Reasons to be excited about current efforts in glycoproteomics

Nicholas M. Riley
SUMS Seminar Series, October 1, 2020

@riley_nm1

Casalino et al., ACS Central Science, 2020, 10.1021/acscentsci.0c01056
EVERY CELL HAS A GLYCOCALYX

collection of glyco-conjugates at the cell surface

EM of a myocardial capillary stained with alcian blue

(GLYCO)BIOLOGY AT THE CELL SURFACE

Kuo et al., Nature Physics, 2018, 14: 658–669
DE-GLYCOPETIDES: EASIER, BUT AT A PRICE
INTACT GLYCOPEPTIDES: MICROHETEROGENEITY

Paucimannose | High Mannose | Complex | Hybrid
---|---|---|---

PO₃

N-terminus | pSerine | pThrreonine | N-Glycosylation | C-terminus

Glucose (Hex) | Galactose (Hex) | N-Acetylglucosamine (HexNAc) | Fucose (Fuc) | Neuraminic Acid (Sialic) (NeuAc)
MAPPING THE GLYCOPROTEOME

https://futuretravel.today/bay-area-2050-the-bart-metro-map-ab83b22d3d8b
THE RISE OF O-GLYCOPROTEASES

Isolating Glycoproteins
Generating Glycopeptides
Enrichment & Separations
Ionization
Tandem MS (MS/MS)
Interpreting Spectra
Quant and Context

O-glycoproteases
OpeRATOR: broadly active on glyco-S and -T

Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO)

Weiming Yang, Minghui Ao, Yingwei Hu, Qing Kay Li & Hui Zhang

Deciphering Protein O-Glycosylation: Solid-Phase Chemoenzymatic Cleavage and Enrichment

Shuang Yang, Philip Onigman, Wells W. Wu, Jonathan Sjogren, Helen Nyhlen, Rong-Fong Shen, and John Cipollo
MS-AMENABLE O-GLYCOPEPTIDES

Minimum Cleavage Motif

Malaker, Pedram, Bertozzi et al., PNAS 116(15) 7278-7287 (2019)
An enzymatic toolkit for selective proteolysis, detection, and visualization of mucin-domain glycoproteins

D. Judy Shon\textsuperscript{a}, Stacy A. Malaker\textsuperscript{b}, Kayvon Pedram\textsuperscript{b}, Emily Yang\textsuperscript{a}, Venkatesh Krishnan\textsuperscript{b}, Oliver Dorigo\textsuperscript{b}, and Carolyn R. Bertozzi\textsuperscript{a,c,1}
O-GLYCOPROTEASE ENRICHMENTS
O-GLYCOPROTEASE ENRICHMENTS

GlycoCatch: Genovis (OpeRATOR)
ENZYMES FOR ENRICHMENTS IN GLYCO

GlycoCatch: Genovis (OpeRATOR)

SiaFind Lectenz Kits (Engineered Sialidases)
Chemical Glycoproteomics

Krishnan K. Palaniappan and Carolyn R. Bertozzi

1Verily Life Sciences, 269 East Grand Ave, South San Francisco, California 94080, United States
2Department of Chemistry and Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, United States
Bioorthogonal Chemoenzymatic

Bioorthogonal Metabolic

MCP
MOLECULAR & CELLULAR PROTEOMICS

A Pragmatic Guide to Enrichment Strategies for Mass Spectrometry-based Glycoproteomics

Riley, Bertozzi, Pitteri, Mol. & Cell. Prot, 2020
Field Asymmetric Ion Mobility Spectrometry (FAIMS)

What are we missing by using hydrophilic enrichment? Improving bacterial glycoproteome coverage using total proteome and FAIMS analysis.

Ameera Raudah Ahmad Izaham¹ Ching-Seng Ang⁴, Shuai Nie¹, Lauren E. Bird¹, Nicholas A. Williamson² and Nichollas E. Scott¹*
FAIMS

https://yost.chem.ufl.edu/research/faims/
What are we missing by using hydrophilic enrichment? Improving bacterial glycoproteome coverage using total proteome and FAIMS analysis.

Ameera Raudah Ahmad Izaham¹ Ching-Seng Ang², Shuai Nie², Lauren E. Bird³, Nicholas A. Williamson² and Nicholas E. Scott¹*
Glycopeptide variable window SWATH for improved data independent acquisition glycoprotein analysis

Chun Zhou*, Benjamin L. Schulz**
Glyco-DIA: a method for quantitative O-glycoproteomics with in silico-boosted glycopeptide libraries

Zilu Ye, Yang Mao, Henrik Clausen and Sergey Y. Vakhrushev
<table>
<thead>
<tr>
<th>Fragmentation type</th>
<th>Resonance activation CID</th>
<th>Beam-type CID / HCD</th>
<th>Electron transfer / higher-energy dissociation (EThcD)</th>
<th>Conventional ETD / ECD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical instrument platform</td>
<td>Ion trap (linear, 3D)</td>
<td>Q-TOFs, Orbitraps</td>
<td>Orbitrap Fusion/Lumos</td>
<td>Ion traps, Orbitraps, FT-ICR</td>
</tr>
<tr>
<td>Typical bond cleavages and fragment formation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Glycopeptide information | Glycan identification*  
Peptide mass identification  
Partial glycan identification | Peptide identification** | Peptide identification**  
Site identification  
Partial glycan identification | Peptide identification  
Site identification  
Glycan mass identification |

Thaysen-Andersen et al., MCP, 2016, 15: 10.1074/mcp.0115.057638, 1773–1790
GLYCOPEPTIDE FRAGMENTATION

Fragmentation methods affect instrumentation decisions

HCD

sceHCD

ETD

EThcD

Orbitrap Exploris 480 (Thermo)
timsTOF Pro (Bruker)
Orbitrap Eclipse (Thermo)
solariX XR (Bruker)
Synapt G2-Si (Waters)
FRAGMENTATION SCORECARD

<table>
<thead>
<tr>
<th>Acquisition Speed</th>
<th>ETD</th>
<th>EThcD35</th>
<th>EThcD25</th>
<th>HCD20</th>
<th>HCD25</th>
<th>HCD30</th>
<th>HCD35</th>
<th>HCD40</th>
<th>sceHCD25+15</th>
<th>sceHCD30+10</th>
<th>sceHCD30+18</th>
<th>sceHCD35+5</th>
<th>sceHCD35+15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptide Fragmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycan Fragmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Localization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use for N-glycopeptides</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use for O-glycopeptides</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Optimal Dissociation Methods Differ for *N*- and *O*-Glycopeptides

Nicholas M. Riley, Stacy A. Malaker, Marc D. Driessen, and Carolyn R. Bertozzi*
**GLYCOINFORMATICS**

- Isolating Glycoproteins
- Generating Glycopeptides
- Enrichment & Separations
- Ionization
- Tandem MS (MS/MS)
- Interpreting Spectra
- Quant and Context

---

**Byonic**

- Available through Stanford (Protein Metrics)
- GUI with viewing options
- Flexible with fragmentation methods
- Compatible with N- and O-glycopeptides

**pGlyco2.0**

- Free (Academic software)
- GUI with limited (but still usable) options
- Designed for sceHCD (more coming soon)
- Generally only works for N-glycopeptides
RECENT BYONIC UPGRADES

Peak Filtering, Peak Annotation, and Wildcard Search for Glycoproteomics

Abhishek Roushan, Gary M. Wilson, Doron Kletter, K. Ilker Sen, Wilfred Tang,
Yong J. Kil, Eric Carlson, Marshall Bern*
Current Canon:
Peptide-centric
DB generation

O-glycopeptide from CD43 (leukosialin)

\[ \text{TGSLEPSSGASGPQVSSVK} \]

7 potential O-glycosites, allow up to 3 to be modified:

different theoretical spectra for each PTM considered
VARIABLE MOD SEARCHING

0-glycopeptide from CD43 (leukosialin)

TGSLEPSGGASGPQVSSVK

7 potential O-glycosites,
allow up to 3 to be modified:

12 common O-glycans

>60,000 combinations to consider...for one peptide
Identify peptide candidates without pre-determined modifications

MS/MS Spectrum

Theoretical Peptide Mass

Measured Mass

mass (Da)
Open database searching enables the identification and comparison of bacterial glycoproteomes without defining glycan compositions prior to searching.

Ameera Raudah Ahmad Izaham\textsuperscript{1} and Nichollas E. Scott\textsuperscript{1,\#}
Fast and Comprehensive N- and O-glycoproteomics analysis with MSFragger-Glyco

Daniel A. Polasky¹, Fengchao Yu¹, Guo Ci Teo¹, Alexey I. Nesvizhskii*¹,²
Fast and Comprehensive N- and O-glycoproteomics analysis with MSFragger-Glyco

Daniel A. Polasky¹, Fengchao Yu¹, Guo Ci Teo¹, Alexey I. Nesvizhskii*¹,²

<table>
<thead>
<tr>
<th>Search</th>
<th>Compositions searched</th>
<th>GlycoPSMs</th>
<th>Unique glycopeptides</th>
<th>Unique glycoproteins</th>
<th>Unique glycosites</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSFragger</td>
<td>182</td>
<td>44,187</td>
<td>2,822</td>
<td>1,070</td>
<td>2,133</td>
</tr>
<tr>
<td>Riley 2019</td>
<td>182</td>
<td>24,099</td>
<td>1,803</td>
<td>771</td>
<td>1,545</td>
</tr>
</tbody>
</table>

b

Glycosites

Riley 2019

Liu 2017

794

576

673

220

239

47

9

c

Compositions per Site

16%

32%

27%

12%

14%

8%

16%

16%

32%

Riley 2019

MSFragger-Glyco

d

Sites per Protein

1

2

3

4-5

>5

Riley 2019

MSFragger-Glyco
Identify peptide candidates without pre-determined modifications

**MS/MS Spectrum**

**Theoretical Peptide Mass**

**Measured Mass**

$\Delta m$

mass (Da)

Collision-based dissociation

Higher-energy collisional dissociation (HCD)

**Total glycan mass:**

$\text{H4N4}$

Scan #: 4027, m/z: 1113.4940, z: 3
O-PAIR SEARCH FOR O-GLYCOPEPTIDES

1. Protein database
2. Fragment index
3. O-glycan database
4. O-glycan groups
5. Identify peptide candidates
   - Theo. Peptide Mass
   - Meas. Mass
   - Δm
   - Best Scoring Peptide Candidates (p1, p2, ..., pn)
     - $M(\text{precursor}) = M(p1) + \Delta m_1$
     - $M(\text{precursor}) = M(p2) + \Delta m_2$
     - $\vdots$
     - $M(\text{precursor}) = M(p_n) + \Delta m_n$
6. Match Δm to glycan groups
   - $M(p_1)$
   - $\Delta m$
   - $M(g_1)$
7. O-glycopeptide Candidates
   - Total Glycan Mass
O-PAIR SEARCH FOR O-GLYCOPEPTIDES

Combining fragment-indexed Open Search with glycan group delta masses and graph theory localization

1. Protein database
2. Fragment index
3. O-glycan database
4. O-glycan groups

5. Identify peptide candidates
   - Theo. Peptide Mass
   - Meas. Mass
   - $\Delta m$
   - Mass (Da)
   - $M(\text{precursor}) = M(p1) + \Delta m_1$
   - $M(\text{precursor}) = M(p2) + \Delta m_2$
   - $\cdots$
   - $M(\text{precursor}) = M(p_n) + \Delta m_n$

6. Match $\Delta m$ to glycan groups

7. O-glycopeptide Candidates
   - Total Glycan Mass

8. Localization
   - N1
   - A1
   - B1
   - AB1
   - N2
   - A2
   - B2
   - AB2
   - N3
   - A3
   - B3
   - AB3
   - N4
   - A4
   - B4
   - AB4

9. Fine scoring and FDR
   - Score
   - Localization Probability
   - +
   - +
   - +

More information:
bioRxiv Pre-Print: https://www.biorxiv.org/content/10.1101/2020.05.18.102327v1: provisionally accepted at Nature Methods
**O-PAIR SEARCH PERFORMANCE**

<table>
<thead>
<tr>
<th>Glycans per peptide</th>
<th>Byonic Time (min)</th>
<th>O-Pair Time (min)</th>
<th>Speed Improvement</th>
<th>Localized Glycosites (Byonic / O-Pair)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2glycans</td>
<td>16.26</td>
<td>0.36</td>
<td>45.24</td>
<td>33 / 76</td>
</tr>
<tr>
<td>3glycans</td>
<td>862.76</td>
<td>0.40</td>
<td>2,161.09</td>
<td>41 / 120</td>
</tr>
<tr>
<td>4glycans</td>
<td>DNF</td>
<td>0.52</td>
<td>NA</td>
<td>NA / 134</td>
</tr>
</tbody>
</table>

**Mucins**
- MUC16
- CD43
- Podocalyxin
- PSGL-1

-Riley et al., JPR, 19 (8): 3286–3301 (2020) -

**GlycoPSMs**

- Byonic
- O-Pair

**Localized**
- All
- (Level 1 and 1b)
Recent Advances in Analytical Approaches for Glycan and Glycopeptide Quantitation

Daniel G. Delafield¹ and Lingjun Li¹,²,*
GLYCOPROTEIN QUANT: NOT STRAIGHTFORWARD

LC-MS/MS

Isolating Glycoproteins
Generating Glycopeptides
Enrichment & Separations
Ionization
Tandem MS (MS/MS)
Interpreting Spectra
Quant and Context

MCP
MOLECULAR & CELLULAR PROTEOMICS

Calculating glycoprotein similarities from mass spectrometric data

LVPVPITN(N)ATLDQITGK

William Hackett¹ and Joseph Zaia¹²
Multiplexed Comparative Analysis of Intact Glycopeptides Using Electron-Transfer Dissociation and Synchronous Precursor Selection Based Triple-Stage Mass Spectrometry

Hongbin Zhu, Chen Qiu, Connie M. Gryniewicz-Ruzicka, David A. Keire, and Hongping Ye

Cite This: Anal. Chem. 2020, 92, 7547–7555
SugarQuant: a streamlined pipeline for multiplexed quantitative site-specific N-glycoproteomics

Pan Fang¹, Yanlong Ji²,³, ¹, Ivan Silber⁴,³, Carmen Doebeli¹, Mornchil Ninov⁴, Christof Lenz²,³, Thomas Oellerich²,³,⁵, Kuang-Ting Pan¹,³,⁵, Henning Urlaub²,³,⁵
Quantitative Longitudinal Inventory of the N-Glycoproteome of Human Milk from a Single Donor Reveals the Highly Variable Repertoire and Dynamic Site-Specific Changes

Jing Zhu,1 Yu-Hsien Lin,1 Kelly A. Dingess, Marko Mank, Bernd Stahl, and Albert J. R. Heck*
Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis

Nicholas M. Riley, Alexander S. Hebert, Michael S. Westphall & Joshua J. Coon

Glycoprotein-Glycan Networks
Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis

Nicholas M. Riley, Alexander S. Hebert, Michael S. Westphall & Joshua J. Coon

Glycoprotein-Glycan Networks
Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis

Nicholas M. Riley 1,2,5, Alexander S. Hebert1, Michael S. Westphall1 & Joshua J. Coon1,2,3,4

Glycoprotein- Glycan Networks
Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis

Nicholas M. Riley1,2,5, Alexander S. Hebert1, Michael S. Westphall1 & Joshua J. Coon1,2,3,4

Glycoprotein-Glycan Networks
Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis

Nicholas M. Riley, Alexander S. Hebert, Michael S. Westphall & Joshua J. Coon

Glycoprotein-Glycan Networks
HOW TO REPRESENT DIMENSIONALITY?
HOW TO REPRESENT DIMENSIONALITY?
Examining and Fine-tuning the Selection of Glycan Compositions with GlyConnect Compozitor

Thibault Robin\textsuperscript{1,2,3,4}, Julien Mariethoz\textsuperscript{1,2,5}, and Frédérique Lisacek\textsuperscript{1,2,5,6}
Letter to Glyco-Forum

The GlySpace Alliance: toward a collaborative global glycoinformatics community

Kiyoko F Aoki-Kinoshita¹, Frederique Lisacek², Raja Mazumder³, William S York⁴ and Nicolle H Packer⁵

¹Glycan & Life Science Integration Center (GaLSIC), Faculty of Science and Engineering, Soka University, 1-236 Tangi-machi, Hachioji, Tokyo, Japan, 192-8577, ²Proteome Informatics Group, SIB Swiss Institute of Bioinformatics, Computer Science Department, University of Geneva, route de Drize 7, CH-1227 Geneva Switzerland, and also Section of Biology, University of Geneva, Geneva, Switzerland, ³Department of Biochemistry & Molecular Medicine, and Department of Medicine, School of Medicine and Health Sciences, George Washington University, Ross Hall, 2300 Eye St., NW, Washington, DC 20037, USA, ⁴Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA, and ⁵Department of Molecular Sciences, Faculty of Science & Engineering, Rm 307, Building EBC, Macquarie University, Sydney, NSW 2109, Australia

¹To whom correspondence should be addressed: Tel/Fax: +81-42-691-4116; e-mail: kkiyoko@soka.ac.jp

Received 5 September 2019, Editorial Decision 20 September 2019; Accepted 20 September 2019
Acknowledgements

Carolyn Bertozzi
Bertozzi Group
Allis Chien
Ryan Leib
SUMS
Stanford University Mass Spectrometry
@riley_nm1

Wed-Thurs October 28-29, 2020
SUMS Research Application Symposium