***Supporting Information***

**2D-HELS MS Seq: A general LC-MS-based method for direct and *de novo* sequencing of RNA mixtures with different nucleotide modifications**

**AUTHORS AND AFFILIATIONS:**

Ning Zhang1a,2, Shundi Shi2, Barney Yoo3, Xiaohong Yuan1a, Wenjia Li1b and Shenglong Zhang1a, \*

1a Department of Biological and Chemical Sciences, New York Institute of Technology, New York, NY, 10023, USA

1b Department of Computer Science, New York Institute of Technology, New York, NY, 10023, USA

2 Department of Chemical Engineering, Columbia University, New York, NY, 10027, USA

3 Department of Chemistry, Hunter College, City University of New York, New York, NY, 10065, USA

**EMAIL ADDRESSES OF CO-AUTHORS:**

Ning Zhang (nzhang07@nyit.edu)

Shundi Shi (ss526@columbia.edu)

Barney Yoo (by104@hunter.cuny.edu)

Xiaohong Yuan (xyuan04@nyit.edu)

Wenjia Li (wli20@nyit.edu)

\***CORRESPONDING AUTHOR:**

Shenglong Zhang (szhang21@nyit.edu)



**Figure S1.** 2D-HELS MS sequencing of 5 mixed RNA strands simultaneously using a biotin tag to label the 3´-ends. Original tR was displayed without any normalization.

**Table S1.** LC-MS analysis of 3´-biotin-labeled RNA #1 after streptavidin-aided bead separation followed by subsequent chemical degradation (3´-labeled ladder components of RNA #1, referring to the top curve in Figure 1c).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 19 | 6781.0733 | 305.0413 | C | 6781.0413 | 9.752 | 16819442 | 100 | 4.72 |
| 18 | 6476.0320 | 345.0474 | G | 6475.9924 | 9.717 | 247965 | 84 | 6.11 |
| 17 | 6130.9846 | 305.0413 | C | 6130.9398 | 9.662 | 178841 | 80 | 7.31 |
| 16 | 5825.9433 | 329.0525 | A | 5825.9037 | 9.782 | 510096 | 80 | 6.80 |
| 15 | 5496.8908 | 306.0253 | U | 5496.8566 | 9.383 | 262486 | 99 | 6.22 |
| 14 | 5190.8655 | 305.0413 | C | 5190.8364 | 9.241 | 349988 | 100 | 5.61 |
| 13 | 4885.8242 | 306.0253 | U | 4885.7908 | 9.135 | 356118 | 100 | 6.84 |
| 12 | 4579.7989 | 345.0475 | G | 4579.7738 | 9.109 | 386687 | 100 | 5.48 |
| 11 | 4234.7514 | 329.0525 | A | 4234.7271 | 9.145 | 305380 | 100 | 5.74 |
| 10 | 3905.6989 | 305.0413 | C | 3905.6749 | 8.575 | 145505 | 96 | 6.14 |
| 9 | 3600.6576 | 306.0253 | U | 3600.6373 | 8.420 | 195308 | 100 | 5.64 |
| 8 | 3294.6323 | 345.0474 | G | 3294.6165 | 8.370 | 125991 | 100 | 4.80 |
| 7 | 2949.5849 | 329.0525 | A | 2949.5716 | 8.339 | 106993 | 100 | 4.51 |
| 6 | 2620.5324 | 305.0413 | C | 2620.5193 | 7.492 | 90629 | 100 | 5.00 |
| 5 | 2315.4911 | 305.0413 | C | 2315.4814 | 7.299 | 163692 | 100 | 4.19 |
| 4 | 2010.4498 | 329.0525 | A | 2010.4388 | 7.625 | 279963 | 100 | 5.47 |
| 3 | 1681.3973 | 329.0525 | A | 1681.3891 | 7.354 | 183827 | 100 | 4.88 |
| 2 | 1352.3448 | 329.0526 | A | 1352.3378 | 7.303 | 135065 | 100 | 5.18 |
| 1 | 1023.2922 | 329.0525 | A | 1023.2859 | 7.219 | 106700 | 100 | 6.16 |

Output sequence: CGCAUCUGACUGACCAAAA

**Table S2.** LC-MS analysis of 3´-biotin-labeled RNA #1 after streptavidin-aided bead separation followed by subsequent chemical degradation (5´-unlabeled ladder components of RNA #1, referring to the bottom curve in Figure 1c).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 19 | 6024.8778 | 249.0862 | A | 6024.8483 | 7.664 | 14325731 | 100 | 4.90 |
| 18 | 5775.7916 | 329.0525 | A | 5775.7522 | 7.701 | 457844 | 87 | 6.82 |
| 17 | 5446.7391 | 329.0525 | A | 5446.6965 | 7.411 | 417145 | 100 | 7.82 |
| 16 | 5117.6866 | 329.0525 | A | 5117.6572 | 7.105 | 490290 | 100 | 5.74 |
| 15 | 4788.6341 | 305.0413 | C | 4788.6060 | 6.685 | 728135 | 100 | 5.87 |
| 14 | 4483.5928 | 305.0413 | C | 4483.5657 | 6.428 | 481770 | 100 | 6.04 |
| 13 | 4178.5515 | 329.0525 | A | 4178.5286 | 6.183 | 297514 | 100 | 5.48 |
| 12 | 3849.4990 | 345.0475 | G | 3849.4787 | 5.653 | 518403 | 100 | 5.27 |
| 11 | 3504.4515 | 306.0253 | U | 3504.4331 | 5.238 | 614494 | 100 | 5.25 |
| 10 | 3198.4262 | 305.0413 | C | 3198.4106 | 4.785 | 524613 | 99 | 4.88 |
| 9 | 2893.3849 | 329.0525 | A | 2893.3714 | 4.341 | 373933 | 100 | 4.67 |
| 8 | 2564.3324 | 345.0474 | G | 2564.3219 | 3.458 | 509219 | 100 | 4.09 |
| 7 | 2219.2850 | 306.0253 | U | 2219.2752 | 2.840 | 579139 | 100 | 4.42 |
| 6 | 1913.2597 | 305.0413 | C | 1913.2521 | 2.081 | 466058 | 100 | 3.97 |
| 5 | 1608.2184 | 306.0253 | U | 1608.2123 | 1.375 | 372038 | 80 | 3.79 |
| 4 | 1302.1931 | 329.0525 | A | 1302.1878 | 0.925 | 240613 | 100 | 4.07 |
| 3 | 973.1406 | 305.0413 | C | 973.1367 | 0.765 | 208989 | 100 | 4.01 |
| 2 | 668.0993 | 345.0474 | G | 668.0955 | 0.652 | 26061 | 100 | 5.69 |
| 1 | 323.0519 | 305.0413 | C | NA\* | NA | NA | NA | NA |

\* NA: Not Analyzed. The 350 Da threshold was set to minimize background ions from the elution buffers. Thus, the masses which are smaller than 350 Da were not detected.

Output sequence: CGCAUCUGACUGACCAAAA

**Table S3.** LC-MS analysis of 5´-biotin-labeled RNA #1 (5´-labeled ladder components of RNA #1, referring to the bottom ladder curve in black in Figure 1d).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 19 | 6600.0415 | 249.0862 | A | 6600.0153 | 10.113 | 1468018 | 100 | 3.97 |
| 18 | 6350.9553 | 329.0525 | A | 6350.9006 | 10.094 | 139388 | 80 | 8.61 |
| 17 | 6021.9028 | 329.0525 | A | 6021.8665 | 9.957 | 152155 | 80 | 6.03 |
| 16 | 5692.8503 | 329.0525 | A | 5692.8225 | 9.806 | 122377 | 84 | 4.88 |
| 15 | 5363.7978 | 305.0413 | C | 5363.7567 | 9.594 | 255396 | 100 | 7.66 |
| 14 | 5058.7565 | 305.0413 | C | 5058.7320 | 9.508 | 169499 | 80 | 4.84 |
| 13 | 4753.7152 | 329.0525 | A | 4753.6944 | 9.449 | 121869 | 96 | 4.38 |
| 12 | 4424.6627 | 345.0475 | G | 4424.6389 | 9.204 | 222046 | 100 | 5.38 |
| 11 | 4079.6152 | 306.0253 | U | 4079.5902 | 9.067 | 296271 | 100 | 6.13 |
| 10 | 3773.5899 | 305.0413 | C | 3773.5679 | 8.937 | 249085 | 100 | 5.83 |
| 9 | 3468.5486 | 329.0525 | A | 3468.5308 | 8.838 | 185624 | 100 | 5.13 |
| 8 | 3139.4961 | 345.0474 | G | 3139.4834 | 8.507 | 319911 | 100 | 4.05 |
| 7 | 2794.4487 | 306.0253 | U | 2794.4360 | 8.288 | 380189 | 100 | 4.54 |
| 6 | 2488.4234 | 305.0413 | C | 2488.4134 | 8.073 | 317954 | 100 | 4.02 |
| 5 | 2183.3821 | 306.0253 | U | 2183.3725 | 7.863 | 305479 | 100 | 4.40 |
| 4 | 1877.3568 | 329.0525 | A | 1877.3489 | 7.642 | 222446 | 100 | 4.21 |
| 3 | 1548.3043 | 305.0413 | C | 1548.2982 | 7.088 | 361254 | 100 | 3.94 |
| 2 | 1243.2630 | 345.0474 | G | 1243.2575 | 6.798 | 162972 | 100 | 4.42 |
| 1 | 898.2156 | 305.0413 | C | 898.2105 | 6.880 | 88421 | 100 | 5.68 |

Output sequence: CGCAUCUGACUGACCAAAA

**Table S4.** LC-MS analysis of 5´-biotin-labeled RNA #2 (5´-labeled ladder components of RNA #2, referring to the top ladder curve in red in Figure 1d).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 20 | 6898.0505 | 225.0750 | C | 6898.0210 | 10.014 | 3995416 | 100 | 4.28 |
| 19 | 6672.9755 | 345.0474 | G | 6673.4755 | 10.115 | 92706 | 80 | -74.9 |
| 18 | 6327.9281 | 305.0413 | C | 6327.8894 | 10.117 | 108088 | 80 | 6.12 |
| 17 | 6022.8868 | 329.0525 | A | 6022.8313 | 10.104 | 133027 | 100 | 9.21 |
| 16 | 5693.8343 | 306.0253 | U | 5693.7870 | 9.920 | 68281 | 80 | 8.31 |
| 15 | 5387.8090 | 305.0413 | C | 5387.7785 | 9.850 | 167081 | 80 | 5.66 |
| 14 | 5082.7677 | 306.0253 | U | 5082.7314 | 9.784 | 170198 | 100 | 7.14 |
| 13 | 4776.7424 | 345.0474 | G | 4776.7210 | 9.695 | 114657 | 99 | 4.48 |
| 12 | 4431.6950 | 329.0526 | A | 4431.6685 | 9.629 | 143358 | 92 | 5.98 |
| 11 | 4102.6424 | 305.0412 | C | 4102.6199 | 9.367 | 245033 | 100 | 5.48 |
| 10 | 3797.6012 | 306.0253 | U | 3797.5819 | 9.264 | 184127 | 100 | 5.08 |
| 9 | 3491.5759 | 345.0475 | G | 3491.5567 | 9.131 | 91691 | 100 | 5.50 |
| 8 | 3146.5284 | 329.0525 | A | 3146.5054 | 9.028 | 187937 | 100 | 7.31 |
| 7 | 2817.4759 | 305.0413 | C | 2817.4633 | 8.675 | 288050 | 100 | 4.47 |
| 6 | 2512.4346 | 305.0413 | C | 2512.4233 | 8.509 | 138698 | 100 | 4.50 |
| 5 | 2207.3933 | 305.0413 | C | 2207.3835 | 8.335 | 192998 | 100 | 4.44 |
| 4 | 1902.3520 | 345.0474 | G | 1902.3433 | 8.161 | 149466 | 100 | 4.57 |
| 3 | 1557.3046 | 329.0525 | A | 1557.2976 | 8.042 | 133349 | 100 | 4.49 |
| 2 | 1228.2521 | 306.0253 | U | 1228.2455 | 7.618 | 188828 | 100 | 5.37 |
| 1 | 922.2268 | 329.0525 | A | 922.2213 | 7.434 | 86674 | 100 | 5.96 |

Output sequence: AUAGCCCAGUCAGUCUACGC

**Table S5.** LC-MS analysis of a 1 ψ-containing RNA #6 (ψ unconverted ladder components in the 5´ ladder of RNA #6, referring to the bottom ladder curve in black in Figure 2b).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 20 | 6345.9028 | 265.0811 | G | 6345.9217 | 11.736 | 41088112 | 100 | -2.98 |
| 19 | 6080.8217 | 329.0525 | A | 6080.8255 | 11.769 | 2582596 | 100 | -0.62 |
| 18 | 5751.7692 | 345.0474 | G | 5751.7749 | 11.496 | 2169051 | 100 | -0.99 |
| 17 | 5406.7218 | 306.0253 | U | 5406.7209 | 11.315 | 2126771 | 100 | 0.17 |
| 16 | 5100.6965 | 319.057 | m5C | 5100.6941 | 11.167 | 1149416 | 100 | 0.47 |
| 15 | 4781.6395 | 329.0525 | A | 4781.6402 | 10.970 | 2692877 | 100 | -0.15 |
| 14 | 4452.5870 | 306.0253 | U | 4452.5866 | 10.566 | 5448251 | 100 | 0.09 |
| 13 | 4146.5617 | 306.0253 | U | 4146.5603 | 10.343 | 4115258 | 100 | 0.34 |
| 12 | 3840.5364 | 329.0526 | A | 3840.5352 | 10.141 | 2038738 | 100 | 0.31 |
| 11 | 3511.4838 | 305.0413 | C | 3511.4836 | 9.610 | 1167942 | 100 | 0.06 |
| 10 | 3206.4425 | 305.0412 | C | 3206.4401 | 9.331 | 3422282 | 100 | 0.75 |
| 9 | 2901.4013 | 329.0526 | A | 2901.3988 | 9.067 | 2391922 | 100 | 0.86 |
| 8 | 2572.3487 | 306.0253 | Unconverted ψ | 2572.3468 | 8.328 | 4952174 | 100 | 0.74 |
| 7 | 2266.3234 | 306.0253 | U | 2266.3215 | 7.944 | 4534905 | 100 | 0.84 |
| 6 | 1960.2981 | 345.0474 | G | 1960.2956 | 7.360 | 3437270 | 100 | 1.28 |
| 5 | 1615.2507 | 305.0413 | C | 1615.2481 | 6.693 | 4151449 | 100 | 1.61 |
| 4 | 1310.2094 | 305.0413 | C | 1310.2062 | 5.915 | 1289241 | 87 | 2.44 |
| 3 | 1005.1681 | 329.0525 | A | 1005.1655 | 4.416 | 913589 | 100 | 2.59 |
| 2 | 676.1156 | 329.0525 | A | 676.1140 | 3.321 | 748977 | 100 | 2.37 |
| 1 | 347.0631 | 329.0525 | A | NA\* | NA | NA | NA | NA |

\* NA: Not Analyzed. The 350 Da threshold was set to minimize background ions from the elution buffers. Thus, the masses which are smaller than 350 Da were not detected.

Output sequence: AAACCGUψACCAUUAm5CUGAG

**Table S6.** LC-MS analysis of a 1 ψ-containing RNA #6 (ladder components with CMC-converted ψ in the 5´ ladder of RNA #6, referring to the top ladder curve in red in Figure 2b)

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 20 | 6597.1025 | 265.0811 | G | 6597.1125 | 13.985 | 60627484 | 100 | -1.52 |
| 19 | 6332.0214 | 329.0525 | A | 6332.0201 | 13.979 | 1541470 | 100 | 0.21 |
| 18 | 6002.9689 | 345.0474 | G | 6002.9756 | 13.816 | 2147847 | 89 | -1.12 |
| 17 | 5657.9215 | 306.0253 | U | 5657.9243 | 13.742 | 2608610 | 100 | -0.49 |
| 16 | 5351.8962 | 319.057 | m5C | 5351.8960 | 13.695 | 2110248 | 100 | 0.04 |
| 15 | 5032.8392 | 329.0525 | A | 5032.8400 | 13.633 | 1907945 | 100 | -0.16 |
| 14 | 4703.7867 | 306.0253 | U | 4703.7861 | 13.394 | 4110706 | 88 | 0.13 |
| 13 | 4397.7614 | 306.0253 | U | 4397.7599 | 13.320 | 2867370 | 100 | 0.34 |
| 12 | 4091.7361 | 329.0526 | A | 4091.7361 | 13.283 | 1855682 | 100 | 0.00 |
| 11 | 3762.6835 | 305.0413 | C | 3762.6830 | 12.962 | 2817838 | 100 | 0.13 |
| 10 | 3457.6422 | 305.0412 | C | 3457.6396 | 12.878 | 1149319 | 100 | 0.75 |
| 9 | 3152.6010 | 329.0526 | A | 3152.5974 | 12.934 | 746862 | 100 | 1.14 |
| 8 | 2823.5485 | 557.2251 | Converted ψ | 2823.5455 | 12.380 | 2149383 | 100 | 1.06 |
| 7 | 2266.3234 | 306.0253 | U | 2266.3213 | 7.944 | 4767282 | 100 | 0.93 |
| 6 | 1960.2981 | 345.0474 | G | 1960.2956 | 7.360 | 3433416 | 100 | 1.28 |
| 5 | 1615.2507 | 305.0413 | C | 1615.2481 | 6.694 | 4174772 | 100 | 1.61 |
| 4 | 1310.2094 | 305.0413 | C | 1310.2071 | 5.917 | 806139 | 87 | 1.76 |
| 3 | 1005.1681 | 329.0525 | A | 1005.1655 | 4.416 | 913589 | 100 | 2.59 |
| 2 | 676.1156 | 329.0525 | A | 676.1140 | 3.321 | 743305 | 100 | 2.37 |
| 1 | 347.0631 | 329.0525 | A | NA\* | NA | NA | NA | NA |

\* NA: Not Analyzed. The 350 Da threshold was set to minimize background ions from the elution buffers. Thus, the masses which are smaller than 350 Da were not detected.

Output sequence: AAACCGUψACCAUUAm5CUGAG

**Table S7.** LC-MS analysis of 3´-biotin-labeled RNA #1, showing its ladder components (referring to the ladder curve in black in Figure 3).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 19 | 6781.0733 | 305.0413 | C | 6781.0426 | 9.576 | 35286012 | 100 | 4.53 |
| 18 | 6476.0320 | 345.0474 | G | 6475.9985 | 9.535 | 23351 | 60 | 5.17 |
| 17 | 6130.9846 | 305.0413 | C | 6130.9933 | 9.473 | 50125 | 90 | -1.42 |
| 16 | 5825.9433 | 329.0525 | A | 5825.9244 | 9.634 | 55880 | 80 | 3.24 |
| 15 | 5496.8908 | 306.0253 | U | 5496.8590 | 9.218 | 633795 | 80 | 5.79 |
| 14 | 5190.8655 | 305.0413 | C | 5190.8470 | 9.078 | 849742 | 100 | 3.56 |
| 13 | 4885.8242 | 306.0253 | U | 4885.7976 | 8.976 | 1193120 | 100 | 5.44 |
| 12 | 4579.7989 | 345.0475 | G | 4579.7742 | 8.951 | 1191558 | 100 | 5.39 |
| 11 | 4234.7514 | 329.0525 | A | 4234.7340 | 8.989 | 1196633 | 100 | 4.11 |
| 10 | 3905.6989 | 305.0413 | C | 3905.6808 | 8.420 | 729180 | 100 | 4.63 |
| 9 | 3600.6576 | 306.0253 | U | 3600.6382 | 8.275 | 605689 | 100 | 5.39 |
| 8 | 3294.6323 | 345.0474 | G | 3294.6179 | 8.229 | 935654 | 100 | 4.37 |
| 7 | 2949.5849 | 329.0525 | A | 2949.5713 | 8.210 | 903559 | 100 | 4.61 |
| 6 | 2620.5324 | 305.0413 | C | 2620.5217 | 7.376 | 587699 | 100 | 4.08 |
| 5 | 2315.4911 | 305.0413 | C | 2315.4825 | 7.191 | 700118 | 100 | 3.71 |
| 4 | 2010.4498 | 329.0525 | A | 2010.4378 | 7.527 | 1052796 | 100 | 5.97 |
| 3 | 1681.3973 | 329.0525 | A | 1681.3901 | 7.273 | 714971 | 100 | 4.28 |
| 2 | 1352.3448 | 329.0526 | A | 1352.3387 | 7.230 | 447072 | 100 | 4.51 |
| 1 | 1023.2922 | 329.0525 | A | 1023.2881 | 7.148 | 736463 | 100 | 4.01 |

Output sequence: CGCAUCUGACUGACCAAAA

**Table S8.** LC-MS analysis of 3´-biotin-labeled RNA #2, showing its ladder components (referring to the ladder curve in red in Figure 3).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 20 | 7079.0823 | 329.2088 | A | 7079.0513 | 9.529 | 34343980 | 100 | 4.38 |
| 19 | 6750.0298 | 306.1667 | U | 6749.9875 | 9.259 | 170073 | 78 | 6.27 |
| 18 | 6444.0045 | 329.2088 | A | 6443.9653 | 9.344 | 934361 | 97 | 6.08 |
| 17 | 6114.9519 | 345.2077 | G | 6114.9082 | 9.000 | 176482 | 94 | 7.15 |
| 16 | 5769.9045 | 305.1828 | C | 5769.8590 | 8.867 | 537259 | 80 | 7.89 |
| 15 | 5464.8632 | 305.1828 | C | 5464.8338 | 8.733 | 381043 | 100 | 5.38 |
| 14 | 5159.8219 | 305.1827 | C | 5159.7998 | 8.619 | 939572 | 99 | 4.28 |
| 13 | 4854.7806 | 329.2088 | A | 4854.7556 | 8.734 | 1104050 | 100 | 5.15 |
| 12 | 4525.7281 | 345.2078 | G | 4525.7027 | 8.273 | 799528 | 100 | 5.61 |
| 11 | 4180.6807 | 306.1667 | U | 4180.6575 | 8.047 | 727253 | 100 | 5.55 |
| 10 | 3874.6554 | 305.1828 | C | 3874.6361 | 7.836 | 1007297 | 100 | 4.98 |
| 9 | 3569.6141 | 329.2087 | A | 3569.5985 | 7.960 | 1323892 | 100 | 4.37 |
| 8 | 3240.5616 | 345.2078 | G | 3240.5458 | 7.328 | 854305 | 100 | 4.88 |
| 7 | 2895.5141 | 306.1668 | U | 2895.5009 | 6.991 | 838944 | 100 | 4.56 |
| 6 | 2589.4888 | 305.1827 | C | 2589.4785 | 6.639 | 1076014 | 100 | 3.98 |
| 5 | 2284.4476 | 306.1668 | U | 2284.4388 | 6.433 | 1085561 | 100 | 3.85 |
| 4 | 1978.4223 | 329.2088 | A | 1978.4152 | 6.298 | 1224106 | 100 | 3.59 |
| 3 | 1649.3697 | 305.1827 | C | 1649.3632 | 5.150 | 443067 | 100 | 3.94 |
| 2 | 1344.3284 | 345.2078 | G | 1344.3229 | 5.115 | 530069 | 100 | 4.09 |
| 1 | 999.2810 | 305.1827 | C | 999.2764 | 5.258 | 300175 | 100 | 4.60 |

Output sequence: AUAGCCCAGUCAGUCUACGC

**Table S9.** LC-MS analysis of 3´-biotin-labeled RNA #3, showing its ladder components (referring to the ladder curve in green in Figure 3).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 20 | 7088.0826 | 329.0525 | A | 7088.0479 | 9.902 | 18422776 | 100 | 4.90 |
| 19 | 6759.0301 | 329.0525 | A | 6758.9878 | 9.816 | 342458 | 82 | 6.26 |
| 18 | 6429.9776 | 329.0525 | A | 6429.9401 | 9.553 | 297978 | 100 | 5.83 |
| 17 | 6100.9251 | 305.0413 | C | 6100.8860 | 9.162 | 176200 | 80 | 6.41 |
| 16 | 5795.8838 | 305.0413 | C | 5795.8502 | 9.059 | 325811 | 100 | 5.80 |
| 15 | 5490.8425 | 345.0475 | G | 5490.8084 | 9.029 | 561379 | 99 | 6.21 |
| 14 | 5145.7950 | 306.0253 | U | 5145.7640 | 8.927 | 543764 | 100 | 6.02 |
| 13 | 4839.7697 | 306.0253 | U | 4839.7382 | 8.852 | 751511 | 100 | 6.51 |
| 12 | 4533.7444 | 329.0525 | A | 4533.7170 | 8.857 | 916467 | 100 | 6.04 |
| 11 | 4204.6919 | 305.0413 | C | 4204.6726 | 8.273 | 363029 | 100 | 4.59 |
| 10 | 3899.6506 | 305.0413 | C | 3899.6323 | 8.164 | 664338 | 100 | 4.69 |
| 9 | 3594.6093 | 329.0525 | A | 3594.5912 | 8.300 | 1247513 | 100 | 5.04 |
| 8 | 3265.5568 | 306.0253 | U | 3265.5400 | 7.653 | 597972 | 100 | 5.14 |
| 7 | 2959.5315 | 306.0253 | U | 2959.5186 | 7.464 | 985122 | 100 | 4.36 |
| 6 | 2653.5062 | 329.0525 | A | 2653.4963 | 7.431 | 1500526 | 100 | 3.73 |
| 5 | 2324.4537 | 305.0413 | C | 2324.4444 | 6.486 | 663475 | 100 | 4.00 |
| 4 | 2019.4124 | 306.0253 | U | 2019.4039 | 6.101 | 752760 | 100 | 4.21 |
| 3 | 1713.3871 | 345.0474 | G | 1713.3811 | 5.973 | 1299628 | 100 | 3.50 |
| 2 | 1368.3397 | 329.0525 | A | 1368.3335 | 6.144 | 379728 | 100 | 4.53 |
| 1 | 1039.2872 | 345.0474 | G | 1039.2820 | 5.644 | 273139 | 100 | 5.00 |

Output sequence: AAACCGUUACCAUUACUGAG

**Table S10.** LC-MS analysis of 3´-biotin-labeled RNA #4, showing its ladder components (referring to the ladder curve in pink in Figure 3).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 20 | 6954.9836 | 345.0475 | G | 6954.9478 | 9.243 | 16978916 | 100 | 5.15 |
| 19 | 6609.9361 | 305.0412 | C | 6609.8899 | 9.131 | 184784 | 80 | 6.99 |
| 18 | 6304.8949 | 345.0475 | G | 6304.8568 | 9.109 | 510790 | 80 | 6.04 |
| 17 | 5959.8474 | 306.0253 | U | 5959.7956 | 9.056 | 393186 | 90 | 8.69 |
| 16 | 5653.8221 | 329.0525 | A | 5653.7838 | 9.059 | 830821 | 100 | 6.77 |
| 15 | 5324.7696 | 305.0413 | C | 5324.7319 | 8.701 | 496925 | 98 | 7.08 |
| 14 | 5019.7283 | 329.0525 | A | 5019.6982 | 8.848 | 1059427 | 100 | 6.00 |
| 13 | 4690.6758 | 306.0253 | U | 4690.6470 | 8.345 | 581020 | 82 | 6.14 |
| 12 | 4384.6505 | 305.0413 | C | 4384.6245 | 8.185 | 852527 | 100 | 5.93 |
| 11 | 4079.6092 | 306.0253 | U | 4079.5872 | 8.071 | 872930 | 100 | 5.39 |
| 10 | 3773.5839 | 306.0253 | U | 3773.5632 | 7.884 | 880358 | 100 | 5.49 |
| 9 | 3467.5586 | 305.0413 | C | 3467.5339 | 7.639 | 168485 | 97 | 7.12 |
| 8 | 3162.5173 | 305.0413 | C | 3162.4881 | 7.411 | 503294 | 100 | 9.23 |
| 7 | 2857.4760 | 305.0413 | C | 2857.4625 | 7.156 | 851140 | 100 | 4.72 |
| 6 | 2552.4347 | 305.0412 | C | 2552.4231 | 6.920 | 1065610 | 100 | 4.54 |
| 5 | 2247.3935 | 306.0253 | U | 2247.3838 | 6.690 | 1189236 | 100 | 4.32 |
| 4 | 1941.3682 | 306.0253 | U | 1941.3605 | 6.350 | 1445336 | 100 | 3.97 |
| 3 | 1635.3429 | 306.0254 | U | 1635.3384 | 6.009 | 22256 | 85 | 2.75 |
| 2 | 1329.3175 | 329.0525 | A | 1329.3120 | 6.598 | 1296266 | 100 | 4.14 |
| 1 | 1000.2650 | 306.0253 | U | 1000.2606 | 5.604 | 422194 | 100 | 4.40 |

Output sequence: GCGUACAUCUUCCCCUUUAU

**Table S11.** LC-MS analysis of 3´-biotin-labeled RNA #5, showing its ladder components (referring to the ladder curve in light blue in Figure 3).

|  |  |  |
| --- | --- | --- |
| **Theoretical** | **Extracted data file after LC/MS analysis** | **Error** |
| **Fragments** | **Theoretical mass** | **Base mass** | **Base** | **MFE mass** | **tR** | **Volume** | **Quality Score** | **ppm** |
| 21 | 7522.1050 | 345.0475 | G | 7522.0681 | 9.519 | 21361914 | 100 | 4.91 |
| 20 | 7177.0575 | 305.0413 | C | 7176.9933 | 9.405 | 68800 | 60 | 8.95 |
| 19 | 6872.0162 | 345.0474 | G | 6871.9775 | 9.363 | 252280 | 88 | 5.63 |
| 18 | 6526.9688 | 345.0474 | G | 6526.9161 | 9.345 | 403291 | 100 | 8.07 |
| 17 | 6181.9214 | 329.0526 | A | 6181.8847 | 9.425 | 1246921 | 100 | 5.94 |
| 16 | 5852.8688 | 306.0253 | U | 5852.8226 | 9.054 | 263228 | 92 | 7.89 |
| 15 | 5546.8435 | 306.0253 | U | 5546.8116 | 8.935 | 1204009 | 100 | 5.75 |
| 14 | 5240.8182 | 306.0253 | U | 5240.7914 | 8.839 | 944494 | 100 | 5.11 |
| 13 | 4934.7929 | 329.0525 | A | 4934.7693 | 8.917 | 796848 | 100 | 4.78 |
| 12 | 4605.7404 | 345.0474 | G | 4605.7119 | 8.465 | 673185 | 100 | 6.19 |
| 11 | 4260.6930 | 305.0413 | C | 4260.6681 | 8.290 | 729523 | 100 | 5.84 |
| 10 | 3955.6517 | 306.0253 | U | 3955.6308 | 8.107 | 803678 | 100 | 5.28 |
| 9 | 3649.6264 | 305.0413 | C | 3649.6084 | 7.894 | 1056834 | 100 | 4.93 |
| 8 | 3344.5851 | 329.0525 | A | 3344.5687 | 7.990 | 1336987 | 100 | 4.90 |
| 7 | 3015.5326 | 345.0474 | G | 3015.5131 | 7.343 | 882742 | 100 | 6.47 |
| 6 | 2670.4852 | 306.0253 | U | 2670.4731 | 6.959 | 659989 | 100 | 4.53 |
| 5 | 2364.4599 | 306.0253 | U | 2364.4502 | 6.560 | 845446 | 100 | 4.10 |
| 4 | 2058.4346 | 345.0475 | G | 2058.4278 | 6.256 | 752026 | 100 | 3.30 |
| 3 | 1713.3871 | 345.0474 | G | 1713.3811 | 5.973 | 1299628 | 100 | 3.50 |
| 2 | 1368.3397 | 345.0475 | G | 1368.3335 | 6.144 | 379728 | 100 | 4.53 |
| 1 | 1023.2922 | 329.0525 | A | 1023.2881 | 7.148 | 736463 | 100 | 4.01 |

Output sequence: GCGGAUUUAGCUCAGUUGGGA