

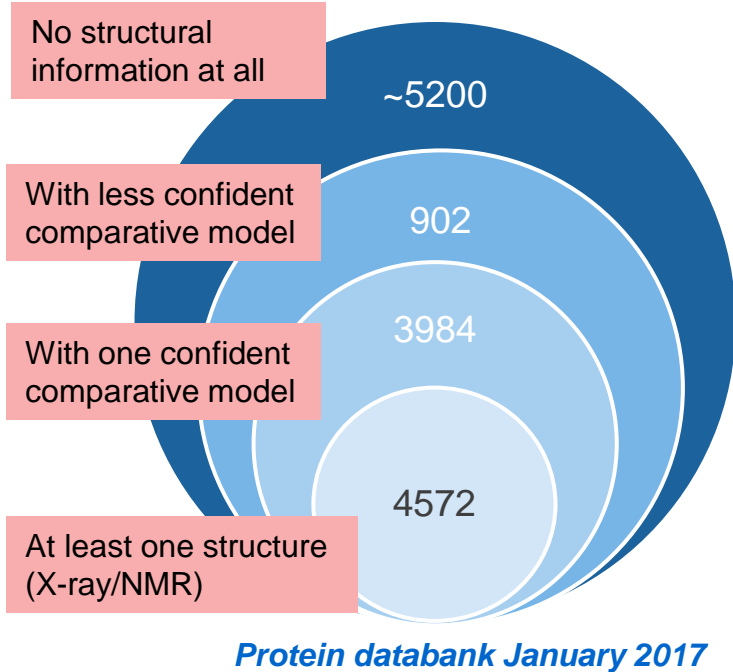


ThermoFisher
SCIENTIFIC

Integrated Solutions for Structural Biology: Introducing a Powerful New Workflow for Cross-linking Studies in Biology

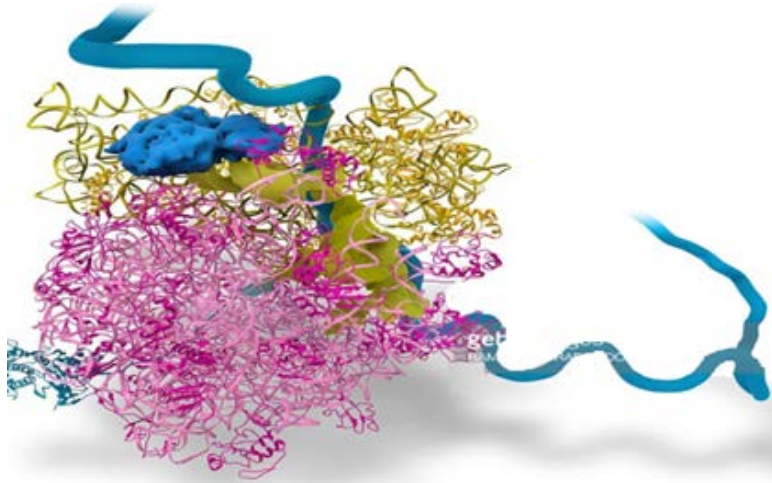
Presenter:
Rosa Viner

Complexity of the Proteome; Layers of the Proteome



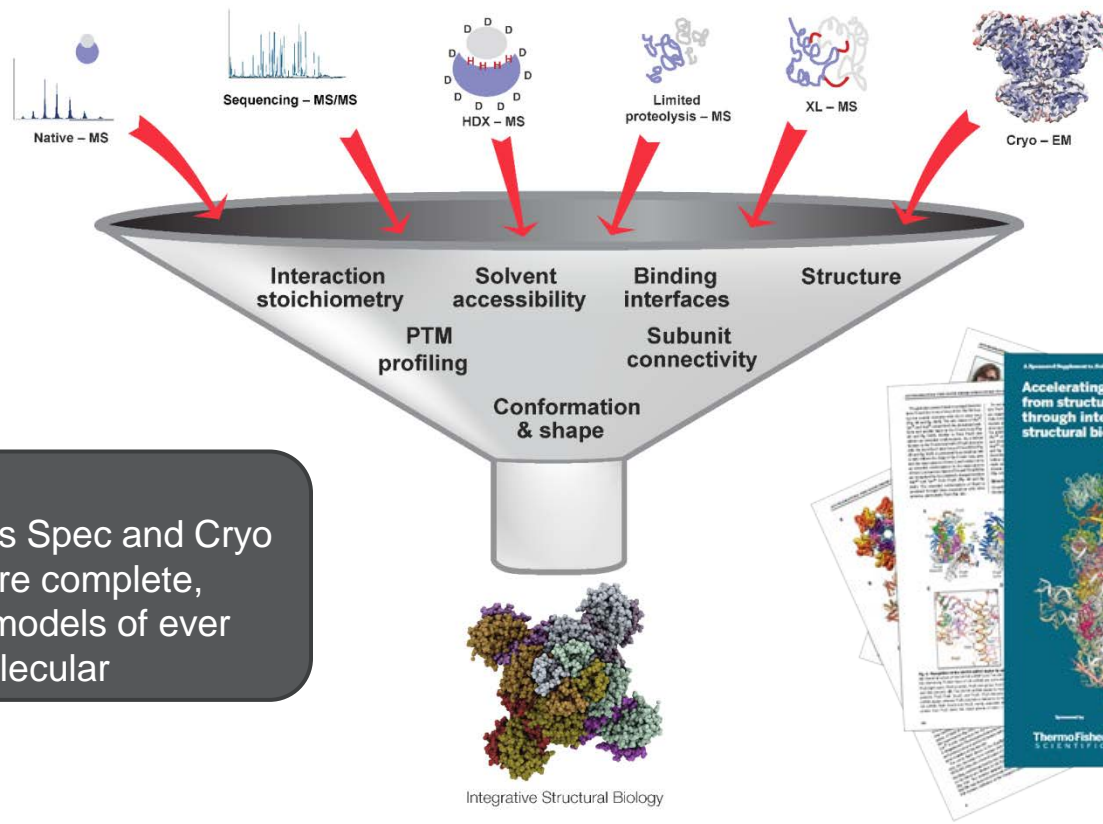
- 5200 protein families (PFAM) with unknown structure
- Recognition of (supra)molecular complexes as the functional subunit of biology in health and disease
- De-constructing and constructing complexes accelerates our understanding of biological function
- Technologies capture structural information for proteins, DNA/RNA, metabolites, drugs

What Questions Does Structural Biology Answer?



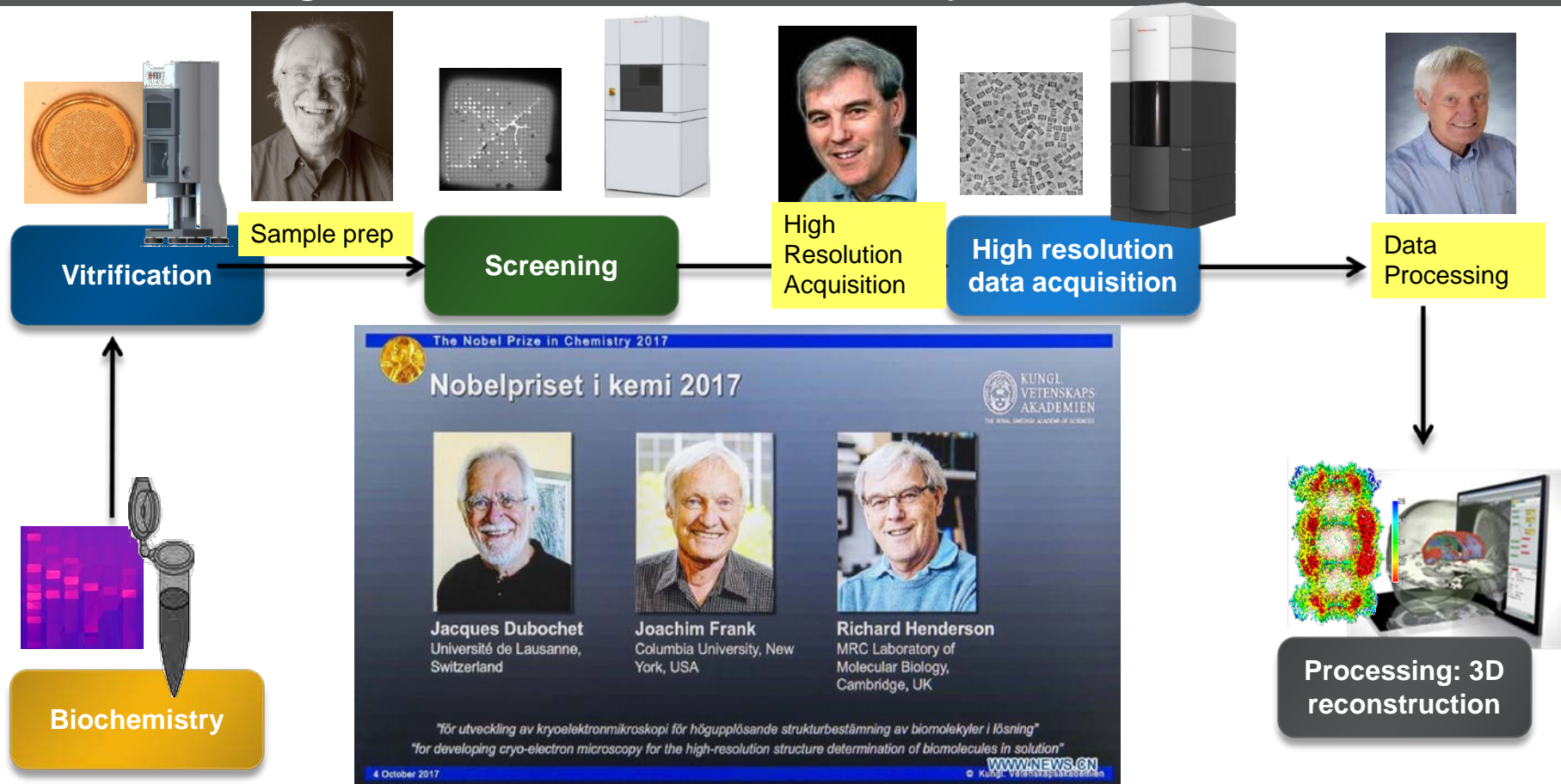
- 3D Structure
- Stoichiometry
- Composition
- Topology (binding partners, cofactors, messengers, etc.)
- Binding affinity
- Dynamics
- Aggregates
- Biological function

Integrative Structural Biology Workflow Requires Many Tools



True Synergy -
Combining Mass Spec and Cryo EM result in more complete, more accurate models of ever larger macromolecular

2017 Breaking News: Nobel Prize for CryoEM SPA Workflow



MS+ FEI Cryo TEM Workflows

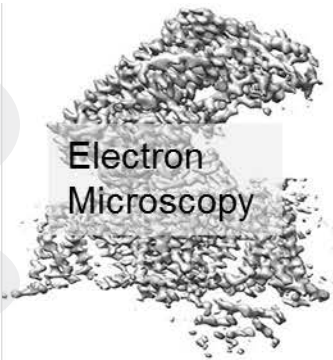
LC/MS/MS

Sample
compositional
homogeneity

Subunit
Composition

Stoichiometry

Native MS



Subunit
topology

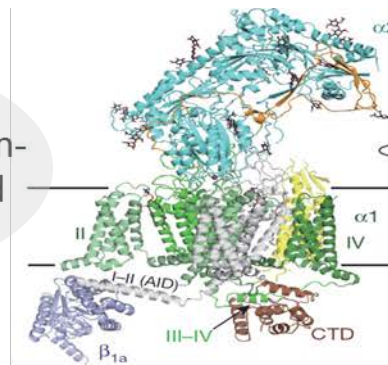
Atomic model
building

HDX, XL-MS

Native MS

Identification of non-
protein associated
molecules


LC/MS/MS



Integrative Structural Biology Workflow Requires Many Tools

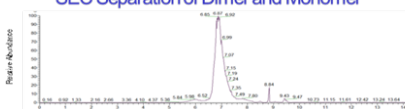
Thermo Scientific Native Mass Spectrometry Workflow

**CE-MS
WAX-MS
SEC-MS**




MabPac SEC-1

SEC Separation of Dimer and Monomer

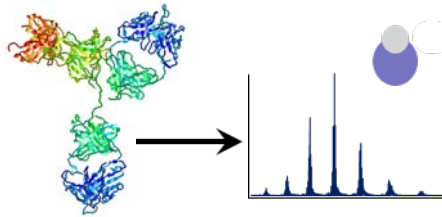


**Orbitrap
Analysis**

**Exactive EMR Q Exactive
BioPharma
UHMR**

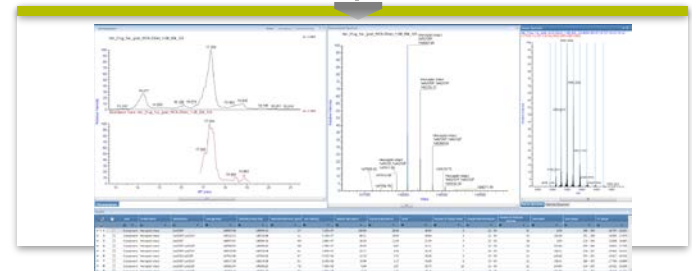


m/z up to 80 kDa



pH=7 Native - MS

**Thermo Scientific™
BioPharma Finder™ Software**



Since 2013

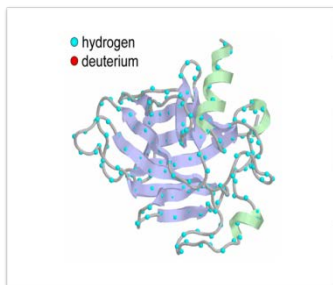
Thermo Scientific Hydrogen/Deuterium Exchange MS Workflow



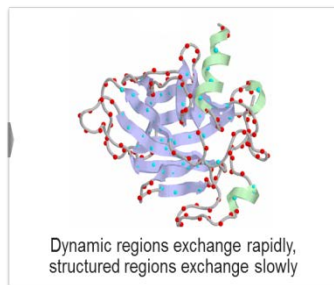
Orbitrap
Analysis



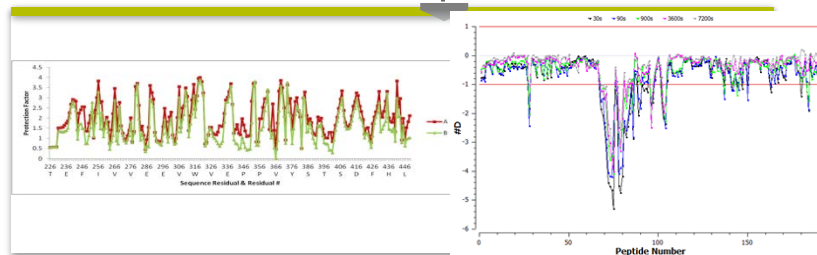
H/D-X PAL™ & Chronos



D₂O
time

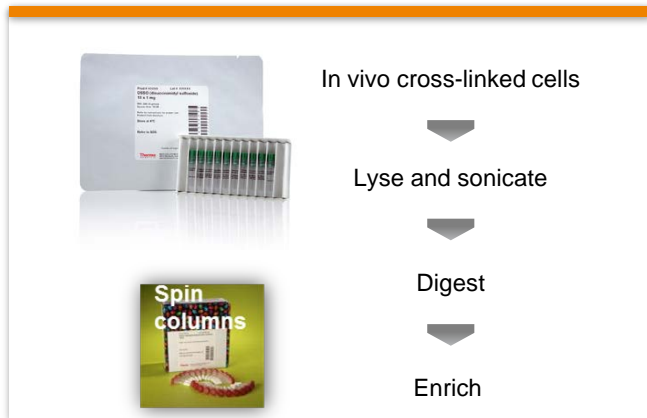


Thermo Scientific™
BioPharma Finder™ Software

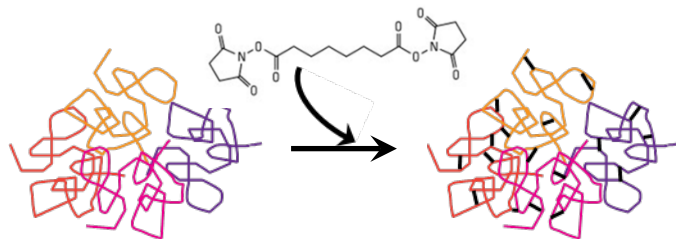


Since 2014

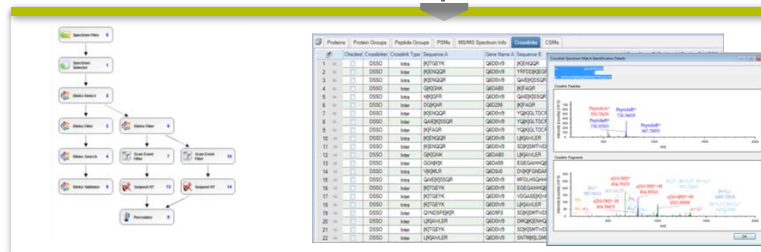
Thermo Scientific Cross-linking Mass Spectrometry Workflow



Orbitrap
Analysis



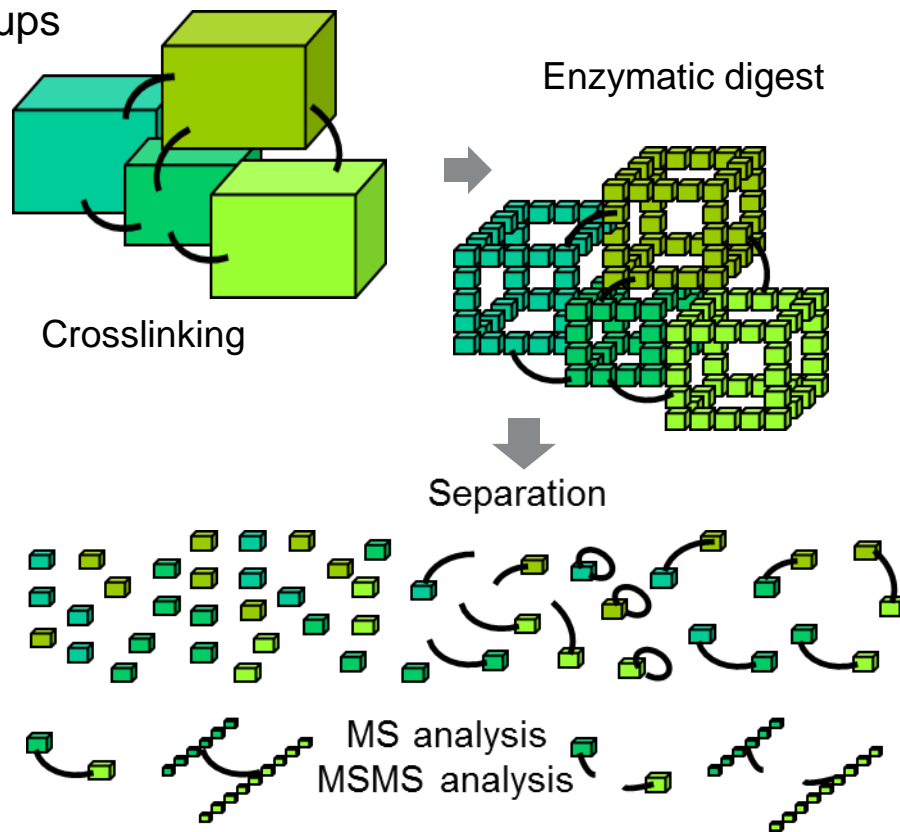
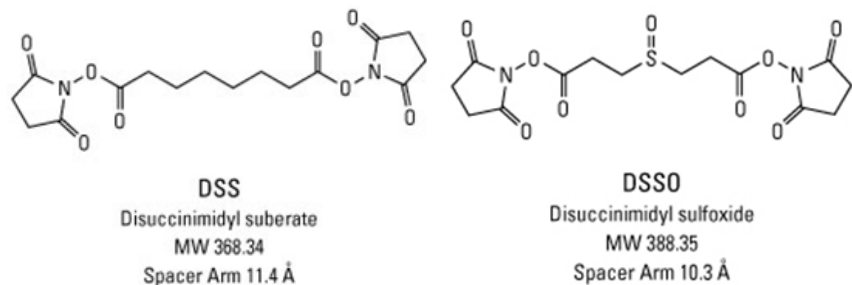
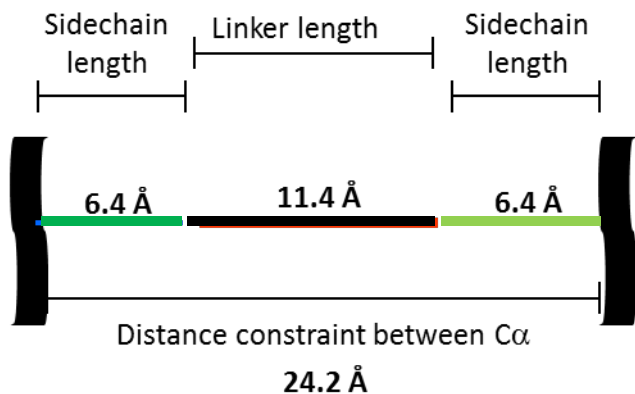
Thermo Scientific™ Proteome Discoverer™ 2.2 XlinkX node



Since 2016

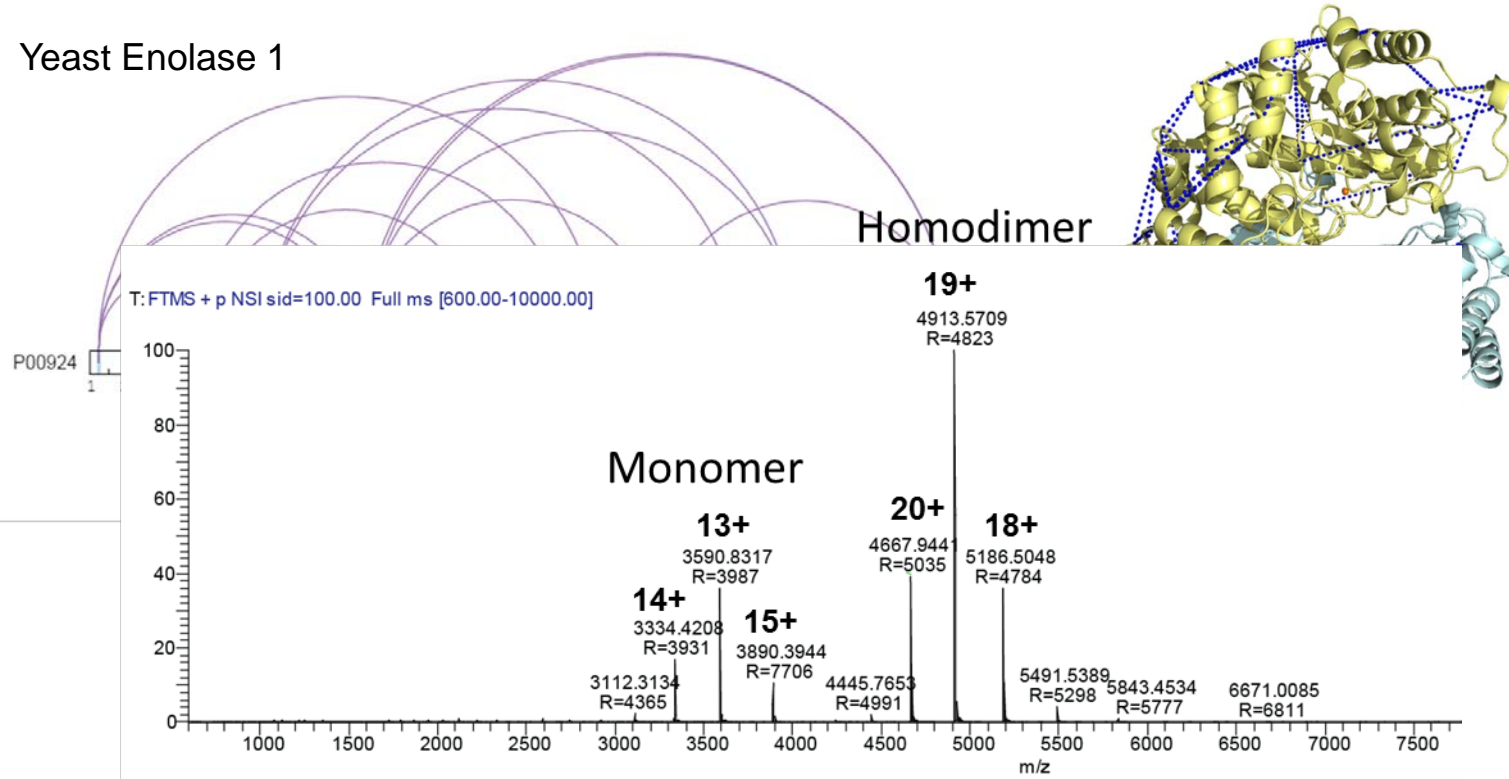
Introduction of Cross-Linking Mass Spectrometry

Distance information between two cross-linked groups

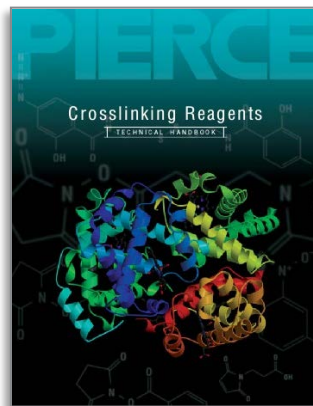
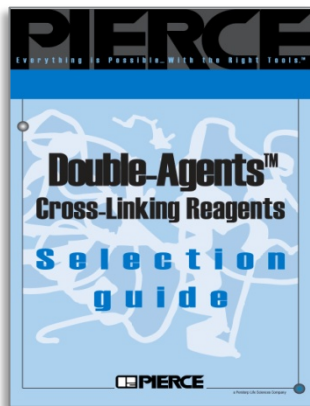
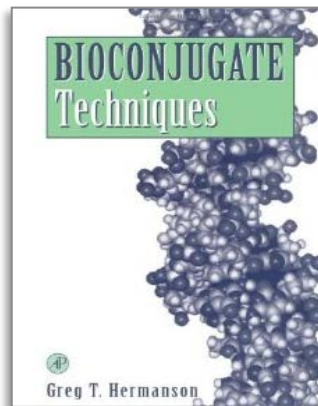


Cross-Linking Measures Distance Constraints

Yeast Enolase 1



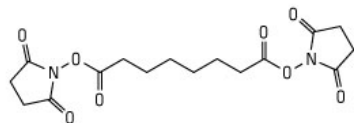
Pierce Portfolio of Chemical Cross-linking Reagents



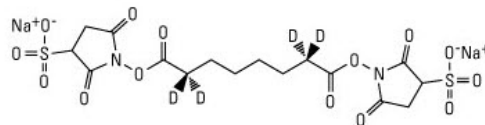
1994



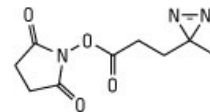
2016



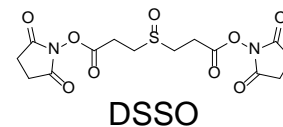
DSS
Disuccinimidyl suberate
MW 368.34
Spacer Arm 11.4 Å



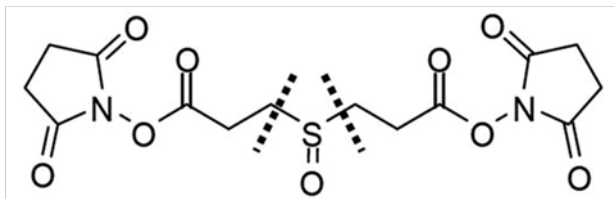
BS3-d4
Bis(sulfosuccinimidyl) 2,2,7,7 suberate-d4
MW 576.45; Crosslink Mass 142.09
Spacer Arm 11.4 Å



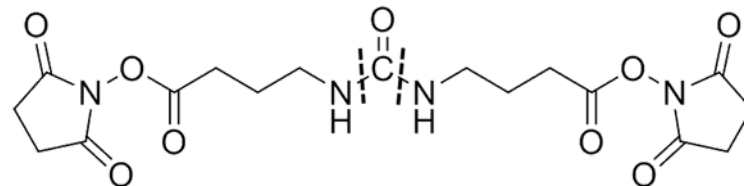
SDA
(NHS-Diazirine)
MW 225.20
Spacer Arm 3.9 Å



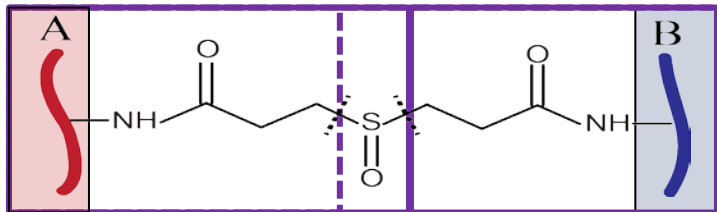
MS-Cleavable Crosslinkers Facilitate MS Analysis



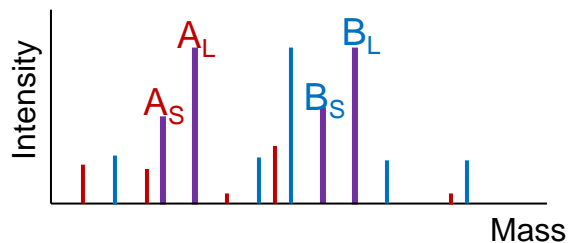
Disuccinimidyl Sulfoxide (DSSO)



Disuccinimidyl Dibutyric Urea (DSBU)



MS2



$$\Delta m = m_L - m_S = m_{A-L} - m_{A-S} = m_{B-L} - m_{B-S}$$

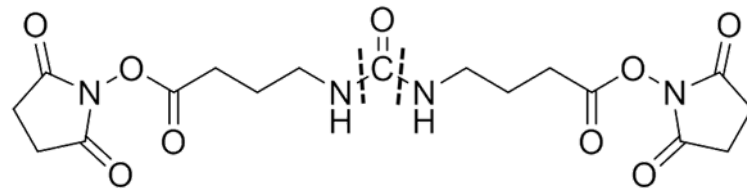
$$m_A = m_{A-L} - m_L = m_{A-S} - m_S$$

$$m_B = m_{A-B} - m_L = m_{B-S} - m_S$$

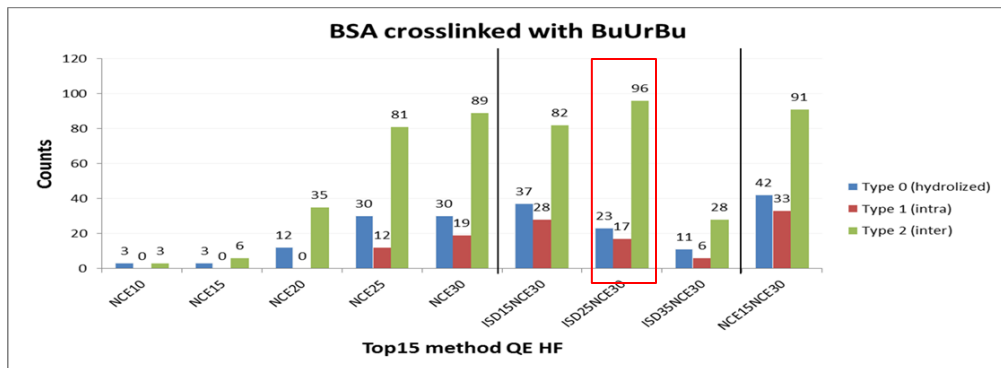
Kao, A et al, MCP, 2011;
Muller, M et al, Anal.Chemistry, 2010

Crosslinking Mass Spectrometry Workflow: MS Analysis

Top15 method	comment
NCE 10	
NCE 15	
NCE 20	
NCE 25	
NCE 30	
ISD15-NCE30	<i>In-Source</i> CID 15 eV
ISD25-NCE30	<i>In-Source</i> CID 25 eV
ISD35-NCE30	<i>In-Source</i> CID 35 eV
NCE15-NCE30	Two Top 15 experiments in one method alternating with NCE 15 and NCE 30



Disuccinimidyl Dibutyric Urea (DSBU)

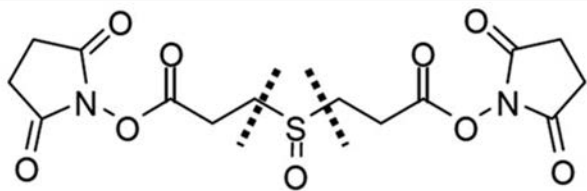


MeroX



Available since May 2017!

Crosslinking Mass Spectrometry Workflow: MS Analysis



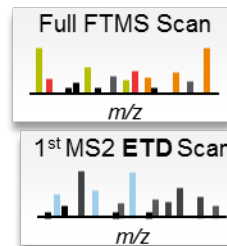
Disuccinimidyl Sulfoxide (DSSO)



Available since November 2016!

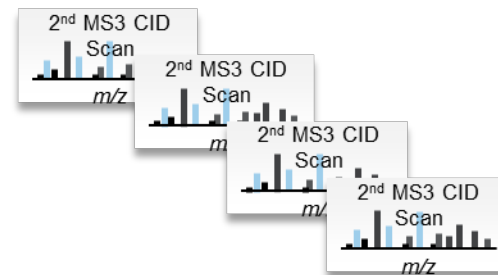
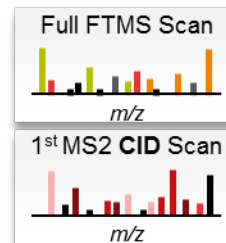
Tune 2.1

- System Templates
 - Glycans
 - Metabolites
 - Peptides-ID
 - Peptides-PTM
 - Peptides-Quan
 - Peptides-Xlink
 - Peptides-Xlink
 - Cleavable - CID-ETcD
 - Cleavable - MS2-MS3
 - Noncleavable - ETcD**
 - Custom Templates
 - My Experiments

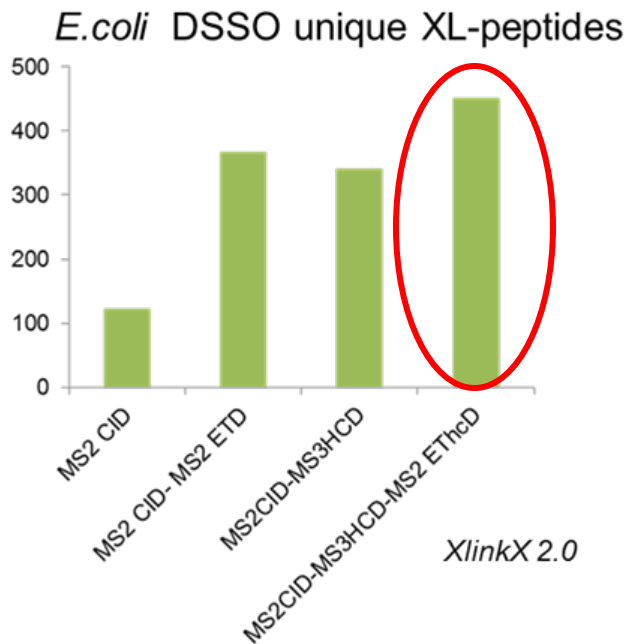


Tune 2.1

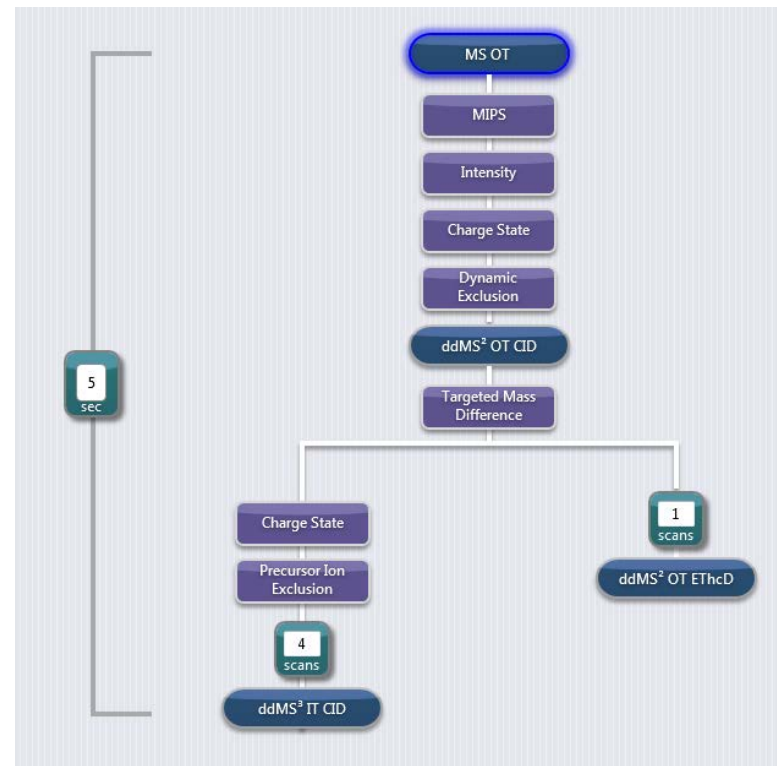
- System Templates
 - Glycans
 - Metabolites
 - Peptides-ID
 - Peptides-PTM
 - Peptides-Quan
 - Peptides-Xlink
 - Peptides-Xlink
 - Cleavable - CID-ETcD
 - Cleavable - MS2-MS3**
 - Noncleavable - ETcD
 - Custom Templates
 - My Experiments



Crosslinking Mass Spectrometry Workflow: MS Analysis

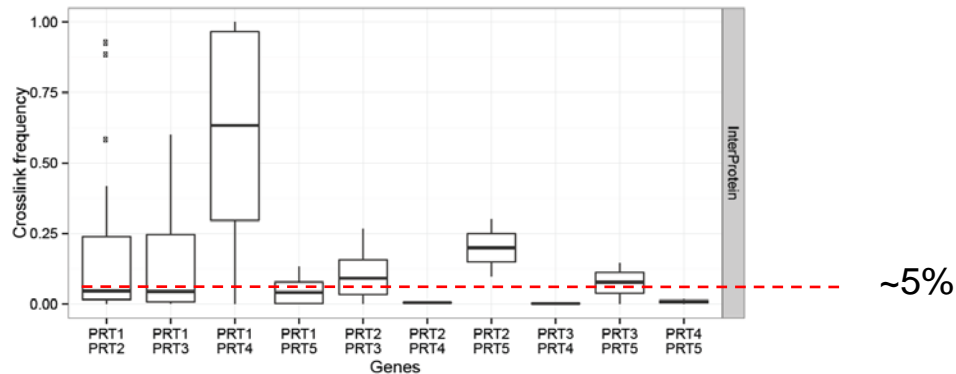


New Templates in Tune 3.0!



Liu F. et al. (2017.) Nat. Commun. doi: 10.1038/ncomms15473

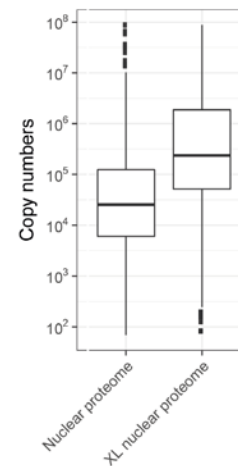
Enrichment Improves Crosslinked Peptides Identification



Crosslink reaction efficiency

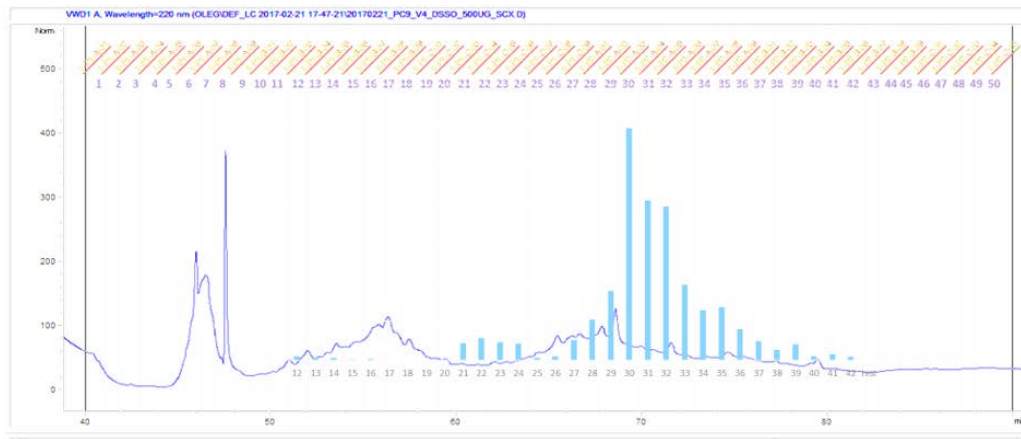
Roughly 5% of inter protein, cross linkable lysine pairs actually get crosslinked.

Not yet reaching the full depth



Still room for grabbing more detail from the interatomic studies.

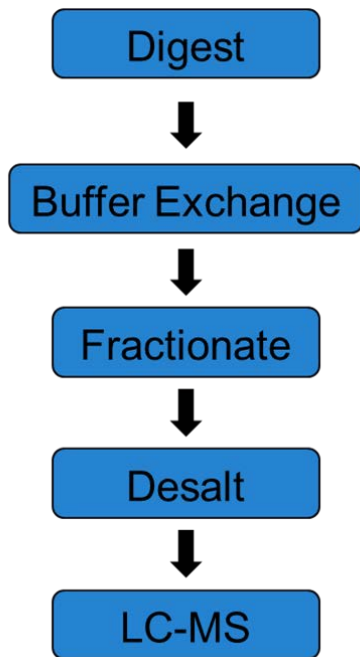
SCX fractionation



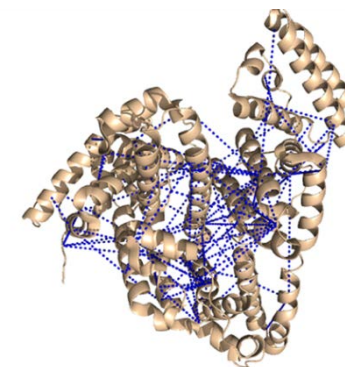
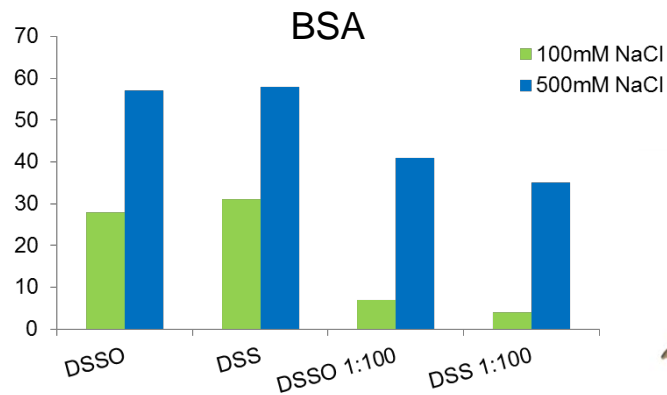
Slide courtesy of Richard Scheltema, Utrecht University

Enrichment Improves Crosslinked Peptides Identification

Strong Cation Exchange (SCX)

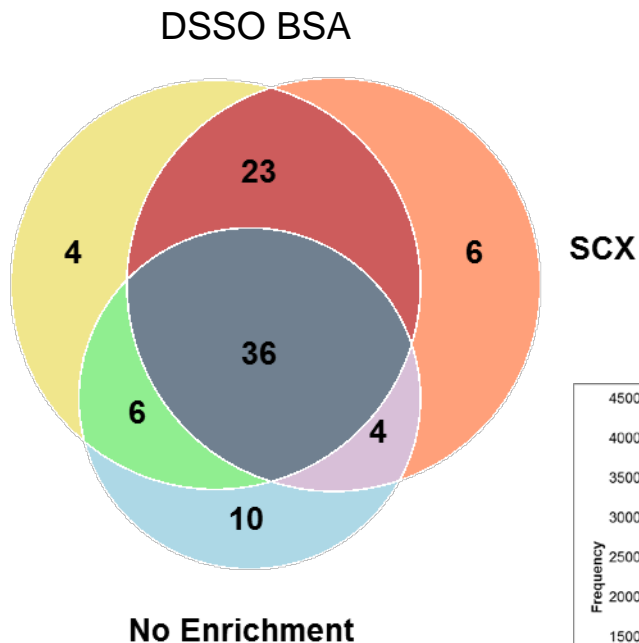
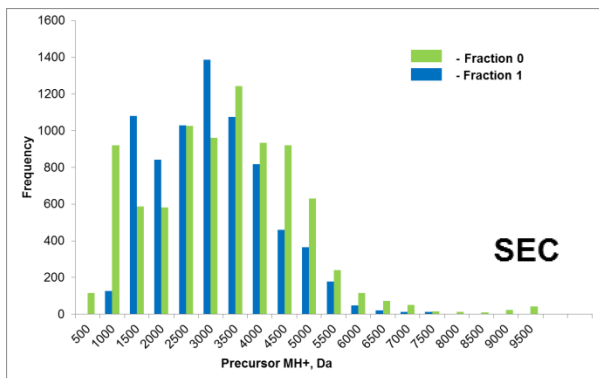


Coming Soon !



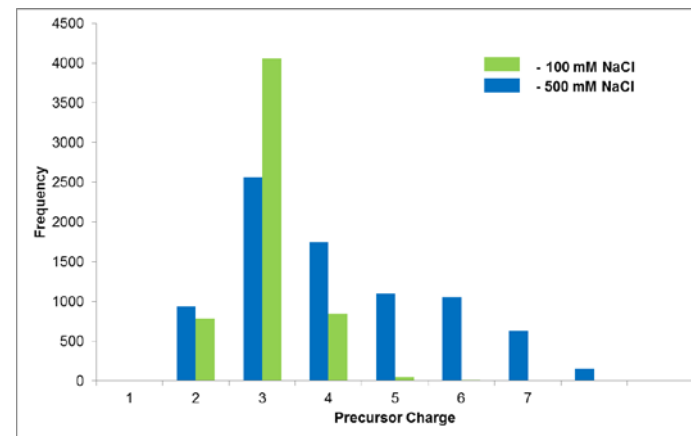
Viner et al, ASMS 2017, ThP069

Enrichment Improves Crosslinked Peptides Identification



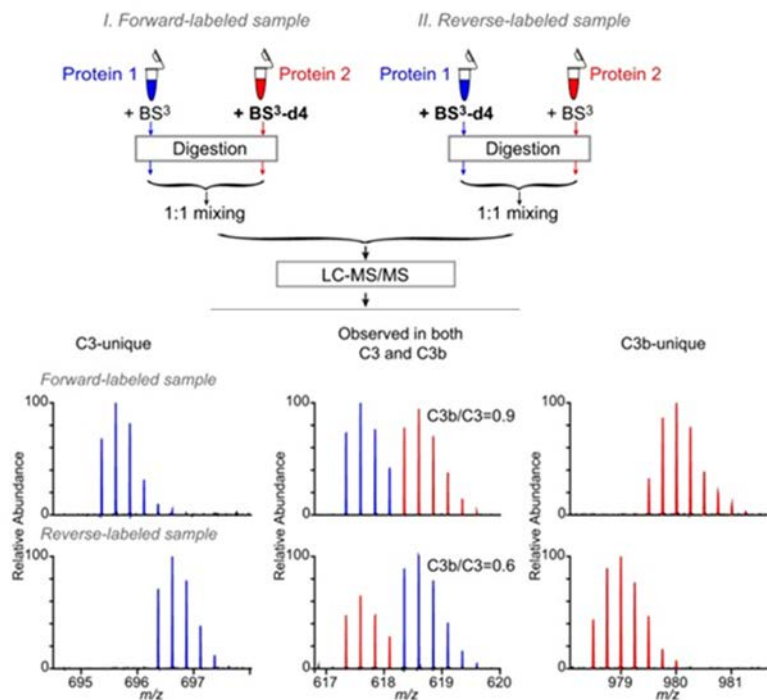
SEC Superdex Peptide PC column (GE Health)

Leitner A et al. 2012. Mol Cell Proteomics, 11(3):1-12.



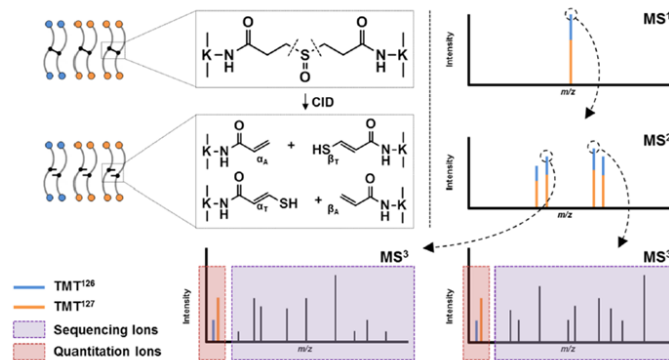
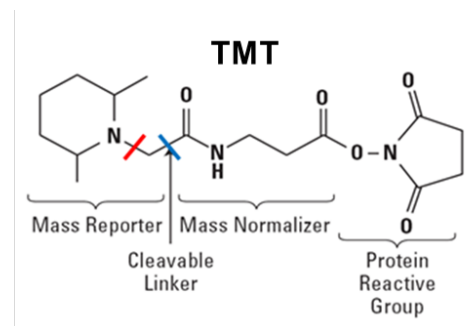
Quantitative XL-MS

QCLM: Isotopically Labeled XL



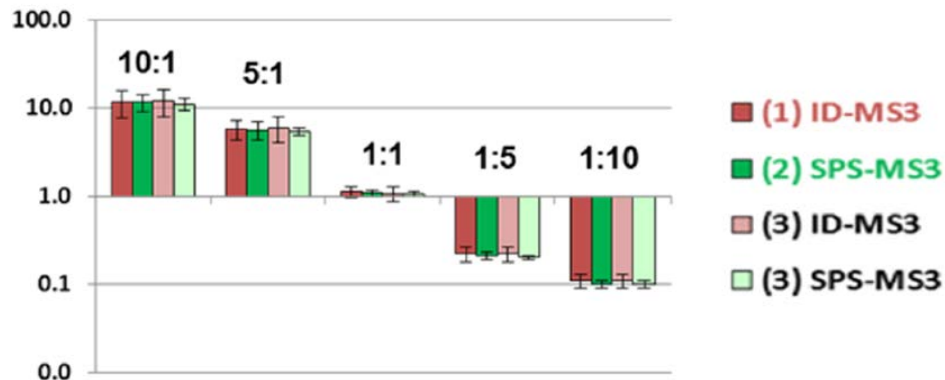
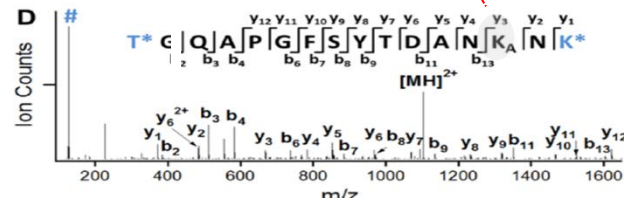
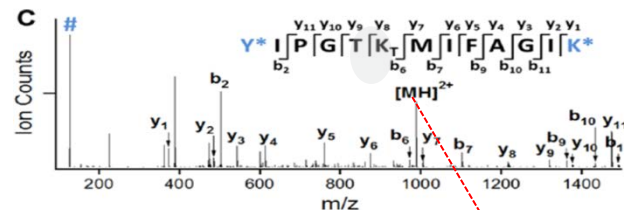
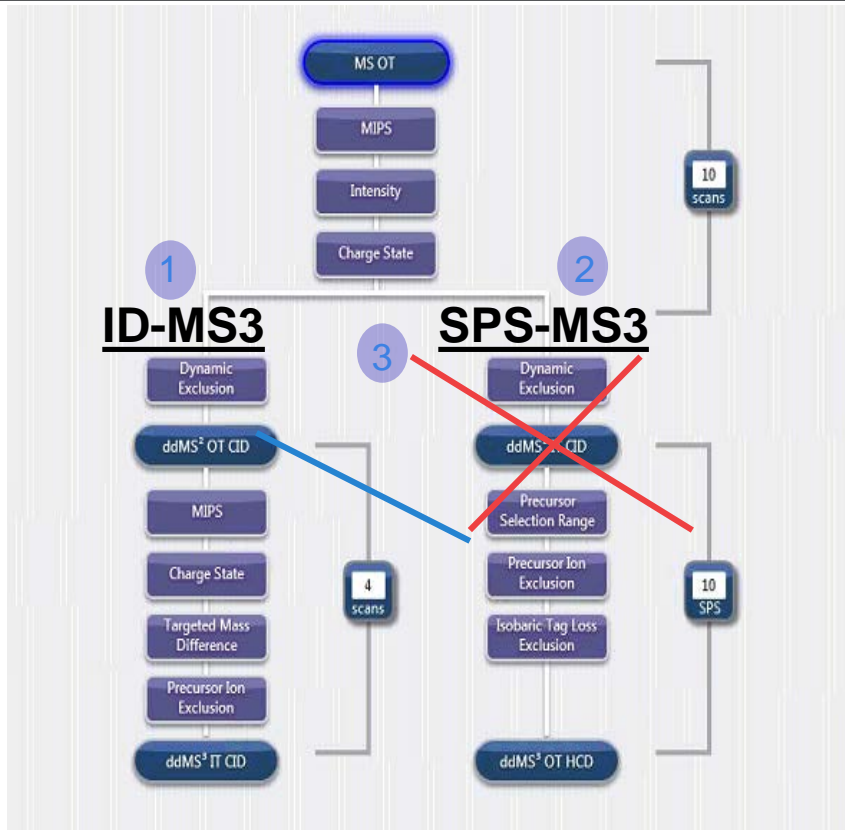
Fisher, L et al (2016)MCP, 2769-2778

QMIX: Isobaric Mass Tags+XL-MS



Yu, C et al (2016)Anal. Chem, 10301-10308

QMIX: Isobaric Mass Tagging +XL-MS

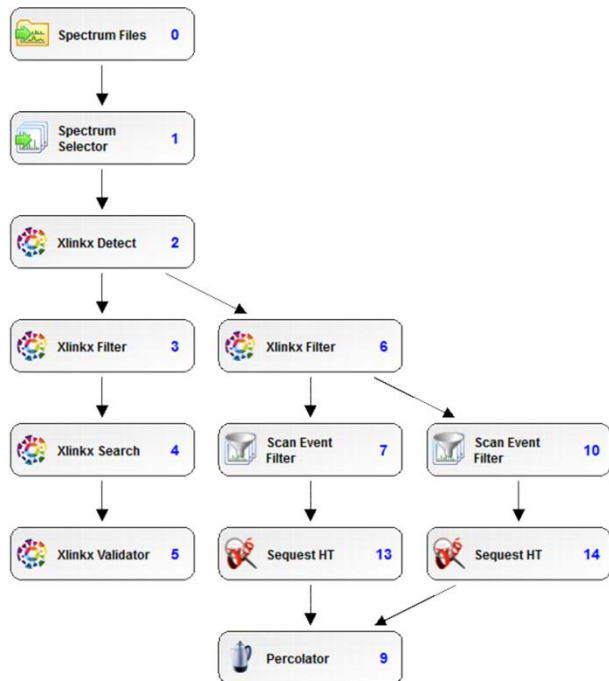


Tune 3.0 Available Now!

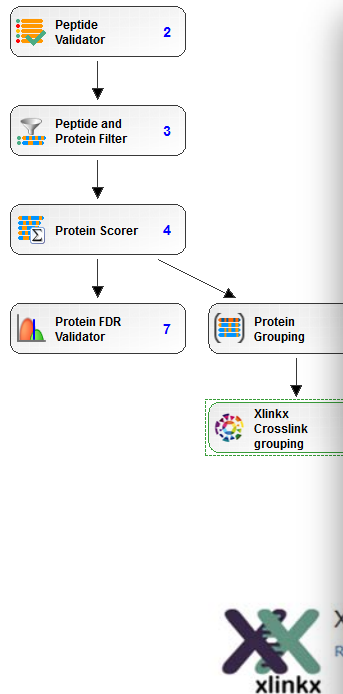
Yu, C et al(2016)Anal.Chem,10301-10308

Thermo Scientific Proteome Discoverer 2.2 XlinkX* Node

Processing Workflow



Consensus Workflow

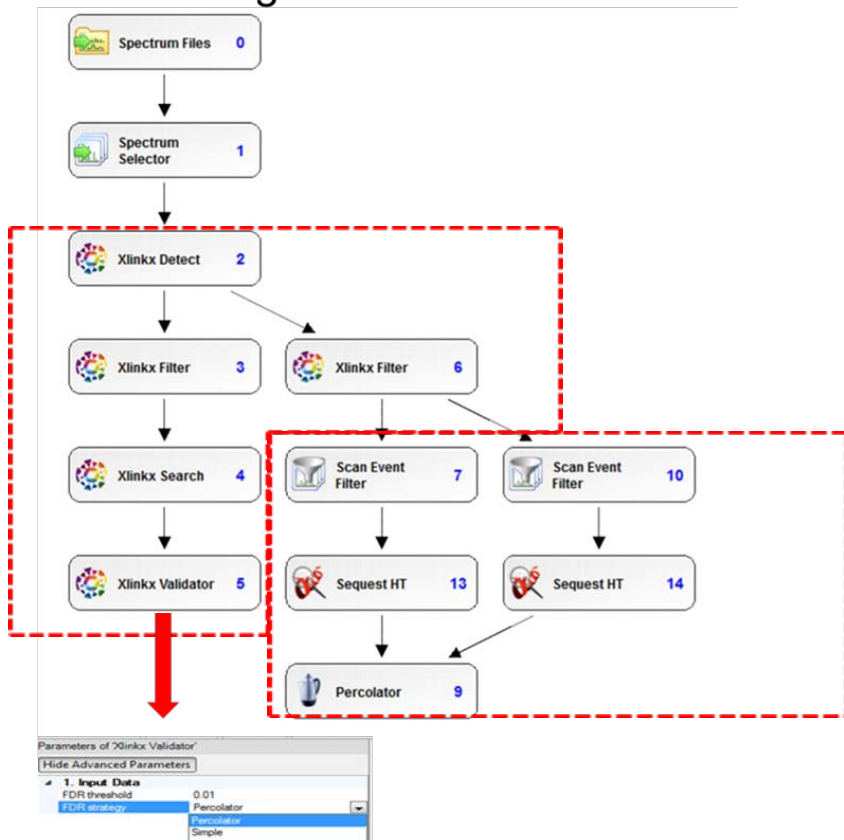


*Liu F. et al. (2015).

Liu F. et al. (2017.) Nat. Commun. doi: 10.1038/ncomms15473

XLinkX Node: Search

Processing Workflow



Definition of Crosslinks

- Addition of any linker (K-K)

Modification: DSSO (158.00376)

Is Crosslink

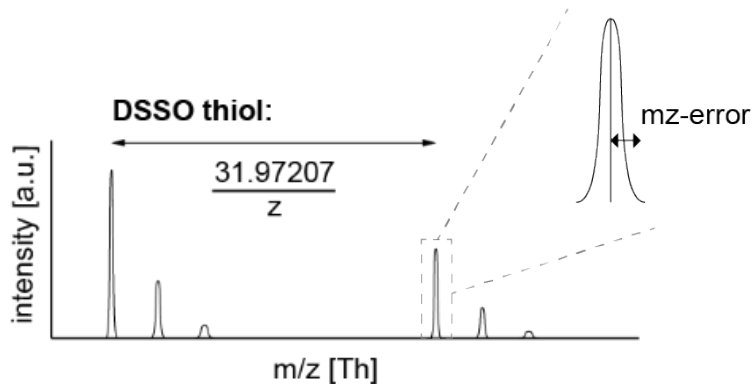
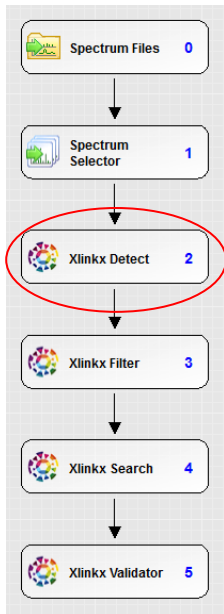
Crosslink fragments
(empty for non-clearable modifications)

Name	Type	Abbreviation	Substitution	Delta Monoisotopic Mass	Delta Average Mass	Substitution
alkene	Fragment	alkene		54.01956485	54.0474940995445	CD0H1Q2
thiol	Fragment	thiol		85.98263536	86.1135788010545	CD0H1Q2S

Database search

- Any PTM defined in PD
- Full complexity databases
- Fragmentation: HCD/CID, ETD, EThcD
- Acquisition: MS2, MS2-MS2, MS2-MS3
- **Percolator supported XL FDR control**
- TMT quantitation*

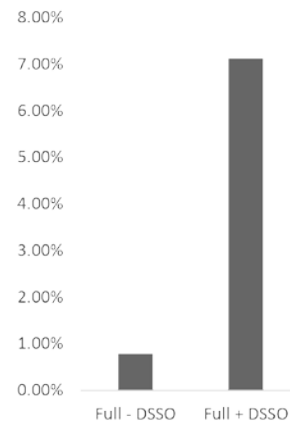
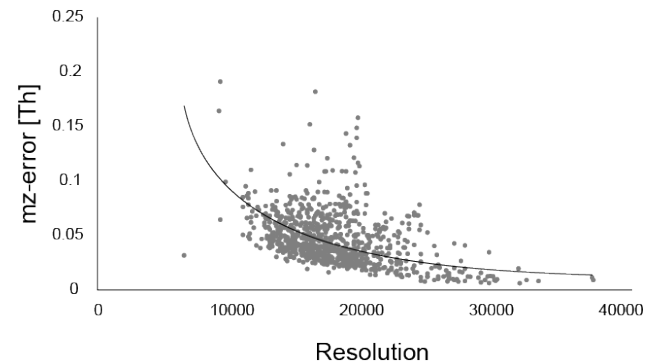
XlinkX Node: Identification-Reporter Peak Detection



Algorithm:

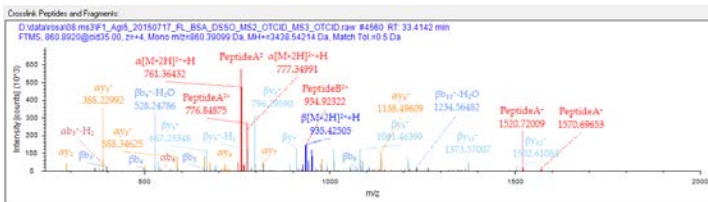
- Deisotope + TopX filter spectrum
- Determine mz-error all de-isotoped peaks
- Matching peaks of same charge based on difference +/- mz-errors
- Prioritize on higher intensity

Selectivity: ~10% FDR

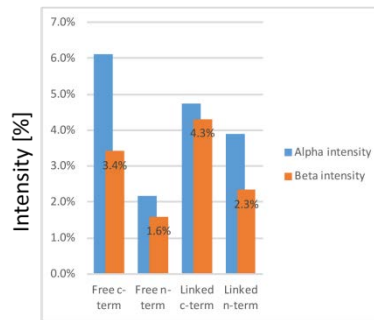
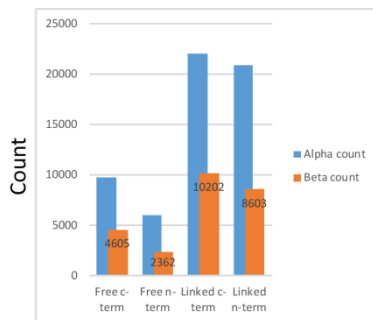
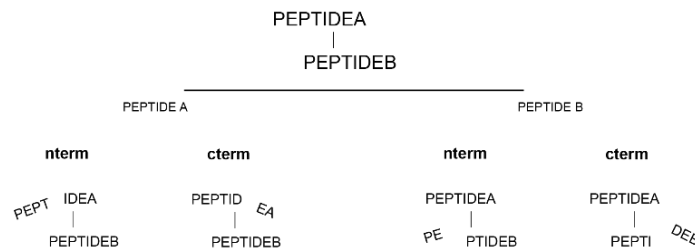


Slide courtesy of Richard Scheltema, Utrecht University

XlinkX Node: Identification - MS2



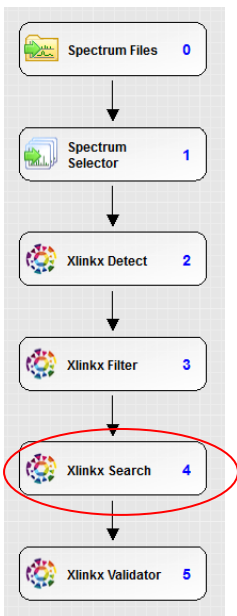
MS2 CID – Which fragments can we expect?!



Algorithm:

- Deisotope + TopX filter
- We match both sets of fragments
- Target / decoy search – score is a now standard likelihood calculation as developed by Olson et al; however it's fragment intensity weighted
- The lowest score is taken to prevent situations where 1 peptide is well covered and the other unidentifiable

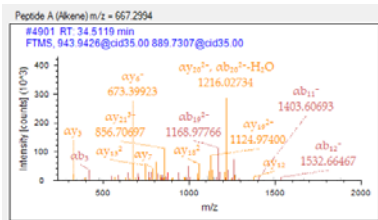
Slide courtesy of Richard Scheltema, Utrecht University



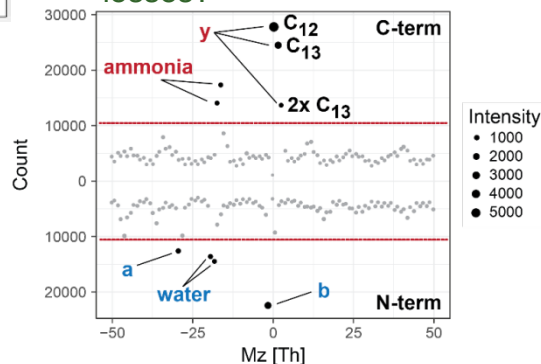
XlinkX Node: Identification - MS3

MS3 HCD – Which fragments can we expect?!

Slide courtesy of Richard Scheltema, Utrecht University

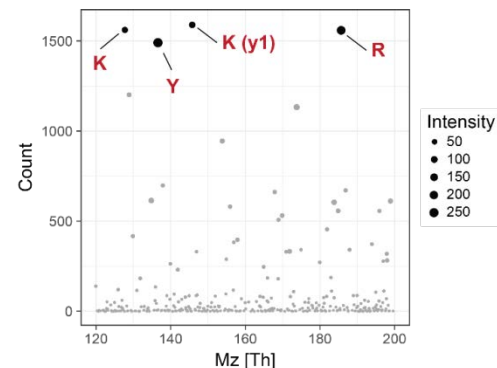


Frequent mass differences from backbone cleavage
Potential specific neutral losses?



Frequently detected masses

Potential immonium ions?



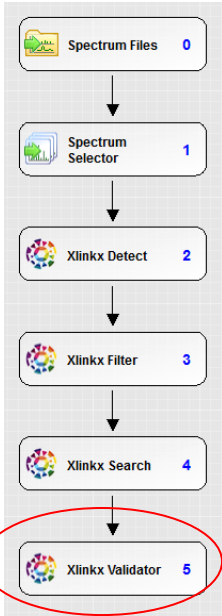
Algorithm:

- **No** deisotoping + **but do** TopX filter
- We match both with standard settings (a[1-2], b, and y)
- Target / decoy search – score is a standard likelihood calculation as developed by Olson et al

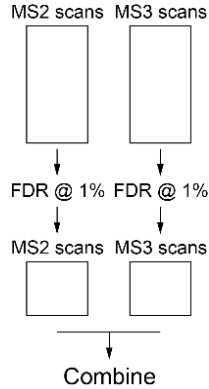
Spectral binning; Kelstrup et al; JPR; 2014

Precision mapping of the metabolome; Breitling et al; Trends in Biotech; 2006

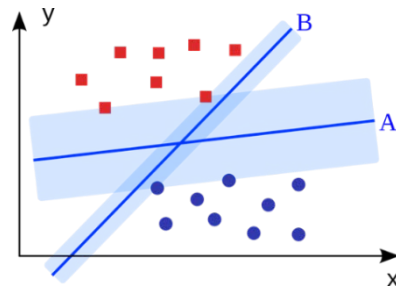
XlinkX Node: Identification - MS2-MS3 FDR Control



FDR processes MS2 and MS3 independently

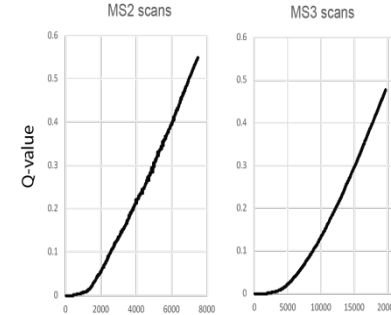


Percolator based FDR



+30% identifications...

Score based FDR (Simple)



Slide courtesy of Richard Scheltema, Utrecht University

MS3 features

1. Score
2. Score ratio
3. Mono isotopic mass
4. Charge
5. Delta mass
6. Absolute delta mass
7. Missed cleavages
8. Variable modification ratio
9. Unique variable modifications
10. Number modifiable residues
11. Number modified residues
12. Total intensity
13. Matched intensity
14. Matched intensity percentage
15. Median fragment mass error
16. Median fragment mass ppm error
17. Fraction ions matched [ion-series]
18. Matched ion intensity [ion-series]

MS2 features

1. Score
2. Score ratio
3. Mono isotopic mass
4. Charge
5. Delta mass
6. Charge [A / B]
7. Mono isotopic mass [A / B]
8. Number matches [A / B]
9. Sequence coverage [A / B]
10. Missed cleavages [A / B]
11. Fraction ions matched [free / linked]
12. Median fragment mass error
13. Total intensity
14. Matched intensity
15. Crosslink position

Percolator; Käll et al; Nat meth; 2007

Data Analysis Statistics

Data set

1. SCX fractionated full lysate (11 fractions)
2. 2hr gradients – mass difference triggered
3. Total **211,916** MS2 scans
4. Total **296,560** MS3 scans

Fasta

1. Uniprot human
2. Total 44,110 protein sequences

Computer

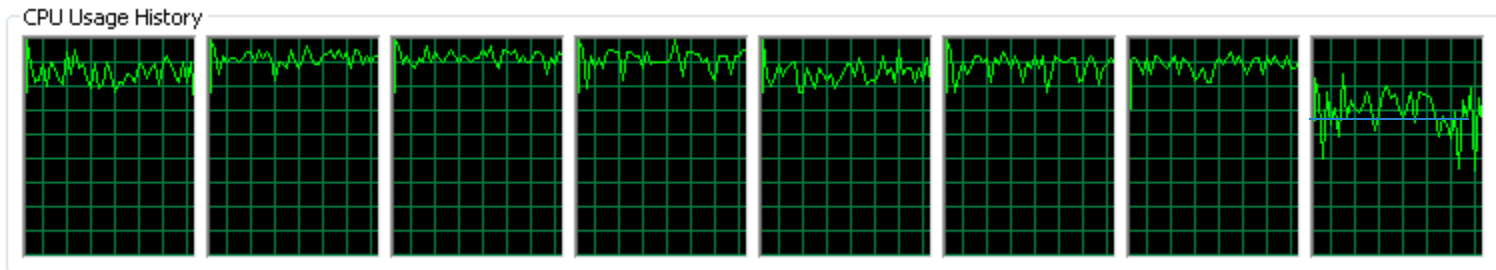
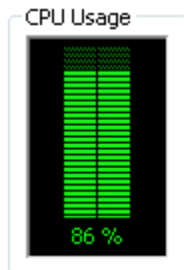
1. 8-core intel i7-4790 processor @ 3.6Hz (64-bit)
2. 16Gb internal memory
3. 222Gb SSD drive

Max memory used:	3.9 Gb
CPU's used:	7
# identified MS2 scans	1764 (of 7474)
# identified MS3 scans	2835 (of 19698)
# crosslink spectrum matches	1761
# crosslinks	1198

Step	Time [hr]
Collect scans	0.57
Detect reporter peaks	0.18
Filter scans	0.02
Generate database	0.37
Search spectra	0.28
Collect results	0.05
FDR control	0.02
Grouping	0.08

Crosslink search

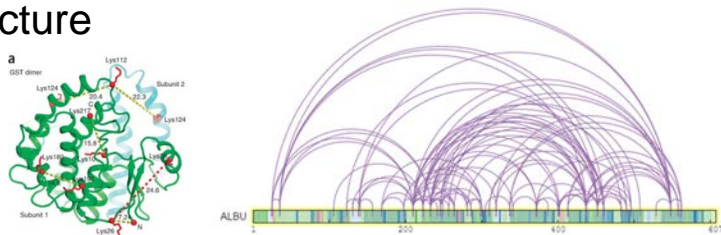
Left for OS tasks



Slide courtesy of Richard Scheltema, Utrecht University

XLinkX Node: Results

- Protein Structure

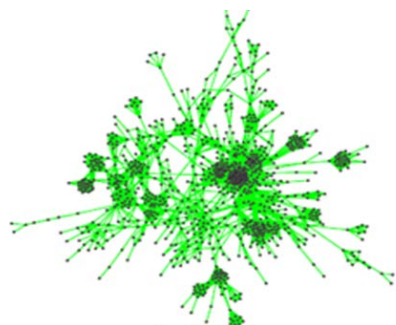


xiNET

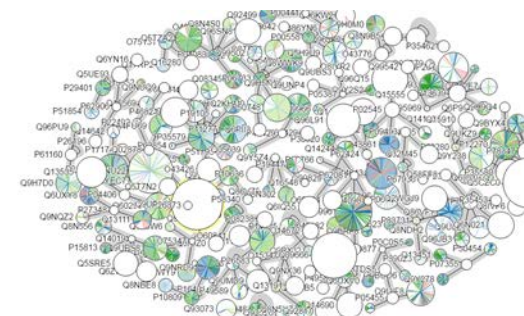
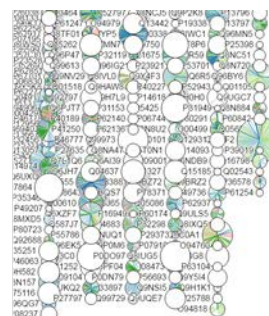
<http://crosslinkviewer.org/>

- Protein Complex

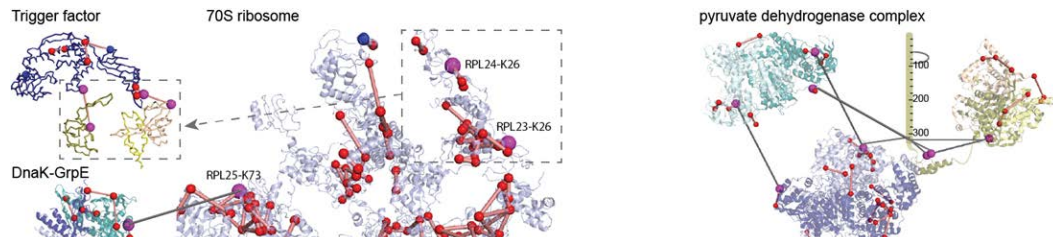
P49588	Q8N488	Q8WB9	O00154
O00264	P62258	P20618	Q58FF7
P62249	P25705	Q9HCC0	Q99436
Q95347	Q99436	Q8TAA3	P08238
P13630	P28070	Q16698	P35232
Q00341	ASD8V7	Q95171	Q99623
Q9UG63	P10609	P50990	Q9LJK0
O00231	O14782	P40227	
Q8WB9	P48556		
Q13085			



- Interactomics

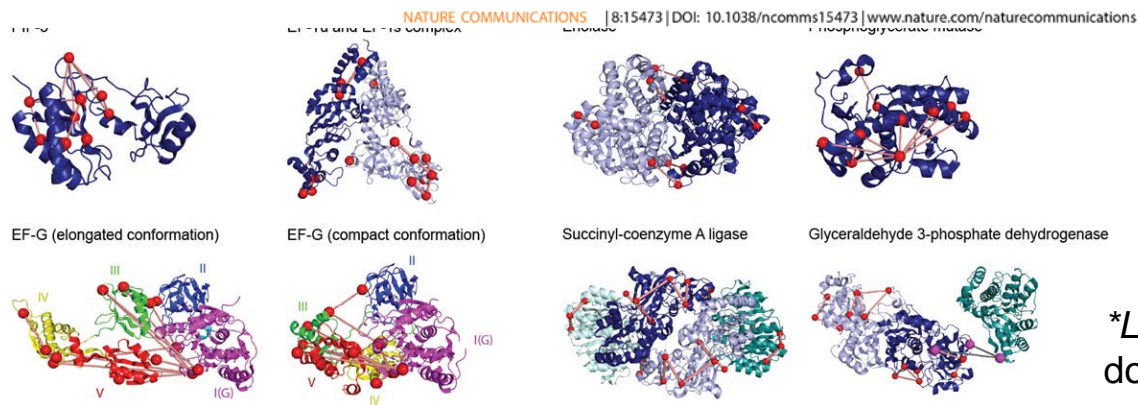


Proteome-wide XL-MS application (*E.coli*)- 1158 Unique XL*



Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification

Fan Liu^{1,2,*}, Philip Lossl^{1,2,*}, Richard Scheltema^{1,2}, Rosa Viner³ & Albert J.R. Heck^{1,2}



*Liu F. et al. (2017.) Nat. Commun.
doi: 10.1038/ncomms15473

Acknowledgements



Ryan Bomgarden
Chris Etienne
Sergei Snovida
Kay Opperman
Mathias Muller
Derek Bailey
Kai Fritze
Bernard Delanghe
Torsten Ueckert
Shijun Li
David Horn
Andreas Huhmer
Terry Zhang
Romain Huguet
Chris Mullen
Alexander Makarov
Linda van Driel
Huiling Liu
Jenny Ho
Vlad Zabrouskov



Universiteit Utrecht

Fan Liu
Richard Scheltema
Albert Heck



Clinton Yi
Craig Gutierrez
Lan Huang