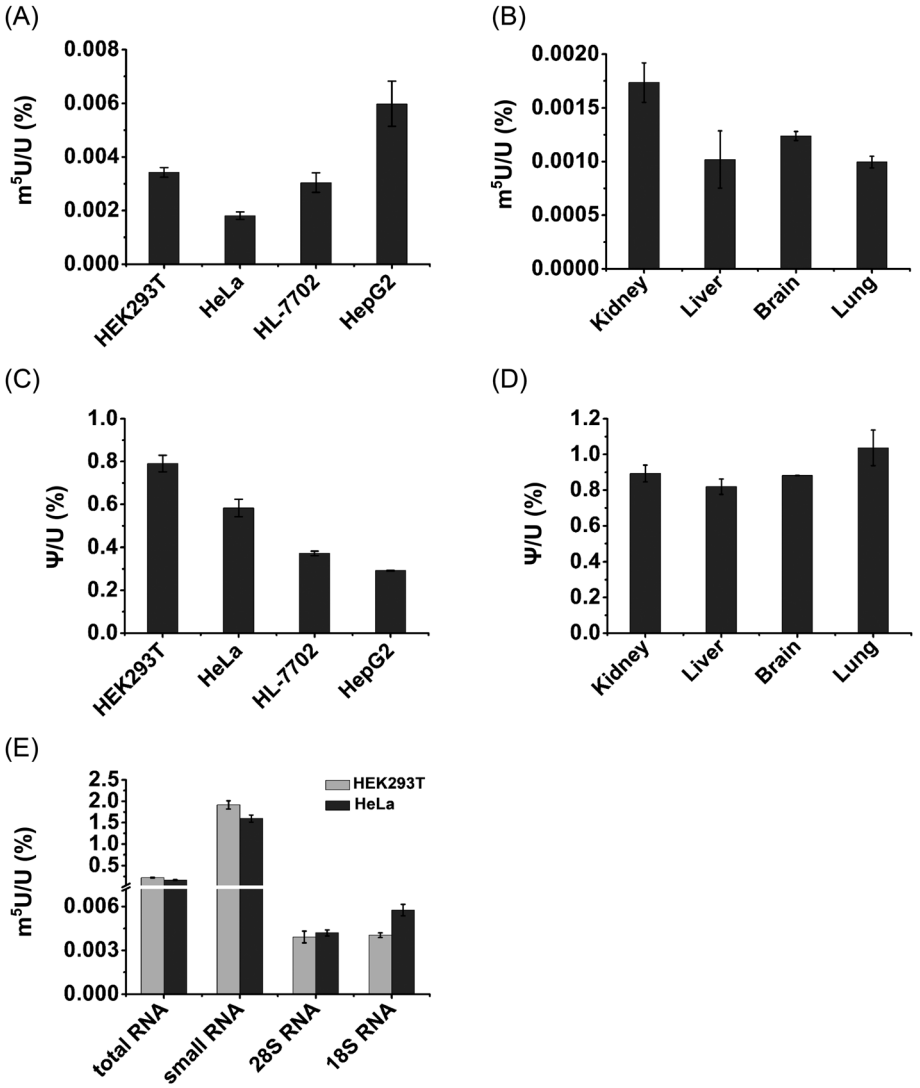

Fig. 9 Measured contents of m^5U in different cells, tissues and RNA species. (A) Measured contents of m^5U in mRNA from various human cell lines. (B) Measured contents of m^5U in mRNA from various mouse tissues. (C) Measured contents of Ψ in mRNA from various human cell lines. (D) Measured contents of Ψ in mRNA from various mouse tissues. (E) Measured contents of m^5U in different RNA species from HEK293T and HeLa cells. These data represent the mean and standard deviation of results acquired from three independent experiments.

the mRNA stability and translation efficiency.[55,56] Although the potential role of m^5U in mRNA remains elusive, we speculate that m^5U in mRNA may affect its interaction with the ribosome as some other modifications. m^5U could also lead to the alteration of RNA secondary structures and recruit m^5U-specific binding proteins, and eventually impact translation.

Evaluation of the level of m^5U in mRNA upon siRNA knockdown or overexpression of *TRMT2A*

In *E. coli*, TrmA catalyzes the formation of m^5U in tRNA.[57] In *Saccharomyces cerevisiae*, Trm2p is responsible for the formation of m^5U in tRNA.[58] TrmA and Trm2p catalyze methyl group transfer from *S*-adenosyl-methionine (SAM) to the C5 of U54 of tRNA.[59] In mammals, tRNA methyltransferase 2 homologs (TRMT2A and TRMT2B) are predicted to be the homologs of yeast Trm2p.[60] The predicted methyltransferase domain of TRMT2A has an identity of 36% with that of Trm2p. TRMT2A contains a RNA recognition motif (RRM) in the N-terminus. However, the shared identity between TRMT2B and Trm2p is 28%, and the TRMT2B protein does not have an RRM motif.[60] This indicates that TRMT2A is the likely mammalian homolog of Trm2p.[40] In this respect, here we examined whether TRMT2A was responsible for depositing m^5U in mRNA upon siRNA knockdown or overexpression of the *TRMT2A* gene.

The real-time quantitative PCR and western blotting results demonstrated the efficient knockdown or overexpression of TRMT2A was achieved (Fig. S14 in the ESI†). Then we analyzed the level of m^5U in small RNA and mRNA upon siRNA knockdown or overexpression of *TRMT2A*. The results showed a significant decrease of m^5U in small RNA ($p = 2.0 \times 10^{-5}$, Fig. 10A), indicating that TRMT2A may be responsible for the formation of m^5U in small RNA. However, no obvious increase

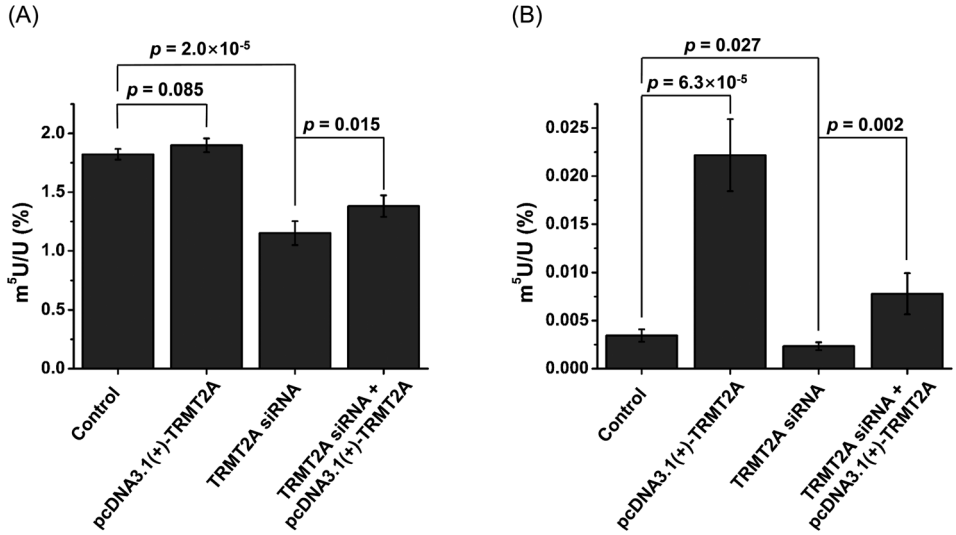

Fig. 10 Effect of overexpression and siRNA knockdown of *TRMT2A* on the contents of m^5U in small RNA and mRNA. (A) Contents of m^5U in small RNA upon overexpression or siRNA knockdown of *TRMT2A*. (B) Contents of m^5U in mRNA upon overexpression or siRNA knockdown of *TRMT2A*. Two-side unpaired *t*-test was performed. These data represent the mean and standard deviation of results acquired from three independent experiments.

of m^5U in small RNA was observed for the overexpression of *TRMT2A* ($p = 0.085$, Fig. 10A). We reason that the increased amount of TRMT2A does not affect the homeostasis of m^5U in small RNA in cells, which was also reported by a recent study.[40] The rescue experiment showed the certain recovery of m^5U in small RNA (Fig. 10A). As for the m^5U in mRNA, we observed a significant decrease ($p = 0.027$) and significant increase ($p = 6.3 \times 10^{-5}$) of m^5U in mRNA upon siRNA knockdown and overexpression of TRMT2A, respectively (Fig. 10B). The rescue experiment showed the efficient recovery of m^5U in mRNA (Fig. 10B). The results demonstrated that TRMT2A should be responsible for the formation of m^5U in mRNA.

Deciphering the distribution and dynamic change of m^5U in mRNA will uncover the underlying mechanism of m^5U. The transcriptome-wide mapping of m^5U could be achieved by affinity pulldown of the m^5U-containing RNA fragment followed by high-throughput sequencing. In addition, some modifications in RNA can lead to reverse transcription arrest or misincorporation of dNTP, which has been utilized as a characteristic signature for locating modifications by subsequent sequencing analysis.[61] By virtue of this strategy of reverse transcription induced arrest or misincorporation of dNTP, people may evolve enzymes that are suitable for probing m^5U and develop method for mapping m^5U at single-base resolution. Moreover, the exploration of the m^5U-specific binding proteins may reveal the readers of m^5U and promote the understanding of the functional roles of m^5U.

Conclusions

In summary, we established a method of CMCT labelling coupled with LC-ESI-MS/MS analysis for the sensitive determination of uridine modifications in RNA. The introduction of positively charged quaternary ammonium from CMCT can greatly increase the detection sensitivities of uridine modifications. Using the developed method, we identified the universal existence of m^5U in mRNA of cultured mammalian cells and tissues. In addition, the stable isotope tracing monitored by mass spectrometry revealed that the methylation group of m^5U was from SAM. The identification of demethylases and reader proteins of m^5U will facilitate understanding the functions of m^5U in the future.

Data availability

Gene expression by RNA-seq is available in the Gene Expression Omnibus (GEO) database at NCBI with accession number GSE135049.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgements

The work is supported by the National Key R&D Program of China (2017YFC0906800) and the National Natural Science Foundation of China [21672166, 21635006, 21721005, 21728802].

References

1 K. Chen, B. S. Zhao and C. He, *Cell Chem. Biol.*, 2016, **23**, 74–85.
2 C. Luo, P. Hajkova and J. R. Ecker, *Science*, 2018, **361**, 1336–1340.
3 M. Frye, B. T. Harada, M. Behm and C. He, *Science*, 2018, **361**, 1346–1349.

4 I. A. Roundtree, M. E. Evans, T. Pan and C. He, *Cell*, 2017, **169**, 1187–1200.
5 P. Boccaletto, M. A. Machnicka, E. Purta, P. Piatkowski, B. Baginski, T. K. Wirecki, V. de Crecy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm and J. M. Bujnicki, *Nucleic Acids Res.*, 2018, **46**, D303–D307.
6 T. Liu, C. J. Ma, B. F. Yuan and Y. Q. Feng, *Sci. China: Chem.*, 2018, **61**, 381–392.
7 J. Jiang, H. Seo and C. S. Chow, *Acc. Chem. Res.*, 2016, **49**, 893–901.
8 T. Pan, *Cell Res.*, 2018, **28**, 395–404.
9 J. Song and C. Yi, *ACS Chem. Biol.*, 2017, **12**, 316–325.
10 Y. Chen, T. Hong, S. Wang, J. Mo, T. Tian and X. Zhou, *Chem. Soc. Rev.*, 2017, **46**, 2844–2872.
11 B. S. Zhao, I. A. Roundtree and C. He, *Nat. Rev. Mol. Cell Biol.*, 2017, **18**, 31–42.
12 A. J. Fisher and P. A. Beal, *Curr. Opin. Struct. Biol.*, 2018, **53**, 59–68.
13 X. Li, X. Xiong and C. Yi, *Nat. Methods*, 2016, **14**, 23–31.
14 D. Arango, D. Sturgill, N. Alhusaini, A. A. Dillman, T. J. Sweet, G. Hanson, M. Hosogane, W. R. Sinclair, K. K. Nanan, M. D. Mandler, S. D. Fox, T. T. Zengeya, T. Andresson, J. L. Meier, J. Coller and S. Oberdoerffer, *Cell*, 2018, **175**, 1872–1886.
15 H. Lian, Q. H. Wang, C. B. Zhu, J. Ma and W. L. Jin, *Trends Cancer*, 2018, **4**, 207–221.
16 Y. Motorin and M. Helm, *Wiley Interdiscip. Rev.: RNA*, 2011, **2**, 611–631.
17 T. T. Lee, S. Agarwalla and R. M. Stroud, *Structure*, 2004, **12**, 397–407.
18 S. Hur, R. M. Stroud and J. Finer-Moore, *J. Biol. Chem.*, 2006, **281**, 38969–38973.
19 B. Felden, K. Hanawa, J. F. Atkins, H. Himeno, A. Muto, R. F. Gesteland, J. A. McCloskey and P. F. Crain, *EMBO J.*, 1998, **17**, 3188–3196.
20 J. T. Kealey, X. Gu and D. V. Santi, *Biochimie*, 1994, **76**, 1133–1142.
21 A. J. Westermann, S. A. Gorski and J. Vogel, *Nat. Rev. Microbiol.*, 2012, **10**, 618–630.
22 L. Xu, X. Liu, N. Sheng, K. S. Oo, J. Liang, Y. H. Chionh, J. Xu, F. Ye, Y. G. Gao, P. C. Dedon and X. Y. Fu, *J. Biol. Chem.*, 2017, **292**, 14695–14703.
23 Q. Y. Li, B. F. Yuan and Y. Q. Feng, *Chem. Lett.*, 2018, **47**, 1453–1459.
24 M. D. Lan, B. F. Yuan and Y. Q. Feng, *Chin. Chem. Lett.*, 2019, **30**, 1–6.
25 B. Chen, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2019, **91**, 743–756.
26 B. Liu, X. Liu, W. Lai and H. Wang, *Anal. Chem.*, 2017, **89**, 6202–6209.
27 R. Yin, J. Mo, M. Lu and H. Wang, *Anal. Chem.*, 2015, **87**, 1846–1852.
28 B. Zhao, Y. Yang, X. Wang, Z. Chong, R. Yin, S. H. Song, C. Zhao, C. Li, H. Huang, B. F. Sun, D. Wu, K. X. Jin, M. Song, B. Z. Zhu, G. Jiang, J. M. Rendtlew Danielsen, G. L. Xu, Y. G. Yang and H. Wang, *Nucleic Acids Res.*, 2014, **42**, 1593–1605.
29 Y. Tang, J. M. Chu, W. Huang, J. Xiong, X. W. Xing, X. Zhou, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2013, **85**, 6129–6135.
30 B. F. Yuan, *Methods Mol. Biol.*, 2017, **1562**, 33–42.
31 B. L. Qi, P. Liu, Q. Y. Wang, W. J. Cai, B. F. Yuan and Y. Q. Feng, *TrAC, Trends Anal. Chem.*, 2014, **59**, 121–132.
32 H. Y. Zhang, J. Xiong, B. L. Qi, Y. Q. Feng and B. F. Yuan, *Chem. Commun.*, 2016, **52**, 737–740.
33 H. Hong and Y. Wang, *Anal. Chem.*, 2007, **79**, 322–326.
34 K. G. Patteson, L. P. Rodicio and P. A. Limbach, *Nucleic Acids Res.*, 2001, **29**, E49.
35 S. Schwartz, D. A. Bernstein, M. R. Mumbach, M. Jovanovic, R. H. Herbst, B. X. Leon-Ricardo, J. M. Engreitz, M. Guttman, R. Satija, E. S. Lander, G. Fink and A. Regev, *Cell*, 2014, **159**, 148–162.
36 X. Li, P. Zhu, S. Ma, J. Song, J. Bai, F. Sun and C. Yi, *Nat. Chem. Biol.*, 2015, **11**, 592–597.
37 M. D. Lan, J. Xiong, X. J. You, X. C. Weng, X. Zhou, B. F. Yuan and Y. Q. Feng, *Chem.–Eur. J.*, 2018, **24**, 9949–9956.
38 G. Augusti-Tocco and G. L. Brown, *Nature*, 1965, **206**, 683–685.
39 J. M. Chu, C. B. Qi, Y. Q. Huang, H. P. Jiang, Y. H. Hao, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2015, **87**, 7364–7372.
40 Y. H. Chang, S. Nishimura, H. Oishi, V. P. Kelly, A. Kuno and S. Takahashi, *Biochem. Biophys. Res. Commun.*, 2019, **508**, 410–415.
41 C. Seto, J. S. Grossert, D. S. Waddell, J. M. Curtis and R. K. Boyd, *J. Am. Soc. Mass Spectrom.*, 2001, **12**, 571–579.
42 M. Antoine and J. Adams, *J. Am. Soc. Mass Spectrom.*, 1992, **3**, 776–778.
43 C. Cheng and M. L. Gross, *Mass Spectrom. Rev.*, 2000, **19**, 398–420.
44 M. Cydzik, M. Rudowska, P. Stefanowicz and Z. Szewczuk, *J. Am. Soc. Mass Spectrom.*, 2011, **22**, 2103–2107.
45 I. Behm-Ansmant, M. Helm and Y. Motorin, *J. Nucleic Acids*, 2011, **2011**, 408053.
46 N. W. Ho and P. T. Gilham, *Biochemistry*, 1967, **6**, 3632–3639.
47 T. Carell, C. Brandmayr, A. Hienzsch, M. Muller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs and M. Wagner, *Angew. Chem., Int. Ed.*, 2012, **51**, 7110–7131.
48 M. Taoka, Y. Nobe, Y. Yamaki, K. Sato, H. Ishikawa, K. Izumikawa, Y. Yamauchi, K. Hirota, H. Nakayama, N. Takahashi and T. Isobe, *Nucleic Acids Res.*, 2018, **46**, 9289–9298.
49 M. Duechler, G. Leszczynska, E. Sochacka and B. Nawrot, *Cell. Mol. Life Sci.*, 2016, **73**, 3075–3095.
50 T. M. Carlile, M. F. Rojas-Duran, B. Zinshteyn, H. Shin, K. M. Bartoli and W. V. Gilbert, *Nature*, 2014, **515**, 143–146.
51 B. J. Landgraf, E. L. McCarthy and S. J. Booker, *Annu. Rev. Biochem.*, 2016, **85**, 485–514.
52 D. Su, C. T. Chan, C. Gu, K. S. Lim, Y. H. Chionh, M. E. McBee, B. S. Russell, I. R. Babu, T. J. Begley and P. C. Dedon, *Nat. Protoc.*, 2014, **9**, 828–841.
53 M. A. Machnicka, K. Milanowska, O. Osman Oglou, E. Purta, M. Kurkowska, A. Olchowik, W. Januszewski, S. Kalinowski, S. Dunin-Horkawicz, K. M. Rother, M. Helm, J. M. Bujnicki and H. Grosjean, *Nucleic Acids Res.*, 2013, **41**, D262–D267.

54 B. Molinie, J. Wang, K. S. Lim, R. Hillebrand, Z. X. Lu, N. Van Wittenberghe, B. D. Howard, K. Daneshvar, A. C. Mullen, P. Dedon, Y. Xing and C. C. Giallourakis, *Nat. Methods*, 2016, **13**, 692–698.
55 J. Mauer, X. Luo, A. Blanjoie, X. Jiao, A. V. Grozhik, D. P. Patil, B. Linder, B. F. Pickering, J. J. Vasseur, Q. Chen, S. S. Gross, O. Elemento, F. Debart, M. Kiledjian and S. R. Jaffrey, *Nature*, 2017, **541**, 371–375.
56 F. Liu and C. He, *J. Biol. Chem.*, 2017, **292**, 14704–14705.
57 T. Ny and G. R. Bjork, *J. Bacteriol.*, 1980, **142**, 371–379.
58 M. E. Nordlund, J. O. Johansson, U. von Pawel-Rammingen and A. S. Bystrom, *RNA*, 2000, **6**, 844–860.
59 J. Urbonavicius, G. Jager and G. R. Bjork, *Nucleic Acids Res.*, 2007, **35**, 3297–3305.
60 W. L. Towns and T. J. Begley, *DNA Cell Biol.*, 2012, **31**, 434–454.
61 S. Schwartz and Y. Motorin, *RNA Biol.*, 2017, **14**, 1124–1137.