Poster # P-II_0566

Thermo Fisher S C I E N T I F I C

Protein–RNA crosslinking combined with MS becomes quantitative at peptide and amino-acid level

Yi He¹, Eugen Damoc², Luisa M. Welp^{3,4}, Olexandr Dybkov³, Aleksandar Chernev³, Timo Sachsenberg^{5,6}, Monika Raabe³, Arslan Siraj⁵, Oliver Kohlbacher^{5,6,7}, Henning Urlaub^{3,4}, Rosa Viner¹, Daniel Hermanson¹

Thermo Fisher Scientific, ¹San Jose, ²Bremen; ³Bioanalytical Mass Spectrometry Group, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany; ⁴Bioanalytics Group, Institute of Clinical Chemistry, University Medical Center Göttingen, 37075 Göttingen, Germany; ⁵Institute of Bioinformatics, University of Tübingen, 72076 Tübingen, Germany; ⁶Applied Bioinformatics, Dept. of Computer Science, University of Tübingen, 72076 Tübingen, Germany; ⁷Institute of Translational Bioinformatics, University Hospital Tübingen, 72076 Tübingen, Germany;

Abstract

Purpose: Develop complete data processing workflows for quantification of Protein-RNA crosslinked samples

Methods: *E. coli* ribosomes were exposed to UV or chemical crosslinkers and/or labeled with TMT6plex. Separation was achieved using Thermo Scientific[™] Vanquish[™] Neo LC system and crosslinked samples were detected on Thermo Scientific™ Orbitrap Exploris[™] 480 or Thermo Scientific[™] Astral[™] mass spectrometers and data was analyzed using using the NuXL node in Thermo Scientific[™] Proteome Discoverer[™] v3.1 software.

Results

Workflows in the NuXL node

NuXL robustly identifies crosslinked peptides and amino acids from XL-MS data offering optimized search settings for a broad spectrum of crosslinking agents (UV, 4SU, 6SG, formaldehyde (FA), 1,2:3,4-diepoxybutane (DEB) and mechlorethamine (NM)) for both RNA- and DNA-protein crosslink samples (Table1). Further, combined with other downstream analysis nodes, the NuXL node can provide TMT quantitation and LFQ of crosslinked peptides(Figure 2).

Results

TMT quantitation

We investigated the feasibility of quantitative UV XL-MS with different amounts of E. coli ribosomes. Samples were exposed to UV light and then split in 2:1 ratio, labeled with TMT6plex in triplicates (Figure 3A). Crosslinked samples were analyzed on OT Exploris 480 (Figure 3B-D).

Figure 3. TMT quantitation of UV crosslinked *-E.coli* ribosomes : Experimental set up (A). PD Analysis of 50S L14 Y(U)AGVGDIIK peptide Annotated MS2 spectrum (B) Identification and Protein Level Quan (C). Peptide Level Quantitation (D)

Α.

Figure 6. Example of Annotated MS2 spectrum of a NM induced RNA-peptide crosslink from OT-Astral.

The MS2 spectrum view in Proteome Discoverer of the peptide GEILGGmAAVEQPEKPAAQPK crosslinked with CGU+C5H9N1

Spectrum View		_	×
110 m/z 1060.1064 R	24.60 Charge 3		
100	1061.054		
90			

Results: TMT quantitation and LFQ can be achieved accurately using the NuXL node in Proteome Discoverer v3.1 software for UV- or chemically-induced crosslinks.

Introduction

Mass Spectrometry can resolve UV- and chemically induced protein-RNA crosslinking at the peptide and amino-acid level. Our search engine NuXL⁴ improves the identification of protein-RNA UV/chemical crosslinks at amino-acid resolution, irrespective of sample complexity (isolated complexes or cellular entities). Yet the quantification of crosslinks remains a challenge for sensitive and precise detection of abundance changes among peptides crosslinked under different cellular conditions/states. UV crosslinks various amino acids to mainly uracil, whereas chemical crosslinking connects specific amino acids to guanosine and adenosine. Accurate MS-based quantification requires monitoring of the same crosslinked species derived from different states, and this can be hampered by changes in crosslinking patterns. Here, we explored a label-free(LFQ) or Tandem Mass Tags(TMT) quantitation approaches to analyze UV/chemical crosslinks using the NuXL node in Proteome Discoverer (PD) software.

- RNA/DNA protein complexes are essential for many cellular processes like DNA replication, DNA repair, transcription, splicing, RNA maturation, translation control etc.
- Defects in nucleic acid processing proteins are linked to severe diseases.
- Determination of the sites of interactions between proteins and RNA and/or DNA is required.

Table 1. Supported protocols.

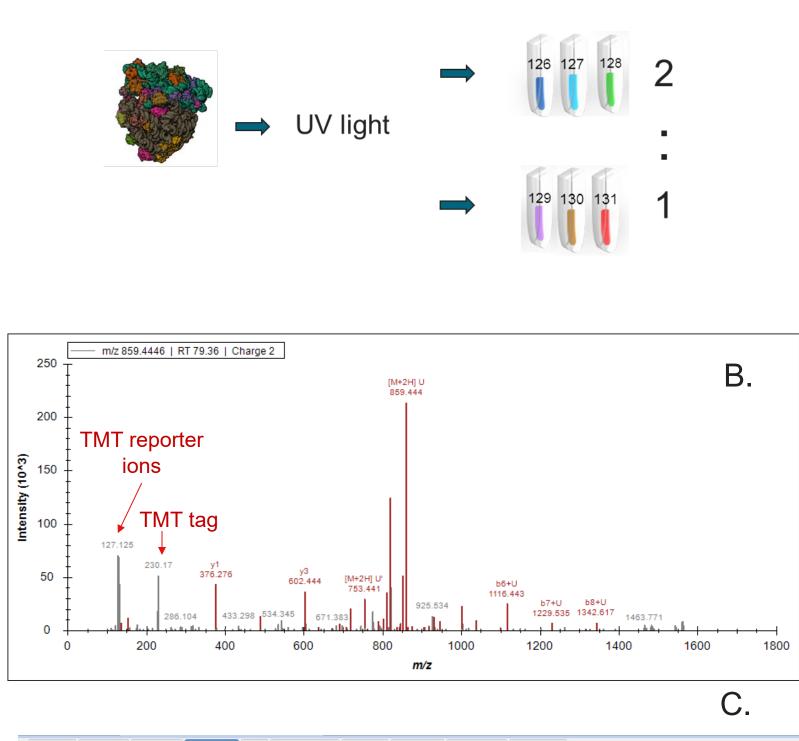
	UV XL	UV XL (4SU)	UV XL (6SG)	DEB	NM	FA
RNA			\checkmark		\checkmark	\checkmark
DNA						

DEB: Diepoxybutane, NM: nitrogen mustard (mechlorethamine), 4SU: 4 Thiouridine, 6SG: 6-Thioguanosine, FA: Formaldehyde

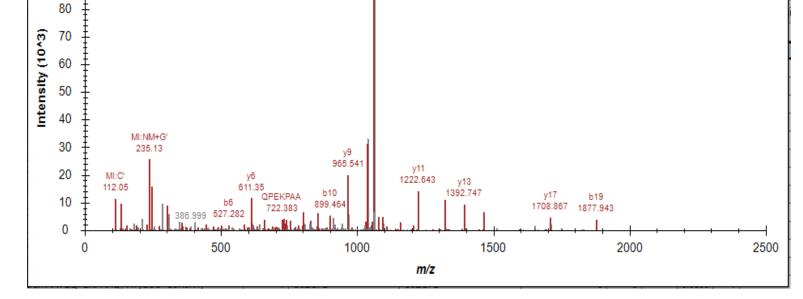
Crosslinked peptides are identified by the NuXL search node, regular peptides by the SEQUEST search engine. The Reporter lons Quantifier node is calculating abundances from TMT labeled peptides. The Minora node is used for LFQ for all the peptides.

Figure 1. Presets available in the NuXL node

,	1. General		
	Protein database	210302_ecoli_K1	2_SwissProt_4389prot.fasta
	CPU cores	4	
1	2. Cross-links		
	Length	2	
	Presets	RNA-UV (UCGA)	
1	3. Peptide identification		
	Precursor mass tolerance	6 ppm	DNA-DEB
	Fragment mass tolerance	20 ppm	DNA-DEB Extended
	Charge low	2	DNA-FA
	Charge high	5	DNA-FA Extended
	Peptide length min	6	DNA-NM
	Peptide length max	40	DNA-NM Extended
	Enzyme	Trypsin/P	DNA-UV
	Scoring	include fragment	
	Missed cleavages	1	none
	Static N-terminal modification	None	RNA-DEB
	Static C-terminal modification	None	RNA-DEB Extended
	1. Static modification	None	RNA-FA
	Max. number of dynamic modifications	2	RNA-FA Extended RNA-NM
	1. Dynamic N-terminal modification	None	RNA-NM Extended
	1. Dynamic C-terminal modification	None	



9	Proteir	s 💡	Protein Gro	ups Peptid	le Groups PSM	ls and NuXL	NuXLs	MS/MS Spectru	m Info	Quan Spec	ctra I	nput Files	9	Specialized Trace	s Study In	formation								
	1	Checked	Show Spectru	m Confidence	Identifying Node	PSM Ambiguity	Annotate	d Sequence	Modi	fications		# Proteins	Mast	er Protein Accessions	Protein Acce	ssions	# Missed	Cleavages	Charge	∆Score	ΔCn	Rank	Search Engine Rank	m/z [Da] 🗕
9	4 +⊐		Show Spectrum		NuXL (A5)	Unambiguous	[R].yAGV	GDIIk.[I]	N-Te	erm(TMT6plex);	; Y1(U); I	K 1	POAD	DY3	P0ADY3			0	2		0.0000	1	1	859.444
9	5 👳		Show Spectrum		NuXL (A5)	Unambiguous	[K].aADM	ITGADIEAMTR.[S]	N-Te	erm(TMT6plex);	: A1(AUL	J 1	C6E	E36	C6EE36			0	3		0.0000	1	1	880.977
9	6 ⊕		Show Spectrum		NuXL (A5)	Unambiguous	[K].aADM	ITGADIEAMTR.[S]	N-Te	erm(TMT6plex);	A1(AUL	J1	CGE	E36	C6EE36			0	3		0.0000	1	1	880.978
9	7 ⊕		Show Spectrum		NuXL (A5)	Unambiguous	[K].aADM	TGADIEAMTR.[S]	N-Te	erm(TMT6plex);	; A1(Oxio	d: 1	C6E	E36	C6EE36			0	3		0.0000	1	1	886.311
9	8 👳		Show Spectrum		NuXL (A5)	Unambiguous	[K].aADM	ITGADIEAMTR.[S]	N-Te	erm(TMT6plex);	A1(Oxio	d 1	C6E	E36	C6EE36			0	3		0.0000	1	1	886.309
9	9 👳		Show Spectrum		NuXL (A5)	Unambiguous	[R].aTLG	EVGNAEHmLR.[V	N-Te	erm(TMT6plex);	: M12(AU	J 1	C6E	GE9	C6EGE9			0	3		0.0000	1	1	794.004
9	k 0 ⊕		Show Spectrum		NuXL (A5)	Unambiguous	[K].aADN	ITGADIEAMTR.[S]	N-Te	erm(TMT6plex);	; A1(Oxio	di 1	C6E	E36	C6EE36			0	3		0.0000	1	1	886.311
9	1 👳		Show Spectrum		NuXL (A5)	Unambiguous	[K].fHAG	ANVGcGR.[D]	N-Te	erm(TMT6plex);	; C9(CU)) 1	C6E	H42	C6EH42			0	3		0.0000	1	1	649.586
9	2 -		Show Spectrum		NuXL (A5)	Unambiguous	[R].aTLG	eVGNAEHMLR.[V]	N-Te	erm(TMT6plex);	: E5(AU)	1	C6E	GE9	C6EGE9			0	3		0.0000	1	1	794.004
9	3 ⊕		Show Spectrum		NuXL (A5)	Unambiguous	[R].aTLG	EVGNaEHMLR.[V]	N-Te	erm(TMT6plex);	: A1(Oxio	d: 1	C6E	GE9	C6EGE9			0	3		0.0000	1	1	799.335
9	4 ⊕		Show Spectrum		NuXL (A5)	Unambiguous	[R].aTLG	EVGNAEHmLR.[V	N-Te	erm(TMT6plex);	: M12(AU	J 1	C6E	GE9	C6EGE9			0	3		0.0000	1	1	794.004
9	15 ⊕		Show Spectrum		NuXL (A5)	Unambiguous	[K].aADM	ITGADIEAMTR.[S]	N-Te	erm(TMT6plex);	: A1(AUU	J_ 1	C6E	E36	C6EE36			0	3		0.0000	1	1	880.980
9	6 👳		Show Spectrum		NuXL (A5)	Unambiguous	[R].aTLg	EVGNAEHMLR.[V]	N-Te	erm(TMT6plex);	: A1(Oxio	d 1	C6E	GE9	C6EGE9			0	3		0.0000	1	1	799.336 🗸
					10																			÷.
<u>^)</u>	Hide As	sociated	Tables																					
Pr	oteins	? 1	Protein Groups	Peptide G	àroups MS/MS	Spectrum Info	Quan S	pectra Found	Modific	cations Pre	cursors	Unknow	n Mod	difications										
	ŧ.	Checke	d Master Acco	ession Descrip	tion			Coverage [6] Seq	uence Coverag	ge #F	Peptides # P	SMs #	# Unique Peptides # /	AAs MW [kDa] calc. pl Se	core S 👻 🕂	# Peptides	+ # R	azor Pept	ides Abi	undanc	e Ratios 🛛 🕂 🖌	Abundance F
1	-12		V POA	DY3 50S rib	osomal protein L14	[OS=Escherichia	coli (strain	K12 41%				6	70	6	123 13.5	10.42	85.64	2	5		0		0.454	
																							-	
																							$\boldsymbol{\nu}$.	



Conclusions

- The NuXL node embedded in PD 3.0 or 3.1 enables analysis of DNA-/RNA-protein UV- and chemically-induced crosslink data.
- TMT and label-free quantitation of crosslinked peptide-RNA/DNA oligonucleotide can be achieved accurately and reproducibly using precursor or reporter ion quantification PD nodes.
- All major *E.coli* ribosomes proteins and RNA-protein crosslinked peptides were quantified using TMT6plex at expected ratio of 2.1 at amino-acid resolution.
- Proteome Discoverer 3.0/3.1 software with NuXL node was able to confidently identify and quantify peptides using precursor based LFQ. Its performance was validated using UV or chemical crosslinker E.coli ribosome.

Availability

The NuXL node is released for Proteome Discoverer 3.0 and 3.1. The binary installer, documentation, and example data are available at https://openms.de/applications/nuxl/.

- Standard methods like X-ray crystallography, NMR, or cryo-EM are the golden standard to study the compositions of RNA and/or DNA protein complexes but can be challenging for large and dynamic complexes and cannot be applied on entire cells.
- Crosslinking mass spectrometry (XL-MS) offers a straightforward method to identify proteins and protein domains, which interact with RNA and/or DNA on the molecular and atomic level not only in isolated complexes but also in entire cellular systems.

Materials and methods

Sample preparation

Different amounts of *E. coli* ribosomes were exposed to UV or chemical crosslinkers. Crosslinked samples were digested, purified using TiO₂ and/or labeled with TMT6plex³.

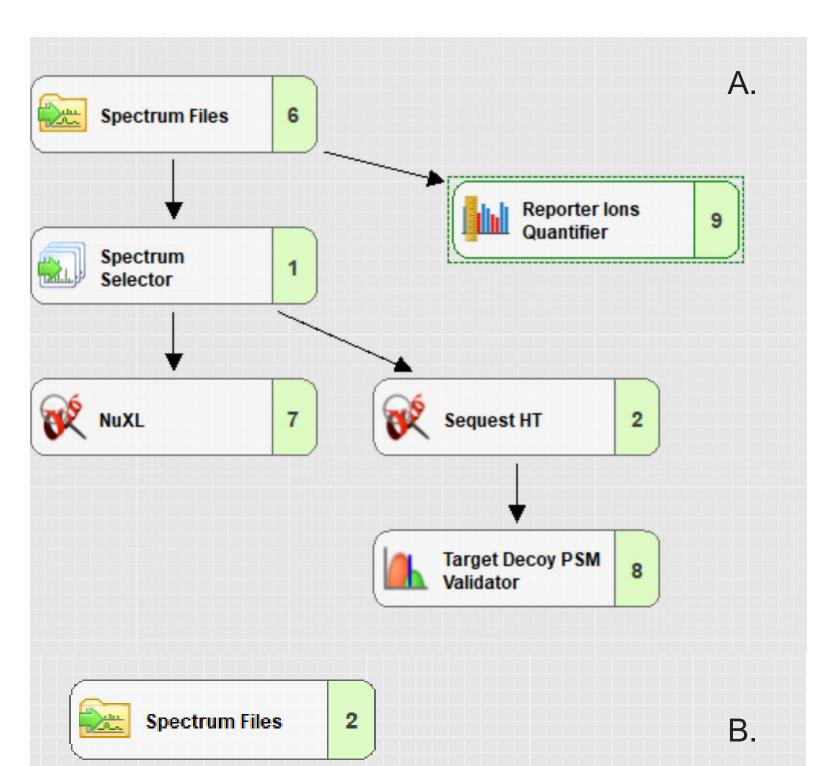
LC-MS/MS methods

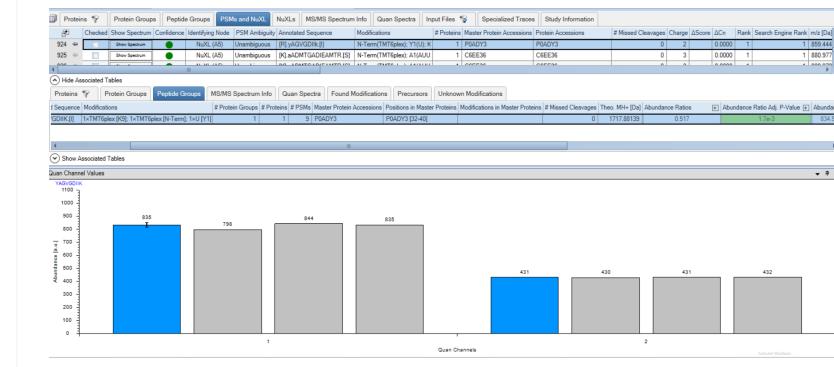
Liquid chromatography was performed using an Thermo Scientific[™] UltiMate[™] -3000 RSLC nanosystem or a Vanquish Neo UHPLC system. Peptides were eluted over a 60 min 6-45 or 50% gradient (A: water, 0.1% formic acid; B: 80% acetonitrile, 0.1% formic acid) at 300 nL/min flow rate. The crosslinked samples were analyzed on Orbitrap Exploris 480 MS or Astral mass spectrometers in DDA mode. The following MS settings were used- Exploris 480- MS1 scan range, 350–1600 m/z; MS1 resolution, 120,000 FWHM; AGC target MS1, 1E6; maximum injection time MS1, 60 ms; isolation window, 1.6 m/z; collision energy, 30 or 33(TMT)%; charge states, 2+ to 6+; dynamic exclusion, 9 s; Top 20 most abundant precursors were selected for fragmentation; MS2 resolution, 30,000, AGC target MS2, 1e5; maximum injection time MS2, 120 ms. OT Astral: MS1 scan range, 400–1400 *m/z*; MS1 resolution, 180,000 FWHM; AGC target MS1, 500%; maximum injection time MS1, 5 ms; isolation window, 1.6 m/z; collision energy, 32%; charge states, 2+ to 6+; dynamic exclusion, 10 s; Top speed 1 sec for fragmentation; MS2, AGC target MS2, 100%; maximum injection time MS2, 20 ms.

1. Dynamic modification	None	RNA-UV (6SG)
2. Dynamic modification	None	RNA-UV (U)
3. Dynamic modification	None	RNA-UV (UCGA)
4. Dynamic modification	None	RNA-UV Extended (4SU)
5. Dynamic modification	None	RNA-UV Extended (6SG)
		RNA-UV Extended (U)
		RNA-UV Extended (UCGA)

RNA-UV (4SU

Figure 2. Processing workflow including (A)TMT Quantitation and (B) Label Free Quantitation (LFQ) and NuXL Consensus workflow(C).





Label-Free quantitation

We investigated the feasibility of LFQ (Figure 4) using the NuXL node with UV- or chemical crosslinking of *E.coli* ribosomes. Samples were analyzed on OT Astral using DDA approach in duplicates (Figures 5&6).

Figure 4. PD result file showing the LFQ of two technical replicates of UV induced RNA-peptide crosslink.

		ein Groups Confidence	Peptide Groups 😴 PSMs Sequence	and NuXL MS/MS Spectrum Info			Specialized Traces		Consensus Features	Abundanas Patis	Abundan	Group
i≇ 191 -⊨	Checked	Conlidence		1×U-H2O1 [K11]		C6EGF7		0.981	9.8e-1	0.00	6.275e4	6.154e
192 -=			RDDEVIVLTGKDK VKFEVK	1×0-H2O1 [K1]		C6EH42	1793.83170 1055.48091	1.105	9.8e-1	0.00	8.385e5	9.264
192 -		•	VTKPEAGHFAK	1×0-H2O1 [K/V]		C6EGE6	1819.72006	0.964	9.8e-1	0.00	0.305e5	1.891
193 ~≕ 194 -⊨			RVMCIK	1×A0+H2O1 [NV] 1×Oxidation [M3]; 1×U [C/R]		POADY3	1019.72006	0.968	9.8e-1	0.00	5.659e5	5.476
194 -=		•	RGDKSMALR	1×Oxidation [M3]; 1×0 [C/R]		P04D13 P02359	2031.69295	0.988	9.8e-1	0.00	5.706e6	5.689
196 -=		•	VKTLLNEK	1×AAU [E/K/N/V]		C6EGF4	1926.71842	1.099	9.8e-1	0.00	3.540e6	3.889
197 -			RGDKSMALR	1×AAU [E/K/N/V] 1×GU [G/K/M/R/S]; 1×Oxidation [M6]; 1	21		1702.64043	0.993	9.8e-1	0.00	6.602e7	6.557
198 +=			VTKPEAGHFAKAGVEAGR	1×GU-H2O1 [T2]		C6EGE6	2476.04425	0.961	9.8e-1	0.00	1.942e5	1.865
199 -=			VKNVLSSGK	1×U-H2O1 [KV]		C6EGE6	1237.58241	0.975	9.8e-1	0.00	4.370e5	4.259
200 -=			AAAQKAFNEMQPIVDR	1×0-H2OT [KV] 1×Oxidation [M10]; 1×U-H2O1 [A1]		C6EB30	2110.92635	1.089	9.8e-1	0.00	4.370e5	2.029
200			QHVPVFVTDEMVGHK	1×Oxidation [M10]; 1×0-H2OT [A1] 1×Oxidation [M11]; 1×UUC-H3N1 [E/M		P0A7U3	2656.93403	1.104	9.8e-1	0.00	4.559e5	5.035
201 -			HKANLTAQINK	1×U-H2O1 [K2]		C6EB30	1543.72645	0.997	9.8e-1	0.00	4.3344e5	4.331
202 -=		•	HHITADGYYR	1×0-H201 [K2]		P0A7N4	1901.66400	0.991	9.8e-1	0.00	2.473e7	2.450
203 ⁻ → 204 -⇒			HHITADGYYR	1×00 [A/H/Y]		P0A7N4	1556.61657	0.968	9.8e-1	0.00	1.106e8	1.070
204 -		•	QHVPVFVTDEMVGHK	1×CUU [Q1]		P0A7IN4 P0A7U3	2657.96566	1.080	9.8e-1	0.00	1.725e5	1.86
205 -=		•	GKPFAPLLEK	1×U-H2O1 [K2]	4		1405.67631	1.086	9.8e-1	0.00	7.209e5	7.832
206		•	GDKSMALR	1×0-H201 [K2]		P02359	1405.67631	0.988	9.8e-1	0.00	7.537e6	7.44
207 -=			GDKSMALR	1×AGU [D/G/M/3] 1×Oxidation [M5]; 1×GU [G/M]		P02359	1562.53423	0.957	9.8e-1	0.00	2.248e6	2.151
			GDKSMALK		3	P02305	1562.53425	0.357	5.0e-1	0.00	2.24000	2.15
) Show A	ssociated T	ables										
an Channe	el Values					→ ∓ X	Chromatogram Tra	ces View				
VKNVLS	SSGK						\bigcirc	<u> </u>				
500 - 500 -			437030	425933			400 -		18.284			

References

- 1.Röst HL, Sachsenberg T, Aiche S, Bielow C, Weisser H, ... & Kohlbacher O. OpenMS: a flexible open-source software platform for mass spectrometry data analysis. Nat Methods. 2016 Aug 30;13(9):741-8.
- 2. Stützer, A., Welp, L. M., Raabe, M., Sachsenberg, T., Kappert, C., Wulf, A., ... & Urlaub, H. (2020). Analysis of protein-DNA interactions in chromatin by UV induced cross-linking and mass spectrometry. Nature Communications, 11(1), 1-12
- 3. Kramer, K., Sachsenberg, T., Beckmann, B. M., Qamar, S., Boon, K. L., Hentze, M. W., & Urlaub, H. (2014). Photo-crosslinking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. Nature Methods, 11(10), 1064-1070.
- 4.Welp,L.M., Sachsenberg,T., Wulf, A., Chernev, A., Horokhovskyi, Y., Neumann, P., Pašen, M., Siraj, A., Raabe, M., Johannsson, S., et al. (2024). Chemical crosslinking extends and complements UV crosslinking in analysis of RNA/DNA nucleic acid-protein interaction sites by mass spectrometry. bioRxiv 2024.08.29.610268

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement Nº 956148.

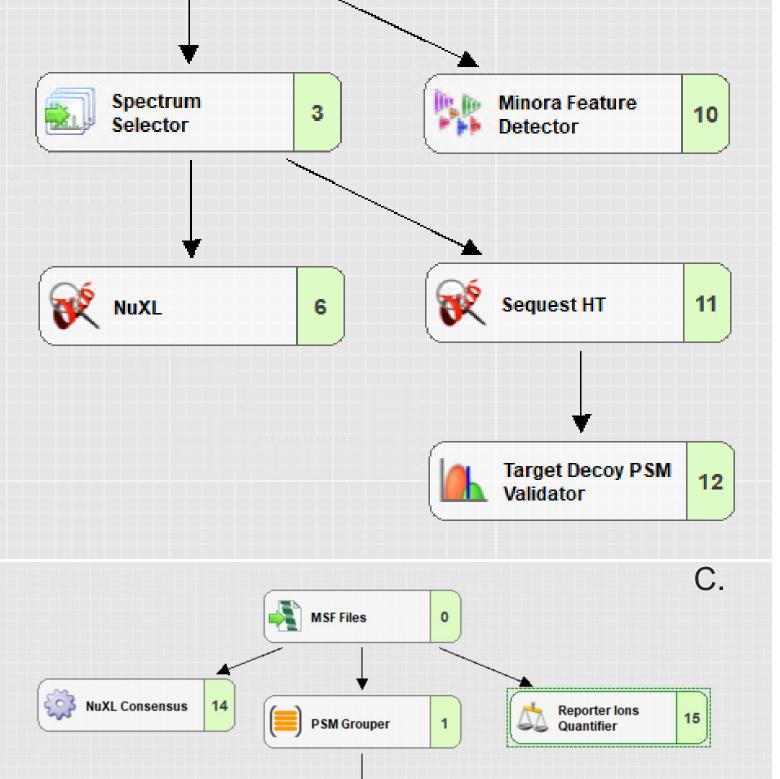
Trademarks/licensing

© 2024 Thermo Fisher Scientific Inc. All rights reserved. SEQUEST is a trademark of the University of Washington. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

Data analysis

The NuXL node in PD 3.1, based on the OpenMS software^{1,3} package, is used for analyzing XL-MS data from UV or chemically crosslinked protein–RNA samples. The NuXL node is designed to identify UV- and chemically-induced crosslinks by taking individual MS fragmentation behavior into account. Respective presets (Figure 1) have been assembled based on our own *in vitro* and *in* vivo datasets including human nucleosomes, HeLa cells and *Escherichia coli* cells as well as on published data. We calculate match-odds and subscores and employ semi-supervised score calibration using Percolator. Moreover, entrapment experiments have been employed using manually curated data ensuring proper false discovery rate (FDR) control.

Quantitation of crosslinking data was accomplished using the NuXL node in Proteome Discoverer 3.1 software together with the Reporter lons Quantifier node (for TMT quantitation) or the Minora Feature Detector node (for LFQ). Detailed templates are shown in Figure 2. LFQ of crosslinked peptide-RNA oligonucleotides was measured by their intensities in MS1.



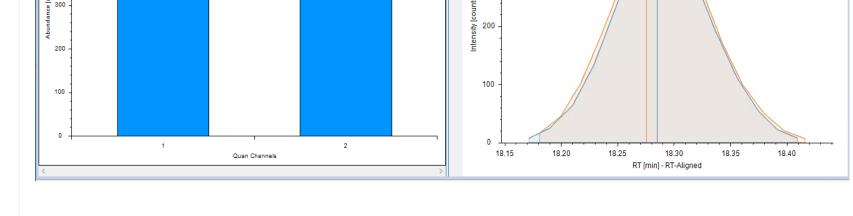
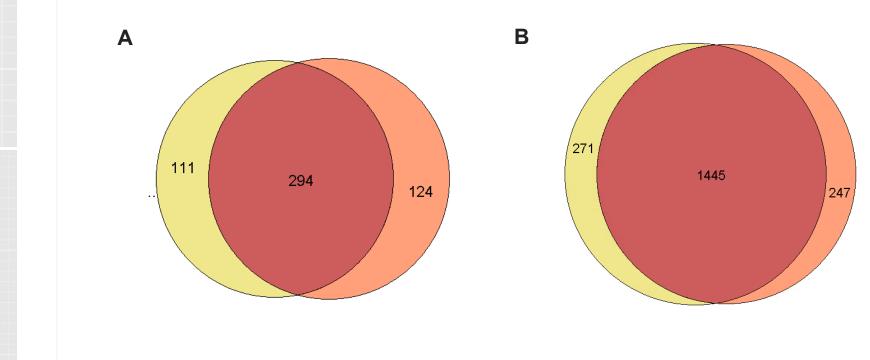


Figure 5. Venn diagram showing the overlap of unique identified RNA crosslinks in duplicates from (A) UV- or (B) NM samples in OT Astral DDA runs













Scan QR code

Science at a scan Scan the QR code on the right with your mobile device to download this and many more scientific posters.



Learn more at thermofisher.com/us/en/home/industrial/mass-spectrometry/proteomics-mass-spectrometry/proteinstructure-analysis-mass-spectrometry /crosslinking-mass-spectrometry.html