Supporting Information

Addressing the determinants of ambiguous assignments incurred by the mass spectrometry sequence analysis of large modified RNAs

Daniele Rollo,[1] Thomas Kenderdine,[1] Ghazaleh Yassaghi,[1] Cole Angell,[1] and Daniele Fabris[1,2*]

[1]*Department of Chemistry, University of Connecticut, Storrs, CT 06269, U.S.A.*

[2]*RiboDynamics LLC, 98 Wyneding Hill Rd., Manchester, CT 06040, U.S.A.*

**Corresponding author: dan.fabris@uconn.edu*

Table of Contents

Experimental

Materials and sample preparation. The samples examined in the study included an RNA construct mimicking the 5'-untranslated region (5'-UTR) of the HIV-1 genome prepared in house according to standard in-vitro transcription techniques, enhanced green fluorescent protein (eGFP) mRNA purchased from Ribopro (Oss, Netherlands), and human 18S ribosomal RNA (18S rRNA) extracted from HeLa cell lysates and purified by size-exclusion chromatography under non-denaturing conditions (**Table S1**). More specifically, total RNA collected from HeLa cells was submitted to two separate chromatographic steps, the first on a Cytiva (Marlborough, MA) HiPrep Sephacryl S-400 HR fast protein liquid chromatography column (FPLC) with 150 mM ammonium acetate (pH 7.0) as the elution buffer. Fractions enriched with 18S rRNA were then further purified on a Phenomenex (Torrance, CA) PolymerX high-performance liquid chromatography column using the same buffer.[1] The integrity of 18S rRNA was confirmed by agarose gel electrophoresis. Extensive desalting of the RNA substrates was carried out by ultrafiltration against 100 mM ammonium citrate (pH adjusted to 7.0) and RNase-free water in Millipore (Burlington, MA) Amicon centrifugal filters.

Enzymatic digestion. RNase T1 was purchased from Sigma-Aldrich (St. Louis, MO) and used with no further purification. Initial test conditions involved using 2,000 units of RNase T1 per 0.1 nmol of 5'-UTR (corresponding to a 4:1 enzyme-to-substrate, E/S, molar ratio) and incubating for 30 min at 37°C in 150 mM ammonium acetate with pH adjusted to 7.0.[2–4] The solution composition was later adjusted to 100 units of RNase T1 per 0.1 nmol of 5'-UTR (1:5 E/S ratio) and 5 units of RNase T1 per 0.1 nmol of 5'-UTR (1:100 E/S ratio) under otherwise identical incubation conditions. The impact of temperature on endonuclease activity was investigated on solutions prepared at 1:100 E/S ratio and incubated at 25, 37, and 60°C. The effects of incubation duration were tested on solutions with the same composition incubated at 25°C for 30 min, 60 min, 120 min, and 10 h. It typically took less than 60 s to introduce a sample in a nanospray emitter, load it on the stage, and start the acquisition. All reported reaction times accounted for this interval. Each sample was added with 20% v/v formamide immediately before analysis to increase the putative charging of the larger products in the mixture. Time-course analyses involved mixing enzyme, substrate, and formamide, and then immediately loading the sample into the nanospray emitter with no prior incubation. No adverse effects on enzymatic activity were noted when comparing results obtained in the absence/presence of formamide. Initial time-course experiments involved the MS analysis of a 1:100 E/S mixture of 5'-UTR and RNase T1 at 25ºC, which was extended for over 180 min of uninterrupted acquisition. In subsequent experiments, the amount of enzyme was adjusted according to the moles of G rather than those of RNA sample present in solution to enable unbiased comparisons of the results obtained from the various substrates. These determinations were carried out at a fixed 1:200 enzyme-to-guanine (E/G) ratio at 25ºC for 240 min.

Mass spectrometry. All MS determinations were performed in the negative ion mode on a Bruker (Billerica, MA) solariX XR Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer equipped with a 7T superconducting magnet. Static nanospray was achieved by using quartz emitters pulled on a Model P-2000 (Sutter Instruments, Novato, CA) pipet puller to obtain tip diameters in the sub-micrometer range, which we have proven capable of minimizing the formation of cation adducts.[5] A ~5 µL sample solution was loaded into the emitter and a stainless-steel wire was inserted into the back end to provide typical electrospray voltages ranging from -0.7 to -1.0 kV. Source conditions (i.e., desolvation voltage, spray voltage, etc.) were finely adjusted by monitoring the incidence of ammonium adducts and water clusters. Typical resolution exceeding 200k at 1,200 *m/z* and accuracy of 500 ppb or better were obtained by using external calibration. Collision induced dissociation (CID) determinations were performed by selecting precursor ions in the mass-selective quadrupole (Q) with a 1 *m/z* window, followed by activation in the RF-only collision cell (q) containing approximately 7.2×10^{-6} mbar of Ar submitted to 12-32 V collision voltage, depending on precursor's size. MS/MS determinations during time-course analysis were carried out by timing the experiments according to the information obtained from a prior survey analysis. No advanced acquisition technique (i.e., DDA or DIA) was employed in the study. Acquisition time was adjusted on a sample-to-sample basis to obtain the best possible signal to noise ratio (S/N) through signal averaging.

Data interpretation. The reported MS spectra were generated by signal-averaging data that were either acquired in regular accumulation mode, or extracted from a dataset in which the individual scans were saved separately as a function of time. The latter was accomplished by using the ensuing total ion current (TIC) trace to select the time range corresponding to the desired snapshot. These raw data were initially processed by the Bruker (Billerica, MA) DataAnalysis 5.3 software to generate a list of monoisotopic *m/z* values that were deconvoluted from the corresponding isotopic envelopes by the SNAP algorithm. The experimental values were subsequently matched within a tolerance window of ±5 ppm to predicted *m/z* values calculated from the initial analyte sequence. Predictions and matching operations were carried out by the RiboDynamics (Manchester, CT) SeqRead software, which allowed for up to 30 missed cleavages and a maximum charge state of -30. Any experimental entry with a match within the selected tolerance was considered as a legitimate digestion product. However, only those that matched just one unique prediction were considered as unambiguous. The respective counts and sizes were used to calculate the metrics included in the report.

The interpretation of MS/MS fragmentation data was carried out by using the SeqRead software to match deconvoluted experimental masses with putative predictions for both terminal and internal fragments.[6] Any experimental entry that could be matched to two or more predictions was discarded as ambiguous. Only those that exhibited a unique match were aligned into overlapping ladders, which generated possible walks covering the precursor's sequence. Coverage metrics were assessed according to the requirement that any given nucleotide be unambiguously identified by the detection of at least two consecutive fragments in the ladders, which differed by its specific incremental mass. Based on this *n+1* criterion, the sequence coverage was calculated as the percentage of nucleotides in the precursor sequence, which were confirmed by the detection of consecutive fragments in the series.[6]

The stoichiometry of a given RNA post-transcriptional modification (rPTM), which is also defined as the percentage of occupancy exhibited by a given modification site, is typically calculated from the abundances of the respective modified and unmodified versions of the RNA of interest. In the case of 18S rRNA, these values were estimated from the abundances of modified/unmodified versions of the digestion product bearing the rPTM. As explained in detail in ref. 7–9, the intensities of the desired signals expressed in arbitrary units were divided by the respective charge states to normalize their values and make subsequent calculations self-consistent. The values obtained from modified and unmodified versions were then used to calculate the respective molar fractions or relative proportions.

Digestion simulations. Simulations of endonuclease digestions were accomplished by using the RiboDynamics (Manchester, CT) SeqRead software to predict all possible products obtainable when treating a desired RNA substrate with a specific cleavage reagent. The predictions were carried out by drawing from a database containing over 230 nucleic acid building blocks of either natural or synthetic origin. In the case of RNase T1, the predictions accounted for possible 3'-ends consisting of regular and cyclic phosphates, as well as hydroxyl groups. For each test sample, the algorithm generated an initial list of products that included their sequences, indexes, masses, and number of missed cleavages (MCs). The initial list was filtered to extract unique mass values, which were counted as total number of products. The initial list was then parsed to identify and count isomeric sequences sharing identical base compositions/masses, or isobaric products possessing exact masses that differed from one another by a desired tolerance expressed in ppm. The fact that each entry was tested against all the others in the list contributed to the expansion of the number of isobars detected within a given tolerance. The number of mis-cleaved products was calculated in cumulative fashion by counting all species that contained all possible numbers of MCs up to the desired value.

References

1. Chionh, Y.H.; Ho, C.H.; Pruksakorn, D.; Ramesh Babu, I.; Ng, C.S.; Hia, F.; McBee, M.E.; Su, D.; Pang, Y.L.J.; Gu, C.; Dong, H.; Prestwich, E.G.; Shi, P.Y.; Preiser, P.R.; Alonso, S.; Dedon, P.C. "A multidimensional platform for the purification of non-coding RNA species" *Nucl. Acids Res*. **2013**, 41(17), e168.

2. Crain, P.F. "Preparation and enzymatic hydrolysis of DNA and RNA for mass spectrometry." *Methods Enzymol*. **1990**, 193, 782-790.

3. Kowalak, J.A.; Bruenger, E.; McCloskey, J.A. "Posttranscriptional Modification of the Central Loop of Domain V in Escherichia coli 23 S Ribosomal RNA (∗)" *J. Biol. Chem*. **1995**, 270(30), 17758-17764.

4. Crain, P.F.; McCloskey, J.A. "Applications of mass spectrometry to the characterization of oligonucleotides and nucleic acids." *Curr. Opin. Biotechnol*. **1998**, 9(1), 25-34.

5. Kenderdine, T.; Xia, Z.; Williams, E.R.; Fabris, D. "Submicrometer Nanospray Emitters Provide New Insights into the Mechanism of Cation Adduction to Anionic Oligonucleotides" *Anal. Chem*. **2018**, 90(22), 13541-13548.

6. Kenderdine, T.; McIntyre, W.; Yassaghi, G.; Rollo, D.; Bunkowski, A.; Goerlach, L.; Suckau, D.; Tremintin, G.; Greig, M.; Bell, C.; Fabris, D. "Integrating Internal Fragments in the Interpretation of Top-Down Sequencing Data of Larger Oligonucleotides" *J. Am. Soc. Mass Spectrom*. **2023**, 34(10), 2296-2307.

7. Hagan, N.; Fabris, D. "Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between HIV-1 nucleocapsid protein and RNA stem-loops hairpins of the HIV-1 Psi-recognition element." *Biochem. (ACS)* **2003**, 42(36), 10736-10745.

8. Turner, K.B.; Hagan, N.A.; Fabris, D. "Inhibitory effects of archetypical nucleic acid ligands on the interactions of HIV-1 nucleocapsid protein with elements of Psi-RNA." *Nucl. Acids Res*. **2006**, 34(5), 1305-1316.

9. Stephenson, W.; Asare-Okai, P.N.; Chen, A.A.; Keller, S.; Santiago, R.; Tenenbaum, S.A.; Garcia, A.E.; Fabris, D.; Li, P.T.X. "The Essential Role of Stacking Adenines in a Two-Base-Pair RNA Kissing Complex" *J. Am. Chem. Soc*. **2013**, 135(15), 5602-5611.

Tables

Table S1: Nucleic acid samples employed in the study (continues in the next page).

Name	Size (nt)	Mass (u)	Sequence (5' → 3')	G content
HIV-1 5'-UTR	364	117,866.9	GGUCUCUCUGGUUAGACCAGAUCUGAGCCUGGGAGCUCUCUGGCUAACUAGGGAACCCACUGCUUAAGCCUCAAUAAAGCUUGCCUUGAGUGCUCAAAGUAGUGUGUGCCCGUCUGUUGUGUGACUCUGGUAACUAGAGAUCCCUCAGACCCUUUUAGUCAGUGUGGAAAAUCUCUAGCAGUGGCGCCCGAACAGGGACUUGAAAGCGAAAGUAAAGCCAGAGGAGAUCUCUCGACGCAGGACUCGGCUUGCUGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACUGGUGAGUACGCCAAAAAUUUUGACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCGGUAUUAA GUCG	115 (32%)
eGFP mRNA	758	246,929.3	GGGAGACGCCGCCACCAUGGCGAGCAAGGGAGAGGAGCUGUUCACAGGGGUGGUGCCAAUACUGGUGGAGCUGGACGGAGACGUAAACGGACACAAGUUCAGCGUGAGUGGAGAGGGAGAGGGAGAUGCAACAUACGGAAAGCUG[1]ACACUGAAGUUCAUAUGCACCACAGGAAAGCUGCCAGUGCCAUGGCCAACACUGGUACAACACUGACAUACGGAGUGCAGUGCUUCAGCCGGUACCCAGACCACAUGAAGCAGCACGACUUCUUCAAGAGUGCAAUGCCAGAAGGAUACGUGCAGGAGCGGACAAUAAGUUUCAAGGACGACGGAAACUACAAGACACGGGCAGAGGUGAAGUUCGAGGGAGACACACUGGUGAACCGGAUAGAGCUGAAGGGAAUAGACUUCAAGGAGGACGGAAACAUACUGGGGCACAAGCUGGAGUACAACUACAACAGCCACAACGUGUAUAUAACGGCAGACAAGCAGAAGAACGGAAUAAAGGCAAACUUCAAGAUACGGCACAACAUAGAGGACGGAAGCGUGCAGCUGGCAGACCACUACCAGCAGAACACACCAAUAGGAGACGGACCAGUGCUGCUGCCAGACAACCACUACCUGAGCACACAGAGUGCACUGAGCAAAGACCCAAACGAGAAGCGGGAUCACAUGGUGCUGCUGGAGUUCGUGACAGCAGCAGGGAUAACUCUGGGAAUGGACGAGCUGUACAAGUAAGGAGAAGAGAAGGAAGAGAAAA	234 (31%)

Name	Size (nt)	Mass (u)	Sequence (5’ → 3’)	G content
Human 18S rRNA	1,869	603,634.4	UACCUGGUUGAUCCUGCCAGUAGCAUAmUGCUUGΨCΨCAAAGAUUAAGCCAUGCAUGUCUAAGUACGCACGGCCGGUACAGUGAAACUGCGAAΨGGCUCAmUUAAAΨCAGΨUAUGGUUmCCΨUUmGGUCGCUCGCUCCUCUCCCACUUGGAUAACUGUGGUAAmUUCUAGAmGCUAAUmACmAUGCCGACGGGCGCUGACCCCCUUCGCGGGGGGGAΨGCGUGCAΨUUAUCAGAUCAAAACCAACCCGGUCAGCCCCUCUCCGGCCCCGGCCGGGGGGCGGGCGCCGGCGGCUUUGGUGACUCΨAGAUAACCUCGGGCCGAUCGCACGCCCCCCGUGGCGGCGACGACCCAUUCGAACGUCUmGCCCUAUCAACUUUCGAUGGUAGUCGCCGUGCCUACCAUGGUGACCACGGGΨGACGGGGAAUCAGGGUUCGAUUmCCGGAGAGmGGAGCCUGAGAAACGGCUACCACAUCmCAAGGAmAGGCAGCAGGCGCGCAmAAUUACCCACUCCCGACCCGGGGAGmGUAmGUGACmGAAAAAUAACAAUACAGGACUCUUUCGAGGCCCUGUAAUUGGAAUGAGUCCACUΨUAAAmUCCUUUAACGAGGAmUCCAUUGGAGGmGCAAGUCΨGGUGCCAGCAGCmCGCGGUmAAUUCCAGCUCCAAUAGmCGUAΨAΨUAAAGUUGCUGCAGUUAmAAAAGCUCGUAGΨUGmGAΨCUUGGGAGCGGGCGGGCGGUCCGCCGCGAGGCGAGCCACCGCCCGUCCCCGCCCCUUGCCUCUCGGCGCCCCCUCGAUGCUCUUAGCUGAGUGUCCCGCGGGGCCCGAAGCmGUmUΨACUUUGAAAAAAΨΨAGAGUGΨUCAAAGCAGGCCCGAGCCGCCUGGAUACCGCAGCUAGGAAΨAAΨGmGAAUAGGACCGCGGUUCUAUUUUGUUGGUΨUUCGGAACUGAGGCCAUGAUΨAAGAGGGACGGCCGGGGGCAUUCGUAUUGCGCCGCUAGAGGUGAAAUΨCUUGGACCGGCGCAAGACGGACCAGAGCGAAAGCAUUΨGCCAAGAAUGUUUUCAUUAAUCAAGAAmCGAAAGUCGGAGGΨΨCGAAGACGAΨCAGAUACCGUCGUAGUUCCGACCAΨAAACGAUGCCGACCGGCGAUGCGGCGGCGUUAUUCCCAUGACCCGCCGGGCAGCΨUCCGGGAAACCAAAGUCUUUGGGUUCCGGGGGGAGUAΨGGΨUGCAAAGCUGAAACUUAAAGGAAUUGACGGAAGGGCACCACCAGGAGUGGAGCCΨGCGGCΨUAAUUΨGACm^{1}acp^{3}ΨCAACACGGGAAACCUCACCCGGCCmCGGACACGGACAGGAUmUGACAGAUUGAUAGCUCUUUCUCGAUUCCGUGGGUGGΨmGGmUGCAUGGCac4CGUUCUUAGUΨGGUGGAGCGAUUUGUCUGGΨUAAUUCCGAUAACGAAmCGAGACUCmUGGCAUGCUAACUAGUUACGCGACCCCCGAGCGGUCGGCGUCCCCCAACUUmCUΨAGmAGGGACAAGUGGCGUUCAGCCACCCGAGAUUGAGCAAUAACAGmGUCUGUGAUGCCCUUAGAUGUCCGGGGCUGCACGCGCGCUACACUGACUGGCUCAGCGUGUGCCUACCCUACGCCGGCAGGCGCGGGUAACCCGUUGAACCCCAUUCGUGAUGGGGAUCGGGGAUUGCAAUUAUΨCCCCAUGAACGAGm^{7}GAAUΨCCCAGUAAGUGCGGGUCAUAAGCUUmGCGUUGAUUAmAGUCCCUGCCCUUΨGUACACACCGCmCCGUCGCUACUACCGAUUGGAUGGUUUAGUGAGGCCCUCGGAUCGGCCCCGCCGGGGUCGGCCCACGGCCCUGGCGGAGCGCUGAGAAGACGGUCGAACUUmGACUAUCUAGAGGAAGUAAAAGUCGUAm^{6}ACAAGGUUUCac4CGUAGGUGm^{6,6}Am^{6,6}ACCUGCGGAAGGAUCAUUA	539 (29%)

Average molecular masses are reported. The G-content was used to determine the appropriate enzyme-to-substrate ratio for time-course analyses (see *Enzymatic digestion* section). The putative secondary structures folded by these RNAs are provided in **Figure S1**.

Table S2. Metrics obtained from the time-course analysis of 5'-UTR digestion with RNase T1 at 1:100 E/S ratio and 25ºC (**Figure 3**).

Parameter	**60 min**	**120 min**	**180 min**	**Total**
Total observed products	280 ±43	386 ±85	343 ±62	701 ±133
Unambiguous (%)	90 ±18 (32%)	102 ±19 (26%)	84 ±14 (24%)	252 ±50 (36%)
Average size (nt)	28.1 ±2.5	21.9 ±0.7	19.2 ±3.7	28.5 ±4.0
Average MC	5.8 ±1.0	3.8 ±0.7	3.4 ±0.6	5.5 ±0.6
Mapping coverage (%)	81 ±7	73 ±6	64 ±8	92 ±8

Total number of observed products was calculated by counting any signal that matched a prediction; the number of unambiguous products counted only those that did not have isomers. Average size and missed cleavages (MC) were calculated exclusively for unambiguous products.

Table S3: Unambiguous products identified by time-course analysis of 5'-UTR digestion with RNase T1 at 1:100 E/S ratio and 25ºC **(Figure 3)** (continues in the next two pages).

Experimental mass (u)	Predicted mass (u)	Digestion product
2502.28	2502.28	[3:10]
4115.48	4115.47	[3:15]
9255.17	9255.15	[3:31]
8153.12	8153.09	[11:35]
1304.16	1304.16	[12:15]
3878.53	3878.53	[16:27]
18954.54	18954.49	[21:79]
29488.99	29488.86	[21:122]
17363.33	17363.30	[26:79]
16689.23	16689.20	[28:79]
9645.29	9645.27	[33:62]
15083.04	15083.00	[33:79]
9318.24	9318.23	[34:62]
14755.99	14755.97	[34:79]
16000.13	16000.09	[34:83]
2178.25	2178.25	[36:42]
8626.14	8626.12	[36:62]
14063.88	14063.85	[36:79]
11540.57	11540.56	[44:79]
16369.13	16369.15	[52:102]
1938.26	1938.26	[63:68]
5437.73	5437.73	[63:79]
6699.88	6699.87	[63:83]
3535.48	3535.50	[69:79]
1567.18	1567.18	[84:88]
5140.68	5140.67	[84:99]
6771.87	6771.86	[84:104]
7422.95	7422.94	[84:106]
2248.32	2248.31	[93:99]
10615.34	10615.32	[107:139]
2522.31	2522.31	[109:116]
26982.50	26982.45	[113:196]
975.11	975.11	[117:119]
23738.07	23738.02	[117:190]
1896.24	1896.23	[124:129]
18572.44	18572.40	[124:181]
19568.54	19568.51	[124:184]

Experimental mass (u)	Predicted mass (u)	Digestion product
6121.81	6121.82	[130:148]
5758.78	5758.76	[131:148]
8901.16	8901.14	[131:158]
10186.33	10186.31	[131:162]
15351.99	15351.98	[131:178]
16331.17	16331.12	[131:181]
16982.21	16982.18	[131:183]
23156.09	23156.02	[131:202]
2835.37	2835.37	[140:148]
7262.93	7262.91	[140:162]
3518.47	3518.47	[168:178]
9041.25	9041.24	[168:195]
11322.54	11322.53	[168:202]
8157.12	8157.12	[182:206]
6173.85	6173.84	[184:202]
16611.32	16611.30	[184:234]
5828.81	5828.79	[185:202]
16266.26	16266.25	[185:234]
5178.71	5178.70	[187:202]
1655.26	1655.26	[191:195]
1656.24	1656.24	[213:217]
3941.56	3941.56	[213:224]
4615.67	4615.66	[213:226]
7122.98	7122.96	[213:234]
8102.11	8102.10	[213:237]
9081.26	9081.25	[213:240]
2507.31	2507.30	[227:234]
3486.44	3486.44	[227:237]
4465.60	4465.58	[227:240]
1608.21	1608.22	[242:246]
10068.39	10068.36	[242:272]
7459.02	7459.01	[248:270]
8151.13	8151.12	[248:272]
1326.21	1326.20	[267:270]
5109.67	5109.66	[295:310]
6723.90	6723.88	[295:315]
3842.51	3842.50	[299:310]

Experimental mass (u)	Predicted mass (u)	Digestion product
1632.23	1632.23	[311:315]
4913.66	4913.67	[311:325]
4320.60	4320.61	[316:328]
4875.64	4875.63	[347:361]
2268.30	2268.29	[355:361]

Experimental monoisotopic masses were obtained from the data displayed in **Figure 3** by using the SNAP algorithm included in the Bruker DataAnalysis software (see *Data interpretation*). The predicted monoisotopic masses were calculated by RiboDynamics SeqRead and then matched to the experimental ones by using a ±5 ppm tolerance window. The unambiguously assigned products covered 92% of the 5’-UTR sequence.

Table S4: Unambiguous products identified by time-course analysis of 5'-UTR digestion with RNase T1 at 1:200 E/G and 25ºC (**Figure S7**) (continues in the next two pages).

Experimental mass (u)	Predicted mass (u)	Digestion product
6418.83	6418.80	[1:20]
8010.03	8009.99	[1:25]
2484.28	2484.27	[3:10]
7319.93	7319.90	[3:25]
9255.20	9255.15	[3:31]
24683.30	24683.20	[3:79]
1304.17	1304.16	[12:15]
16354.20	16354.20	[12:62]
23335.18	23335.10	[16:88]
24660.38	24660.27	[16:92]
17363.40	17363.31	[26:79]
1279.17	1279.17	[28:31]
16689.27	16689.20	[28:79]
24661.36	24661.26	[28:104]
16345.20	16345.13	[33:83]
21485.91	21485.80	[33:99]
9300.27	9300.22	[34:62]
14738.02	14737.95	[34:79]
2178.25	2178.25	[36:42]
8626.17	8626.13	[36:62]
14063.91	14063.85	[36:79]
11540.61	11540.56	[44:79]
22075.02	22074.93	[44:112]
10186.37	10186.41	[52:83]
16369.22	16369.15	[52:102]
2796.45	2796.44	[54:62]
8296.16	8296.13	[54:79]
14717.01	14716.94	[54:99]
16268.25	16268.17	[54:104]
1858.30	1858.29	[63:68]
6958.96	6958.95	[69:90]
1585.20	1585.19	[84:88]
6771.90	6771.86	[84:104]
7440.98	7440.95	[84:106]
8074.05	8074.01	[84:108]
9334.21	9334.18	[84:112]
975.11	975.11	[117:119]

Experimental mass (u)	Predicted mass (u)	Digestion product
1896.24	1896.23	[124:129]
2249.31	2249.30	[131:137]
5758.79	5758.77	[131:148]
8901.19	8901.14	[131:158]
15352.05	15351.97	[131:178]
16331.19	16331.11	[131:181]
20874.82	20874.73	[131:195]
3509.48	3509.47	[138:148]
2835.38	2835.37	[140:148]
3142.40	3142.38	[149:158]
10572.36	10572.35	[149:181]
7368.01	7368.01	[159:181]
3518.48	3518.47	[168:178]
7404.03	7403.99	[168:190]
11322.57	11322.53	[168:202]
3885.54	3885.52	[179:190]
14715.00	14715.07	[179:223]
2906.39	2906.38	[182:190]
8175.11	8175.13	[182:206]
24722.35	24722.37	[182:257]
6173.87	6173.84	[184:202]
20849.86	20849.89	[184:247]
5828.82	5828.79	[185:202]
5178.73	5178.70	[187:202]
1637.25	1637.25	[191:195]
20812.93	20812.83	[191:254]
14676.11	14676.05	[203:247]
1638.24	1638.23	[213:217]
3941.58	3941.56	[213:224]
4615.68	4615.66	[213:226]
7122.99	7122.96	[213:234]
9081.29	9081.25	[213:240]
4200.57	4200.55	[222:234]
9983.27	9983.29	[224:254]
2507.31	2507.30	[227:234]
4465.60	4465.58	[227:240]
6418.84	6418.85	[227:246]
8902.20	8902.19	[227:254]

Experimental mass (u)	Predicted mass (u)	Digestion product
19246.63	19246.60	[227:285]
14040.89	14040.93	[235:277]
16331.29	16331.27	[241:290]
1590.22	1590.21	[242:246]
5805.79	5805.77	[248:265]
7459.04	7459.01	[248:270]
8133.15	8133.11	[248:272]
22162.06	22161.98	[248:315]
8478.21	8478.24	[252:277]
21487.98	21487.97	[252:317]
5240.78	5240.75	[255:270]
5914.88	5914.85	[255:272]
14760.00	14760.00	[266:310]
18388.46	18388.51	[266:321]
1246.24	1246.24	[267:270]
16351.18	16351.20	[276:325]
3824.52	3824.50	[299:310]
1614.22	1614.22	[311:315]
3646.50	3646.51	[311:321]
3018.43	3018.42	[332:340]
2250.29	2250.28	[355:361]

Experimental monoisotopic masses were obtained from the data displayed in **Figure S4** by using the SNAP algorithm included in the Bruker DataAnalysis software (see *Data interpretation*). The predicted monoisotopic masses were calculated by RiboDynamics SeqRead and then matched to the experimental ones by using a ±5 ppm tolerance window. The unambiguously assigned products covered 92% of the 5'-UTR sequence.

Table S5: Unambiguous products identified by time-course analysis of eGFP mRNA digestion with RNase T1 at 1:200 E/G and 25ºC (**Figure S8**) (continues in the next two pages).

Experimental mass (u)	Predicted mass (u)	Digestion product
4826.69	4826.67	[6:20]
7459.08	7459.05	[6:28]
7786.12	7786.09	[6:29]
2892.41	2892.40	[12:20]
5863.75	5863.73	[38:55]
8723.14	8723.11	[38:64]
4907.64	4907.62	[41:55]
17924.41	17924.46	[48:102]
2289.36	2289.35	[91:97]
4530.64	4530.63	[91:104]
1609.21	1609.20	[98:102]
23190.30	23190.25	[105:175]
3381.49	3381.48	[107:116]
22250.14	22250.09	[111:178]
998.14	998.14	[126:128]
23536.28	23536.19	[139:211]
27392.87	27392.77	[139:223]
1919.27	1919.26	[146:151]
8007.11	8007.08	[146:170]
2532.31	2532.30	[155:162]
5084.69	5084.66	[155:170]
2897.41	2897.42	[163:171]
3525.47	3525.46	[179:189]
8672.20	8672.18	[185:211]
7081.99	7081.97	[190:211]
3854.52	3854.51	[191:202]
6736.95	6736.92	[191:211]
8985.27	8985.23	[191:218]
2248.32	2248.31	[212:218]
2593.37	2593.36	[212:219]
8376.14	8376.10	[220:245]
9604.31	9604.28	[224:253]
1896.24	1896.23	[229:234]
9894.37	9894.41	[229:259]
1318.19	1318.19	[235:238]
4553.65	4553.66	[246:259]
9291.30	9291.26	[246:274]

Experimental mass (u)	Predicted mass (u)	Digestion product
19714.80	19714.76	[246:306]
5734.76	5734.75	[257:274]
14551.04	14550.96	[257:301]
14834.10	14834.05	[257:302]
3471.44	3471.43	[264:274]
1952.32	1952.33	[302:307]
5822.81	5822.78	[305:322]
2601.39	2601.38	[308:315]
2244.28	2244.28	[316:322]
3253.49	3253.48	[331:340]
7192.06	7192.04	[331:352]
16341.25	16341.29	[348:397]
15362.09	15362.16	[351:397]
12815.83	12815.78	[366:404]
14486.07	14486.01	[377:420]
13471.89	13471.87	[380:420]
1656.25	1656.24	[400:404]
3230.47	3230.45	[421:430]
12645.82	12645.76	[421:459]
15198.20	15198.13	[421:467]
27869.92	27869.97	[432:517]
10950.60	10950.55	[434:467]
18733.71	18733.68	[440:497]
4480.64	4480.63	[446:459]
28433.10	28433.05	[446:533]
2884.39	2884.38	[470:478]
2314.35	2314.35	[499:505]
3535.51	3535.50	[507:517]
3229.48	3229.47	[524:533]
3903.58	3903.57	[524:535]
3581.46	3581.47	[544:554]
4560.62	4560.61	[544:557]
13206.90	13206.84	[544:584]
3910.53	3910.52	[546:557]
3259.46	3259.44	[548:557]
7731.11	7731.08	[548:571]
18730.63	18730.60	[551:608]

Experimental mass (u)	Predicted mass (u)	Digestion product
3492.50	3492.49	[558:568]
8664.27	8664.25	[558:584]
12895.82	12895.85	[569:608]
4192.63	4192.61	[572:584]
4519.67	4519.65	[572:585]
27836.95	27836.87	[588:673]
5460.75	5460.72	[592:608]
9893.36	9893.33	[592:622]
8280.12	8280.09	[597:622]
11201.57	11201.52	[597:631]
9594.37	9594.33	[602:631]
4432.63	4432.61	[609:622]
7354.06	7354.04	[609:631]
26416.77	26416.65	[623:703]
2265.34	2265.34	[625:631]
4536.66	4536.68	[643:656]
7191.06	7191.06	[643:664]
2881.44	2881.42	[648:656]
5462.70	5462.68	[666:682]
3582.46	3582.45	[675:685]
2938.42	2938.41	[690:698]
7468.04	7468.01	[690:712]
12405.74	12405.69	[696:733]
2860.37	2860.36	[704:712]

Experimental monoisotopic masses were obtained from the data displayed in **Figure S5** by using the SNAP algorithm included in the Bruker DataAnalysis software (see *Data interpretation*). The predicted monoisotopic masses were calculated by RiboDynamics SeqRead and then matched to the experimental ones by using a ±5 ppm tolerance window. The unambiguously assigned products covered 93% of eGFP mRNA sequence.

Table S6: Unambiguous products identified by time-course analysis of 18S rRNA digestion with RNase T1 at 1:200 E/G and 25ºC (**Figure S9**) (continues in the next six pages).

Experimental mass (u)	Predicted mass (u)	Digestion products	Modifications
23782.20	23782.09	[7:80]	Am27, Ψ34, Ψ36
11549.54	11549.51	[17:52]	Am27, Ψ34, Ψ36
4176.54	4176.53	[21:33]	Am27
6730.89	6730.86	[21:41]	Am27, Ψ34, Ψ36
1920.25	1920.25	[57:62]	-
18383.50	18383.42	[57:113]	Ψ93, Am99, Ψ105, Ψ109
5853.82	5853.79	[71:88]	
8980.21	8980.25	[81:108]	Ψ93, Am99, Ψ105
4143.57	4143.55	[96:108]	Am99, Ψ105
2513.30	2513.29	[115:122]	Um116, Ψ119, Um121
16215.18	16215.14	[127:177]	Am159, Am166, Um172, Cm174
9964.37	9964.39	[147:177]	Am159, Am166, Um172, Cm174
2875.38	2875.36	[157:165]	-
706.13	706.13	[166:167]	-
3217.46	3217.45	[168:177]	Um172, Cm174
23517.26	23517.14	[168:240]	Um172, Cm174, Ψ210, Ψ218
20644.85	20644.75	[178:241]	Ψ210, Ψ218
4350.59	4350.59	[201:213]	Ψ210
5001.68	5001.66	[201:215]	Ψ210
3166.40	3166.39	[216:225]	Ψ218
18383.58	18383.50	[226:282]	-
18746.65	18746.57	[226:283]	-
20296.81	20296.71	[226:288]	-
20641.84	20641.77	[226:289]	-
21292.89	21292.84	[226:291]	-
22058.91	22058.91	[266:333]	Ψ296
8168.11	8168.11	[267:291]	-
1568.17	1568.16	[284:288]	-
2225.29	2225.29	[292:298]	Ψ296
3204.44	3204.43	[299:308]	-
12869.79	12869.74	[299:338]	-

Experimental mass (u)	Predicted mass (u)	Digestion products	Modifications
15705.18	15705.10	[299:347]	-
4744.66	4744.64	[313:327]	-
5790.79	5790.80	[334:351]	-
7066.96	7066.96	[334:355]	Um354
4143.57	4143.56	[339:351]	-
4692.60	4692.58	[356:370]	-
19817.70	19817.61	[374:434]	Ψ406, Um428
9349.28	9349.28	[384:412]	Ψ406
17977.41	17977.41	[384:438]	Ψ406, Um428, Gm436
1918.27	1918.28	[398:403]	-
7925.10	7925.10	[398:421]	Ψ406
10118.45	10118.40	[404:434]	Ψ406, Um428
20644.86	20644.90	[441:504]	Cm462, Am468, Am484
4464.65	4464.63	[453:466]	Cm462
1017.17	1017.17	[468:470]	Am468
10611.49	10611.47	[475:507]	-
5362.75	5362.73	[483:499]	Am484
10014.43	10014.40	[483:513]	Am484, Gm509, Am512
31100.24	31100.16	[505:600]	Gm509, Am512, Cm517, Ψ572, Am576, Am590
1012.16	1012.15	[511:513]	Am512
993.16	993.16	[516:518]	Cm517
6524.98	6524.96	[516:535]	Cm517
10357.41	10357.44	[516:547]	Cm517
5186.78	5186.76	[519:534]	-
12851.76	12851.71	[519:558]	-
13196.82	13196.76	[519:559]	-
2813.33	2813.32	[536:544]	-
1921.24	1921.23	[553:558]	-
6611.86	6611.83	[566:586]	Ψ572, Am576
10176.35	10176.30	[566:597]	Ψ572, Am576, Am590
18743.67	18743.59	[589:646]	Am590, Gm601, Ψ609, Cm621, Um627, Gm644
2570.35	2570.33	[603:610]	Ψ609
20362.65	20362.75	[611:673]	Cm621, Um627, Gm644, Gm649, Ψ651, Am668

Experimental mass (u)	Predicted mass (u)	Digestion products	Modifications
2616.40	2616.39	[618:625]	Cm621
987.16	987.16	[621:623]	Cm621
11564.53	11564.50	[624:659]	Um627, Gm644, Gm649, Ψ651
2892.40	2892.39	[627:635]	Um627
11543.55	11543.50	[627:662]	Um627, Gm644, Gm649, Ψ651
3232.43	3232.42	[647:656]	Ψ651
12218.68	12218.62	[647:684]	Ψ651, Am668, Ψ681, Gm683
23125.10	23125.07	[657:727]	Am668, Ψ681, Gm683
28578.81	28578.74	[674:762]	Ψ681, Gm683
1334.18	1334.17	[681:684]	Ψ681, Gm683
3849.53	3849.55	[685:696]	-
5189.71	5189.73	[685:700]	-
9153.23	9153.20	[685:712]	-
9803.33	9803.29	[685:714]	-
7924.10	7924.08	[691:714]	-
20297.76	20297.71	[692:754]	-
16088.22	16088.15	[695:744]	-
17978.40	17978.38	[697:752]	-
10134.44	10134.41	[713:744]	-
4823.66	4823.67	[717:731]	-
3236.47	3236.46	[718:727]	-
16039.08	16039.05	[722:772]	-
2522.32	2522.31	[745:752]	-
8836.13	8836.10	[745:772]	-
16632.22	16632.15	[745:796]	-
6596.91	6596.88	[752:772]	-
2481.33	2481.32	[755:762]	-
11289.49	11289.46	[773:807]	Cm797, Um799, Ψ801
2829.31	2829.31	[799:807]	Um799, Ψ801
19816.72	19816.71	[820:880]	Ψ822, Ψ863, Ψ866, Gm867
3243.45	3243.45	[846:855]	-
11706.65	11706.67	[846:881]	Ψ863, Ψ866, Gm867

Experimental mass (u)	Predicted mass (u)	Digestion products	Modifications
1638.24	1638.23	[869:873]	-
5745.70	5745.69	[874:891]	-
10881.32	10881.29	[874:907]	-
5400.65	5400.63	[875:891]	-
3121.33	3121.32	[882:891]	-
4241.57	4241.58	[903:915]	-
3643.52	3643.51	[922:932]	-
8677.20	8677.16	[935:961]	-
28187.86	28187.86	[957:1043]	Ψ966, Ψ1004, Am1031
2861.35	2861.35	[962:970]	Ψ966
9731.37	9731.33	[962:991]	Ψ966
28186.87	28186.78	[962:1048]	Ψ966, Ψ1004, Am1031, Ψ1045
3587.52	3587.51	[972:982]	-
6524.96	6524.94	[972:991]	-
11544.54	11544.50	[994:1029]	Ψ1004
21291.93	21291.90	[1006:1071]	Am1031, Ψ1045, Ψ1046, Ψ1056
4742.59	4742.57	[1015:1029]	-
16650.25	16650.28	[1045:1096]	Ψ1045, Ψ1046, Ψ1056, Ψ1081
8634.16	8634.14	[1060:1086]	Ψ1081
12848.76	12848.71	[1060:1099]	Ψ1081
23836.20	23836.14	[1069:1142]	Ψ1081, Ψ1136
10641.48	10641.44	[1072:1104]	Ψ1081
3211.47	3211.46	[1077:1086]	-
10069.38	10069.39	[1077:1107]	Ψ1081
13192.80	13192.74	[1090:1130]	-
5187.78	5187.76	[1093:1108]	-
21290.89	21290.80	[1093:1158]	Ψ1136
5992.78	5992.75	[1111:1129]	-
10162.32	10162.36	[1111:1142]	Ψ1136
2562.35	2562.36	[1122:1129]	-
24672.36	24672.28	[1122:1197]	Ψ1136, Ψ1177
6470.93	6470.90	[1132:1151]	Ψ1136
5491.79	5491.76	[1135:1151]	Ψ1136
2929.46	2929.44	[1143:1151]	-
3966.53	3966.52	[1160:1171]	-
18855.63	18855.63	[1176:1233]	Ψ1177, Ψ1232

Experimental mass (u)	**Predicted mass (u)**	**Digestion products**	**Modifications**
10135.46	10135.42	[1180:1210]	-
18855.62	18855.53	[1188:1245]	Ψ1232, Ψ1238, Ψ1244
1615.21	1615.20	[1199:1203]	-
10117.50	10117.47	[1199:1229]	-
28184.84	28184.94	[1212:1298]	Ψ1232, Ψ1238, Ψ1244, m^1acp^3Ψ1248, Cm1272, Um1288
10641.44	10641.39	[1213:1245]	Ψ1232, Ψ1238, Ψ1244
18744.70	18744.64	[1223:1280]	Ψ1232, Ψ1238, Ψ1244, m^1acp^3Ψ1248, Cm1272
12902.79	12902.76	[1230:1269]	Ψ1232, Ψ1238, Ψ1244, m^1acp^3Ψ1248
29670.14	29669.99	[1234:1325]	Ψ1238, Ψ1244, m^1acp^3Ψ1248, Cm1272, Um1288
3302.52	3302.51	[1246:1255]	m^1acp^3Ψ1248
7790.16	7790.13	[1246:1269]	m^1acp^3Ψ1248
8135.21	8135.18	[1246:1270]	m^1acp^3Ψ1248
6469.92	6469.92	[1256:1275]	Cm1272
18380.53	18380.46	[1256:1312]	Cm1272, Um1288
20196.80	20196.73	[1256:1318]	Cm1272, Um1288
14922.00	14922.07	[1257:1302]	Cm1272, Um1288
5416.78	5416.76	[1258:1274]	Cm1272
23519.22	23519.15	[1258:1330]	Cm1272, Um1288, Um1326, Gm1328
7839.12	7839.12	[1271:1294]	Cm1272, Um1288
3284.50	3284.49	[1276:1285]	-
8833.16	8833.14	[1291:1318]	-
3095.35	3095.34	[1303:1312]	-
4012.52	4012.51	[1319:1330]	Um1326, Gm1328
20383.66	20383.58	[1319:1381]	Um1326, Gm1328, ac4C1337, Ψ1347, Ψ1367
3956.52	3956.51	[1323:1334]	Um1326, Gm1328
8831.14	8831.11	[1328:1354]	Gm1328, ac4C1337, Ψ1347
2203.25	2203.24	[1339:1345]	-
4444.54	4444.52	[1362:1375]	Ψ1367
6387.82	6387.79	[1362:1381]	Ψ1367
2837.35	2837.34	[1367:1375]	Ψ1367

Experimental mass (u)	Predicted mass (u)	Digestion products	Modifications
12889.71	12889.72	[1367:1406]	Ψ1367, 1383Am, Cm1391
11563.56	11563.61	[1376:1411]	Ψ1367, 1383Am, Cm1391
4536.62	4536.64	[1412:1425]	-
20296.83	20296.81	[1429:1491]	Um1442, Ψ1445, Gm1447, Gm1490
8677.21	8677.18	[1430:1456]	Um1442, Ψ1445, Gm1447
9328.29	9328.26	[1430:1458]	Um1442, Ψ1445, Gm1447
14113.93	14113.89	[1430:1473]	Um1442, Ψ1445, Gm1447
21293.97	21293.88	[1430:1495]	Um1442, Ψ1445, Gm1447, Gm1490
13189.76	13189.76	[1476:1516]	Gm1490
11291.50	11291.47	[1482:1516]	Gm1490
2201.28	2201.28	[1501:1507]	-
12851.74	12851.68	[1501:1540]	-
8367.13	8367.10	[1511:1536]	-
8674.20	8674.20	[1515:1541]	-
9919.37	9919.36	[1516:1546]	-
6070.83	6070.80	[1518:1536]	-
9291.24	9291.22	[1518:1546]	-
2530.34	2530.33	[1529:1536]	-
8021.05	8021.07	[1551:1575]	-
3445.46	3445.45	[1553:1563]	-
14201.89	14201.89	[1574:1617]	-
4444.55	4444.57	[1585:1598]	-
15354.06	15354.00	[1585:1632]	Ψ1625
3469.47	3469.46	[1588:1598]	-
15705.16	15705.09	[1588:1636]	Ψ1625
11893.61	11893.57	[1612:1648]	Ψ1625, m^{7}G1639, Ψ1643
11548.58	11548.53	[1613:1648]	Ψ1625, m^{7}G1639, Ψ1643
6002.77	6002.74	[1614:1632]	Ψ1625, m^{7}G1639
4716.61	4716.59	[1618:1632]	Ψ1625
5200.75	5200.73	[1633:1648]	m^{7}G1639, Ψ1643
13633.80	13633.81	[1633:1674]	m^{7}G1639, Ψ1643, Um1668, Am1678
5549.75	5549.73	[1649:1665]	-

Experimental mass (u)	Predicted mass (u)	Digestion products	Modifications
16632.24	16632.16	[1655:1706]	Um1668, Am1678, Ψ1692, Cm1703
8830.13	8830.12	[1666:1693]	Um1668, Am1678, Ψ1692
19438.38	19438.45	[1672:1732]	Am1678, Ψ1692, Cm1703
4050.49	4050.47	[1681:1693]	Ψ1692
22057.89	22057.81	[1681:1749]	Ψ1692, Cm1703
2178.25	2178.25	[1687:1693]	Ψ1692
7924.10	7924.06	[1694:1718]	Cm1703
14921.95	14921.92	[1719:1764]	-
7099.95	7099.95	[1733:1754]	-
9966.39	9966.36	[1733:1763]	-
8662.20	8662.16	[1750:1776]	-
16178.28	16178.24	[1750:1799]	-
8461.17	8461.14	[1759:1784]	-
16178.26	16178.18	[1765:1814]	Um1804
5549.77	5549.74	[1771:1787]	-
5878.81	5878.79	[1772:1789]	-
23514.24	23514.16	[1778:1849]	-
3299.46	3299.45	[1780:1789]	-
12334.73	12334.73	[1780:1817]	Um1804
23520.33	23520.24	[1790:1861]	Um1804, m^6A1832, ac4C1842, $m^{6,6}$A1850, $m^{6,6}$A1851
15725.17	15725.25	[1815:1862]	m^6A1832, ac4C1842, $m^{6,6}$A1850, $m^{6,6}$A1851
1967.29	1967.28	[1821:1826]	-
2286.35	2286.34	[1830:1836]	m^6A1832
2198.31	2198.31	[1837:1843]	ac4C1842
2833.39	2833.39	[1838:1846]	ac4C1842
3891.48	3891.46	[1838:1849]	ac4C1842
6454.90	6454.92	[1838:1857]	ac4C1842, $m^{6,6}$A1850, $m^{6,6}$A1851
1993.34	1993.33	[1850:1855]	$m^{6,6}$A1850, $m^{6,6}$A1851
3973.62	3973.61	[1850:1861]	$m^{6,6}$A1850, $m^{6,6}$A1851
2148.33	2148.32	[1863:1869]	-

Experimental monoisotopic masses were obtained from the data displayed in **Figure S6** by using the SNAP algorithm included in the Bruker DataAnalysis software (see *Data interpretation*). The predicted monoisotopic masses were calculated by RiboDynamics SeqRead and then matched to the experimental ones by using a ±5 ppm tolerance window. Position and identity of modified bases identified within each product are reported in the fourth column. The unambiguously assigned products covered 98% of the 18S rRNA sequence.

Table S7: Estimated stoichiometries of post-transcriptional modifications observed by time-course analysis of 18S rRNA digestion with RNase T1.

Modification	**Modified product**	**Modified mass (u)**	**Mod. ion intensity**	**Unmodified product**	**Unmodified mass (u)**	**Unmod. ion intensity**	**Stoich. (%)**
Cm462	[453:466]	4464.65	6,429,966	[453:466]	4450.62	1,677,741	79
Am1031	[1030:1041]	3973.62	2,521	[1015:1044]	9721.21	58,086	4
Cm1391	[1367:1406]	12889.72	7,330	[1376:1393]	5821.81	37,330	16
Gm1490	[1430:1495]	21294.01	8,530	[1467:1490]	7740.07	1,080	73
	[1460:1491]	10358.42	11,994				

Experimental monoisotopic masses are reported for both unmodified and modified versions of selected digestion products. Unbiased determinations were ensured by considering unmodified/modified products detected at the same time point of the reaction course (i.e., 180’). The reported ion intensities expressed in arbitrary counts were obtained by dividing the detected intensities by the respective charge states as described in the ***Data interpretation*** section.

Table S8: Sequence coverage obtained by MS/MS determination of selected digestion products of 5'-UTR.

Digestion product	Size (nt)	Precursor mass (u)	Sequence coverage (%)
[3:31]	29	9255.16	100
[34:62]	29	9300.22	100
[36:112]	77	24660.18	72
[107:139]	33	10597.31	100
[131:202]	72	23156.06	96
[196:259]	64	20828.82	94
[248:272]	25	8133.11	100
[274:289]	16	5289.72	81
[291:310]	20	6434.84	100
[295:310]	16	5109.66	100
[299:329]	31	10104.11	100
[334:361]	18	9195.21	100

Fragmentation data were obtained by CID analysis as described in the *Mass spectrometry* section and interpreted as described in *Data interpretation*. The coverages obtained from the individual precursor ions were combined to calculate that the results covered 92% of the entire 5'-UTR sequence.

Table S9: Sequence coverage obtained by MS/MS determination of selected digestion products of eGFP mRNA (continues in the next page).

Digestion product	Size (nt)	Precursor mass (u)	Sequence coverage (%)
[4:28]	25	8115.15	91
[6:28]	23	7441.04	100
[56:64]	9	2859.38	100
[78:123]	46	15197.12	92
[91:106]	16	5163.69	100
[126:142]	17	5557.79	95
[139:154]	16	5228.74	100
[155:170]	16	5084.66	100
[172:211]	40	12895.75	95
[191:218]	28	8985.23	89
[219:274]	56	18012.33	81
[246:274]	29	9291.26	92
[257:301]	45	14550.96	87
[305:322]	18	5822.78	100
[316:350]	35	11330.64	96
[331:340]	10	3235.47	100
[341:350]	10	3282.47	100
[354:364]	11	3590.46	100
[365:389]	25	8181.13	97
[400:412]	13	4192.57	100
[421:459]	39	12645.77	100
[434:467]	34	10950.55	100
[470:478]	9	2884.38	100
[479:487]	9	2979.45	80
[483:493]	11	3619.54	100
[483:497]	15	4927.73	89
[499:505]	7	2296.33	100
[499:533]	35	11329.59	88
[507:522]	16	5131.71	100
[524:533]	10	3211.46	100
[544:584]	41	13206.84	100
[572:640]	69	22248.07	97
[625:647]	23	7492.06	100
[648:658]	11	3555.52	100
[665:689]	25	8107.03	95
[690:712]	23	7468.01	97

Digestion product	Size (nt)	Precursor mass (u)	Sequence coverage (%)
[720:737]	18	5861.82	100
[741:748]	8	2680.40	100

Fragmentation data were obtained by CID analysis as described in the *Mass spectrometry* section and interpreted as described in *Data interpretation*. The coverages obtained from the individual precursor ions were combined to calculate that the results covered 84% of the entire eGFP mRNA sequence.

Table S10: Sequence coverage obtained by MS/MS determination of selected digestion products of 18S rRNA (continues in the next page).

Digestion product	Size (nt)	Precursor mass (u)	Sequence coverage (%)	Confirmed modifications
[17:52]	36	11549.51	100	Am27, Ψ34, Ψ36
[21:56]	36	11550.49	100	-
[48:80]	33	10642.43	100	-
[168:177]	10	3235.46	100	Um172, Cm174
[181:200]	20	6391.84	100	-
[201:215]	15	5001.66	100	Ψ210
[216:225]	10	3166.39	100	-
[226:255]	30	9507.29	98	Ψ218
[226:291]	66	21292.71	52	-
[292:298]	7	2225.29	100	Ψ296
[299:347]	49	15705.10	97	-
[356:370]	15	4692.58	100	-
[386:403]	18	5749.76	100	-
[453:466]	14	4446.62	100	Cm462
[483:499]	17	5362.73	100	Am484
[519:558]	40	12851.71	99	-
[566:597]	32	10194.31	100	Ψ572, Am576, Am590
[590:620]	31	10116.38	100	Gm601, Ψ609
[618:690]	73	23519.09	82	Cm621, Um627, Ψ649, Ψ651, Am668, Gm683
[627:662]	36	11543.49	99	Gm644
[685:714]	30	9803.29	97	-
[685:772]	88	28185.62	70	-
[745:796]	52	16650.15	78	-
[773:807]	35	14676.28	100	Cm797, Um799, Ψ801
[789:821]	33	10705.43	100	Ψ814, Ψ815
[829:838]	10	3258.46	100	-
[833:878]	46	14924.05	99	Ψ863, Ψ866, Gm867
[874:891]	18	5745.68	100	-
[903:915]	13	4223.57	100	-
[926:942]	17	5555.74	100	-
[962:1029]	68	21949.92	97	Ψ966
[1000:1040]	41	13190.73	98	Ψ1004, Am1031

Digestion product	Size (nt)	Precursor mass (u)	Sequence coverage (%)	Confirmed modifications
[1030:1037]	8	2654.42	100	-
[1042:1099]	58	18744.53	100	Ψ1045, Ψ1046, Ψ1056, Ψ1081
[1093:1158]	66	21479.80	99	Ψ1136
[1160:1184]	25	8210.02	100	Ψ1174, Ψ1177
[1180:1210]	31	10135.41	67	-
[1222:1229]	8	2707.38	100	-
[1230:1321]	92	29669.94	71	Ψ1232, Ψ1238, Ψ1244, Cm1272
[1246:1255]	10	3302.51	100	m^{1}acp^{3}Ψ1248
[1281:1325]	45	14548.85	100	Um1288
[1339:1352]	14	4439.55	100	Ψ1347
[1367:1406]	40	12889.73	94	Ψ1367, Am1383, Cm1391
[1412:1420]	9	2849.40	100	-
[1430:1456]	27	8677.09	100	Um1442, Ψ1445, Gm1447
[1430:1495]	66	22814.86	83	Gm1490
[1501:1563]	63	20166.60	90	-
[1537:1598]	62	19883.61	79	-
[1588:1598]	11	3469.46	100	-
[1601:1617]	17	5621.73	100	-
[1618:1632]	15	4716.59	100	Ψ1625
[1633:1648]	16	5200.74	100	m^{7}G1639, Ψ1643
[1640:1680]	41	13188.76	100	Um1668
[1672:1706]	35	11098.43	100	Am1678, Ψ1692
[1710:1744]	35	11288.43	99	-
[1733:1754]	22	7099.94	100	-
[1765:1814]	50	16178.17	87	Um1804
[1821:1826]	6	1967.28	100	-
[1830:1836]	7	2286.34	100	m^{6}A1832
[1838:1861]	24	7865.07	96	m^{6,6}A1850, m^{6,6}A1851
[1863:1869]	7	2148.32	100	-

Fragmentation data were obtained by CID analysis as described in the *Mass spectrometry* section and interpreted as described in *Data interpretation*. The coverages obtained from the individual precursor ions were combined to calculate that the results covered 76% of the entire 18S rRNA sequence, which confirmed the position of 64 out of 91 (70%) rPTMs.

Figures

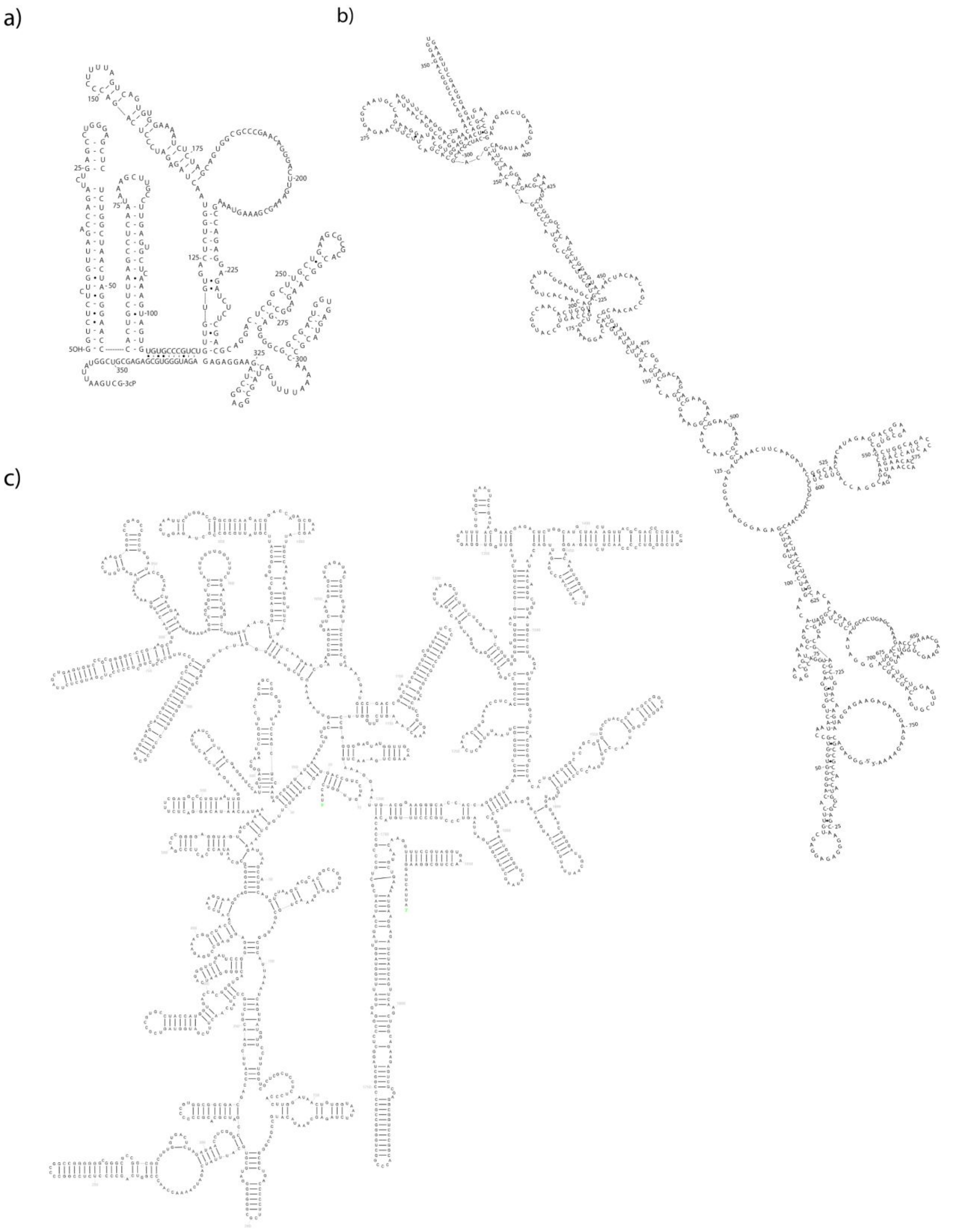

Figure S1: Putative secondary structures of the test samples considered in the study, i.e. **a)** HIV-1 5'-UTR, **b)** eGFP mRNA, and **c)** human 18S rRNA. The actual sequences and associated statistics are included in **Table S1**.

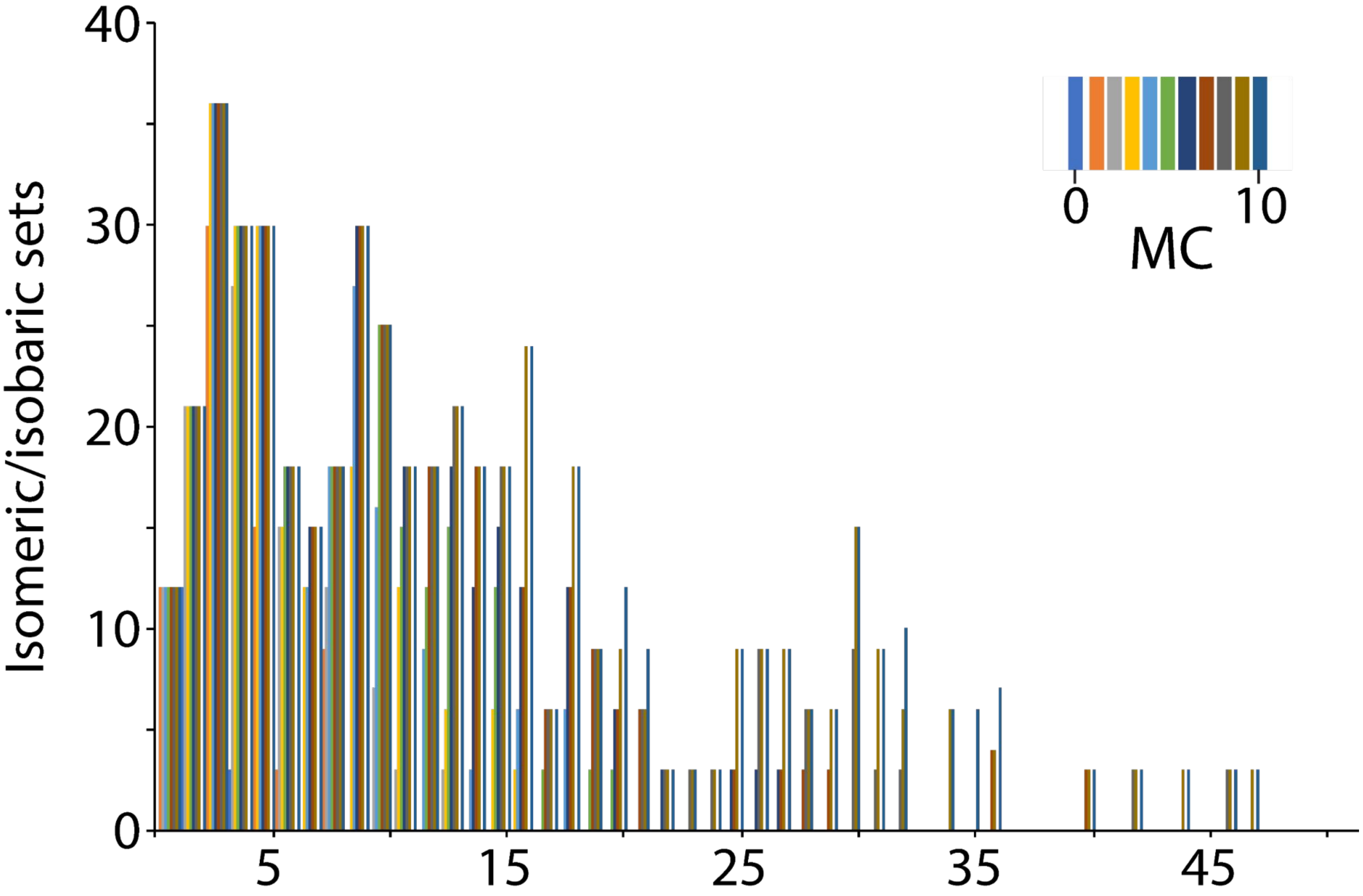

Figure S2: Combined number of isomeric and ±5 ppm isobaric sets as a function of size.

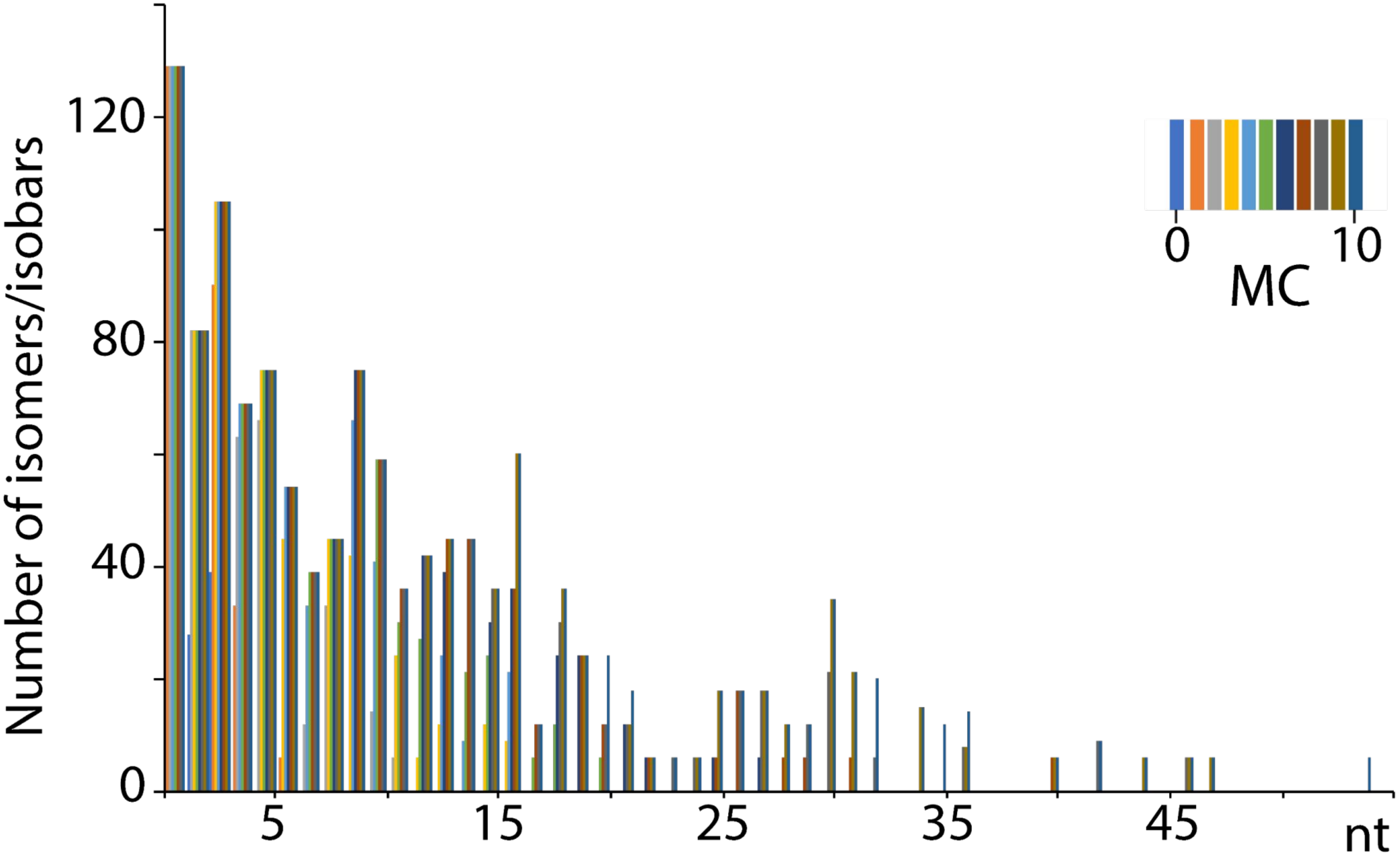

Figure S3: Combined number of isomers and ±5 ppm isobars as a function of size.

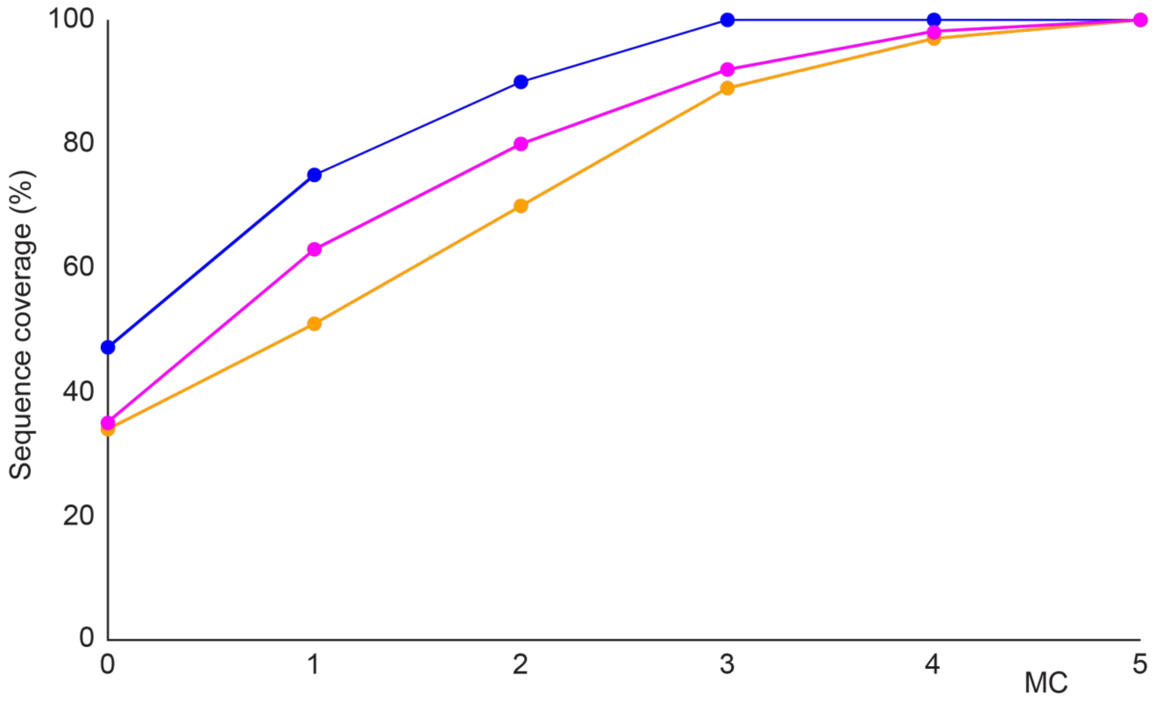

Figure S4: Mapping coverages predicted from RNase T1 digestions of HIV-1 5'-UTR (blue), eGFP mRNA (orange), and 18S rRNA (magenta) as a function of missed cleavages (MC). The percentage values were calculated after discarding all predicted products sharing the same mass (i.e. isomers).

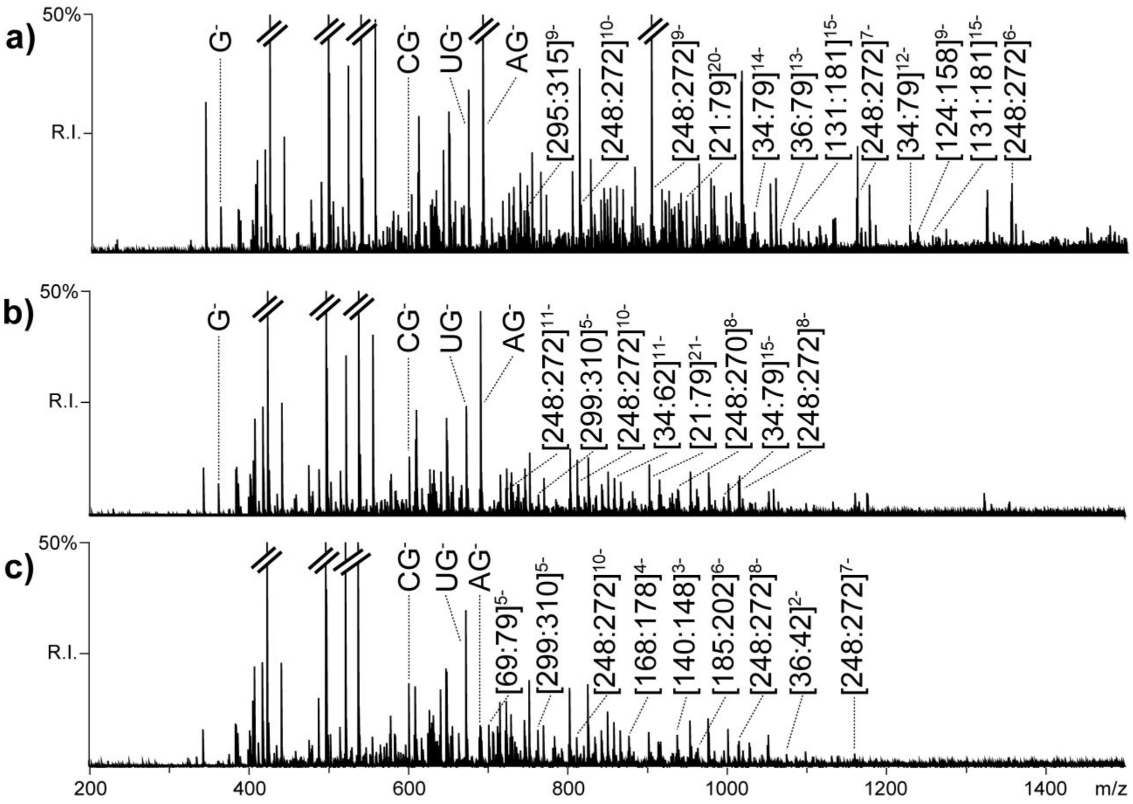

Figure S5: Direct infusion nanospray mass spectra of a digestion mixture obtained upon 30 min incubation of 5'-UTR with RNase T1 at 1:100 E/S ratio in 150 mM ammonium acetate at pH 7.0 at variable temperatures, i.e. **a)** 25°C, **b)** 37°C, and **c)** 60°C (see *Enzyme digestion* section and **Table 1b**). Products identified by a range were unambiguously assigned, whereas those identified by a sequence could be attributed to alternate sections of the initial substrate.

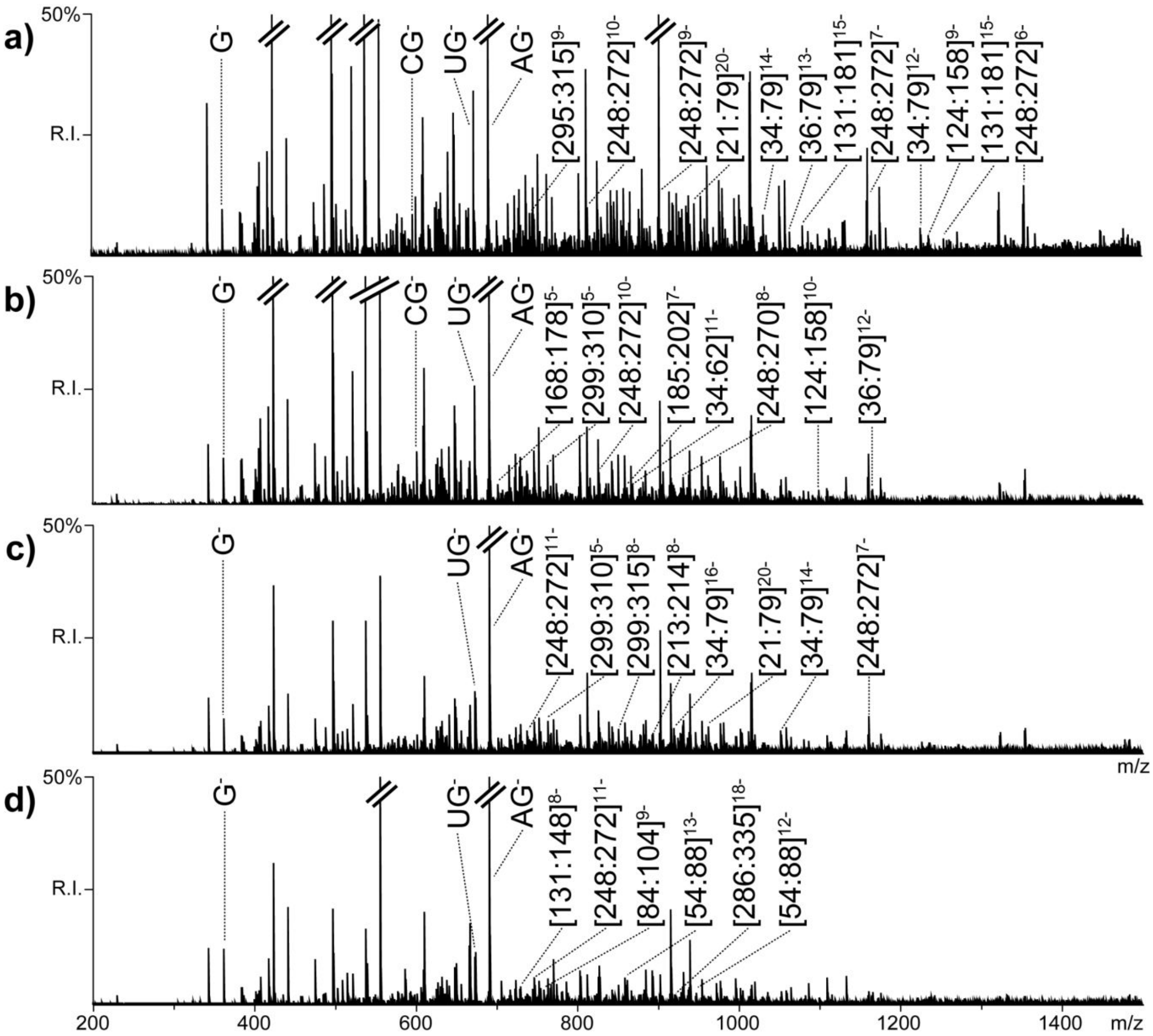

Figure S6: Direct infusion nanospray mass spectra of a digestion mixture obtained by incubating 5'-UTR with RNase T1 at a 1:100 E/S ratio in 150 mM ammonium acetate at pH 7.0 and 25°C for variable times, i.e. **a)** 30 min, **b)** 60 min, **c)** 120 min and **d)** 10 h (overnight, see *Enzyme digestion* section and **Table 1c**).

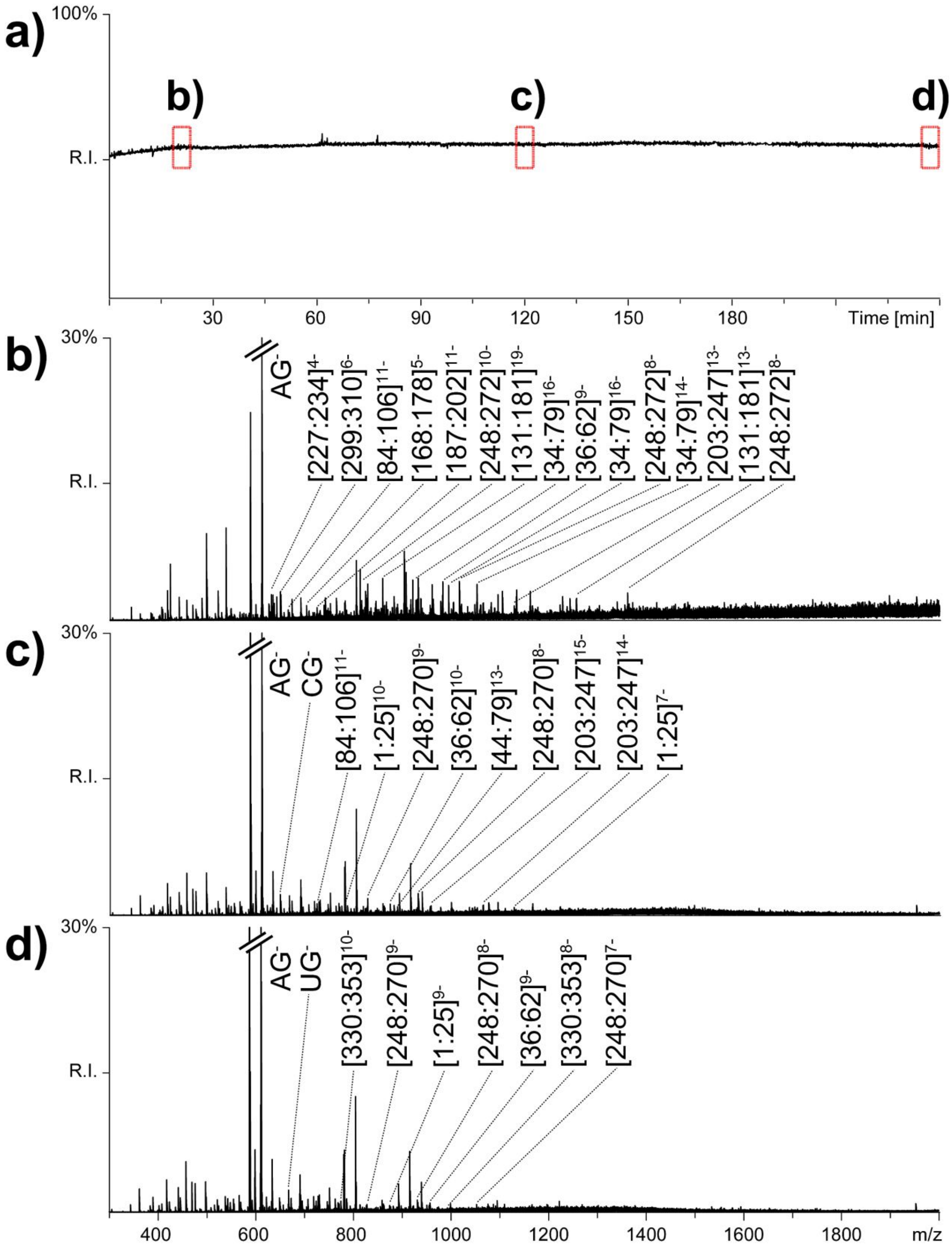

Figure S7: Time-course analysis of 5'-UTR digestion with RNase T1 at 1:200 E/G and 25ºC (see *Enzyme digestion* section and **Figure 4**). **a)** Total ion current (TIC) recorded continuously over a 240 min reaction time. Panels **b**) through **d**) are signal-averaged mass spectra extracted at the time points indicated in panel **a)**. Products identified by a range were unambiguously assigned, whereas those identified by a sequence could be attributed to alternate sections of the initial substrate.

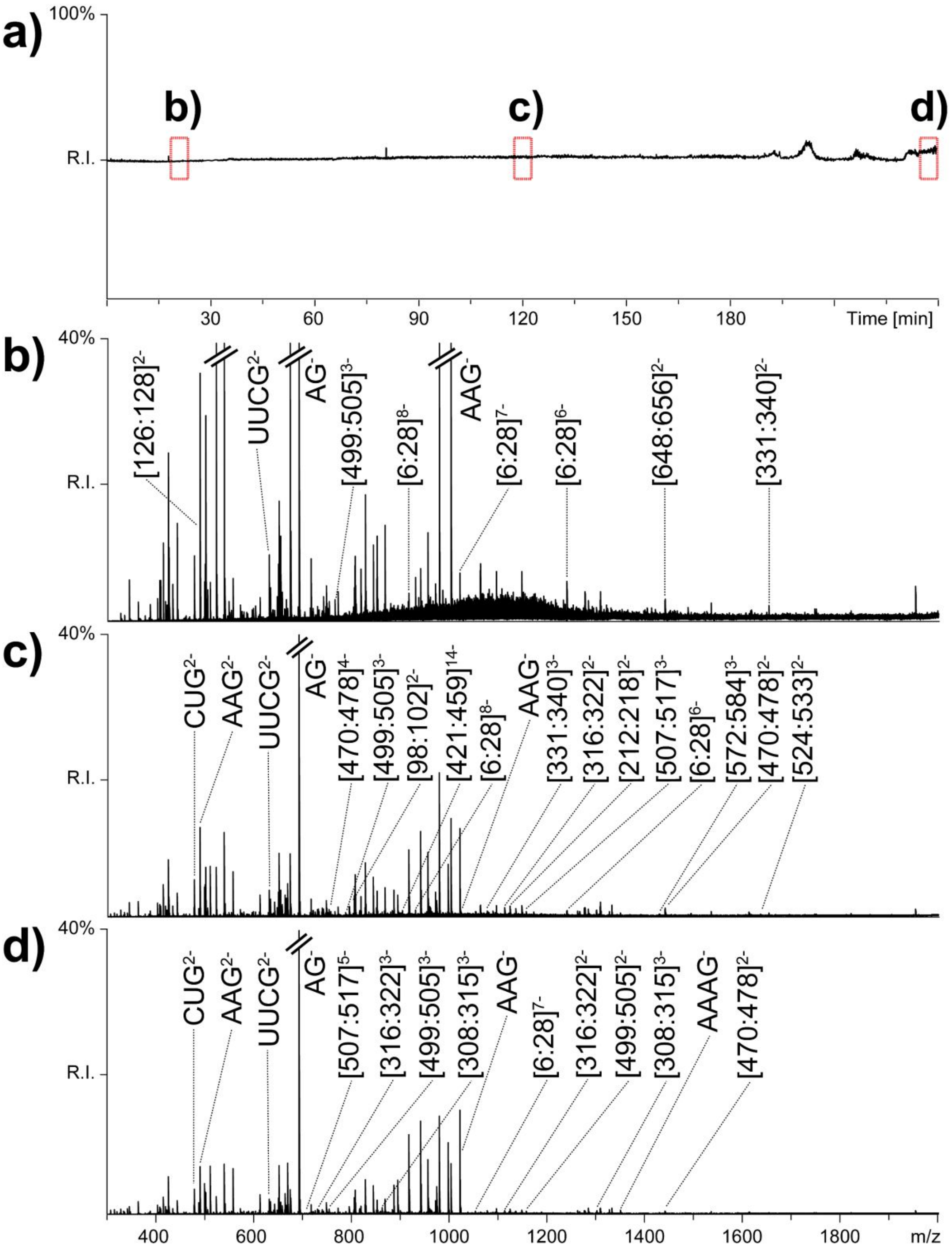

Figure S8: Time-course analysis of eGFP mRNA digestion with RNase T1 at 1:200 E/G and 25ºC (see *__Enzyme digestion__* section and **Figure 4**). **a)** Total ion current (TIC) recorded continuously over a 240 min reaction time. Panels **b**) through **d**) are signal-averaged mass spectra extracted at the time points indicated in panel **a)**.

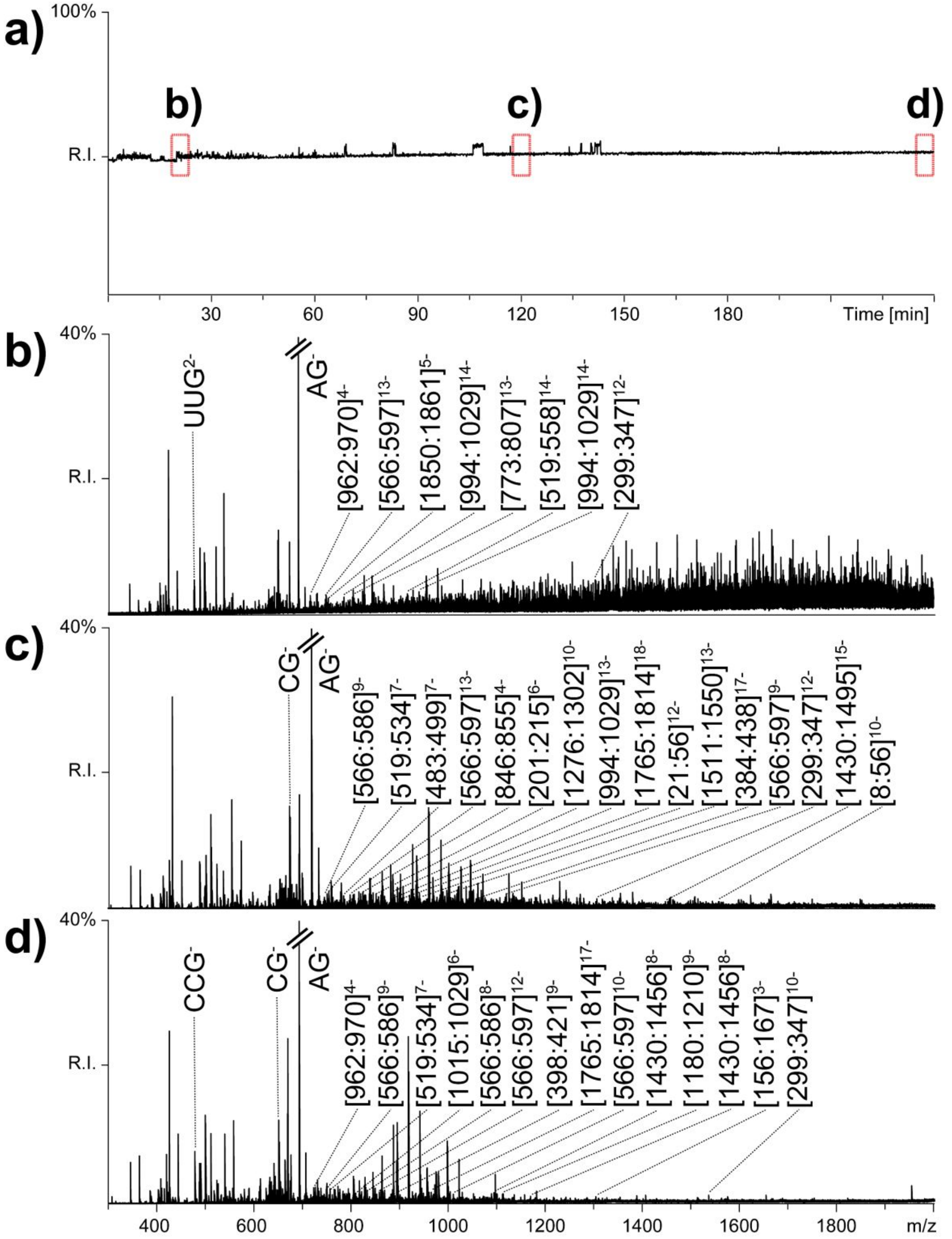

Figure S9: Time-course analysis of 18S rRNA digestion with RNase T1 at 1:200 E/G and 25ºC (see *Enzyme digestion* section and **Figure 4**). **a)** Total ion current (TIC) recorded continuously over a 240 min reaction time. Panels **b**) through **d**) are signal-averaged mass spectra extracted at the time points indicated in panel **a)**.