This year the interest group meeting focused on discussing general workflows, current status and future perspectives of the three areas covered in this group.

Three contributors gave 15-minute technique-oriented presentations followed by discussion of their topics with the audience. The presenters covered a broad spectrum of each technique and provide useful information for experts and new comers. Ample time was devoted to an engaging discussion of each topic in turn. New and experienced researchers were able to ask questions and offer comments and suggestions. In general, participants found the presentations informative and discussions interesting. Questions were contributed about equally from the floor and electronically.

The contributors were

- Prof. Jim Bruce University of Washington
- Prof. Joshua Sharp, University of Mississippi
- Dr. Mikklos Guttman, University of Washington

Some of the major points of discussion were:

- Current workflows of cross-linking mass spectrometry in studying protein-protein interactions especially in living cells
- Experimental conditions for successful cross-linking reactions
- Data analysis and visualization of cross-link data
- General aspects of FPOP experiments
- Applications of FPOP in probing protein interactions
- New approaches for HX-MS experiments

The slides presented follow this report.
Cross-linking – what where and how

Jim Bruce
University of Washington

• General workflows for XL-MS
• Strengths and weaknesses of approaches
• New developments and applications
• Current challenges and advice for newcomers

New covalent bonds hold **structural** information

Cross-linked interactors can be found without identification of cross-linked sites.
What question is to be pursued?

- Purified protein or complex
- Complex mixture
- Cells, tissues, plants, animals
- Quantitation?

Workflow choices become more limited as sample complexity increases

Cross-linking mass spec: general workflows

I) Choose chemistry: Amine reactive, photoactivatable, zero length

A) Frequency of amino acid/reactive group

B) Reaction Specificity
   
   XXXKXXXK/R

C) Solubility
   Protein conc. in cells ≈ 300mg/ml
   Even multiple rounds of cross-linking causes little loss in bacterial cell viability
Cross-linking mass spec: general workflows

II) Cross-linking reaction conditions

A) pH – usually neutral 7-8
B) Buffers used – choice based on XL reactive groups, 170mM Na₂PO₄, pH 8.0 frequent choice
C) Temperature, 4°C, RT?
D) Typical XL time – 30 Min.
E) Xlinker concentration? In vivo, as high as possible [ ]final mM, single protein, titration to determine

Cross-linking mass spec: general workflows

III) Sample preparation

A) Get rid of unreacted cross-linker? If so, how?
   If cells, wash with PBS, All can use FASP, Gel, TCA ppt., others
B) Digestion: Enzyme choice? For lysine-reactive XL, trypsin is common choice
   XXXKXXXK/R
C) Enrichment? Strong cation exchange (most x-linkers), affinity enrichment (tag needed)
Cross-linking mass spec: general workflows

IV) Mass spectrometry analysis

A) Single protein – accurate mass LC/MS/MS, selection of higher charge states
B) More complex samples benefit from MS cleavable cross-linkers – LC/MS/MS/MS, LC MS/MS ETD/HCD, others

V) Database search

A) Single or few proteins, non-cleavable linkers – xQuest, P-link, others, chimeric MS\(^2\) spectra must be searched. Database grows with \(n^2\), where \(n\) is # of peptides.
B) Complex samples, cleavable linkers, MS\(^3\) data – mass relationships enable unambiguous peptide mass ID, existing proteome search engines- Sequest, Mascot, etc. – MS/MS of single peptides. Advanced search tools XlinkX ID from combined ETD and HDC spectra use mass difference of long and short arm, others.
Cross-linking mass spec: general workflows

VI) Then what?

A) Single or few proteins with pdb files, manually visualize links with favorite viewer, jmol, molsoft, NGL, Docking: Patchdock, Haddock, iDock, others. Distance constraint: empirically determine

B) Complex samples; need network viewer: cytoscape, xiNet, Xlink-DB, many others. In general, most PPI networks tools are not designed with Xlink data in mind.

C) Capabilities are emerging

XLinkDB 2.0 Database and Tools to Store, Visualize, and Predict Protein Interaction Topologies
Upload Your Crosslink Data

Upload New Data:

File upload: Browse... No file selected
Choose organism: A. baumanni
Experiment name:
Lab name:
I want my published data to be available to the public?

Upload

Cross-link data text file

<table>
<thead>
<tr>
<th>Peptide 1</th>
<th>Protein 1</th>
<th>Pos 1</th>
<th>Peptide 2</th>
<th>Protein 2</th>
<th>Pos 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LASSLTTKASSFK</td>
<td>Q0DB77</td>
<td>7</td>
<td>DQWTVYKDQKELPPKL</td>
<td>Q9DB855</td>
<td>9</td>
</tr>
<tr>
<td>NAFFGAPLTVGKQFK</td>
<td>Q0T473</td>
<td>8</td>
<td>EYLDGSQGKDPMIGR</td>
<td>Q07417</td>
<td>11</td>
</tr>
<tr>
<td>FFGGDPITAGKVPFK</td>
<td>Q6K341</td>
<td>10</td>
<td>VALSPAVNGKTVQIDFGVNSVYVESEAGASK</td>
<td>Q61941</td>
<td>12</td>
</tr>
<tr>
<td>IVPENELIQGRKAK</td>
<td>Q6ZC36</td>
<td>10</td>
<td>GMLVETSVLDFSEGIR</td>
<td>Q9DB855</td>
<td>2</td>
</tr>
<tr>
<td>ANVAKFGVLKDFEXX</td>
<td>Q6ZC32</td>
<td>13</td>
<td>LASLSEEKPASHRFR</td>
<td>Q6ZC32</td>
<td>6</td>
</tr>
<tr>
<td>TLLKQDDNGFVGR</td>
<td>Q6ZM31</td>
<td>3</td>
<td>DSIFSNLQDGKYOQEK</td>
<td>Q6ZM31</td>
<td>13</td>
</tr>
<tr>
<td>TPQDKDGAVKVEK</td>
<td>P5Q544</td>
<td>11</td>
<td>KGI/VNEQFLIQDR</td>
<td>P5Q544</td>
<td>0</td>
</tr>
<tr>
<td>KVDHELQAEIER</td>
<td>Q6ZM36</td>
<td>0</td>
<td>QAEAEIESQTYNLSKFR</td>
<td>Q6ZM36</td>
<td>14</td>
</tr>
<tr>
<td>GIFTISQEQMNK</td>
<td>Q6ZM31</td>
<td>9</td>
<td>KGMSIVQVQREPGSIK</td>
<td>Q6ZM31</td>
<td>0</td>
</tr>
<tr>
<td>VQGLQDQPVLALVPYPIYK</td>
<td>Q6ZB20</td>
<td>8</td>
<td>GQKVLDSQAPKVPQPETGTR</td>
<td>P64680</td>
<td>2</td>
</tr>
<tr>
<td>LADDYGGKIQQQR</td>
<td>P48962</td>
<td>16</td>
<td>VKLQVQHASQK</td>
<td>P48962</td>
<td>1</td>
</tr>
<tr>
<td>FNPEDHSTGKDGR</td>
<td>Q6ZK60</td>
<td>10</td>
<td>XPKLEAQAEDELPR</td>
<td>Q6ZK60</td>
<td>0</td>
</tr>
</tbody>
</table>

Representation of Uploaded Crosslinked Peptides in XLinkDB

Upload Cross-links to XLinkDB

Dock Cross-links

Compute intra-protein cross-link distances

Compute inter-protein cross-link distances
View Your Crosslink Data: Cytoscape

Explore protein-protein interactions
View Your Crosslink Data with JSmol

JSmol Viewer

View single crosslink

View all crosslinks

NGL Viewer

View Your Crosslink Data with NGL
View Your Crosslink Data with NGL

NGL Viewer

Predicted protein interaction topology with idock

ACTC1-ACTN4 Interface

Generate PRM Transitions of Crosslink

NGL Viewer

Generate PRM Transitions of Crosslink
Covalent Labeling: Experimental Considerations
Joshua Sharp
Assistant Professor, Department of BioMolecular Sciences
The University of Mississippi School of Pharmacy

Covalent Labeling

- **Stable** (usually) chemical modification of amino acid side chains
- **Topographical analysis**: Generally thought to report on changes in surface accessibility
- **Differential analysis**: usually compare a polypeptide sequence in two different structural states; report on relative differences

Wang et al., Structure (2011) 19, 1138-1148
Experimental Overview

1. Chemical Probe Selection
   a) Selective
   b) Broad
2. Labeling
   a) Exposure control
      a) Labeling reagent concentration
      b) Labeling time
      c) Labeling temperature
   b) HOS maintenance
3. Protease Selection and Digestion
4. Quantification
   a) Peptide level
   b) Amino acid resolution
5. Data Interpretation
Chemical Probe Selection

- **Selective Probes**: Target specific functional groups
  - Almost 50 year history in protein chemistry
  - Targeted analysis, best for testing specific hypotheses (e.g., primary amine labeling to identify charge-charge interactions in heparin binding, Ori et al. Mol Cell Proteomics (2009) 8, 2256-2265)
  - Relatively simple chemistry
  - Easily controlled reagent quantities and reaction times
  - Chemistry must be compatible with maintenance of HOS
  - New innovation in isotopic labeling, new reagents (e.g., Zhang et al. J Am Soc Mass Spectrom (2016) 27, 178-181)

- **Broadly Reactive Probes**: Simultaneously labels multiple amino acid classes
  - Best for general topographical analysis
  - More complex chemistry, more complex analysis
  - Does not label all amino acids equally
  - Reactive group formed in situ; reagent quantities and reaction times more difficult to control
  - Very sensitive to buffer composition
  - New innovation in reagent activation methods, new reagents (e.g., carbene reagents: Manzi et al. (2016) Nat Commun 7, 13288)

Labeling: Maintaining HOS

- Major Issue: Covalent labeling alters the biophysical properties of amino acids
  - They will alter your protein HOS
  - Probing proteins with altered HOS results in artifactual results
  - Three common methods for ensuring HOS is retained

  1. Functional assay
  2. Biophysical assay
  3. Kinetic analysis
Labeling Quantification: Isotopic Labeling

• Problem: Several labeling reagents significantly alter the retention time and ionization efficiency of peptides and/or the digestion pattern of proteases, making relative quantification by ion intensity unreliable.

• Selective probes are often saturatable; based on this, isotopic labeling is possible.

• Protein is labeled in native form, then denatured and all reactive groups labeled with isotopomer.

• Without isotopic labeling, protease must be carefully matched with reactive groups, ionization efficiency differences taken into account.


Labeling Quantification: Amino Acid Resolution

• Problem: Quantitatively resolving how much labeling occurs on each amino acid in a peptide with multiple targets.

• Mostly a problem with broad labeling reagents; usually few labeling site(s) per peptide of specific reagents.

• Peptide modification isomers will usually separate (at least partially) by most peptide chromatography methods.

• In ergodic MS/MS, modification site often alters fragmentation pathways.

• Three basic approaches:
  1. ECD/ETD
  2. UHPLC Resolution
  3. UHPLC/Product Ion Chrom.

Li et al., J Am Soc Mass Spectrom (2013) 24, 1767-1776


Kaur et al., Mol cell Proteom (2015) 14, 1159-1168
Data Interpretation: What Are We Measuring?

- Broadly Reactive Probes: Higher reaction rates, lower activation energies, less effects from “microenvironment”
- As reactivity of target decreases, effect of “microenvironment” (e.g. protein sequence inductive effects) increases

<table>
<thead>
<tr>
<th>Reactivity Level</th>
<th>Reaction Rate Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly Reactive</td>
<td>$k/k_p &gt; 10$</td>
</tr>
<tr>
<td>Moderately Reactive</td>
<td>$4 &lt; k/k_p &lt; 10$</td>
</tr>
<tr>
<td>Poorly Reactive</td>
<td>$k/k_p &lt; 4$</td>
</tr>
</tbody>
</table>

Normalized Protection Factor (NPF) $R_i = \frac{\text{Slope}_{R_i}}{k_p/k_p}$


Wide Reactive Probes

- If we solely probe highly reactive amino acids or if we correct for sequence context effects, we can directly measure $<\text{SASA}>$ by rate of covalent modification
- Accuracy of HR-HRPF $<\text{SASA}>$ measurements are comparable to X-ray crystallography structures when compared to $<\text{SASA}>$ from MD trajectories

$R^2 = 0.62$  
$R^2 = 0.72$  

Quantitative Topography Measurements and Computational Modeling

- Quantitative changes in <SASA> can help in understanding binding events
  Gustavsson et al., Nat Commun (2017) 8, 14135

- Absolute <SASA> measurements can be used to experimentally score models, differentiate high quality models from inaccurate models

Sharp Group

Current Members
- Hao Liu
- Nilofar Khaje
- Dr. Charles Mobley
- Mohammad Riaz
- Dr. Quntao Liang
- Dr. Surendar Tadi
- Dr. Sandeep Misra
- Chelsea Suppinger
- Lindsey Miller
- Joseph Mason
- Delaney Mason
- Sydney Watson

Alumni
- Dr. Boer Xie (St. Jude Children’s Research Hospital)
- Dr. Zixuan Li (NYU School of Medicine)
- Dr. Yulun Chiu (MD Anderson Cancer Center)
- Dr. Xiaoyan Li (Massachusetts General Hospital)
- Dr. Rongrong Huang (Greenwood Genetic Center)
- Dr. Caroline Watson (CDC)
- Dr. Dandan Zhou (CCRC, University of Georgia)
- ViLinh Tran (Emory)
- Jessica Saladino
Acknowledgements

Joshua Sharp

jsharp@olemss.edu

(662) 915-1758

Resource for Integrated Glycotechnology

- Prof. Geert-Jan Boons—GAG synthesis
- Prof. Kelley Moremen—Protein expression
- Prof. James Prestegard—NMR spectroscopy
- Prof. Rob Woods—MD simulations
- Prof. Lianchun Wang—GAG cell biology
- Prof. Jon Amster—Purified GAG sequencing

Funding

- NIGMS P41 GM103390
- R01 RGM096049A
- P20 GM104932

- NSF CHE 1608685

Financial Conflict of Interest Disclosure

- J.S.S. has a substantial ownership interest in GenNext Technologies, Inc., a small company designing and developing standardized and automated technologies for hydroxyl radical protein footprinting for the biopharmaceutical industry
HDX, Covalent Labeling & Crosslinking Interest Group

ASMS 2017

Mike Guttman
University of Washington, Seattle

HDX-MS

H$_2$O $\rightarrow$ D$_2$O
HDX-MS

$\text{H}_2\text{O} \rightarrow \text{D}_2\text{O}$
1. Label and Lock

- Initiate exchange
  - pH ~7, 25°C, native
- Halt exchange (quench)
  - pH 2.5, 0°C, denatured
- Varies with pH & temperature
  - 2 fold change with 0.3 pH units or 7°C
- Minimize back-exchange
  - t_{1/2} ~ 30 min
  - Fast & reproducible

2. Denature & Reduce

- 8M Urea or 2M Guanidine
  - (careful with proteases)
2. Denature & Reduce

- 8M Urea or 2M Guanidine
  - (careful with proteases)
- Disulfide bonds
  - TCEP
  - Online electrochemical reduction
    - Commercially available


2. Digestion

- Limited proteases (pH 2.5)
<table>
<thead>
<tr>
<th>protease</th>
<th>organism</th>
<th>availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepsin</td>
<td>Sus scrofa</td>
<td>✔</td>
</tr>
<tr>
<td>Aspergillopepsin</td>
<td>Aspergillus Saitoi</td>
<td>✔</td>
</tr>
<tr>
<td>Rhizopepsin</td>
<td>Rhizopus sp.</td>
<td>✔</td>
</tr>
<tr>
<td>Nep2</td>
<td>Nepenthes gracilis</td>
<td>✗</td>
</tr>
</tbody>
</table>

- High pressure (12 kpsi)
2. Digestion

• Limited proteases (pH 2.5)

<table>
<thead>
<tr>
<th>protease</th>
<th>organism</th>
<th>availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepsin</td>
<td>Sus scrofa</td>
<td>✓</td>
</tr>
<tr>
<td>Aspergillopsin</td>
<td>Aspergillus Saitoi</td>
<td>✓</td>
</tr>
<tr>
<td>Rhizopepsin</td>
<td>Rhizopus sp.</td>
<td>✓</td>
</tr>
<tr>
<td>Nep2</td>
<td>Nepenthes gracilis</td>
<td>✗</td>
</tr>
</tbody>
</table>

• High pressure (12 kpsi)

• Deglycosylate glycoproteins

3. Resolve peptides (LC)

• UPLC columns
  – Shorter run times
  – Higher resolution
  – Peak capacity

• Ultra low temperatures

• Alternatives – CE-MS

4. MS analysis

• Lower source temperature
• High resolution helps

[Graph showing mass spectra]

4. MS analysis

- Lower source temperature
- High resolution helps
- Ion mobility helps
- Site-specific analysis
  - ETD
  - Top-down

4. Data processing

- Automated peptide ID and deuterium integration
- Many software options
- Simple centroid is most common
4. Data processing

- Automated peptide ID and deuterium integration
- Many software options
- Simple centroid is most common
- Often insufficient
  
  Zhang J., (2013) JASMS

- Alternative approaches
  
  Guttman M. et al., (2013) JASMS
  Zhang J., (2013) JASMS

- Bimodal deconvolution

What can we learn?

- Exchange governed by structure and dynamics
- Which regions are well folded?
- Deuterated control
  - Denaturation & exhaustive deuteration
  - Deuteration after proteolysis
**Interface mapping by HDX**

**Can**
- Localize regions at the interface
- All allosteric effects

**Can’t**
- 3D structural information
- Single residue resolution
- Contact residues

**Remaining challenges**

- Can’t quantitatively compare HDX data collected at different times
  - Exchange conditions
  - Back exchange variation
  - Internal exchange reporter (PPPI)

Remaining challenges

• Can’t quantitatively compare HDX data collected at different times
  – Exchange conditions
  – Back exchange variation
  – Internal exchange reporter (PPPI)

• Limited temporal sampling
  – 10 sec to 10 hours may miss many relevant kinetic regions


Hamuro Y., 2017 JASMS

Remaining challenges

• Can’t quantitatively compare HDX data collected at different times
  – Exchange conditions
  – Back exchange variation
  – Internal exchange reporter (PPPI)

• Limited temporal sampling
  – 10 sec to 10 hours may miss many relevant kinetic regions
• Get the most from your data
  – Look beyond the centroid

Hamuro Y., 2017 JASMS