Supporting Information

eXL-MS: An enhanced Cross-Linking Mass Spectrometry Workflow to Study Protein Complexes.

Martial Rey, Mathieu Dupré, Isabel Lopez-Neira, Magalie Duchateau, Julia Chamot-Rooke*

Mass Spectrometry for Biology Unit, CNRS USR 2000, Institut Pasteur, Paris, France 75015.

AUTHOR INFORMATION

Corresponding Author

* E-mail : julia.chamot-rooke@pasteur.fr

Table of contents **Page**

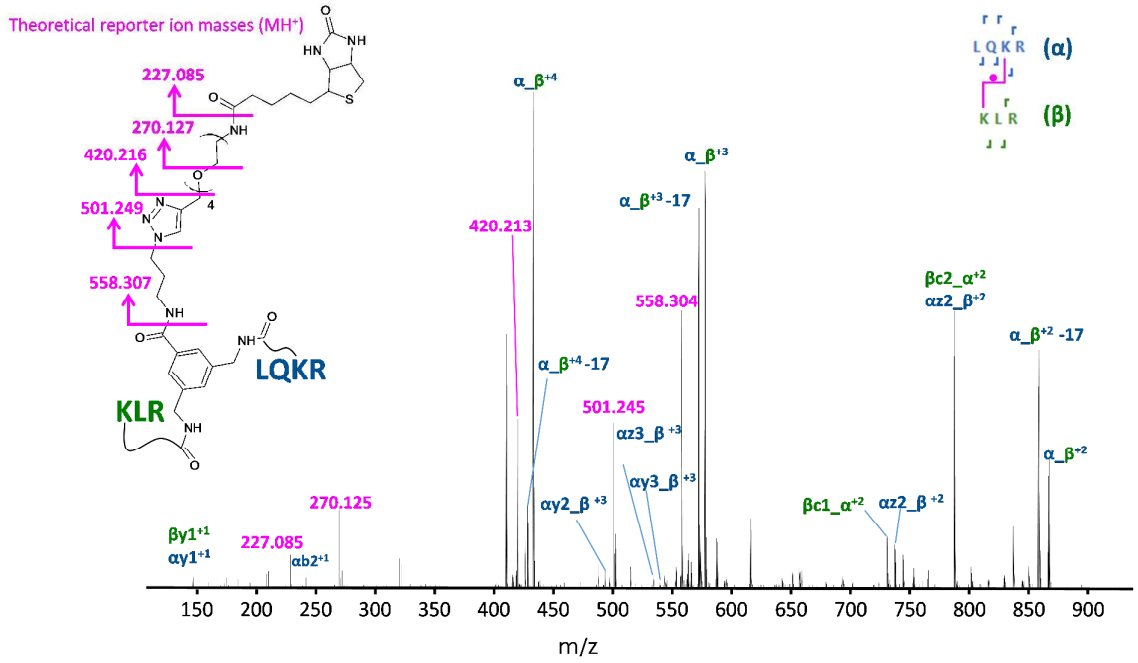
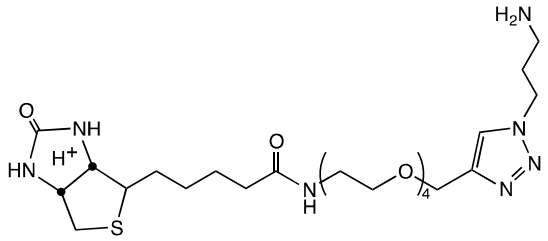
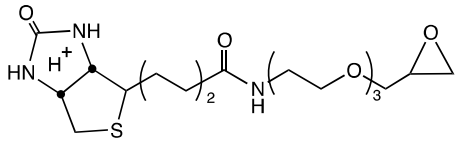


Figure S1

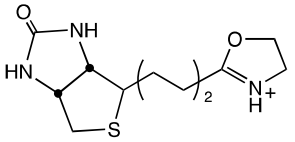
EThcD fragmentation spectrum of LQKR-KLR cross-linked peptide. Identified fragments are annotated in blue (for peptide α fragments), in green (for peptide β fragments) or in pink (reporter ions from the biotin moiety). The peptide fragments are reported on the sequence in the top right corner. The reporter ions are annotated on the chemical structure of the biotinylated cross-linker on the top left corner.



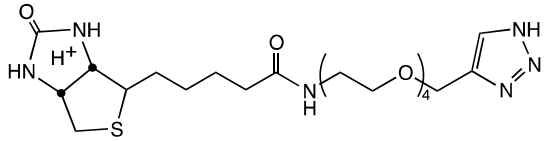
$C_{24}O_6N_7S_1H_{44}^+$
m/z 558.307



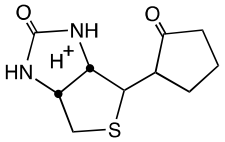
$C_{18}O_6N_3S_1H_{34}^+$
m/z 420.216



$C_{12}O_2N_3S_1H_{20}^+$
m/z 270.127



$C_{21}O_6N_6S_1H_{37}^+$
m/z 501.249



$C_{10}O_2N_2S_1H_{15}^+$
m/z 227.085

Figure S2

Proposed structures for the reporter ions arising from the fragmentation of the biotin moiety of a cross-linked peptide with their respective chemical composition and m/z.

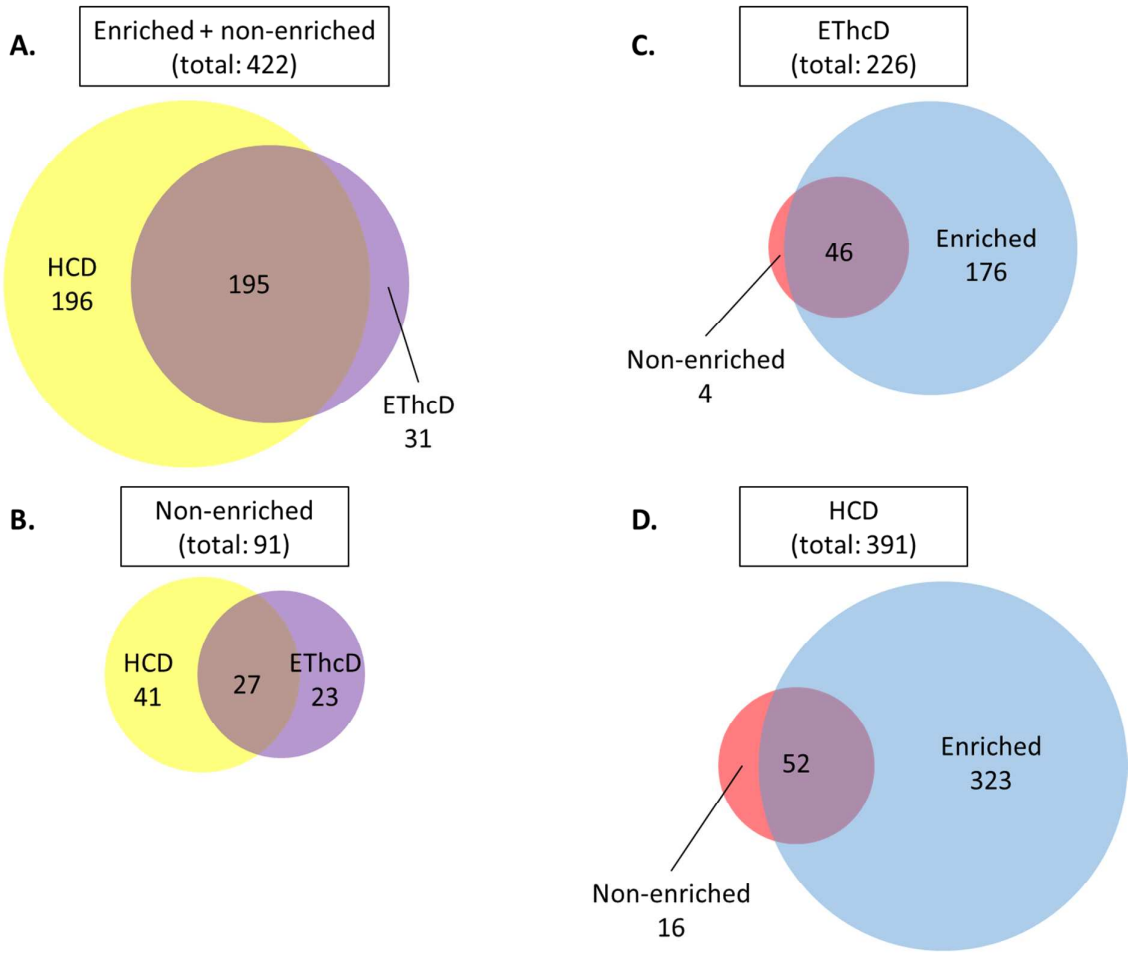


Figure S3

Venn diagram representing the number of identified cross-linked peptides **A.** for the different fragmentation mode: HCD (non-enriched and enriched fraction, yellow) and EThcD (non-enriched and enriched fraction, purple). **B.** without enrichment (HCD in yellow and EThcD in purple). **C.** for EThcD with (blue) or without (red) enrichment. **D.** for HCD with (blue) or without (red) enrichment.

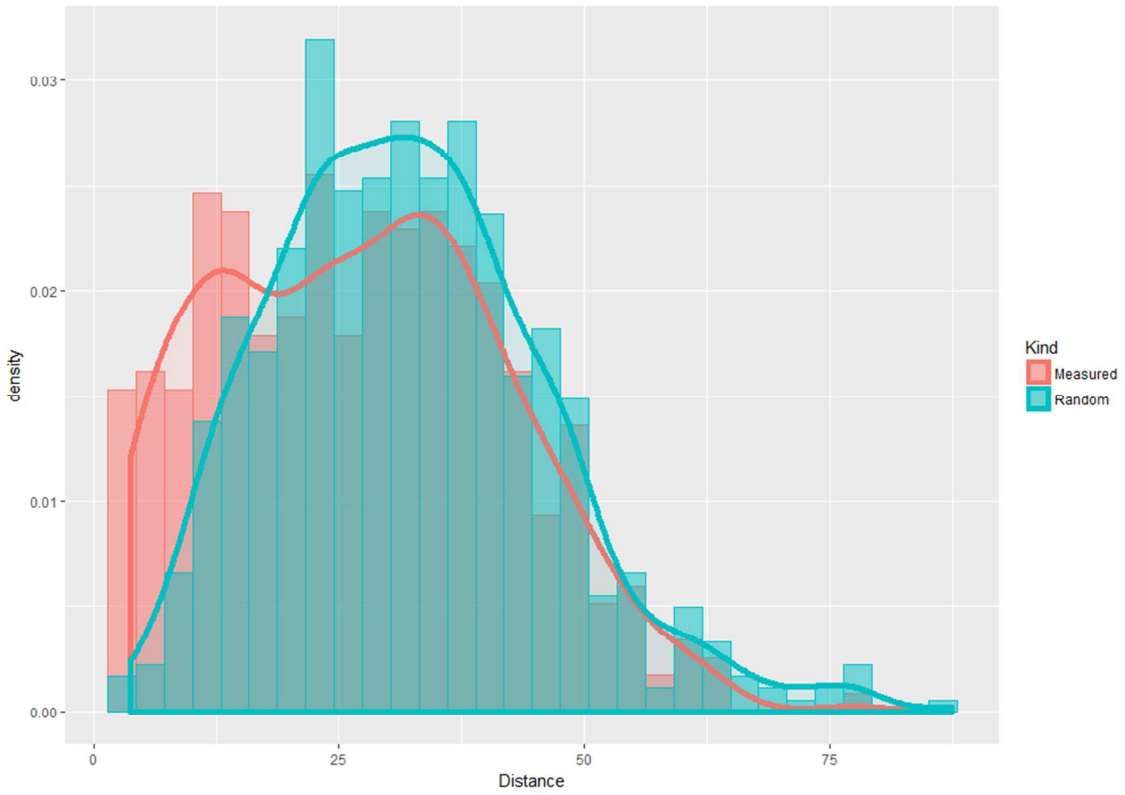


Figure S4

Estimated distributions of measured distances in red and random distances in blue. The displayed densities have been estimated using Gaussian kernels thanks to the geom_density() function in R.
A two-sample Cramér-von Mises test (performed with the CvM.test() function of "RVAideMemoire" R package) rejects the hypothesis of equality of distributions between measured and random distances (p-value < 2.2e-16). The distributions of distances are thus significantly different between measured and random distances.

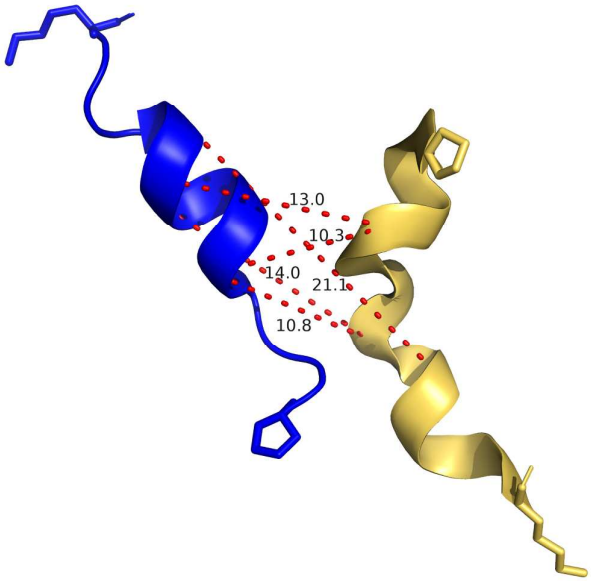


Figure S5

Snapshot of the first 14 amino acids of the CK dimer structure (in pymol, pdb code 1u6r) represented as a cartoon. The first proline and the lysine 14 are shown as sticks. The 5 intermolecular cross-links identified are shown as red dashed lines with the distance in Angstroms labelled in black.

Table S1. See Excel file

Values obtained for ID_{10} (number of identification above a score threshold of 10), S_{50} (average score of the top 50 identification) and $ID_{10}*S_{50}$ for the automatic identification of cross-linked peptides obtained from various HCD, ETD and EThcD LC-MS/MS runs with Mass Spec Studio.

Table S2. See Excel file

List of cross-linked peptides (labelled with their cross-linked site) identified in HCD and EThcD with or without enrichment.

Table S3. See Excel file

Number of cross-linked peptides identified in each replicate for each condition. The column F present the sum of the identified cross-linked peptides for the three replicates per experiment (sample preparation conditions and fragmentation modes). The last column present the sum of the identified cross-linked peptides per sample preparation conditions (SDS-PAGE, eFASP with and without enrichment).

Table S4. See Excel file

Type and percentages of residue linkages obtained with NNP9 for CK.

Table S5. See Excel file

Average and median scores and number of fragments for the MS/MS spectra of identified cross-linked peptides for all eFASP experiment (with or without enrichment, with HCD or EThcD fragmentation).

Table S6. See Excel file

List of minimum Cα-Cα distances (in Å) calculated with Pymol on the CK dimer structure (1U6R) for all cross-linked peptides.