Supplementary Figure 1

Localization of crosslinks detected for the RRM1-EMCVsupE complex.

His62 was detected as the sole crosslinking site and is shown in red on the structure of RRM1 in complex with CU-RNAsup1. The detection of U- and UU-modifications suggests binding of RRM1 to nucleotides 324-326.
Supplementary Figure 2

Segmental labeling of EMCV\textsuperscript{DElinkF}-RNA.

a, strategy for segmental labeling as employed in our study. Notably, only one aliquot of isotope labeled RNA is used independent of the size of the RNA and the number of isotope labeled segments. b, RNase H digestion of EMCV\textsuperscript{DElinkF} with chimSLD (lane 1), chimLinkF (lane 2), chimSLE (lane 3) and all three chimeras simultaneously (lane 5). The full-length RNA is shown in lane 4, an example for splinted RNA-ligation is shown in lane 6 (input) and 7 (ligation). Note that the isotope labeled RNA fragment is kept as limiting substrate in the reaction.
Supplementary Figure 3

NMR titration experiments for RRM1 (top) and RRM2 (bottom).

(a) EMCV\textsuperscript{E} and (b) EMCV\textsuperscript{F}. The chemical shift perturbations between \textsuperscript{1}H\textsuperscript{15}N-signals of the free and the bound RRM\textsubscript{1}s indicate binding. The signals of the UCUUU-pentaloop of EMCV\textsuperscript{F} (right) shift upon addition of RRM\textsubscript{1} (top) and RRM\textsubscript{2} (bottom) due to binding.
Supplementary Figure 4

RRM2 binds preferentially to EMCV^F.

a, Upon titration of RRM12 with EMCV^F only the RRM2 moiety reaches saturation at a 1:1 ratio (EMCV^F:RRM12) and signals overlap almost perfectly with signals of the RRM2-EMCV^F complex (yellow). b, In the presence of SLE, RRM1 and RRM2 bind SLE and SLF, respectively, as indicated by the almost perfect superimposition of spectra collected on RRM1-EMCV^E (green), RRM2-EMCV^F (yellow) and RRM12-EMCV_{EmuF} (blue; right) complexes.
Comparison of fingerprint spectra of subcomplexes with their corresponding signals in the PTBP1-EMCV\textsuperscript{DElinkF} complex.

\textbf{a}, the $^1$H-$^{15}$N-TROSY signals and, \textbf{b}, the ILV-methyl-$^1$H-$^{13}$C-signals of RRM1 (green), RRM2 (yellow) and RRM34 (magenta) in complex with EMCV\textsuperscript{E}, EMCV\textsuperscript{F} and EMCV\textsuperscript{DElink}, respectively, overlap well with signals of the PTBP1-EMCV\textsuperscript{DElinkF} complex (black) which is typical for identical binding of the RRM5s in the subcomplexes.
Supplementary Figure 6

NMR titration of RRM34-EMCV<sub>DElink</sub> with RRM1.

<sup>1</sup>H<sup>15</sup>N resonances of RRM34-EMCV<sub>DElink</sub> are not affected by RRM1 addition (color code as indicated). The RRM1-moiety shows the signature of the RRM1-EMCV<sub>E</sub> complex.
Cyana-based modeling strategy using CLIR-MS/MS-derived restraints.

a, peptides and modification sites are identified by CLIR-MS/MS; the mass of the modification infers on its composition as exemplarily shown for RRM3 crosslinks detected in “D”. b, CLIR-MS/MS data are converted to ambiguous upper proton-proton distance limits (illustrated for simplicity as red, dashed lines between carbon atoms) and, c, fed to Cyana 3.0\(^2\) noeassign. Relevant structural information from published RNA-protein complexes and structure predictions serve to generate complementary ambiguous and unambiguous distance restraints. Calculated atomic resolution models were further refined under amber 12 force-field\(^3\). d, representation of the final CLIR-MS/MS restraints on the model of RRM3 in complex with nucleotides 299-306 of SLD.
Ensembles of the RRM-RNA complexes modeled using CLIR-MS/MS-derived distance restraints.

Shown are the bundles of the 20 final structures of the, a, RRM1-EMCV\textsuperscript{E}, b, RRM2-EMCV\textsuperscript{F}, c, RRM3-EMCV\textsuperscript{D} and, d, RRM4-EMCV\textsuperscript{link} complex. Only nucleotides 354-364, 300-305 and 340-344 are displayed in A, C and D, respectively, as these represent the well-defined binding site. RMSD values (Deviation of coordinates to their mean) are calculated for amino acids 47-153 (RRM1), 181-244 and 255-273 (RRM2) and 336-531 (RRM34 tandem domain) and nucleotides 321-336 (EMCV\textsuperscript{E}), 354-364 (EMCV\textsuperscript{F}), 302-305 (EMCV\textsuperscript{D}) and 341-344 (EMCV\textsuperscript{link}).
Supplementary Figure 9

Recognition of the EMCV-IRES domains D–F by PTBP1-RRMs according to CLIR-MS/MS-based models.

Hydrophobic and specific hydrogen bond interactions are depicted for each binding pocket of the RRMs as derived from the CLIR-MS/MS-based models and compared with the published interactions detected in an RRM1-stemloop complex (pdb: 3N2O), the NMR derived model of RRM2-EMCV and in the complex of RRM34 with ssRNA. Recognition of nucleotides in the reference structures is shown in grey if different. Please note that U360 is looped out and not recognized by RRM2.
Supplementary Figure 10

Comparison of the RRM2-EMCV\textsuperscript{F} models derived from NMR and CLIR-MS/MS.

RRM2 recognizes EMCV\textsuperscript{F} by contacting C358 and U359 in pocket 1 and 2 by base-specific hydrogen bonds in both, a, NMR and, b, CLIR-MS/MS derived models. While U360 does not contact RRM2, U361 stacks on Tyr 267.
**Supplementary Table 1: Calculated mass-shifts induced by crosslinking of RNA moieties**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>324.035867</td>
<td>335.06013</td>
<td>306.025302</td>
<td>317.049566</td>
<td>11.024263</td>
</tr>
<tr>
<td>C</td>
<td>323.051851</td>
<td>335.07315</td>
<td>305.041287</td>
<td>317.062585</td>
<td>12.021298</td>
</tr>
<tr>
<td>G</td>
<td>363.07999</td>
<td>378.076722</td>
<td>345.07435</td>
<td>360.066158</td>
<td>15.0172822</td>
</tr>
<tr>
<td>A</td>
<td>347.063085</td>
<td>362.081808</td>
<td>329.05252</td>
<td>344.071243</td>
<td>15.01872282</td>
</tr>
<tr>
<td>UU</td>
<td>630.061169</td>
<td>652.109696</td>
<td>612.050605</td>
<td>634.099131</td>
<td>22.04852661</td>
</tr>
<tr>
<td>CU</td>
<td>629.077154</td>
<td>652.122715</td>
<td>611.06689</td>
<td>634.112151</td>
<td>23.045615</td>
</tr>
<tr>
<td>GU</td>
<td>669.083802</td>
<td>695.126288</td>
<td>651.07237</td>
<td>677.115723</td>
<td>26.04298613</td>
</tr>
<tr>
<td>AU</td>
<td>653.088387</td>
<td>679.131373</td>
<td>635.077822</td>
<td>661.120809</td>
<td>26.04298613</td>
</tr>
<tr>
<td>CC</td>
<td>628.093138</td>
<td>652.135735</td>
<td>610.08257</td>
<td>634.12517</td>
<td>24.0425964</td>
</tr>
<tr>
<td>GC</td>
<td>668.099286</td>
<td>695.139307</td>
<td>650.08872</td>
<td>677.128743</td>
<td>27.04002102</td>
</tr>
<tr>
<td>AC</td>
<td>652.104372</td>
<td>679.144393</td>
<td>634.093807</td>
<td>661.133828</td>
<td>27.04002102</td>
</tr>
<tr>
<td>GG</td>
<td>708.105434</td>
<td>738.14288</td>
<td>690.09486</td>
<td>720.132315</td>
<td>30.0374565</td>
</tr>
<tr>
<td>AG</td>
<td>692.11052</td>
<td>722.147965</td>
<td>674.09995</td>
<td>704.137401</td>
<td>30.0374565</td>
</tr>
<tr>
<td>AA</td>
<td>676.115605</td>
<td>706.153051</td>
<td>658.10504</td>
<td>688.142486</td>
<td>30.0374565</td>
</tr>
<tr>
<td>UUU</td>
<td>936.086472</td>
<td>969.159262</td>
<td>918.075907</td>
<td>951.148697</td>
<td>33.07278991</td>
</tr>
<tr>
<td>CUU</td>
<td>935.102456</td>
<td>969.172281</td>
<td>917.09189</td>
<td>951.161716</td>
<td>34.0692481</td>
</tr>
<tr>
<td>GUU</td>
<td>975.108604</td>
<td>1012.17585</td>
<td>957.09803</td>
<td>994.165289</td>
<td>37.06724943</td>
</tr>
<tr>
<td>AUU</td>
<td>959.113689</td>
<td>996.180399</td>
<td>941.103125</td>
<td>978.170374</td>
<td>37.06724943</td>
</tr>
<tr>
<td>CCE</td>
<td>934.11844</td>
<td>969.1853</td>
<td>916.10786</td>
<td>951.17435</td>
<td>35.0668597</td>
</tr>
<tr>
<td>GCU</td>
<td>974.124588</td>
<td>1012.18887</td>
<td>956.114024</td>
<td>994.178308</td>
<td>38.06428433</td>
</tr>
<tr>
<td>ACU</td>
<td>958.129674</td>
<td>996.193958</td>
<td>940.119109</td>
<td>978.183394</td>
<td>38.06428433</td>
</tr>
<tr>
<td>GGU</td>
<td>1014.13074</td>
<td>1055.19245</td>
<td>996.120172</td>
<td>1037.18188</td>
<td>41.06170895</td>
</tr>
<tr>
<td>AGU</td>
<td>998.135822</td>
<td>1039.19753</td>
<td>980.125257</td>
<td>1021.18697</td>
<td>41.06170895</td>
</tr>
<tr>
<td>AUA</td>
<td>982.140907</td>
<td>1023.20262</td>
<td>964.130343</td>
<td>1005.19205</td>
<td>41.06170895</td>
</tr>
<tr>
<td>CCC</td>
<td>933.134425</td>
<td>969.198319</td>
<td>969.198319</td>
<td>991.187755</td>
<td>36.0638496</td>
</tr>
<tr>
<td>GCC</td>
<td>973.140573</td>
<td>1012.20189</td>
<td>1012.20189</td>
<td>994.191327</td>
<td>39.06313922</td>
</tr>
<tr>
<td>ACC</td>
<td>957.145668</td>
<td>996.206978</td>
<td>996.206978</td>
<td>1017.196413</td>
<td>39.06313922</td>
</tr>
<tr>
<td>GGC</td>
<td>1013.14672</td>
<td>1055.20546</td>
<td>1055.20546</td>
<td>1057.1949</td>
<td>42.05874384</td>
</tr>
<tr>
<td>AGC</td>
<td>997.151806</td>
<td>1039.21055</td>
<td>1039.21055</td>
<td>1051.19999</td>
<td>42.05874384</td>
</tr>
<tr>
<td>AAC</td>
<td>981.156892</td>
<td>1023.21564</td>
<td>1023.21564</td>
<td>1055.20507</td>
<td>42.05874384</td>
</tr>
</tbody>
</table>
**Supplementary Table 2: Localization of crosslinks**

<table>
<thead>
<tr>
<th>Complex</th>
<th>Isotope labelling scheme</th>
<th>Domain</th>
<th>Peptide</th>
<th>Position of modification</th>
<th>Detected mono-, di- and trinucleotide modifications</th>
<th>Spectral counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRM1:EMCV</td>
<td>uniformly</td>
<td>RRM1</td>
<td>VIHIR</td>
<td>H62</td>
<td>U, UU</td>
<td>8</td>
</tr>
<tr>
<td>PTBP1:EMCV</td>
<td>uniformly</td>
<td>RRM1</td>
<td>VIHIR</td>
<td>H62</td>
<td>UU</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GQPIYIQFSNHK</td>
<td></td>
<td>Y127</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IITFTK</td>
<td></td>
<td>I214, T217</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(LTSLNVK)YNNDK</td>
<td></td>
<td>L263, V265, K266, Y267</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(VK)ILFNKK</td>
<td></td>
<td>L370, F371, N372, K373</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HQNVQLPR</td>
<td></td>
<td>H411, Q412, N413, V414</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DYGNSPLHR</td>
<td></td>
<td>Y430, N432, P434, L435</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GFKFFQK</td>
<td></td>
<td>F487</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FFQKDR</td>
<td></td>
<td>K489</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VSFSKSTI</td>
<td></td>
<td>K528</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(ITLSK)HQNVQLPR</td>
<td></td>
<td>T407, L408, S409, H411, Q412, N413, V414</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GQPYIQFSNHK</td>
<td></td>
<td>P125, Y127, Q129</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(ITLSHGQNVLPR</td>
<td></td>
<td>L408, S409, H411, Q412</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GQPYIQFSNHK</td>
<td></td>
<td>I126, Y127, I128</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(ITLSKGQDQGLTK)</td>
<td></td>
<td>K410, H411, N413, V414, P417</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HQNVQPPREG</td>
<td></td>
<td>P417</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NFQNIIPPSATLHS</td>
<td></td>
<td>G442, Q447, L456, H457, S459, N460</td>
<td></td>
<td>33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NIPPSVSEEDLK</td>
<td></td>
<td>GQPYIQFSNHK</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GQPYIQFSNHK</td>
<td></td>
<td>Y127</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LTLNHYNYNDK</td>
<td></td>
<td>S262, N264, V265, K266, Y267</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FTLSKGQDQGLTPR</td>
<td></td>
<td>H411</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

* Longer sequence due to partial trypsin cleavage
### Supplementary Table 3: CLIR-MS/MS derived intermolecular restraints as used for 3D modeling

#### RRM1-EMCV<sup>U</sup>

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>227</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>228</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>229</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>231</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>233</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>235</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>237</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>239</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>227</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>228</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>229</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>231</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>233</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>235</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>237</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>126</td>
<td>ILE</td>
<td>QD1</td>
<td>239</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

#### UU-modifications

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>227</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>228</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>127</td>
<td>TYR</td>
<td>QE</td>
<td>229</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

#### AU-modifications

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>ILE</td>
<td>QD1</td>
<td>233</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>128</td>
<td>ILE</td>
<td>QD1</td>
<td>235</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>128</td>
<td>ILE</td>
<td>QD1</td>
<td>237</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>128</td>
<td>ILE</td>
<td>QD1</td>
<td>236</td>
<td>RADE</td>
<td>H8</td>
<td>0</td>
</tr>
<tr>
<td>128</td>
<td>ILE</td>
<td>QD1</td>
<td>234</td>
<td>RADE</td>
<td>H8</td>
<td>0</td>
</tr>
</tbody>
</table>

#### GU-modifications

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>89</td>
<td>LEU</td>
<td>QQD</td>
<td>226</td>
<td>RGUA</td>
<td>H8</td>
<td>4</td>
</tr>
<tr>
<td>89</td>
<td>LEU</td>
<td>QQD</td>
<td>227</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>89</td>
<td>LEU</td>
<td>QQD</td>
<td>229</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>89</td>
<td>LEU</td>
<td>QQD</td>
<td>230</td>
<td>RGUA</td>
<td>H8</td>
<td>0</td>
</tr>
<tr>
<td>89</td>
<td>LEU</td>
<td>QQD</td>
<td>231</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

<p>| comment: TIC &lt; 0.1, manually controlled |</p>
<table>
<thead>
<tr>
<th>Res.-no.</th>
<th>Res.-name</th>
<th>Atom name</th>
<th>Res. no.</th>
<th>Res.-name</th>
<th>Atom name</th>
<th>Distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Res.-no.</th>
<th>Res.-name</th>
<th>Atom name</th>
<th>Res. no.</th>
<th>Res.-name</th>
<th>Atom name</th>
<th>Distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>269</td>
<td>ASN</td>
<td>QB</td>
<td>342</td>
<td>URA</td>
<td>H5</td>
<td>3.3</td>
</tr>
<tr>
<td>269</td>
<td>ASN</td>
<td>QB</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>269</td>
<td>ASN</td>
<td>QB</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>269</td>
<td>ASN</td>
<td>QB</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

Comment: TIC < 0.1, manually controlled
<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>342</td>
<td>URA</td>
<td>H5</td>
<td>3.3</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>343</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

**GU-modifications**

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>339</td>
<td>RGUA</td>
<td>H1</td>
<td>3.3</td>
</tr>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>340</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>265</td>
<td>VAL</td>
<td>QQG</td>
<td>346</td>
<td>RGUA</td>
<td>H1</td>
<td>0</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>339</td>
<td>RGUA</td>
<td>H1</td>
<td>3.3</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>340</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>267</td>
<td>TYR</td>
<td>QE</td>
<td>346</td>
<td>RGUA</td>
<td>H1</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>339</td>
<td>RGUA</td>
<td>H1</td>
<td>3.3</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>340</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>ASN</td>
<td>QB</td>
<td>346</td>
<td>RGUA</td>
<td>H1</td>
<td>0</td>
</tr>
</tbody>
</table>

**GUU-modifications**

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>344</td>
<td>URA</td>
<td>H5</td>
<td>3.3</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>345</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>266</td>
<td>LYS</td>
<td>QZ</td>
<td>346</td>
<td>RGUA</td>
<td>H1</td>
<td>0</td>
</tr>
</tbody>
</table>

**comment: TIC < 0.1, manually controlled**

**RRM3-EMCV\^o**

**U-modifications**

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>413</td>
<td>ASN</td>
<td>QB</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>6</td>
</tr>
<tr>
<td>413</td>
<td>ASN</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>409</td>
<td>SER</td>
<td>QB</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>6</td>
</tr>
<tr>
<td>409</td>
<td>SER</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>407</td>
<td>THR</td>
<td>HB</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>407</td>
<td>THR</td>
<td>HB</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>411</td>
<td>HIS</td>
<td>HE1</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>411</td>
<td>HIS</td>
<td>HE1</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>408</td>
<td>LEU</td>
<td>QB</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>408</td>
<td>LEU</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
</tbody>
</table>

**UU-modifications**

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>411</td>
<td>HIS</td>
<td>HE1</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>4</td>
</tr>
<tr>
<td>411</td>
<td>HIS</td>
<td>HD2</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>413</td>
<td>ASN</td>
<td>QB</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>6</td>
</tr>
<tr>
<td>413</td>
<td>ASN</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>412</td>
<td>GLN</td>
<td>QG</td>
<td>596</td>
<td>URA</td>
<td>H5</td>
<td>6</td>
</tr>
<tr>
<td>412</td>
<td>GLN</td>
<td>QG</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td>0</td>
</tr>
<tr>
<td>GU-modifications</td>
<td>res.-no.</td>
<td>res.-name</td>
<td>atom name</td>
<td>res. no.</td>
<td>res.-name</td>
<td>atom name</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>415</td>
<td>GLN</td>
<td>QG</td>
<td>597</td>
<td>URA</td>
<td>H5</td>
<td></td>
</tr>
<tr>
<td>415</td>
<td>GLN</td>
<td>QG</td>
<td>598</td>
<td>RGUA</td>
<td>H8</td>
<td></td>
</tr>
<tr>
<td>415</td>
<td>GLN</td>
<td>QG</td>
<td>598</td>
<td>RGUA</td>
<td>C5</td>
<td></td>
</tr>
</tbody>
</table>

Comment: TIC < 0.1, manually controlled

<table>
<thead>
<tr>
<th>RRM4-EMCView</th>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-modifications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>592</td>
<td>URA</td>
<td>H6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>594</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>595</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>596</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C-modifications</th>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>598</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>599</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>601</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>598</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>599</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>601</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>460</td>
<td>ASN</td>
<td>QB</td>
<td>598</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>460</td>
<td>ASN</td>
<td>QB</td>
<td>599</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>460</td>
<td>ASN</td>
<td>QB</td>
<td>601</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UUU-modifications</th>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>594</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>595</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>596</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>456</td>
<td>LEU</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>594</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>595</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>596</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CU-modifications</th>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>457</td>
<td>HIS</td>
<td>HE1</td>
<td>598</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>459</td>
<td>SER</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>459</td>
<td>SER</td>
<td>QB</td>
<td>598</td>
<td>RCYT</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UUUU-modifications</th>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>442</td>
<td>GLY</td>
<td>QA</td>
<td>594</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>442</td>
<td>GLY</td>
<td>QA</td>
<td>595</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>442</td>
<td>GLY</td>
<td>QA</td>
<td>596</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>442</td>
<td>GLY</td>
<td>QA</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>res.-no.</td>
<td>res.-name</td>
<td>atom name</td>
<td>res. no.</td>
<td>res.-name</td>
<td>atom name</td>
<td>distance* (Å)</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>442</td>
<td>GLY</td>
<td>QA</td>
<td>596</td>
<td>URA</td>
<td>H6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>442</td>
<td>GLY</td>
<td>QA</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>461</td>
<td>ILE</td>
<td>QB</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

CUU-modifications

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>447</td>
<td>GLN</td>
<td>QA</td>
<td>594</td>
<td>URA</td>
<td>H6</td>
<td>4</td>
</tr>
<tr>
<td>447</td>
<td>GLN</td>
<td>QA</td>
<td>595</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
<tr>
<td>447</td>
<td>GLN</td>
<td>QA</td>
<td>596</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
<tr>
<td>447</td>
<td>GLN</td>
<td>QA</td>
<td>597</td>
<td>URA</td>
<td>H6</td>
<td>0</td>
</tr>
</tbody>
</table>

CCU-modifications

<table>
<thead>
<tr>
<th>res.-no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>res. no.</th>
<th>res.-name</th>
<th>atom name</th>
<th>distance* (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>461</td>
<td>ILE</td>
<td>QB</td>
<td>598</td>
<td>RCYT</td>
<td>H6</td>
<td>0</td>
</tr>
<tr>
<td>461</td>
<td>ILE</td>
<td>QB</td>
<td>599</td>
<td>RCYT</td>
<td>H6</td>
<td>0</td>
</tr>
</tbody>
</table>

* A distance of 0 Å is read as ambiguous restraint with respect to and with the value of the previous entry with distance ≠ 0 Å
Supplementary Note 1: Rationale of the CLIR-MS/MS data analysis strategy.

The localization of protein-RNA contact sites by mass spectrometry after UV crosslinking is a particularly challenging problem because many different amino acid residues can be modified by a large number of potential nucleotide adducts (different nucleobases, different lengths and secondary reaction products such as those resulting from loss of water). In contrast, endogenous protein co- and post-translational modifications are typically restricted to a single molecular entity on one or a few sites on a peptide (for example in the case of phosphorylation); or restricted to structures of higher diversity attached to residues that are part of a consensus sequence motif (for example, N-glycosylation on asparagines in an Asn-X-Ser/Thr motif). Even the related technique of protein-protein crosslinking usually takes advantage of chemistries that are specific to one or a few residues.

For bioinformatic analysis, we are therefore faced with a challenge of performing an “all-against-all” search. That means we would need to consider all modification types (a set of more than ten different nucleotide adducts) on all amino acid residues. This problem cannot be addressed by database search software commonly used in proteomics research. Software that can search for unrestricted mass shifts lack the statistical sensitivity because typically only few spectra per modification are found in a data set, making it hard to discriminate them from random false positive hits. Urlaub and co-workers have introduced RNPxl, a data analysis strategy based on the OpenMS framework. However, it relies on a search engine, OMSSA, that is no longer maintained, and it is not straightforward to extend the analysis strategy to segmentally labeled RNA, which offers the unique advantage to restrict possible interaction sites also on the RNA level.

We therefore took advantage of xQuest, a software originally developed for protein-protein crosslinking analysis. A key feature of xQuest is that it is able to combine spectra from differentially isotope labeled peptides, including those resulting from the use of isotope labeled crosslinking reagents or, in this case, differentially labeled adducts. In a first step, xQuest looks for MS/MS spectra of precursors with a mass difference that matches the expected difference between light and heavy forms. Only these spectra are used at following stages of the data analysis, therefore restricting the number of spectra to be searched and reducing the likelihood of making false positive assignments. MS/MS spectra of the light and heavy forms of the modified
peptides are then combined into a consensus spectrum, whereby different fragment ion types (those that contain or do not contain the modification site) are recognized and noise peaks are removed. Finally, the consensus spectra are matched against a database of target proteins. In its current version, xQuest does not allow searching for more than one differential mass shift between light and heavy adduct at a time. However, different adduct masses with the same light/heavy difference, such as +U and +U-H₂O, may be searched simultaneously. Therefore searches for different adducts have to be performed independently, although this process might be automated in future releases.

Localization of the actual modification sites depends on two factors: the quality or information content of the MS/MS spectra and the efficiency of the nuclease digestion. xQuest is able to localize the nucleotide adduct to a particular site if unique diagnostic fragment ions are present in the spectrum. This depends on the intrinsic fragmentation properties of the peptide (the fragments must be formed in the first place) and the performance of the mass analyzer (the fragments must be detected). In the present work, we used collision-induced dissociation and detection in a linear ion trap analyzer and were able to localize the modification site in approximately 80% of cases. However, it is possible that the use of other fragmentation methods, such as electron transfer dissociation, or other analyzer types, such as orbitraps, will result in even better localization rates.

**Supplementary Note 2: Analytical workflow of the CLIR-MS/MS method.**

The process of protein-RNA crosslinking by UV irradiation is characterized by its low efficiency. This creates two main experimental challenges for identifying the cross-linking sites by mass spectrometry. Firstly, after protease and RNase digestion, modified peptides will be present in a large excess of unmodified peptides. To overcome this, we enrich modified peptides by metal oxide affinity chromatography using titanium dioxide as affinity material, as originally proposed by Urlaub and co-workers⁸. Secondly, the absolute amounts of modified peptides are very low. Thus, sufficient starting material is necessary for a successful detection in the mass spectrometer. To obtain maximum sensitivity for MS/MS spectra, we acquired fragment ions in the linear ion trap analyzer of our hybrid instrument. For the regular mapping of RNA adducts on PTBP1, we injected enriched peptides obtained from an equivalent of 62.5 µg amount (approximately 1
nmol) of protein as a starting material. However, we were able to locate all RRMs from significantly less material, e.g. using 1/5th of starting material prior to enrichment (approx. 200 pmol injected) or injecting 1/25th of the amount (equivalent to 40 pmol) of the original sample without any changes to the experimental setup. Further optimization of the enrichment procedure and use of more sensitive instrumentation will enable lower detection limits.

**Supplementary Note 3: Labeling strategy**

The segmental isotope labeling strategy strongly depends on the question at hand and on the RNA sequence. In principle, CLIR-MS/MS can be used to map binding sites without any prior knowledge in a screening-like fashion, however, reducing the size of the RNA sequence of interest by simple deletion experiments might be considered. In our experience, detection of single and di-nucleotide modifications can be expected, in favorable cases also tri-nucleotide modifications. That would correspond to a stretch of up to five nucleotides for which the composition can be determined and – in combination with shorter stretches – its partial sequence. The length of the segments can be designed according to the redundancy of the sequence. In our study, the shortest EMCV-segment comprised 10 nucleotides (EMCV<sub>link</sub>) and the largest 36 nucleotides (EMCV<sup>0</sup>). In case of a very redundant RNA sequence, it might be necessary to label one nucleotide at a time synthetically and ligate it subsequently to the unlabeled RNA segments. Ligation of synthetized RNA to fragments derived from in-vitro transcription and RNase H cleavage is possible.<sup>9</sup>

**Supplementary Note 4: RRM3 crosslinks to "E" might report on the packing of the PTBP1-EMCV<sub>DElinkF</sub> complex**

Previous NMR studies showed that RRM3 and RRM4 stably interact with each other<sup>10</sup> and can only bind simultaneously to an RNA with two pyrimidine tracts if these are separated by a spacer of at least 15 nt<sup>11</sup>. Residues around Pro417 and Phe371 on RRM3 exhibit additional chemical shift perturbations<sup>11</sup>, suggesting an interaction of these regions with the linking RNA moiety. This could explain the crosslinks observed between RRM3 and EMCV<sub>DElinkF</sub> “E” with SLE being the RNA linker between the RRM3 and RRM4 binding sites in EMCV<sub>DElinkF</sub>. These crosslinks might then report on the overall packing of the RNA in this complex. This positioning of SLE agrees also with the reported medium intensity cleavage of SLE by hydroxyl radicals generated at RRM3<sup>12</sup>.
Supplementary Note 5: Mapping of binding sites and modeling

The exact localization of the modification site on the detected peptide and the identity of the modification does not directly allow the mapping of the crosslink throughout the RNA sequence because the location of single nucleotide modifications throughout the isotope labeled RNA segment is often ambiguous. Longer nucleotide modifications (di- and trinucleotides) might be unambiguously mapped on the labeled RNA segment but the location of the crosslink itself remains unresolved throughout the detected di- or trinucleotide modification. However, this information is already very valuable for further structural or functional experiments. Furthermore, we detect modifications of different composition for the same modification site. Assuming a predominant conformation of the RNP and considering the low crosslinking efficiency, this infers not only a longer primary RNA sequence that consists of both modifications but also restricts the crosslink to the moiety present in both modifications. Available structural data, either from homology models of protein domains or from previous studies, helps to refine the mapping of interactions further. For example, RRM3 modifications comprise U and UU. Throughout SLD, only one UU appears in the primary sequence at position 303/304, which is flanked by C and G. However, we detect U-modifications of Thr407, Leu408 and Ser409, which surround the binding pocket 1 of RRM3 and of Asn413, which is located at pocket 2. Thus, RRM3 can accommodate U₃₀₃ in pocket 1 and U₃₀₄ in pocket 2. Steric hindrance induced by G₃₀₁ could disfavor in this particular case binding to C₃₀₂U₃₀₃ as a purine residue would conflict with the previously reported requirement for a pyrimidine at that position. Based on similar considerations, we could exclude U₃₅₁UG₃₅₃ as RRM2-binding site and map RRM2 to the UCUUU-pentaloop of SLF (detected modifications of RRM2: UU, GU and GUU). In the case of RRM1, the modification of Ile128 by AU is the only crosslink (1 spectral count) that is not compatible with detected UU-modification and the proposed low-resolution models. Ile128 can only contact nucleotides that are located 5’ to pocket 1. Consequently, and assuming this is not a false positive assignment at the MS/MS level, this crosslink can only occur if RRM1 binds the region 3’ of A₃₃₁ which is only possible upon melting of the stem of SLE and is probably caused by a very small subpopulation of misfolded RNA. Binding of a syn-G in pocket 1 is the only possibility to explain the observed UU crosslinks to Y127 as this residue contacts the nucleotides 5’ to pocket 1. Binding to the proximal C₃₂₉U₃₃₀ sequence is probably hindered by A₃₃₁ which would be placed then in pocket 3. The structure of RRM1 in
complex with a stem loop RNA (pdb code 2N3O) shows a pyrimidine at that position and the binding pocket 3 is too narrow to accommodate a purine. Furthermore, previous reports suggest that G can indeed be recognized by PTBP1 whereas A is generally disfavoring PTBP1-binding\(^{13}\). In conclusion, high affinity motifs like CU\(^{14}\) could be disfavored by their sequential and secondary structure context which leads here to binding of other, compatible sequences. Mapping of RRM4 to the linker sequence is straightforward: His457 is modified by C, U, CU and UU, which indicates its localization between U\(_{342}\) and C\(_{343}\) and thus fixes C\(_{343}\) in pocket 1.

Our example shows the great value of CLIR-MS/MS derived restraints. Firstly, the CLIR-MS/MS based mapping provides the basis for subsequent 3D modelling of protein-RNA complexes. Secondly, CLIR-MS/MS data combined with anchoring points derived from available structural information facilitate an efficient search for the optimal binding interface and improve convergence of the simulated annealing procedure to define atomic scale models. Finally, CLIR-MS/MS data demonstrate their strength for validation of previously calculated ensembles. Noteworthy, we could extract the binding register of RRM3 and RRM4 from models exclusively based on ambiguous intermolecular restraints derived from CLIR-MS/MS.

**Supplementary Note 6: Crosslinking efficiency and limitations of CLIR-MS/MS**

The crosslinking efficiency depends on three factors, namely the nucleotide, the amino acid and the geometry between both\(^{15-17}\). A quantitative comparison of different domains/peptides crosslinked to the same RNA sequence is therefore misleading. Comparison of crosslinking events occurring for the same domain/peptide with different positions on the RNA is justified as long as the specificity of the domain leads to binding of the same nucleotide in the same position. This implies that comparative studies using the same RNA e.g. under different conditions, for example in presence and absence of a small molecule compound, are generally justified. In case of PTBP1-RRMs, crosslinks occurred at RNA sequence specific positions. The affinity influences indirectly the crosslinking efficiency as it determines the amount of formed complex that can be crosslinked. Furthermore, the multitude of different crosslinking products between protein and RNA dilutes the intensity of individual adducts in the MS which requires higher/larger samples amounts than for protein-protein crosslinks (200 µg for CLIR-MS/MS vs. 50 µg or less in XL-MS\(^{18}\)). However, the amount of material required is small when compared to NMR studies. In addition,
Photoactivatable nucleotide analogs could be incorporated to enhance the crosslinking efficiency as in PAR-CLIP\textsuperscript{19} and to increase sensitivity.

The location of the protein-RNA crosslink is not predictable. However, the crosslinking event itself is reproducible as can be deduced from the repetition of the experiment with sample PTBP1-EMCV\textsuperscript{DlinkF} “D”. The successful identification of crosslinking sites will be challenged by an increase in complexity of the samples. However, the use of differential isotope labeling increases the confidence of identifications as we impose the exact mass shift corresponding to the labeling scheme as an additional criterion for a positive identification. Strikingly, we could reliably identify protein-RNA crosslinks that occurred in U1snRNP for a total of five different proteins without any changes to our crosslinking and enrichment protocol. Furthermore, the identified proteins represent three different families of RNA binding domains, Sm proteins, zinc fingers (U1-C) and RRMs (U1-70K, U1-A) which suggests that CLIR-MS/MS can be applied to various types of RBPs. Future applications in very complex systems like cell lysates will require additional purification steps prior to RNA digestion.

Supplementary Information references