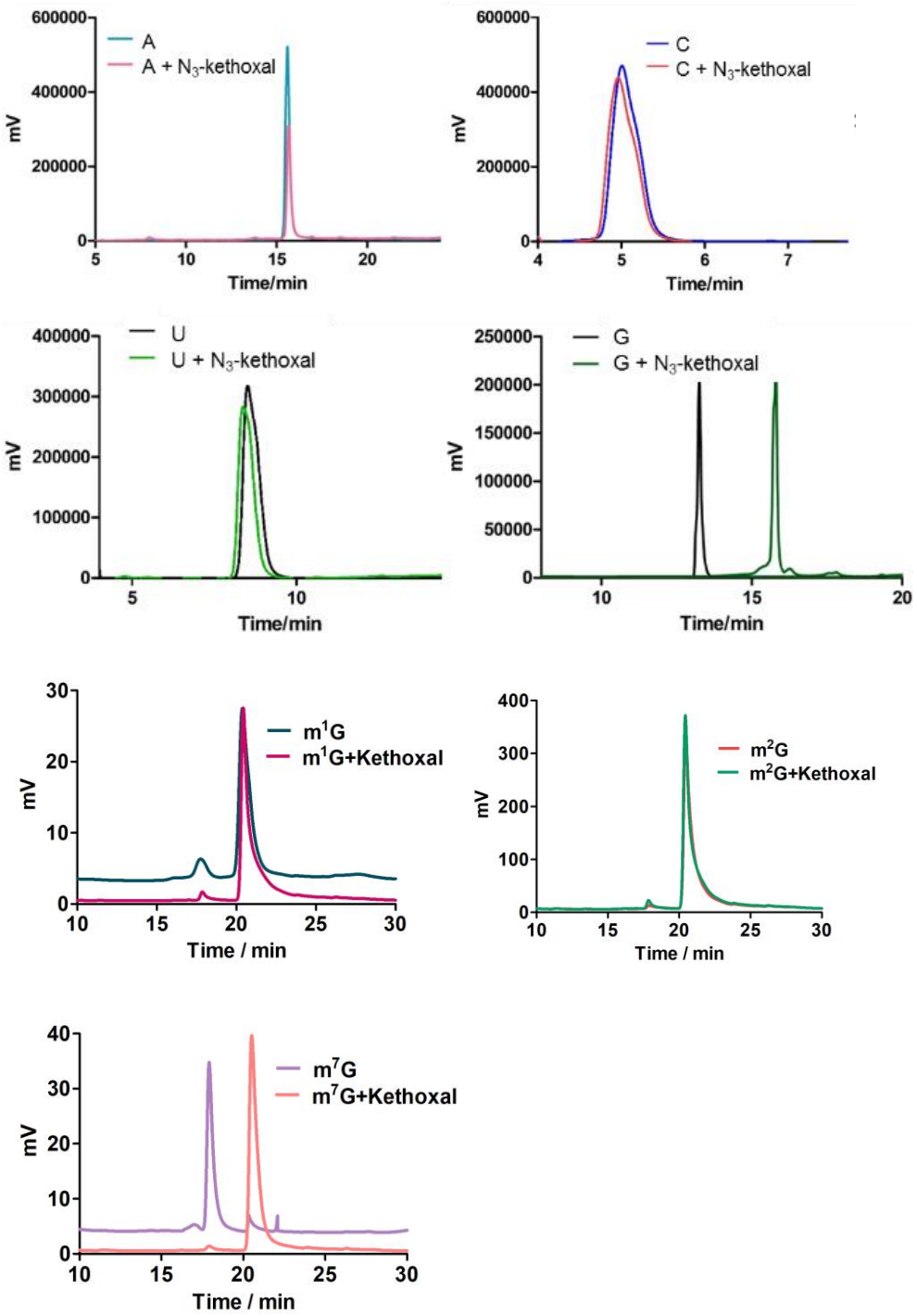
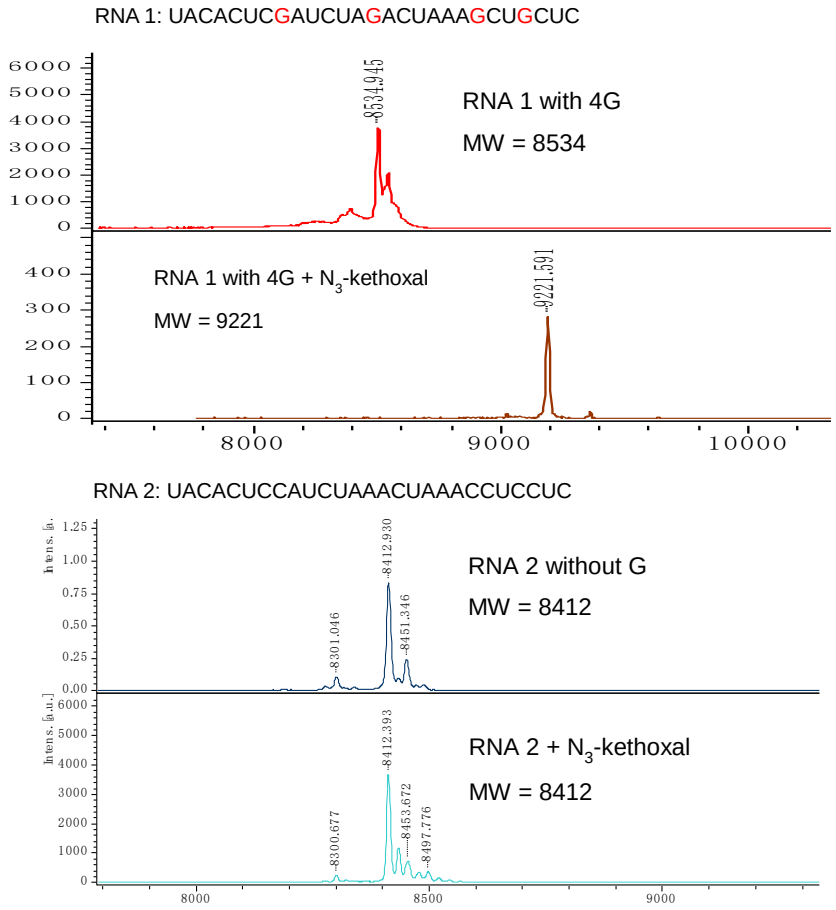
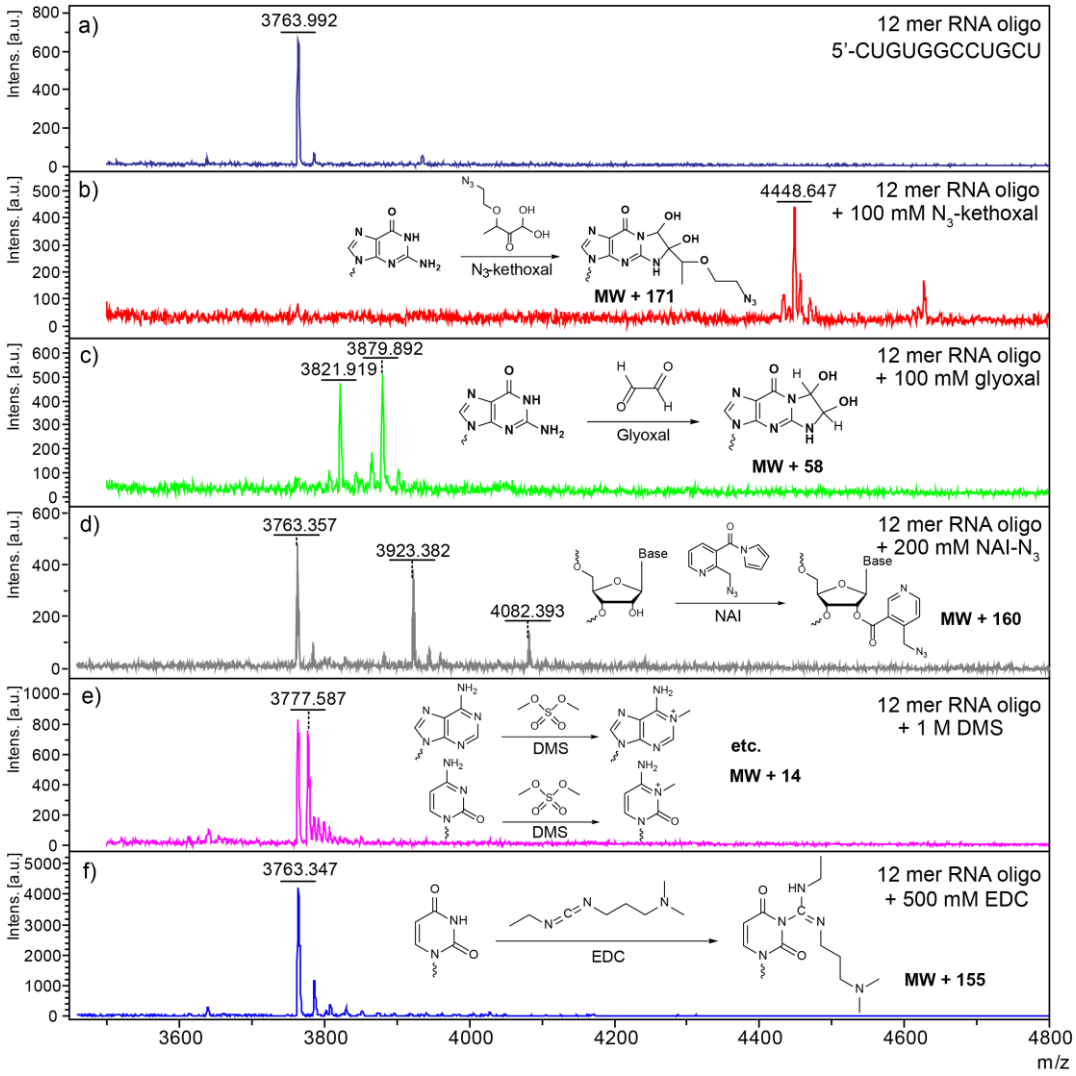
Supplementary Figures



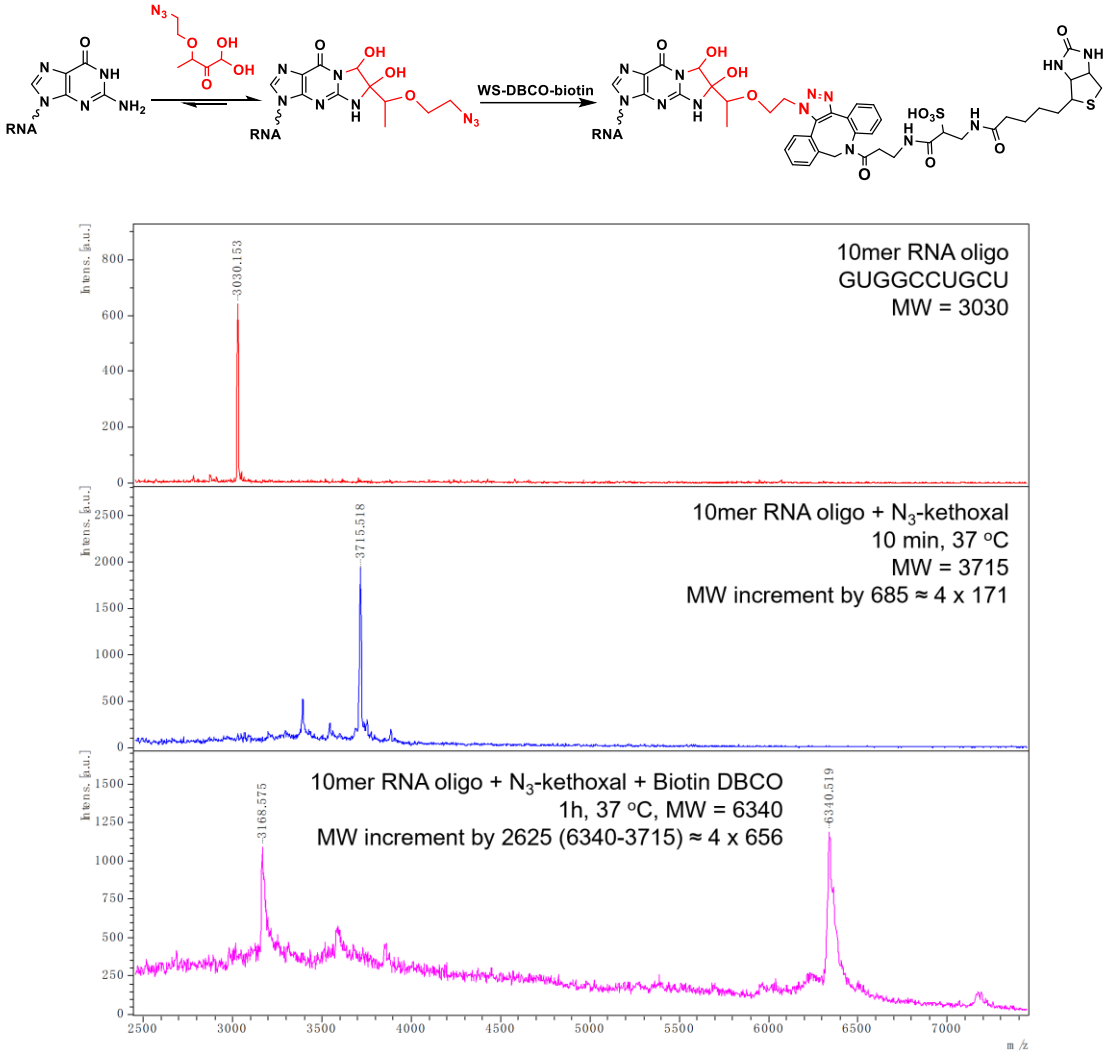
Supplementary Figure 1 | The HPLC results of N_3-kethoxal with RNA nucleosides (A, U, C, G, m^1G, m^2G, m^7G). N_3-kethoxal was inert with A, C, and U. For guanine and its derivatives, only G and m^7G can react with N_3-kethoxal. Both m^1G and m^2G, with methylation on N1 or N2 positions respectively, cannot be labeled by N_3-kethoxal. Experiments were repeated twice with similar results obtained.



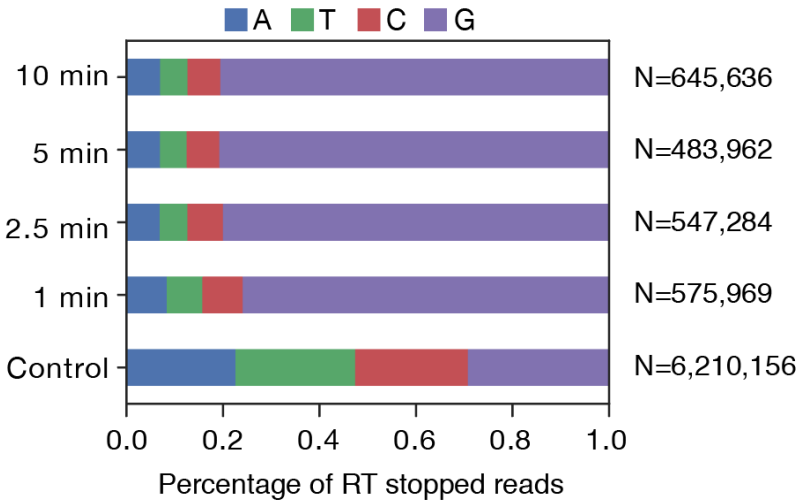
Supplementary Figure 2 | The RNA labeling reactivity of N_3-kethoxal towards RNA probes with or without guanines. RNA probes were incubated with N_3-kethoxal and then purified by BioRad Micro Bio-Spin™ P-6 Gel columns. The reaction products were analyzed by MALDI-TOF-MS. Experiments were repeated twice with similar results obtained.



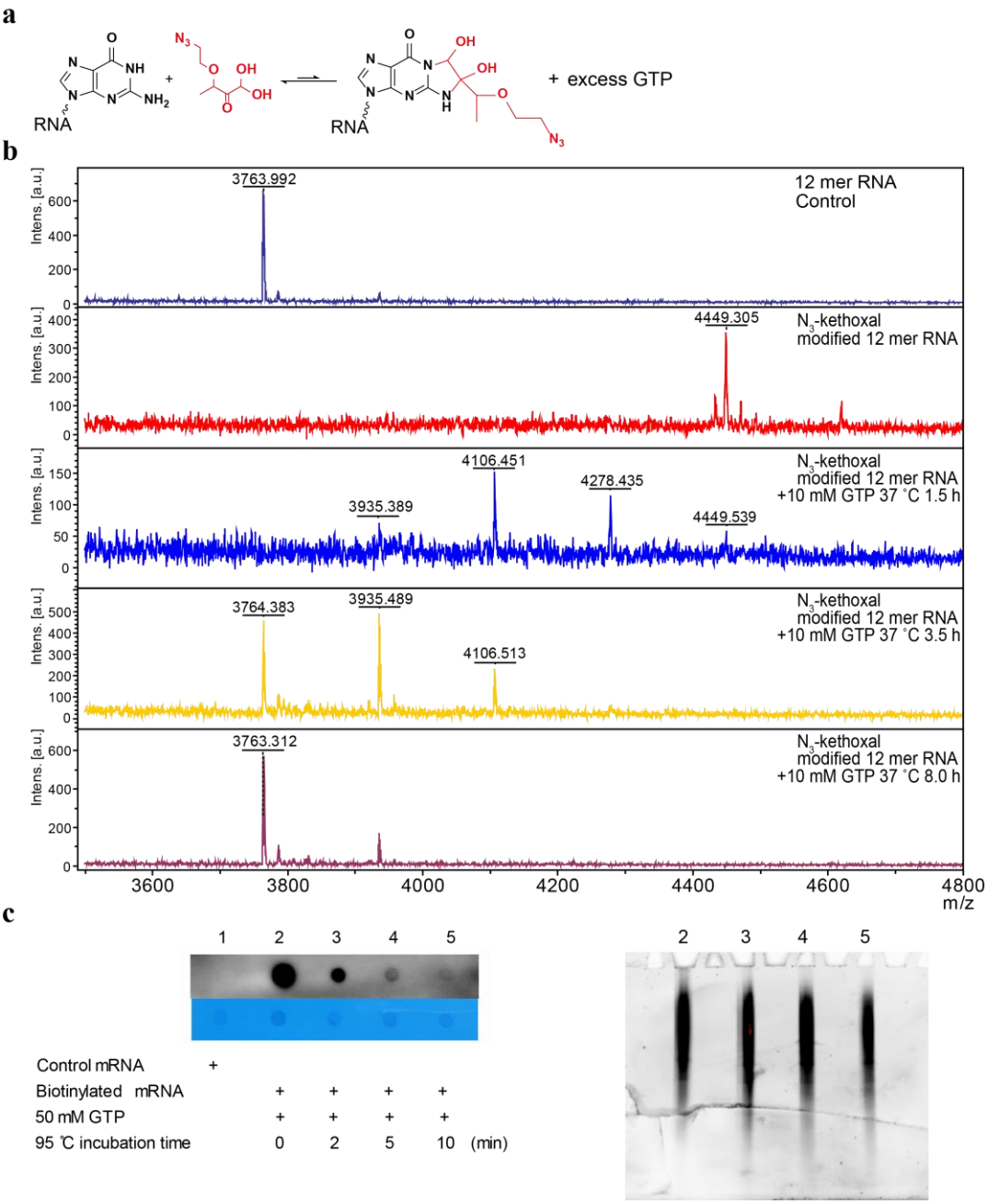
Supplementary Figure 3 | The comparison of RNA labeling reactivity of N_3-kethoxal and other known RNA secondary structure probes. The RNA was isolated by Biorad Micro Bio-Spin™ P-6 Gel Columns and analyzed by MALDI-TOF-MS. (**a**) MALDI–MS analysis of 12 mer RNA oligo (5'-CUGUGGCCUGCU). (**b**) MALDI–MS analysis of 12 mer RNA oligo after reacting with N_3-kethoxal (100 mM). All guanines were modified by N_3-kethoxal with the increment of molecular weight (MW) by 684 (one N_3-kethoxal modification with 171 MW increment). (**c**) MALDI–MS analysis of 12 mer RNA oligo after reacting with glyoxal (100 mM). The MW were 3821 and 3879, which means one or two guanines were labeled by glyoxal. (**d**) MALDI–MS analysis of 12 mer RNA oligo after reacting with icSHAPE probe NAI-N_3 (200 mM). The MWs were 3923 and 4082, which represent addition of one or two NAI-N_3 modifications in RNA, respectively. (**e**) MALDI–MS analysis of 12 mer RNA oligo after reacting with DMS (1 M). The MW of RNA oligo was increased to 3777, which means only one base was labeled by DMS even at high concentration (1 M). (**f**) MALDI–MS analysis of 12 mer RNA oligo after reacting with EDC (500 mM). No obvious MW increment was observed. These results indicate that N_3-kethoxal has a higher labeling activity than other reported RNA secondary structure probes. The experiment was independently repeated twice with similar results obtained. Experiments were repeated twice with similar results obtained.



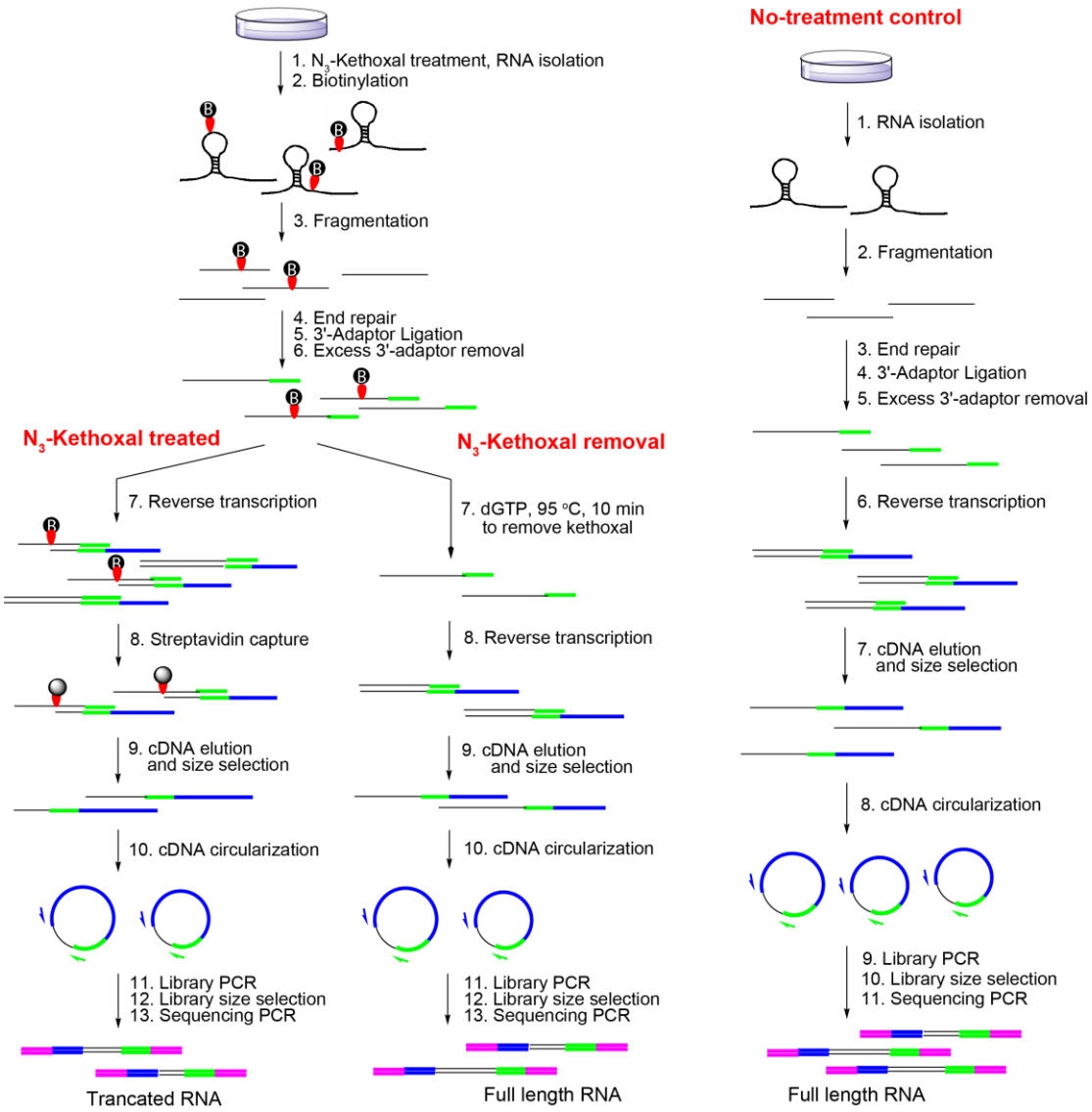
Supplementary Figure 4 | The scheme (upper) and mass spectrum (bottom) analysis of the N_3-kethoxal labeling and biotinylation reaction. 10 mer RNA with four guanines was used in this experiment. The experiment was independently repeated twice with similar results obtained.



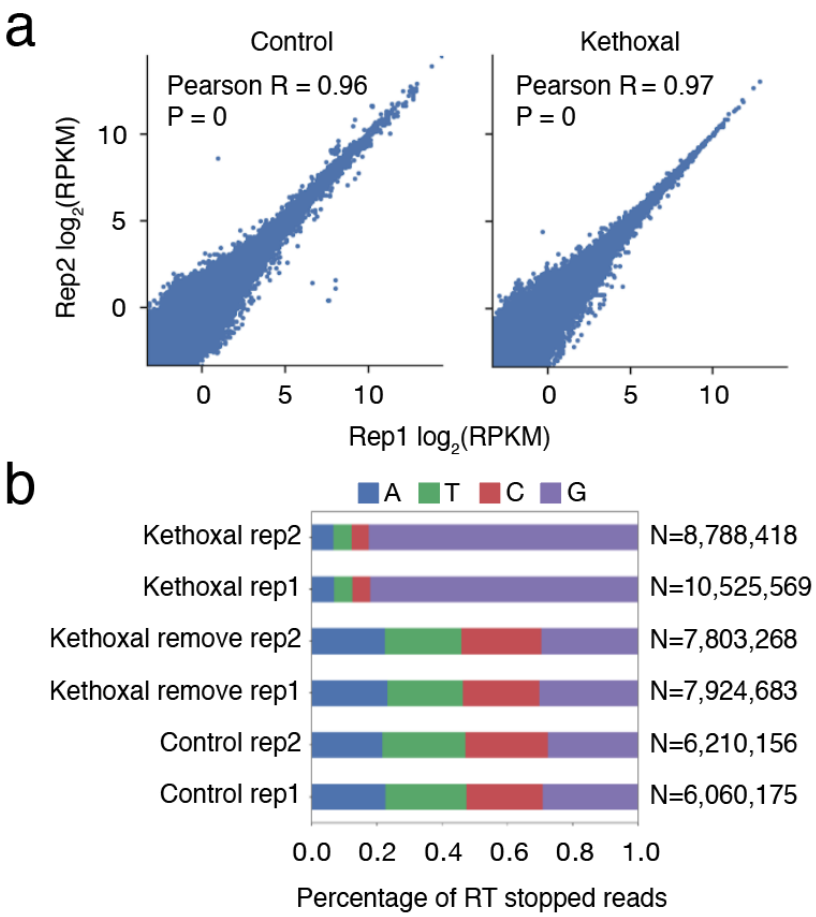
Supplementary Figure 5 | The time-dependence of G labeling by N_3-kethoxal in the cellular state. N_3-kethoxal was added, and then washed away after 1-10 min treatment. Then RNA were isolated by Trizol immediately and Keth-seq was performed. G-stop sites were saturated at 2.5 min.



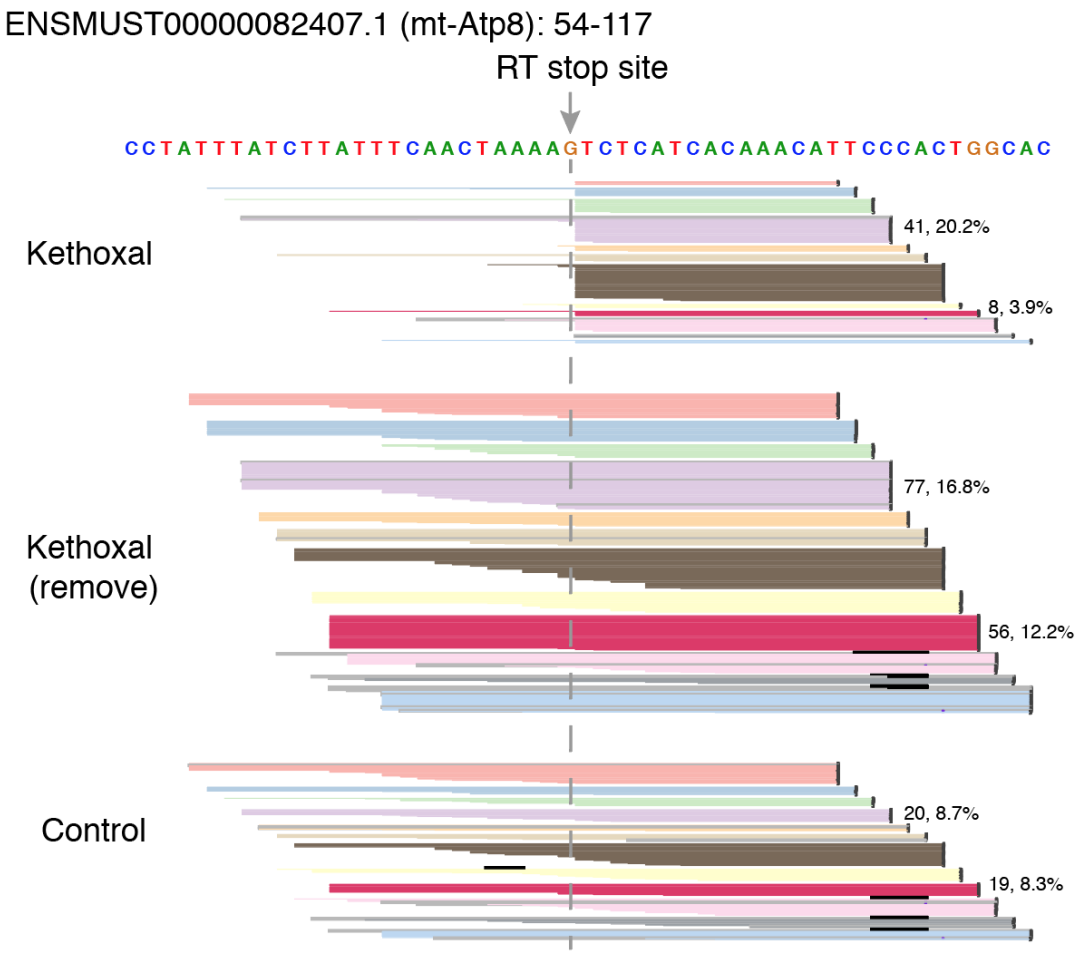
Supplementary Figure 6 | Reversibility of N_3-kethoxal labeling. (**a**) The schemes indicate equilibrium change of N_3-kethoxal-guanine reaction in the presence of excess GTP. The N_3-kethoxal-guanine can be reverted to unmodified RNA in the presence of excess GTP. (**b**) The reversibility of N_3-kethoxal-labeled RNA monitored by MALDI-TOF. The N_3-kethoxal modifications were mostly removed after 1.5 h incubation with 10 mM GTP in neutral buffer at 37 ˚C, and was almost completely removed after 8 h incubation. Experiments were repeated twice with similar results obtained. (**c**) Left: Dot blot result of N_3-kethoxal-labelled mRNA with 50 mM GTP at 95 ˚C in PBS buffer (pH = 7.4) during different incubation periods. Right: denaturing gel electrophoresis of the RNA fragments under different incubation conditions same as the dot blot samples. The length of fragmented RNA was mostly un-affected at 95 ˚C within 10 min in the presence of excess GTP. Experiments were independently repeated twice with similar results obtained.



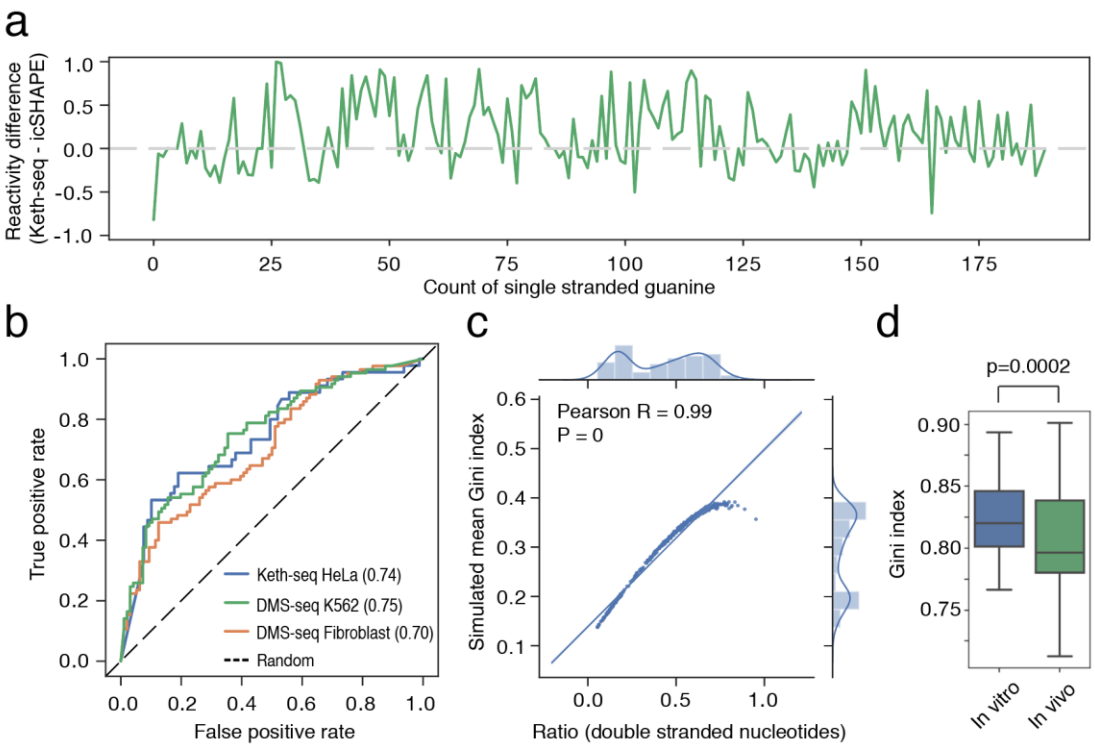
Supplementary Figure 7 | A flow chart showing library preparation procedures of the N_3-kethoxal-treated sample, the N_3-kethoxal-removal sample, and the no-treatment control sample. The detailed protocol is included in Supplementary Note 2.



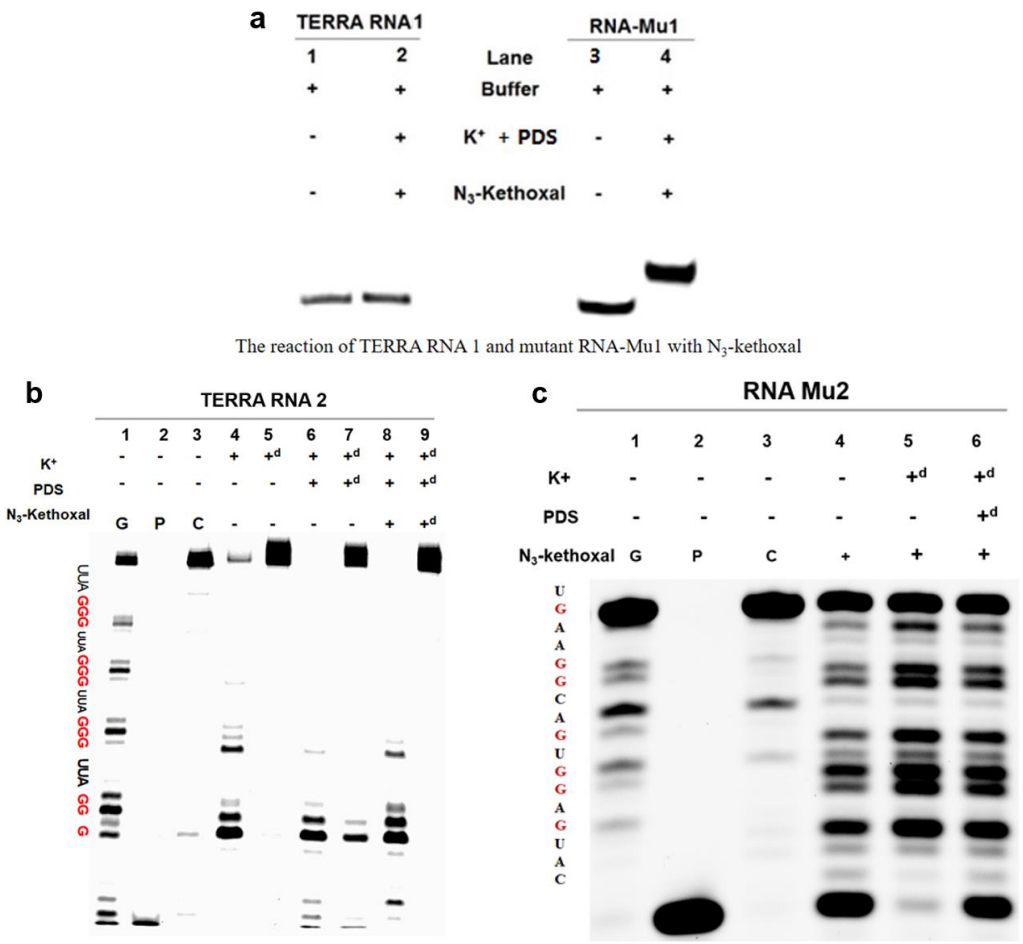
Supplementary Figure 8 | Quality control of Keth-seq. (**a**) RPKM correlation between replicates for N_3-kethoxal (right) and no-treat control (left) sample. 69,594 transcripts were detected in all replicates and were therefore plotted. The R and P denotes Pearson's correlation coefficient and two-tailed p-value. (**b**) RT stopped the reads distribution of replicates for N_3-kethoxal, N_3-kethoxal-remove samples, and no-treat control samples, the N number on the right denotes total stopped reads for each sample.



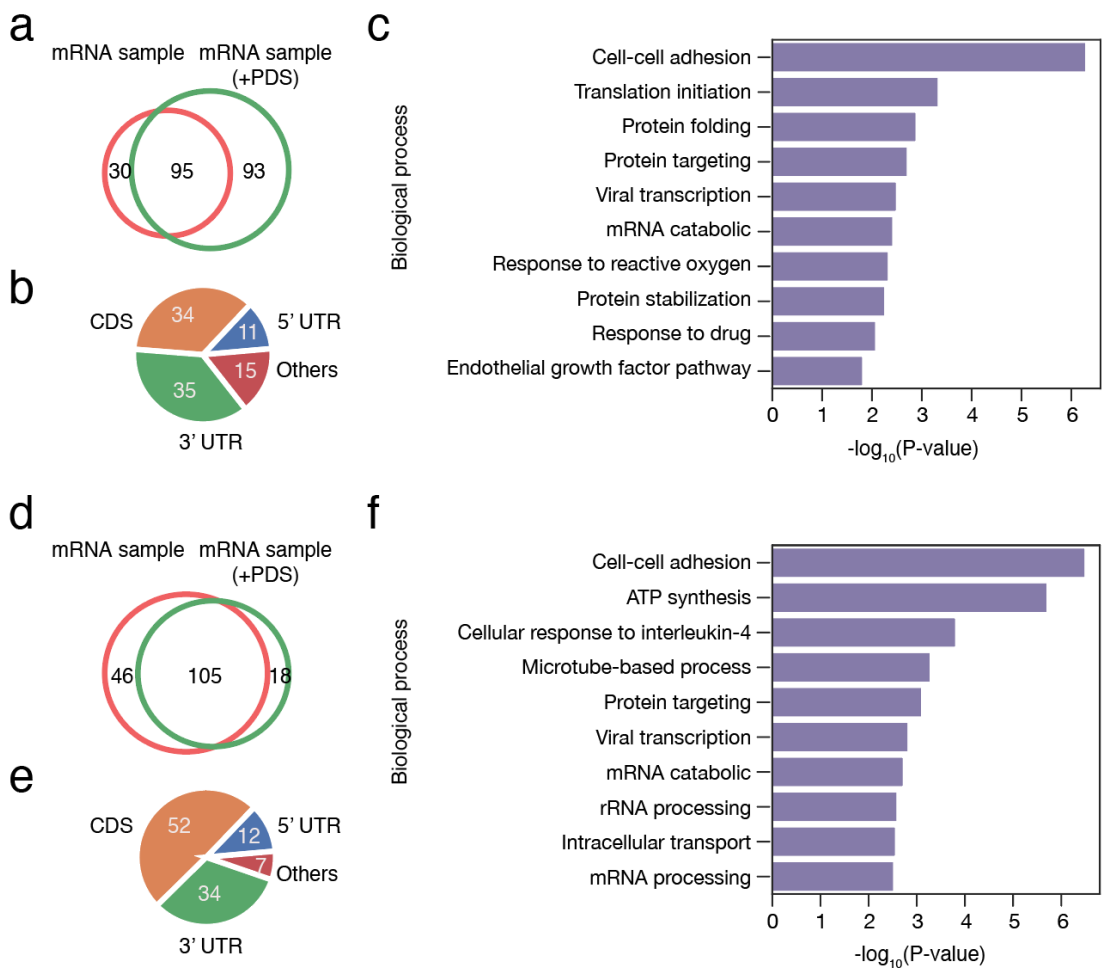
Supplementary Figure 9 | An example of read distribution among different samples. For protein coding RNA mt-Atp8, all mapped reads with 3' positions from 97 to 108 are shown. Reads were grouped and colored by end position. In the N_3-kethoxal sample, most of the reads stopped at guanine, while in N_3-kethoxal-remove and no-treat control samples, the reads stops were evenly distributed, suggesting the high specificity of N_3-kethoxal molecule. For most of the read groups, much longer reads with the same end in the removal sample were observed than in the control (for example, the highlighted cyan and red read groups), indicating the RT stopped sites with more confidence.



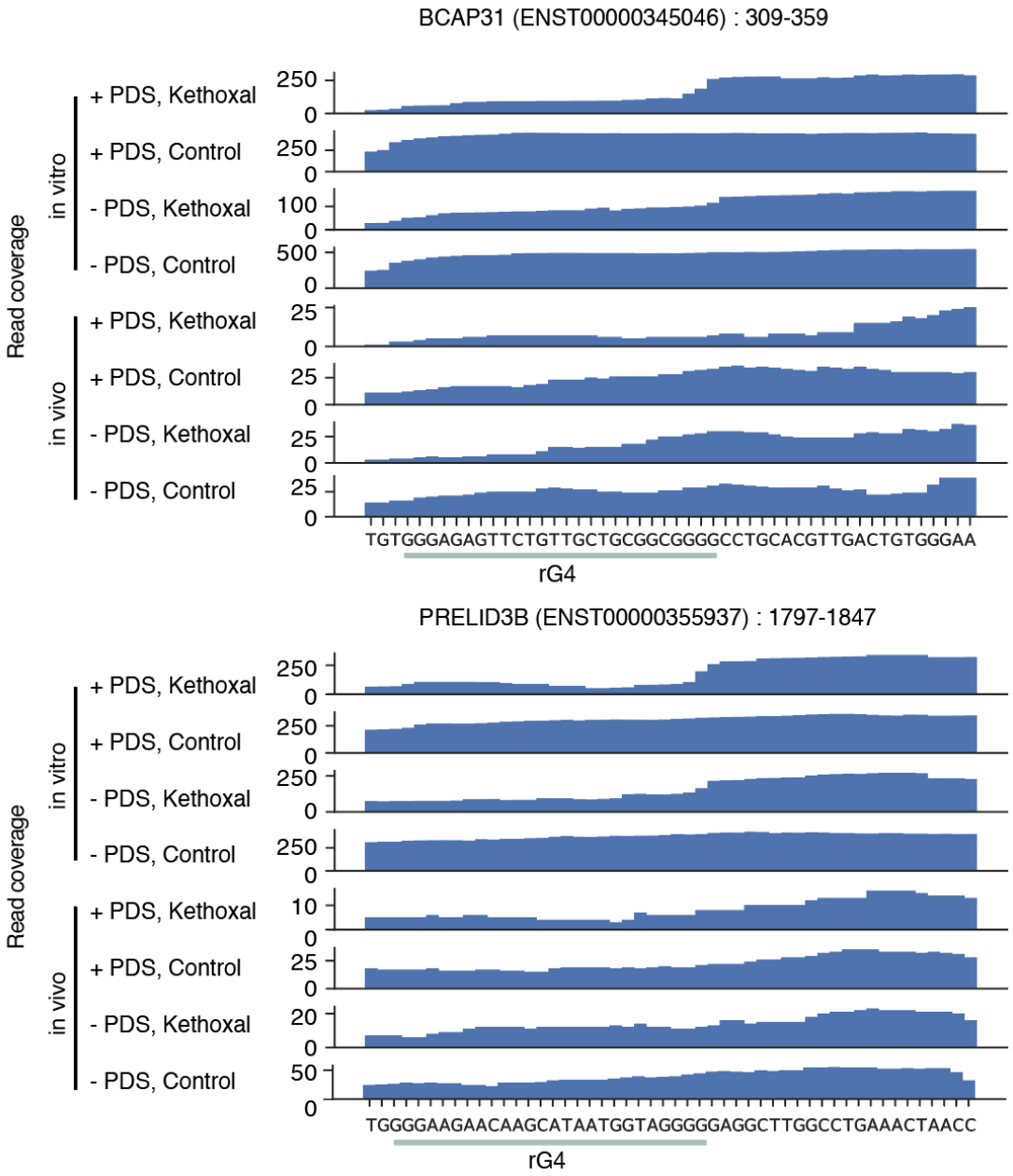
Supplementary Figure 10 | Transcriptome-wide comparason of the behaviour of Keth-seq with DMS-seq and icSHAPE. (**a**) Distribution of reactivity difference of single-stranded guanine for mouse 18S ribosomal RNA between Keth-seq and icSHAPE. The reactivity difference (Keth-seq – icSHAPE) on most guanines are larger than 0, indicating the sensitivity of Keth-seq on guanine bases is higher than icSHAPE. (**b**) Comparison of Keth-seq and DMS-seq on human 18S ribosomal RNA structure, the number in the brackets means AUC. Keth-seq behaves similarly as DMS-seq on the strcuture probing on 18S rRNA (n = 1,869). (**c**) Correlation between the double stranded nucleotides ratio and Gini index by using all mouse RNAs from the Rfam database (n = 614) with known secondary structures. The R and P denotes Pearson's correlation coefficient and two-tailed p-value. (**d**) Gini index distribution of all transcripts for *in vivo* and *in vitro* mESC Keth-seq samples. 78 transcripts were shown in both *in vivo* and *in vitro* samples and were therefore plotted. The p-value is calculated using two-sided t-test. The box spans first to last quartiles. The center line denotes median and whiskers represent 1.5× the interquartile range.



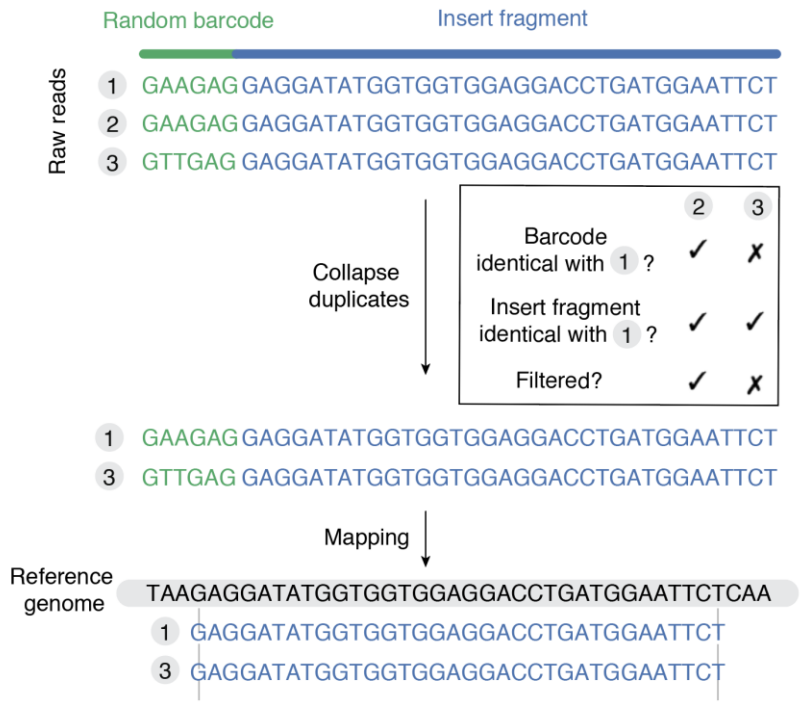
Supplementary Figure 11 | RNA G-quadruplex (rG4) probing by N_3-kethoxal. (**a**) N_3-kethoxal cannot react with rG4 form of TERRA RNA 1 (lane 1 and lane 2) but can modify unstructured RNA oligo RNA-Mu1 (lane 3 and lane 4). The Watson-Crick interface of guanine is the reactive site with N_3-kethoxal, which is blocked in the rG4 structure. TERRA RNA 1 (can form rG4): 5'-UUAGGGUUAGGGUUAGGGUUAGGG. RNA-Mu 1 (cannot form rG4): 5'-UAAGAUUACCGUUAGCGUUUAAAUUUAAAUUUAAA. (**b**) The primer extension assay of TERRA RNA 2 after N_3-kethoxal treatment. In lanes 4 and 6, RTSs caused by rG4 formation were observed, validating that K^+ and PDS can induce the formation of rG4. In lane 8 with N_3-kethoxal treatment, the RTS pattern remains the same as in lanes 4 and 6, indicating that formation of rG4 blocks N_3-kethoxal labeling. After the reaction shown in lane 8 was performed, RNA was purified for RTS assay, with the result shown in lane 9. The full-length cDNA, indicating that rG4 was not labelled by N_3-kethoxal. (**c**) The primer extension assay of RNA Mu 2 after N_3-kethoxal treatment. Similar RTSs were observed in lane 4, 5 and 6, indicating N_3-kethxal labeling without rG4 formation. Lane 1-3 in (b) and (c) indicated G ladder (G), primers (P), and full-length cDNA controls (C), respectively. Lanes annotated with a "d" letter indicate that the RNA was purified to remove K^+, PDS and N_3-kethoxal after the reaction. TERRA RNA 2: UCUCAUUUAGGGUUAGGGUUAGGGUUAGGGUAUCACAAAAAAAAAAAAAAAA AA. Primer 1: 5'-FAM – TTTTTTTTTTTTTTTTT. RNA-Mu 2: UGAAGGCAGUGGAGUACAACAAAUUAACAAACCAA. Primer 2: 5'-FAM – TTGGTTTGTTAATTTGTT. These data collectively show that N_3-kethoxal can be used to detect the formation of rG4 structures.



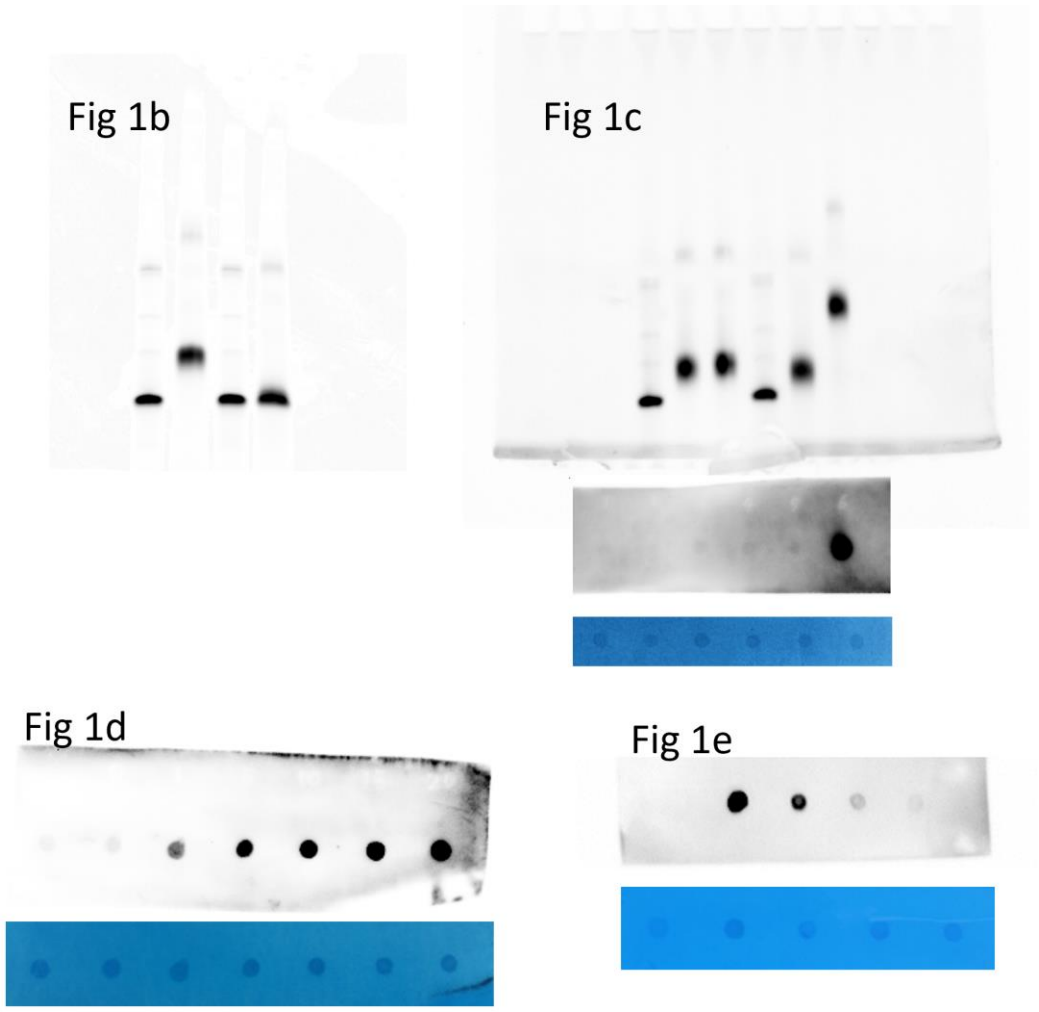
Supplementary Figure 12 | Transcriptome-wide distribution and gene ontology analysis of potential rG4 regions detected by Keth-seq *in vitro* and *in vivo* . (a) Overlap of rG4 region with structure information between native and +PDS *in vitro* samples. (b) Genomic context distribution of common rG4 regions in both native and +PDS *in vitro* samples. (c) Functional annotation of common rG4 regions involved genes (n = 50) in both native and +PDS *in vitro* samples. The P-values are provided by DAVID online tool and calculated by Fisher's exact test. (d) Overlap of rG4 region with structure information between native and +PDS *in vivo* samples. (e) Genomic context distribution of common rG4 regions in both native and +PDS *in vivo* samples. (f) Functional annotation of common rG4 regions involved genes (n = 81) in both native and +PDS *in vivo* samples. The P-values are provided by DAVID online tool and calculated with Fisher's exact test.



Supplementary Figure 13 | Two examples showing coverage tracks around potential rG4 regions of kethoxal and control samples under native or PDS treatment conditions for *in vitro* and *in vivo*, respectively. Reads coverage in +PDS Kethoxal samples are lower than their correspondence control sample, indicating the folding of rG4s. These two potential rG4 regions revealed by Keth-seq are also reported to form rG4 fold *in vitro* by rG4-seq (Kwok et.al., 2016, Nature Method).



Supplementary Figure 14 | Using in-line barcodes to collapse PCR duplicates and detect RT stops accurately. PCR duplication is a frequent problem for ensemble sequencing when many copies of the same loci exist and fragmentation generates many identical short fragments. To solve this problem, barcodes of random hexamer sequences (NNNNNN) were added to the fragments during library construction. These random barcodes serve to distinguish PCR duplicates from real different fragments with identical sequences. As the total number of different barcodes is large (4^6 = 4,096), the chance for two fragments to share the same barcode is low. Thus, the same fragment sequences with different random barcodes are regarded as real different fragments as the results of random fragmentation, while the same fragment sequences with the same random barcodes are PCR duplicates. For instance, the following illustration shows 3 reads including a PCR duplicate and a real different fragment. Reads 1 and 2 share the same barcode and identical insert sequence, and thus are PCR duplicates. Only one of them will be kept in the collapsing step. In contrast, the barcode region of read 3 is different from read 1, thus being retained even though they have the same insert sequences. After pre-processing, read 1 and read 3 will be mapped to the same loci and share identical start/stop positions.



Supplementary Figure 15 | Uncropped images related to Fig. 1. Information about which main figure does each uncropped gel image corresponds to are noted.

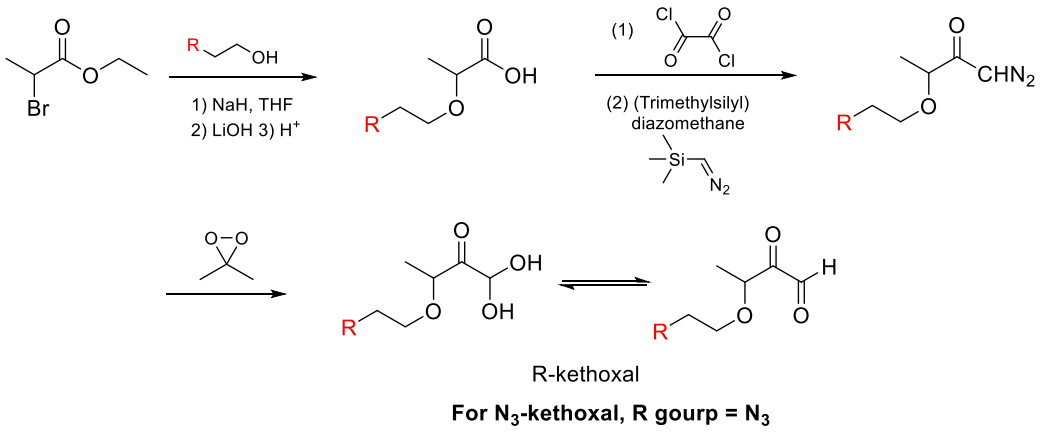
Supplementary Note 1

The synthesis of N_3-kethoxal and compounds characterization

Synthesis of kethoxal derivatives

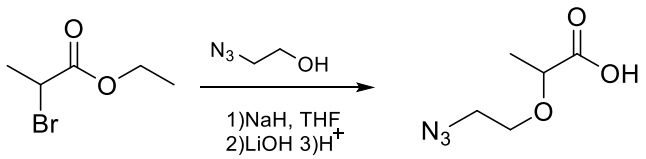
Kethoxal and its analogues were first reported to react with and inactivate the RNA virus since the 1950s[1]. The 1,2-dicarbonyl group of kethoxal showed high specificity to guanine, which make it very useful in the probing of RNA secondary structure. In addition, other kethoxal derivatives, such as kethoxal bis(Thiosemicarbazone) (KTS)[2] displayed promising anticancer activity, bikethoxal[3] demonstrated the ability to cross-link RNA and proteins within intact ribosomal 30S and 50S subunits. However, it is surprising that the synthesis of kethoxal and its derivatives are rarely reported. A review of the literature indicates that kethoxal preparation was mostly based on oxidation by selenium dioxide following purification by vacuum distillation[3,4,5]. This method has several limitations. First, metal oxidation reaction always results in byproducts. Second, the excess selenium was hard to remove. Third, synthesis of kethoxal derivatives with other functional groups is difficult because the reagents with functional groups may not be survived with selenium dioxide under reflux conditions. For example, our study indicates that azide- and thiol-modified kethoxal cannot be prepared by selenium dioxide oxidation. Lastly, vacuum distillation purification is not suitable for kethoxal derivatives with high-molecule weight.

Glyoxal and its analogues are sensitive to air and therefore cannot be purified by chromatography[6]. The mild oxidation of diazoketone by freshly prepared dimethyl-dioxirane (DMD) can produce a glyoxal functional group in quantitative yield[6]. In this study, azide-kethoxal was prepared through a novel synthetic strategy following a three-step synthesis (Scheme S1). The advantage of the synthetic process is its easy-to-operate high yield. What's more, this strategy is also convenient for the preparation of other kethoxal derivatives with various functional groups.



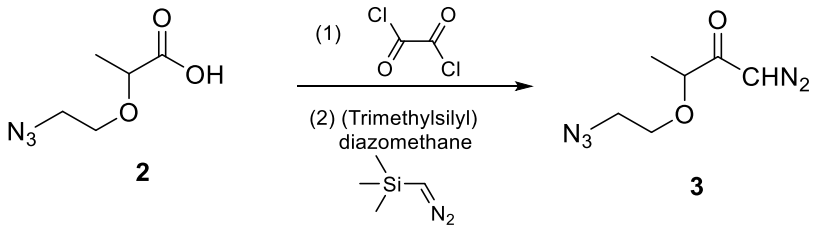
Supplementary scheme 1 | The strategy for the synthesis of kethoxal derivatives from various functional groups (R group)

The synthesis route of N_3-kethoxal

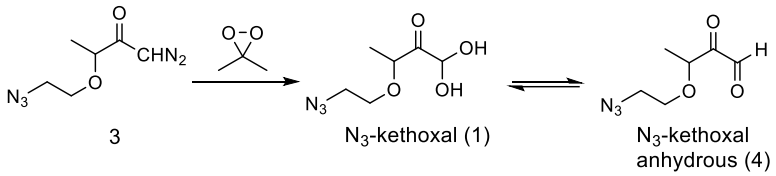


2

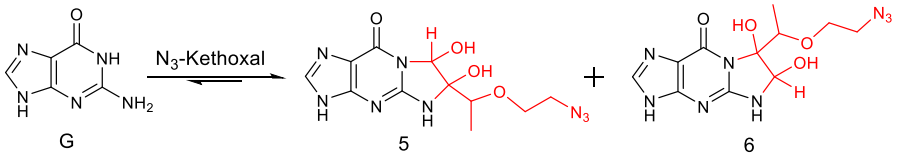
2-(2-azidoethoxy)propanoic acid 2: Sodium hydride (60 % dispersion in mineral oil, 6 g, 0.15 mol) was added to a 250 mL two-necked flask, then anhydrous THF 50 mL was added under N_2 condition. The suspension was vigorously stirred and cooled to 0 ˚C. 2-Azidoenthanol (8.7 g, 0.1 mol) in 20 mL anhydrous THF was added dropwise over 20 minutes. The solution was stirred at an ambient temperature for 15 min, then cooled to 0 ˚C again. Ethyl 2-bromopropionate (27.15 g, 0.15 mol) in 10 mL THF was added dropwise. The reaction mixture was warmed to room temperature and stirred overnight under N_2 atmosphere. 100 mL Water was used to quench the reaction and the resulted mixture was washed by diethyl ether three times (3 × 100 mL). The combined organic layers were dried over anhydrous Na_2SO_4. The crude product was dissolved in 50 mL THF and was added to LiOH aqueous solution (40 mL, 1 M). The mixture was stirred for 16 h at room temperature. THF was removed and HCl (2 M) was added to pH 2. Then, the THF was extracted by diethyl ether three times (3 × 100 mL). The combined organic layers were dried over anhydrous $NaSO_4$. After concentration and silica gel chromatography (ethyl acetate : petroleum ether = 1:7), the product **2** was collected as colorless oil (6.67 g, 26 %). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.09 (q, *J* = 6.9 Hz, 1H), 3.85 (ddd, *J* = 9.8, 5.9, 3.4 Hz, 1H), 3.66 – 3.58 (m, 1H), 3.55 – 3.46 (m, 1H), 3.42 – 3.33 (m, 1H), 1.49 (t, *J* = 9.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 178.5, 74.9, 69.1, 50.7, 18.5. HRMS $C_3H_9N_3O_3^+$ $[M+H]^+$ calculated 160.0717, found 160.0709.



3-(2-azidoethoxy)-1-diazopentane-2-one 3: Under N_2 condition, **2** (1.59 g, 10 mmol) was dissolved in 15 mL anhydrous CH_2Cl_2 and one drop of DMF. Oxalyl chloride (926 μL, 15 mmol) was added to the solution and stirred at room temperature for 2 h. After that, the solvent and excess oxalyl chloride was removed. The residue was dissolved in anhydrous CH_3CN 50 mL, cooled to 0^{o}C, and (Trimethylsilyl)diazomethane solution 2 M in diethyl ether (4 mL, 10 mmol) was added dropwise. The reaction mixture was stirred at 0 ˚C overnight. The solvent was evaporated and silica gel chromatography (ethyl acetate : petroleum ather = 1 : 7) was performed in order to afford product **3** as yellow oil (620 mg, 33.8 %). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.82 (s, 1H), 4.00 – 3.85 (m, 1H), 3.72 – 3.60 (m, 2H), 3.48 – 3.35 (m, 2H), 1.38 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 196.9, 80.9, 68.7, 52.3, 50.9, 18.6. HRMS $C_6H_9N_5O_2^+$ $[M+H]^+$ calculated 184.0829, found 184.0822.



Azido-kethoxal 1 (N_3-kethoxal), or 3-(2-azidoethoxy)-1,1-dihydroxybutan-2-one (4): According to Adam's procedure, the Dimethyldioxirane (DMD) in an acetone solution was prepared. To the compound **3** (183 mg, 1 mmol), 11 mL DMD-acetone was added in several portions. Obvious gas evolution was observed. The reaction mixture was stirred at room temperature until the reaction was complete under TLC minitor to give title compound Azido-kethoxal **1** and its anhydrous form **4** as a yellow oil. ^{1}H NMR (400 MHz, C_6D_6): δ = [9.5 (m) + 5.5 (m), 1H], 4.75 – 3.40 (m, 1H), 3.26-3.18 (m, 2H), 2.81 – 2.71 (m, 2H), 1.36 – 1.12 (m, 3H). ^{13}C NMR (101 MHz, C_6D_6): δ = 197.7, 102.7, 77.1, 68.7, 50.6, 14.2. HRMS $C_6H_9N_3O_3^+$ $[M+Na]^+$ calculated 194.0536, found 194.0555.



To ensure this compound is exactly what we want, we verify the reaction of N_3-kethoxal and guanine. Guanine (100 μM, 2 μL), N_3-kethoxal (1 M in DMSO, 1 μL), sodium cacodylate buffer (0.1 M, pH = 7.0, 1 μL) and 6 μL ddH_2O were added together into 1.5 mL microcentrifuge tube at 37 ℃ for 10 min. HRMS $C_{11}H_{14}N_8O_4^+$ $[M+H]^+$ calculated 323.1216, found 323.1203.

Reference:

1. Staehelin, M. Inactivation of virus nucleic acid with glyoxal derivatives. *Biochimca Biophysica Acta* **31**, 448-454 (1959).

2. Booth, B.A. & Sartorelli, A.C. Synergistic interaction of kethoxal bis(thiosemicarbazone) and cupric ions in sarcoma 180. *Nature* **210**, 104-105 (1966).

3. Brewer, L.A., Goelz, S. & Noller, H.F. Ribonucleic acid-protein crosslinking within the intact Escherichia coli ribosome, utilizing ethylene glycol bis[3-(2-ketobutyraldehyde) ether], a reversible, bifunctional reagent: synthesis and crosslinking within 30S and 50S subunits. *Biochemistry* **22**, 4303-4309 (1983).

4. Tiffany, B.D. et al. Antiviral Compounds. I. Aliphatic Glyoxals, α-Hydroxyaldehydes and Related Compounds. *Journal of the American Chemical Society* **79**, 1682-1687 (1957).

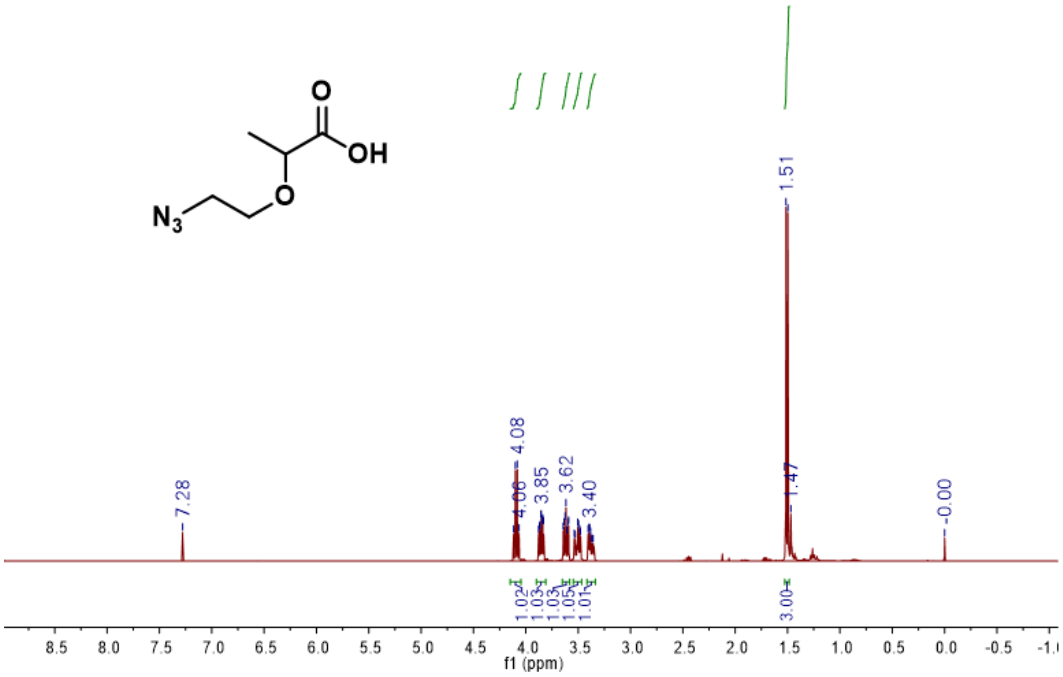
5. Lo, J.M. et al. Synthesis, radiochemical characterization and biodistribution of Tc-99m labeled kethoxal bis(thiosemicarbazone) complexes. *Journal of Labelled Compounds and Radiopharmaceuticals* **44**, S654-S656 (2001).

6. Jiang, B., Liu, J.F. & Zhao, S.Y. Enantioselective synthesis for the antipodes of slagenins B and C: establishment of absolute stereochemistry. *Organic Letters* **3**, 4011-4013 (2001).

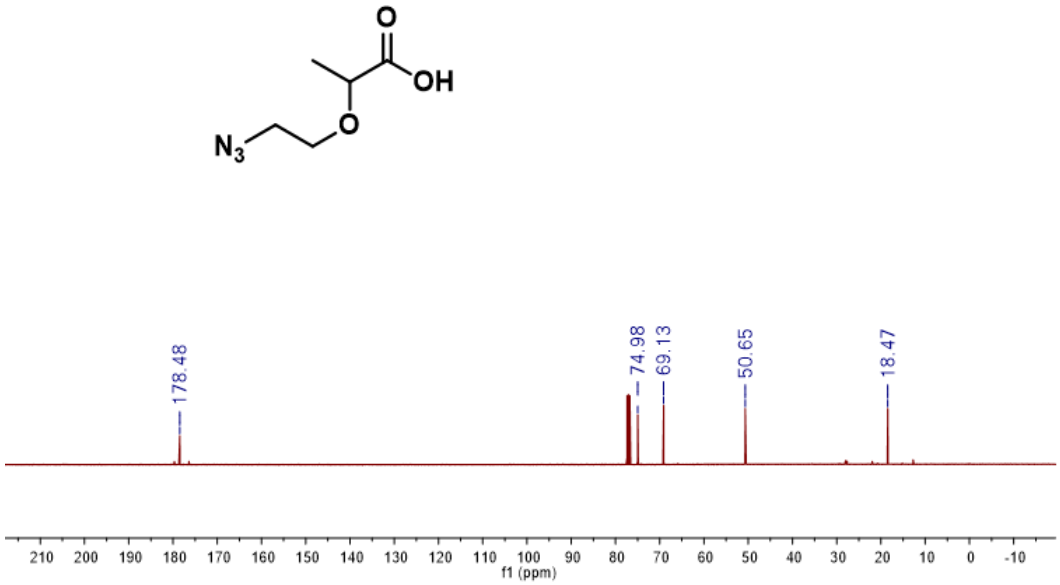
7. Adam, W., Biolas, J. & Hadjiarapoglou. A Convenient Preparation of Acetone Solutions of Dimethyldioxirane. *Chemische Berichte* **124**, 2377 (1991)

8. Spitale, R.C. et al. RNA SHAPE analysis in living cells. *Nature Chemical Biology* **9**, 18 (2012).

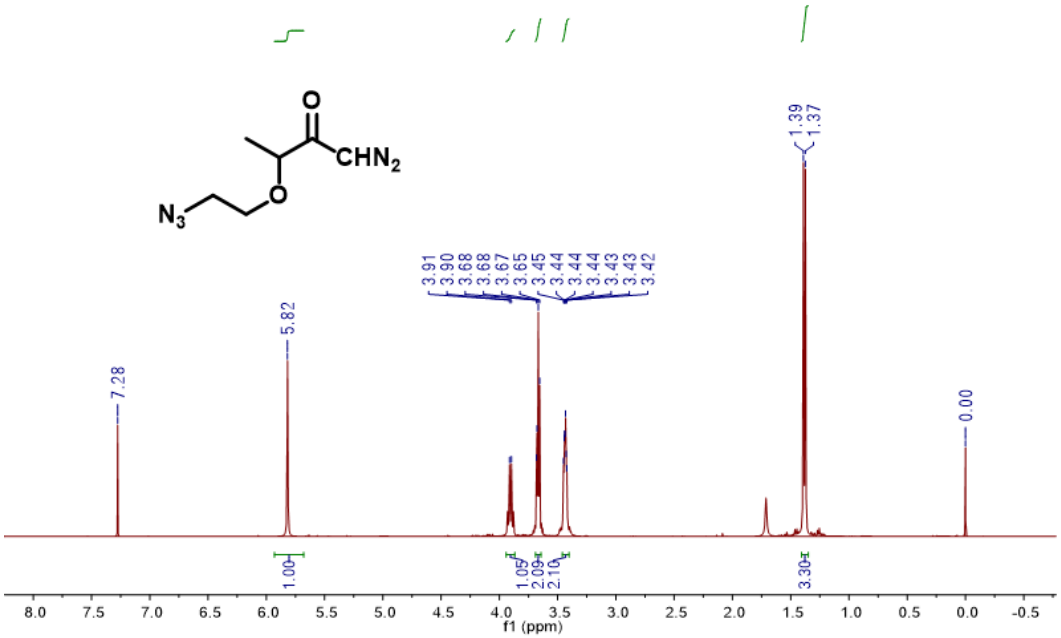
Compounds Characterization



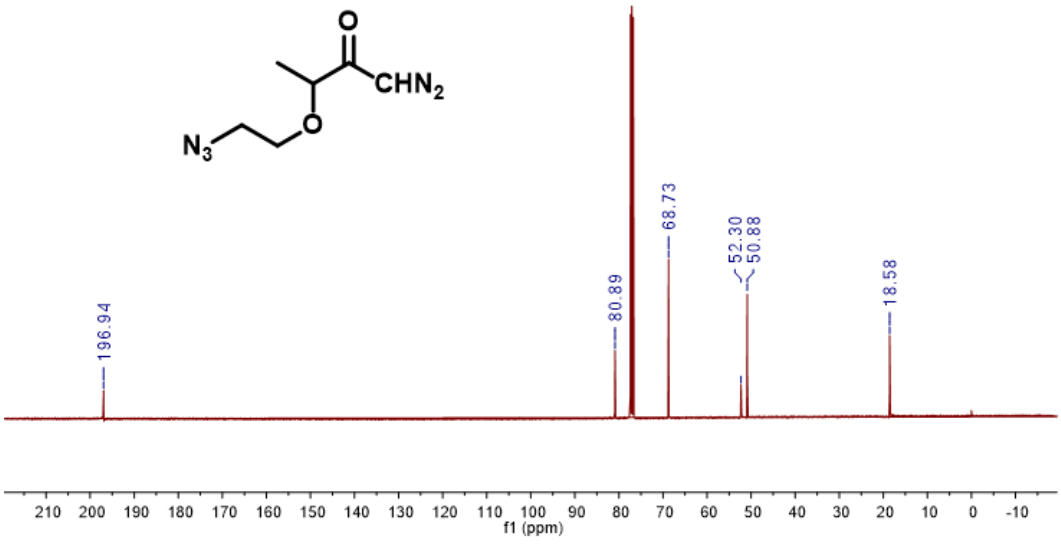
^{1}H NMR spectrum of compound **2** ($CDCl_3$, 400 MHz).



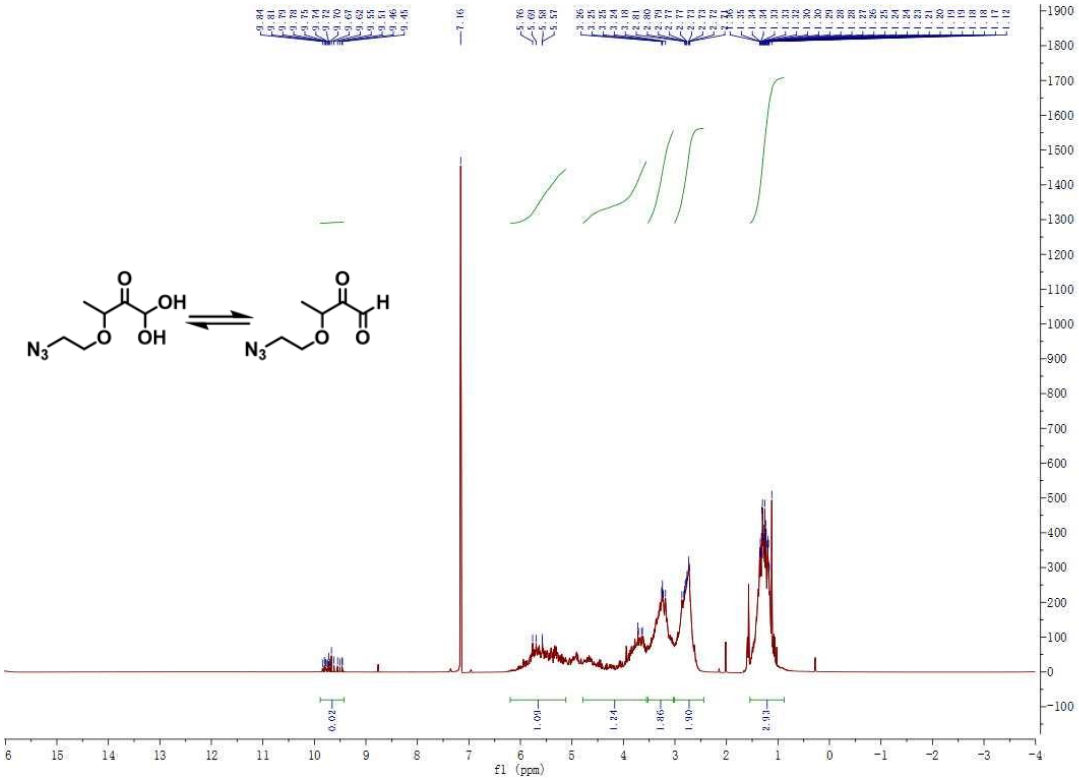
^{13}C NMR spectrum of compound **2** ($CDCl_3$, 400 MHz).



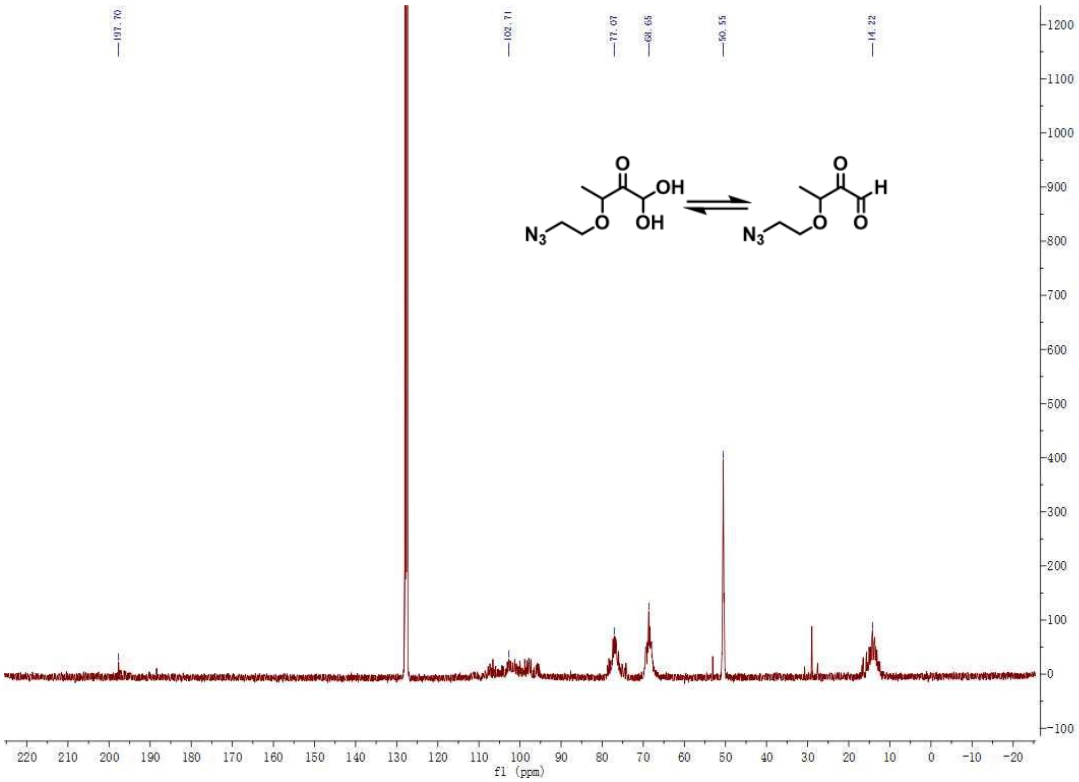
^{1}H NMR spectrum of compound **3** ($CDCl_3$, 400 MHz).



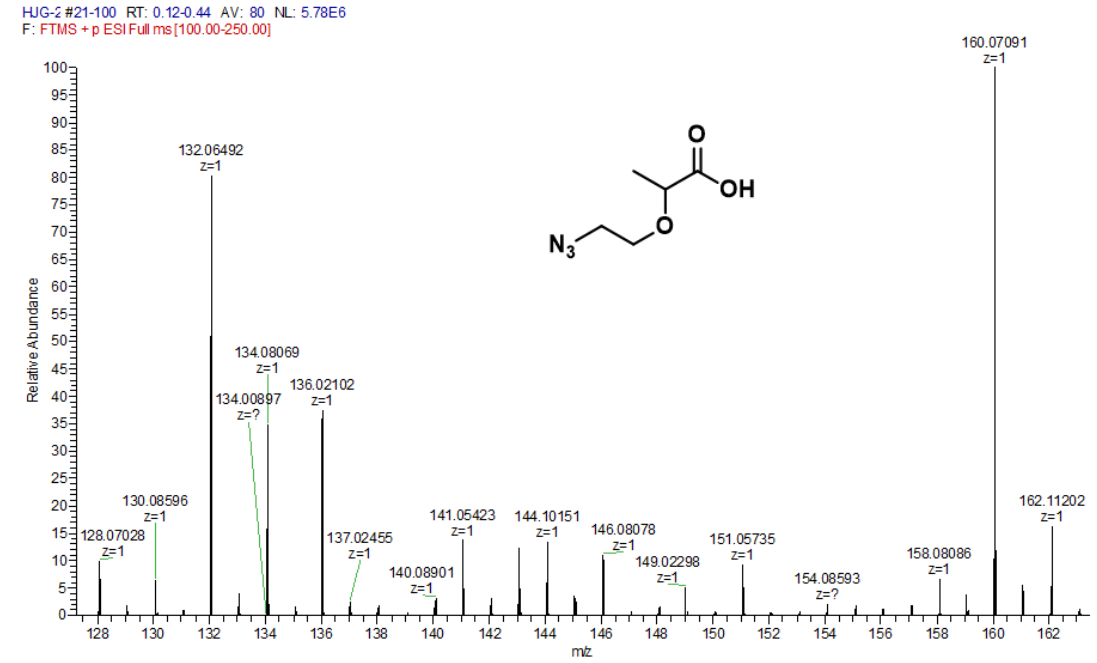
^{13}C NMR spectrum of compound **3** ($CDCl_3$, 400 MHz).



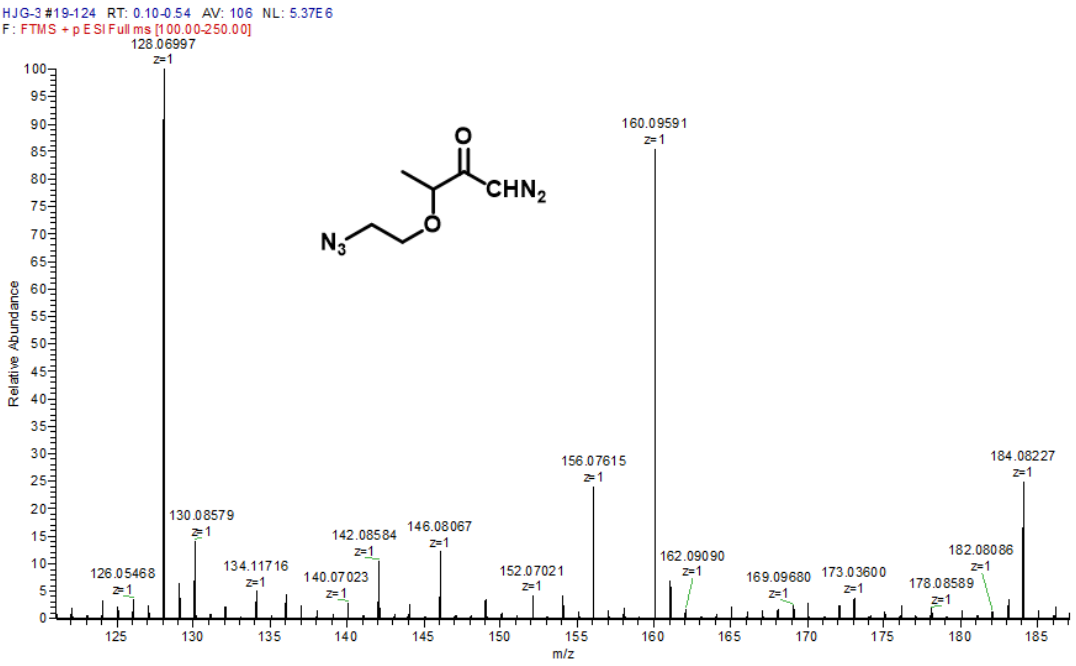
^{1}H NMR spectrum of compound **4** (C_6D_6, 400 MHz).



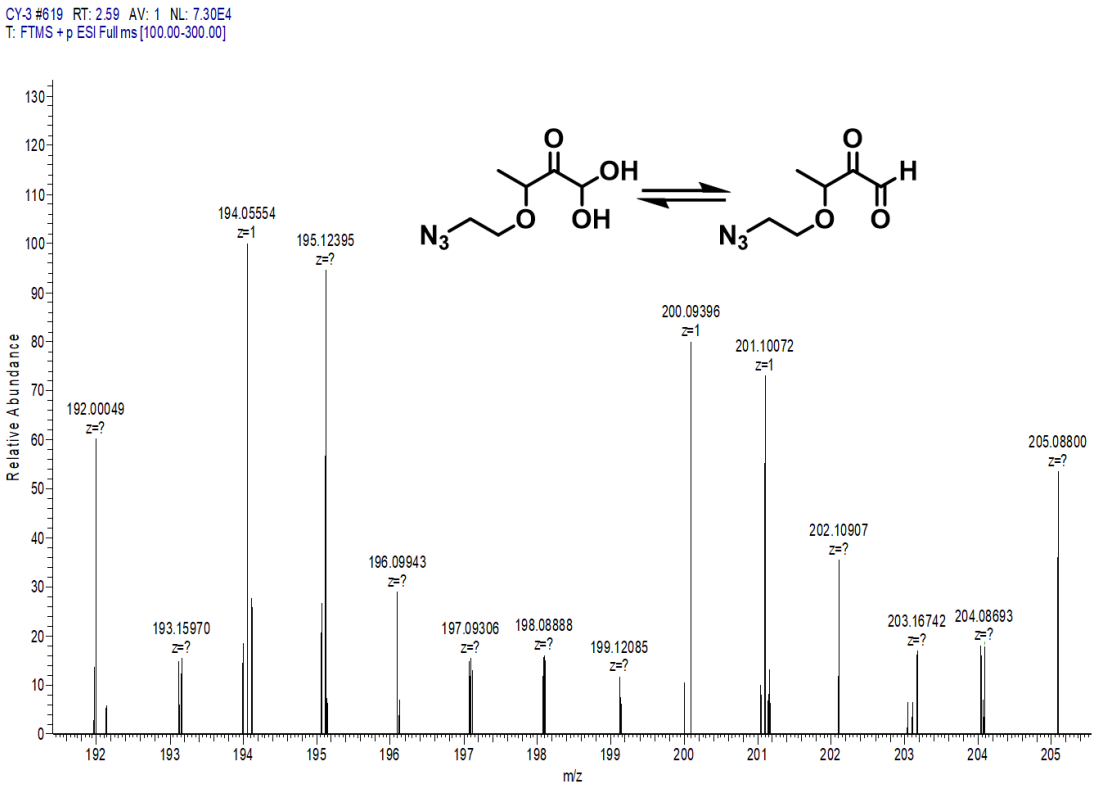
^{13}C NMR spectrum of compound **4** (C_6D_6, 400 MHz).



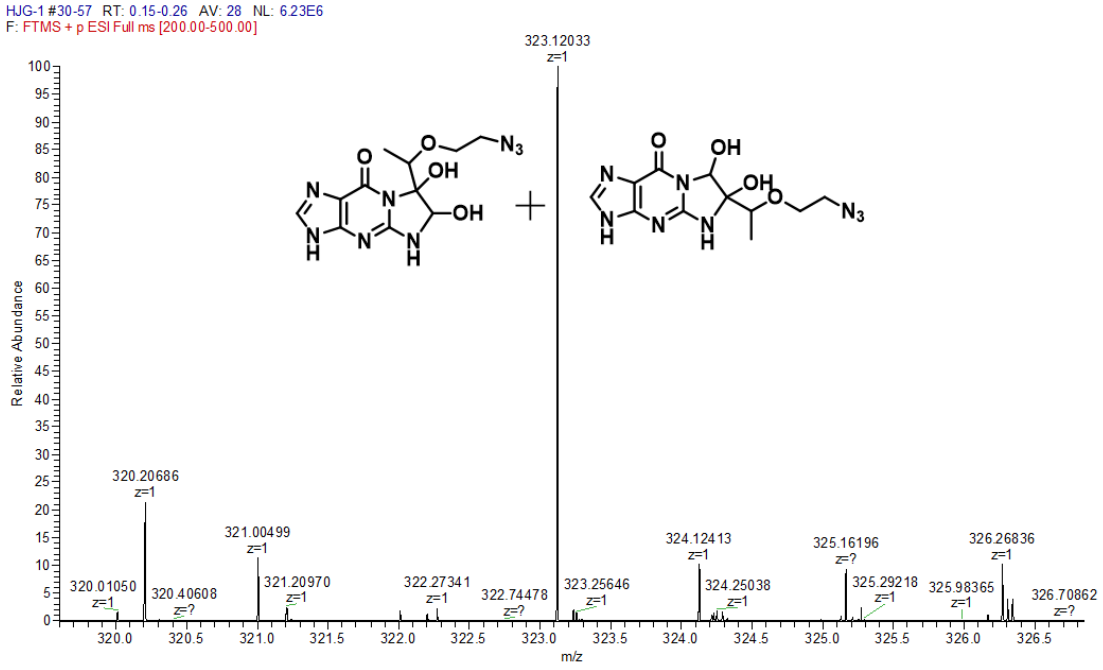
HRMS spectrum of compound **2**.



HRMS spectrum of compound **3**.



HRMS spectrum of compound **4**.



HRMS spectrum of compound **5, 6**.

Supplementary Note 2

The protocol for N_3-kethoxal & N_3-kethoxal-remove library preparation.

1. N_3-Kethoxal *in vivo* labeling

a. Passage mES cell at the previous day in 10 cm plate with 50% confluency.

b. The second day morning, nearly ~80% confluency. Remove 5 mL medium, then add the 10 μL of N_3-kethoxal. Incubate at 37 ˚C for various minutes.

c. Aspirate medium, use PBS wash twice, add 1mL PBS, scrape the cell into a 1.5 mL tube. Spin down at 2000 rpm for 5 min at 4 ˚C. Aspirate the supernatant.

d. Use Qiagen RNeasy plus mini kit to isolate total RNA, then use Dynabeads® mRNA Purification Kit to isolate mRNA from total RNA samples.

2. Biotinylation:

a. Take nearly 2 μg total RNA or mRNA, add 20 mM WS DBCO-Biotin (Click Chemistry tool, A116) in 1.5 mL tube, water bath 37 ˚C for 2 h.

10× PBS	10 μL
500 mM borate buffer	5 μL
20mM WS DBCO-Biotin	5 μL
SUPERase In RNase Inhibitor	2 μL
RNA solution + H_2O	78 μL

Total 100 μL, incubate at 37 ˚C for 2 h.

(borate buffer: 500 mM potassium borate (pH 7.0), pH adjusted by adding potassium hydroxide pellets into 500 mM boric acid)

b. RNA recovery for 100 μL RNA reaction solution. Qiagen RNeasy MinElute Kit.

a) Add 350 μL buffer RLT to 100 μL reaction solution, then add 900 μL 100 % ethanol.

b) Load solution to Qiagen MinElute column, two 500 μL RPE washes, one no-buffer spin to dry the column.

c) Two 50 μL RNase-free water elute the RNA solution.

c. (Optional) Use dot blot to check the efficiency of biotinylation.

3. Fragmentation (Sonication)

a. Transfer RNA solution to Bioruptor NGS 0.65 mL Microtubes. Sonicate 30 cycles with 30 s ON/30 s OFF.

b. Lyophilize to 3 μL and perform ligation.

4. T4 PNK RNA end repair, 3'-End Ligation and 3'-adaptor remove

a. T4 PNK End repair (PCR tube, thermo cycle):

RNA Sample	3 μL
10× T4 PNK buffer:	1 μL
SUPERase In RNase Inhibitor	1 μL
Borate buffer	1 μL
10 mM ATP	1 μL

T4 PNK enzyme	2 μL
FastAP	1 μL

Total 10 μL, incubate at 37 ˚C for 1 h, and perform 3'-end ligation directly.

b. 3'-end ligation:

RNA solution after T4 PNK repair	10 μL
3'-Adaptor 20 μM	1 μL
10×T4RNL2tr buffer	1 μL
100 mM DTT	1 μL
50% PEG8000	6 μL
T4 RNL2tr K227Q	1.5 μL

Total 20.5 μL, incubate at 16 ˚C overnight.

c. Add 29.5 μL H_2O to each sample and purify RNA with Zymo RNA clean & concentrator 5 kit. Elute by 7 μL H_2O twice and get ~13 μL RNA solution

a) Add 100 μL RNA binding buffer to the 50 μL reaction solution, then add 150 μL 100 % Ethanol to mixture, mixed.

b) Load solution to column, 400 μL RNA prep buffer, 700 μL RNA wash buffer, then one no-buffer spin to dry the column.

c) Two 7 μL RNase-free water elute the RNA solution.

d. Excess 3'-Adaptor remove:

RNA solution	13 μL
NEB buffer 2	2 μL
5'-Deadenylase	2 μL
Borate buffer	1 μL

Total 18 μL, incubate at 30 ˚C for 30 min.

Then, add $RecJ_f$ 2 μL, total 20 μL, incubate at 37 ˚C for 1 h.

Add 30 μL H_2O to each sample and purify RNA with Zymo RNA clean & concentrator 5 kit. Elute twice by 7 μL and 6 μL H_2O, respectively, and get ~12 μL RNA solution.

e. Separate RNA solution to two fractions.

(1) Keep 10.5 μL for N_3-kethoxal sample and move to step 6 directly.

(2) Residual 1.5 μL for N_3-kethoxal-remove sample preparation in step 5.

5. N_3-kethoxal-remove sample preparation:

a. Add GTP solution to remove N_3-kethoxal modification and produce N_3-kethoxal-remove sample.

RNA solution	1.5 μL
SUPERase In RNase Inhibitor	1 μL
100 mM GTP	5 μL
H_2O	2.5 μL

Incubate at 95 ˚C for 10 min. Then, add 40 μL H_2O to recover RNA by Zymo RNA clean & concentrator 5 kit. Elute twice with 6 μL H_2O then move to step 6.

a) Add 100 μL RNA binding buffer to 50 μL reaction solution, then add 150 μL 100% Ethanol to mixture, mixed.

b) Load solution to column, 400 μL RNA prep buffer, 700 μL RNA wash buffer, one

no-buffer spin to dry column.

c) Two 6 μL RNase-free water elute the RNA solution.

6. cDNA synthesis

a. Transfer RNA samples to PCR tube. Add 1 μL 5 μM RT primer, mixed.

N_3-kethoxal-remove		N_3-kethoxal	
RNA solution	10.5 μL	RNA solution	10.5 μL
5 μM RT primer	1 μL	5 μM RT primer	1 μL
H_2O	1 μL	borate buffer	1 μL

Total 12.5 μL mixture was heated at 70 ˚C for 5 min in thermo cycler, and then cool slowly to 25 ˚C (1 ˚C per 1 s, 45 steps) and hold at 25 ˚C for 1 min.

b. After primer annealing, the following was added:

5× First Strand Buffer	4 μL
SUPERase In RNase Inhibitor	0.5 μL
100 mM DTT	1 μL
10 mM dNTP mix	1 μL
SuperScript III	1 μL

Total 20 μL mixture was incubate at 25 ˚C for 3 min, 7 min at 42 ˚C, and finally at 52 ˚C for 30 min, hold at 4 ˚C. After cDNA extension, put the mixture on ice or 4 ˚C, do not raise samples above 37 ˚C to avoid denaturing conditions.

7a. Streptavidin capture, cDNA elution (Strepavidin capture ONLY for N_3-kethoxal sample, NOT for N_3-kethoxal-remove sample)

- Biotin Binding Buffer: 100 mM Tris-HCl pH 7.0, 10 mM EDTA, 1 M NaCl
- Biotin Wash Buffer: 10 mM Tris-HCl pH 7.0, 1 mM EDTA, 4 M NaCl, 0.2% Tween
- 10× RNaseH buffer: 500 mM HEPES, 750 mM NaCl, 30 mM $MgCl_2$, 1.25% Sarkosyl, 0.25% Nadeoxycholate, 50 mM DTT

a. Per sample, 20 μL Dynabeads® MyOne™ Streptavidin C1 was washed twice with 1mL Biotin Binding buffer. After wash, resuspend the beads in 10 μL beads binding buffer and 1 μL SUPERase In RNase Inhibitor. Then store on ice until needed.

b. 10 μL of pre-washed beads are added to each Reverse Transcription kethoxal sample (20 μL), and incubated at room temperature for 45 min with rotation.

d. After streptavidin capture, add 100 μL Biotin Wash Buffer and transfer to 1.5 mL tubes. Add additional 400 μL Biotin Wash Buffer (total 500 μL) and invert the tubes four times to mix.

e. Apply the samples to magnet rack. Remove the supernatant. Then, use 500 μL Biotin Wash Buffer to wash four more time. (Total 5 washes).

f. Wash the samples twice using 500 μL 1 x PBS.

g. cDNA elution (for Kethoxal sample)

a). cDNA is eluted by adding following solution:

10× RNaseH buffer	5 μL
RNaseA/T1 cocktail	1 μL
RNaseH	1 μL
50 mM D-biotin	12.5 μL

H_2O 30.5 µL

Total 50 µL solution was incubated at 37 ˚C for 30 min in thermomixer at 1000 r.p.m.

b). Samples are mixed with 1 µL 100 % DMSO, heated to 95 ˚C for 4 min, placed on a magnet rack, and transfer the 50 µL cDNA elution to a new tube.

h. Purify cDNA using DNA Clean & Concentrator-5 Kit with modified method:
   a) 50 µL cDNA elution solution, add 350 µL of DNA Binding Buffer and 350 µL of 100% ethanol.
   b) Continue with purification according to the manufacturer's instructions (200 µL wash buffer wash twice, then empty spin once). Elute cDNA twice with 10 µL H_2O. Get total ~20 µL cDNA solution. Lyophilize to ~5 µL solution.

7b. cDNA elution (ONLY for N_3-kethoxal-remove sample, NOT for N_3-kethoxal sample)

a. cDNA of N_3-kethoxal-remove sample is eluted by adding following solution to 20 µL RT reaction solution:

10 × RNaseH buffer	5 µL
RNaseA/T1 cocktail	2 µL
RNaseH	2 µL
H_2O	21 µL

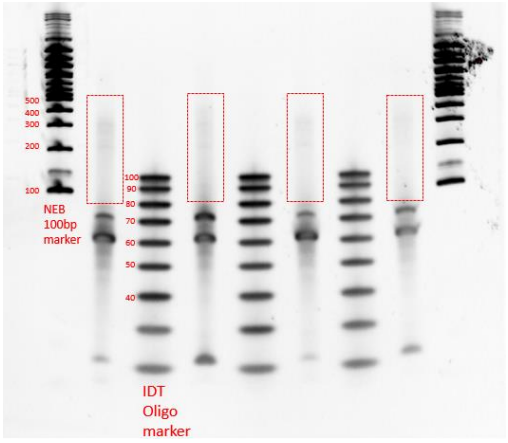
Total 50 µL solution was incubated at 37 ˚C for 30 min.

b. Purify cDNA using DNA Clean & Concentrator-5 Kit with modified method:
   c) 50 µL cDNA elution solution, add 350 µL of DNA Binding Buffer and 350 µL of 100% ethanol. Mix well.
   d) Continue with purification according to the manufacturer's instructions (200 µL wash buffer wash twice, then empty spin once). Elute cDNA twice with 10 µL H_2O. Get total ~20 µL cDNA solution. Lyophilize to ~5 µL solution.

8. cDNA size selection.

a. Add 5 µL TBU 2 × loading dye, and load to 6% TBE-Urea gel for size selection as well as heating samples to 95 ˚C for 2 min, inserting in ice immediately before PAGE separation (180 V, 40 min).

b. Sybr Gold staining, image and cut > 70 nucleotide (70 ~ 500) in the gel (May invisible in gel, see the right picture). Purify the cDNA from the gel.



   a) Transfer each gel slice to a 0.5 mL microcentrifuge tube with a hole punched in the bottom using a sterile needle, and close the tube cap. Place each 0.5 mL tube inside a 1.5 mL tube and centrifuge for 2 min at 12,000 g in a microcentrifuge to shred the gel slices.
   b) Remove and discard the 0.5 mL tubes. To each 1.5 mL collection tube, add:

Nuclease-Free Water	400 µL

5 M ammonium acetate	40 μL
10% SDS	2 μL

c) Gently rock the samples at 50 ℃ more than 3 hours to elute the cDNA from the disrupted gel slices.

d) Transfer the slurry to new 1.5 mL filter tubes and centrifuge for ~2 min at maxi-speed to separate the disrupted gel pieces from the eluted cDNA solution.

e) To each aqueous solution, add 2 μL of Glycogen and 700 μL of 100% isopropanol. Store at -80 ℃ for > 1 hour.

f) Centrifuge the tubes at 4 ˚C for 30 min at > 12,000 g to pellet the cDNA. Wash the pellet with ice-cold 80% ethanol and air-dry. Resuspend the pellet in 14 μL of Nuclease-Free Water.

9. cDNA cyclization.

a. Transfer 14 μL cDNA solution to PCR tube, then add reagents as follows:

cDNA/H_2O	14 μL
10× CircLigaseII buffer	2 μL
1 mM ATP	1 μL
50 mM $MnCl_2$	1 μL
CircLigase I	2 μL

Total 20 μL is incubated at 60 ˚C for 2 h.

b. Purify the CircDNA using DNA Clean & Concentrator-5 Kit with modified method:

a) Add 30 μL H_2O to the cDNA solution to get 50 μL solution, then add 350 μL of DNA Binding Buffer and 350 μL of 100 % ethanol. Mix well.

b) Continue with purification according to the manufacturer's instructions (200 μL wash buffer wash twice, then empty spin once). Elute cDNA twice with 10 μL H_2O. Get total ~20 μL CircDNA solution.

10. Library PCR

a. Using 1 μL CircDNA for qPCR monitoring to optimize the PCR cycle numbers.

RP/RPI short primer 20mer 5 μM	1 μL
CircDNA from step 9	1 μL
H_2O	8 μL
qPCR 2× Master mix	10 μL

We suggest choosing a cycle number, which is located at half of the sigmoid curve.

b. PCR by short primer (RP/RPI short 20 mer).

Phusion MM	25 μL
RP/RPI short primer 20mer 5uM	1 μL
CircDNA solution	5 μL
H_2O	19 μL

Total 50 μL, PCR parameter: 98 ˚C 30s, then 98 ˚C 10s/60 ˚C 30s/72 ˚C 30s each cycle.

c. Purify PCR products using Bio-rad Green spin column.

Lyophilize to 5 μL, add 2 μL DNA 6× loading dye, and load to 6% TBE gel for size selection (180 V, 30 min). Sybr-Gold staining, image and cut > 80 nucleotide (80 ~ 300) in the gel. Purify the cDNA as the protocol in step 8. Resuspend the pellet in 20 μL of

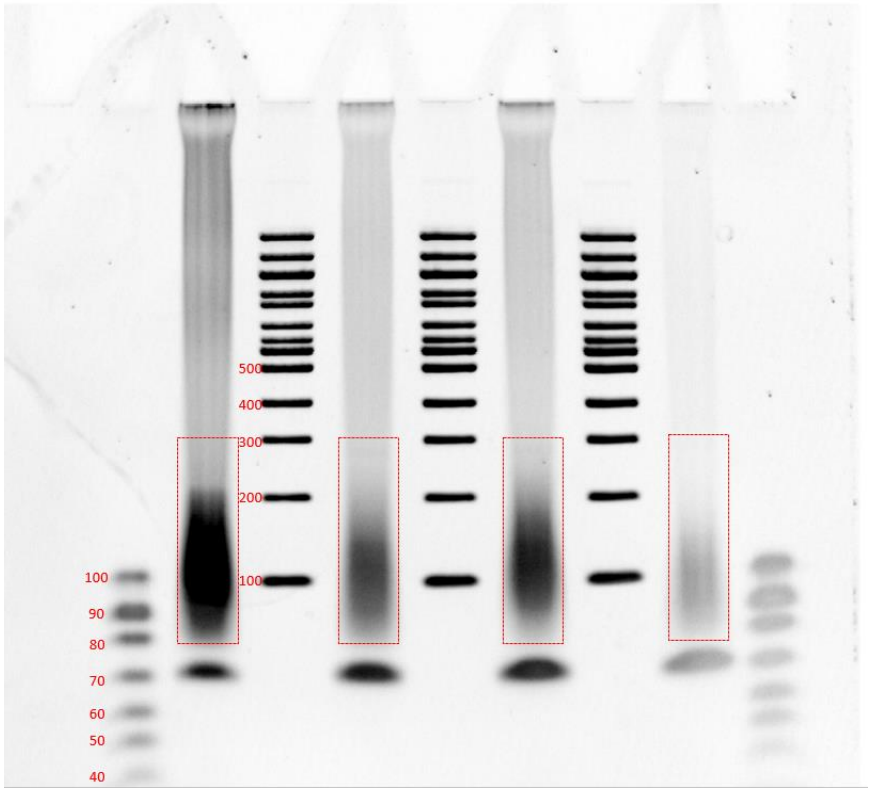
nuclease-Free Water.

d. PCR by Sequencing primer (RP/RPI index X, 3 ~ 5cycle).

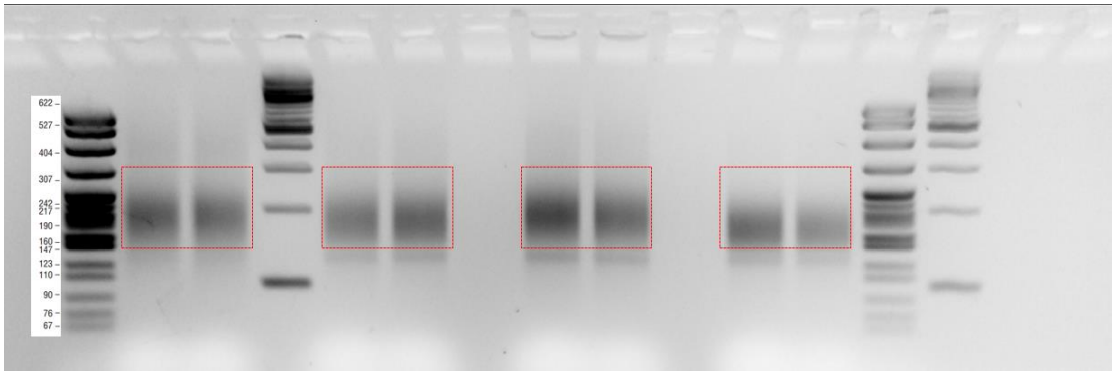
Phusion MM	25 µL
RP primer 10 µM	1 µL
RPI index X 10 µM	1 µL
Purified DNA	20 µL
H_2O	3 µL

Total 50 µL, PCR parameter: 98 ˚C 30s, then 98 ˚C 10s/60 ˚C 30s/72 ˚C 30s, 3~5 cycles.

e. After sequencing PCR, using 3% low melting agarose gel to purify the PCR products (remove primer dimer, 90 V, 45 min) and recovery by Qiagen QIAquick gel extraction Kit. Elute DNA library by 20 µL H_2O.



Short primer PCR TBE gel



Sequencing PCR 3% Agarose Gel