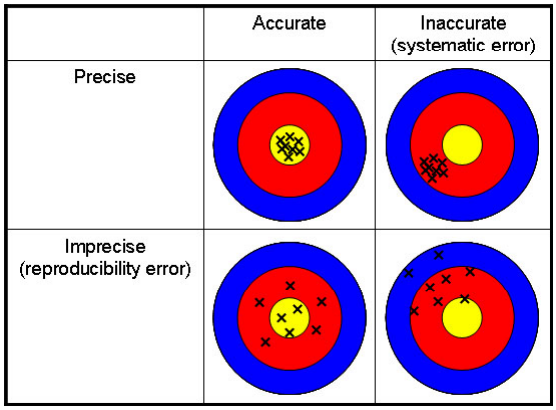
Quantitative Methods for the Analysis of N- and O-glycans

Ron Orlando

Outline

- Introduction
  - Background/Terminology
    - Absolute vs. Relative Quantification
    - Calibration curve, Sensitivity, LOD, LOQ, LOL, etc.
    - Statistics
  - Sources of error
    - Matrix affects, instrument response, sample preparation, etc.
- Quantification
  - Label free
    - Direct MS analysis
    - Normalization
    - Results from multilaboratory studies
  - Isotopic Labeling
    - *in vitro* labeling
    - *in vivo* tagging
- Conclusions

Accuracy vs. Precision



From: http://www.wellesley.edu/Chemistry/Chem105manual/Lab04/lab04.html

Precision

- Reproducibility
- Check by repeating measurements
- Poor precision results from poor experimental technique

Accuracy

- Correctness
- Check by using orthogonal method
- Poor Accuracy results from procedural or equipment flaws.

Statistical Terminology - Precision

- Standard Deviation

$$S = \left(\frac{\Sigma(x_n - x)^2}{(N-1)} \right)^{1/2}$$

- Relative Standard Deviation (RSD) = $s / \bar{x}$
- Coefficient of Variation (CV) = $(s / \bar{x}) \times 100$

Absolute vs. Relative Quantitation

Absolute

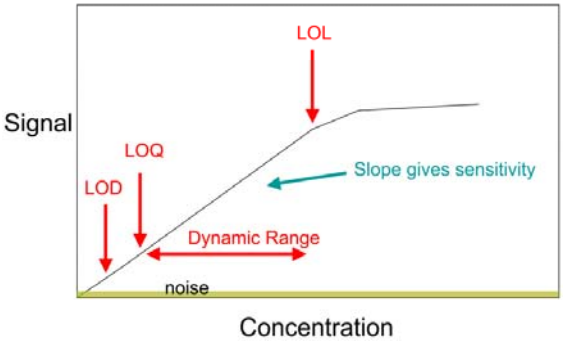
- Determine the amount of an analyte in an unknown.
- Requires standards of known concentration. Since MS response is typically analyte specific, one ideally would like a standard for each analyte.

Relative

- Determine the change in amount of an analyte between multiple samples.

Absolute Quant. - Calibration Curve

The calibration curve provides a reliable way to calculate the uncertainty of an analyte's concentration using statistics of least squares line fit to the data. - DA Skoog *et. al*. (2006) *Principles of Instrumental Analysis*



LOD – Limit of Detection. This limit depends upon the ration of the analytical signal to the size of the fluctuations in the blank signal (noise). 3 is often used as the acceptable signal to noise ratio.

LOQ – Limit of Quantitative measurement. LOQ is generally taken to be 10x the noise level

LOL – Limit of Linearity. Range from LOQ to the concentration at which the calibration curve departs from linearity.

Sensitivity – IUPAC definition – the slope of the calibration curve at the concentration of interest (also called calibration sensitivity).

Several Sources of Error

- Matrix effects
- Variable instrument response
- Variations between instruments
- Differential loss of analyte during sample preparation
- Differences in Ionization Efficiency

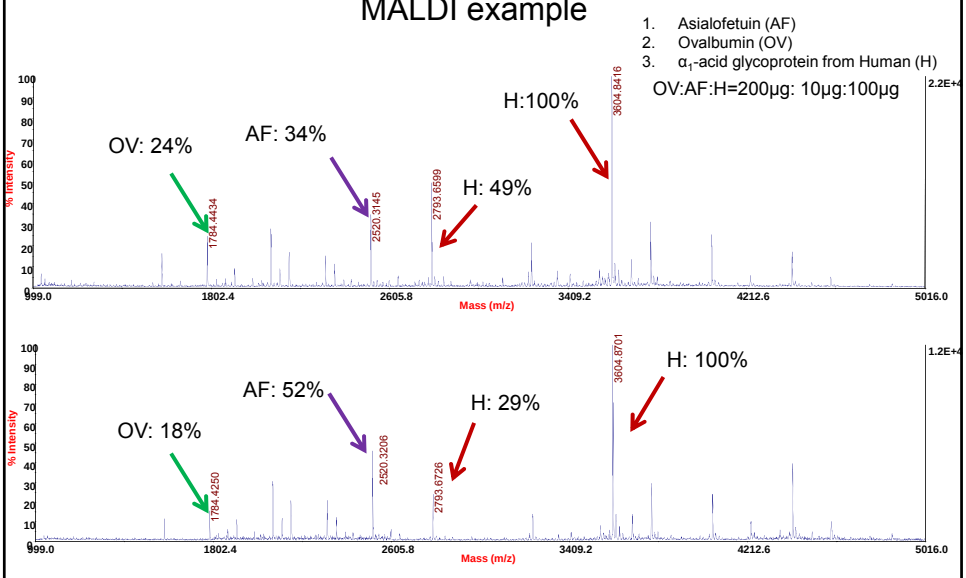
Matrix Effects

Electrospray

- Matrix effects occur when molecules co-eluting with the compound/s of interest alter the ionization efficiency of the electrospray interface. (L. Tang and P. Kebarle, Anal. Chem. 65 (1993), pp. 3654–3668.)
- The chemical nature of a compound has a significant effect on the degree of matrix effects. (R. Bonfiglio, R.C. King, T.V. Olah and K. Merkle, Rapid Commun. Mass Spectrom. 13 (1999), pp. 1175–1185.)

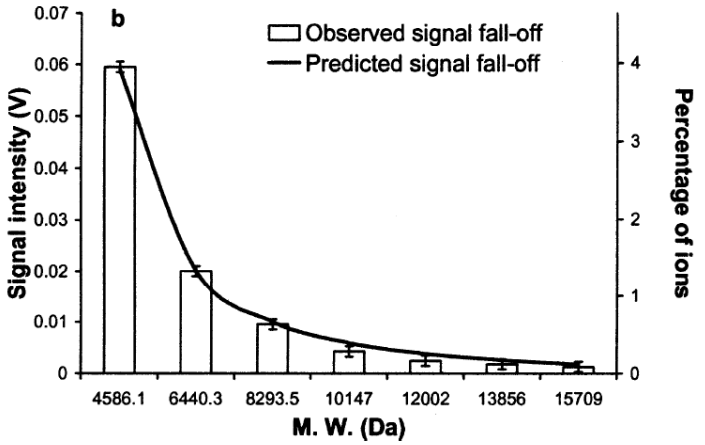
Instrument Variability

MALDI example



Instrumental Issues

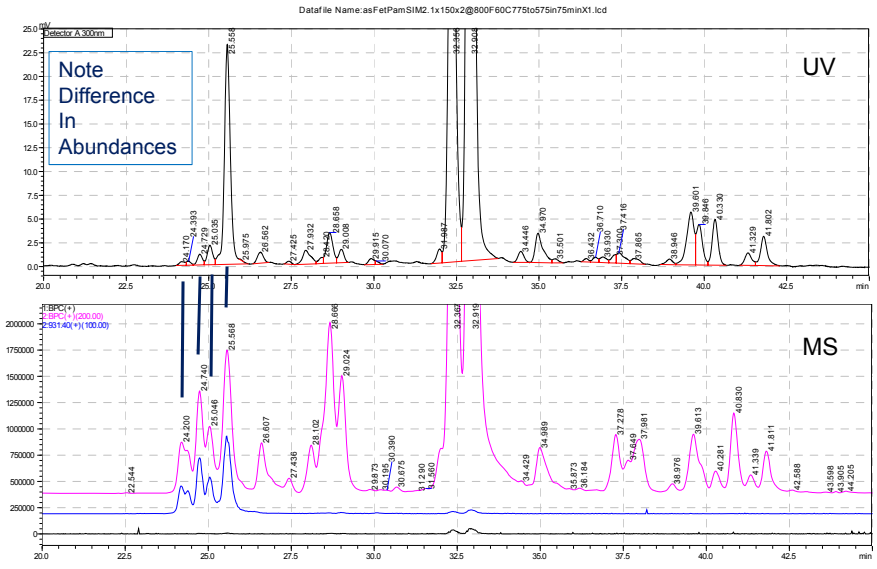
Mass Dependence of Detector Response



X. Chen, M.S. Westphall, L.M. Smith, *Anal. Chem*. 2003, *75*, 5944-5952

Differences in Ionization Efficiency

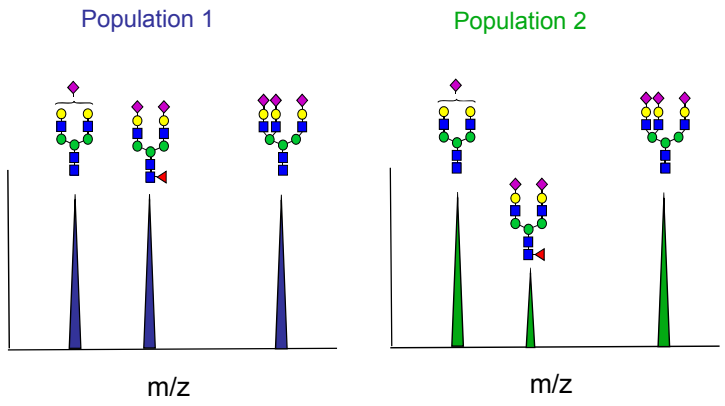
LC-MS of fetuin glycans tagged on reducing terminus



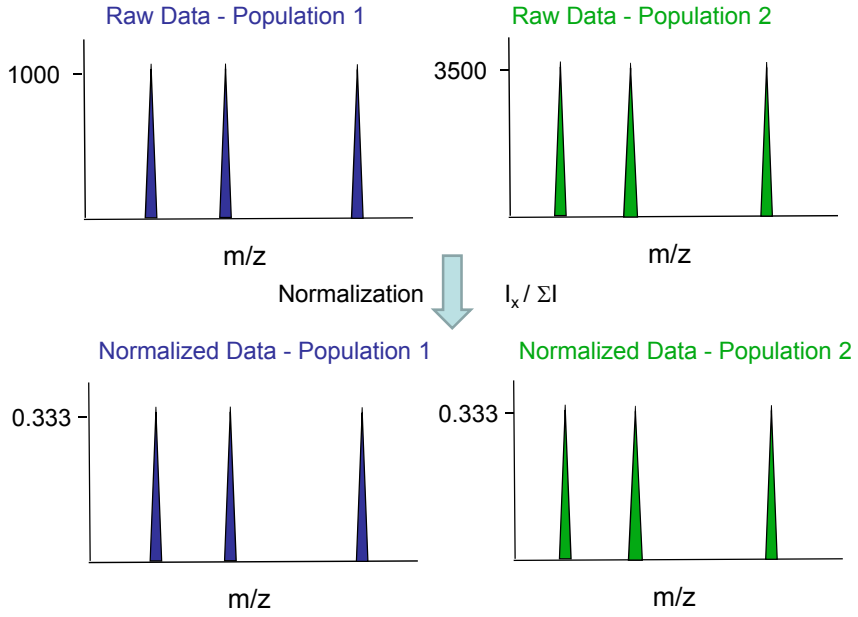
Outline

- Introduction
  - Background/Terminology
    - Absolute vs. Relative Quantification
    - Calibration curve, Sensitivity, LOD, LOQ, LOL, etc.
    - Statistics
  - Sources of error
    - Matrix affects, instrument response, sample preparation, etc.
- **Quantification**
  - **Label free**
    - Direct MS analysis
    - Normalization
    - Results from multilaboratory studies
  - Isotopic Labeling
    - *in vitro* labeling
    - *in vivo* tagging
- Conclusions

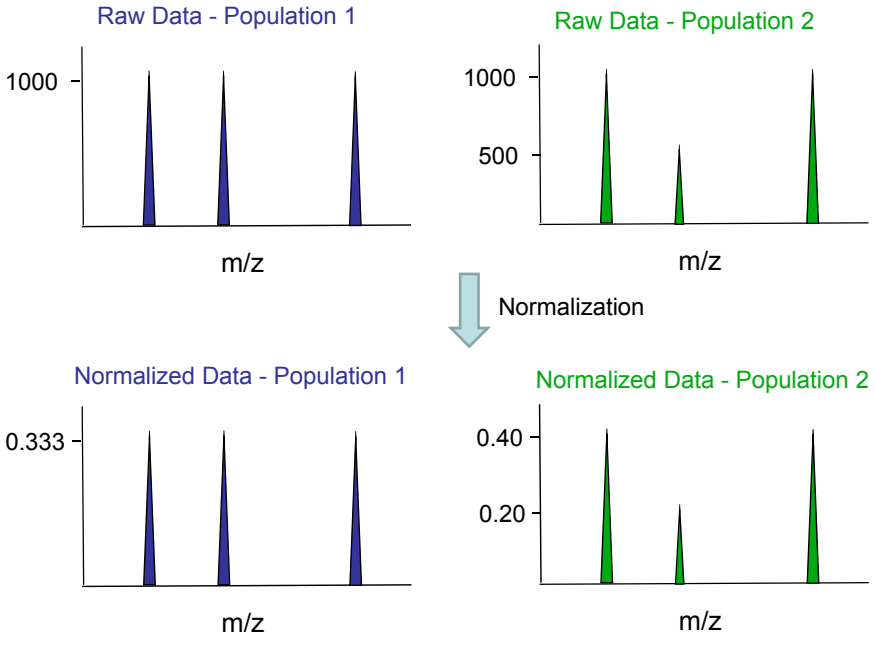
Relative Quantification using Ion Abundance



Normalization is used to compensate for changes in instrument response between experiments.



Normalization can cause problems.



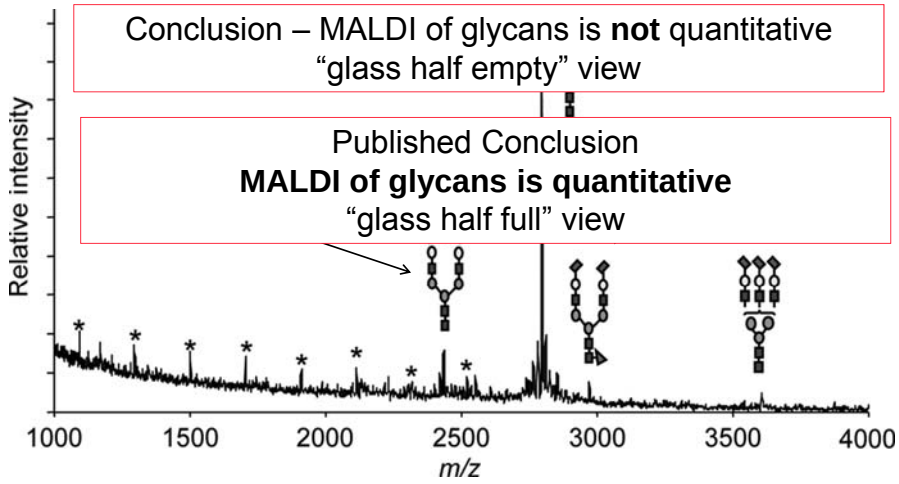
HUPO Study

Y. Wada *et. al.*, *Glycobiology* (*2007*) *17* (4): 411-422.

- Analyzed N-linked glycans released from serum transferrin and an IgG by MALDI, LC/MS and chromatographic analysis.
- "the results of this multi-institutional study indicate that MS-based analysis appears as the efficient method for identification and quantitation of oligosaccharides in glycomic studies ..."

HUPO 2007 – Glycan Quantitation

MALDI-MS of permethylated glycans from sample transferrin



Wada Y et al. Glycobiology 2007;17:411-422

ABRF gPRG 2010:
Quantitative Glycoprotein Study

Used a mixture of glycoproteins that have unique glycans, varied the glycoprotein ratios, thus changing the glycan ratios in a known manner.

Study Participants

35 – Samples requested --- 19 – Data submissions
- 10 Academic – 7 Industry – 2 Vendors
- 14 North America – 4 Asia – 1 Europe
- 10 Core labs

Most (18 of 19) labs released glycans with PNGase F

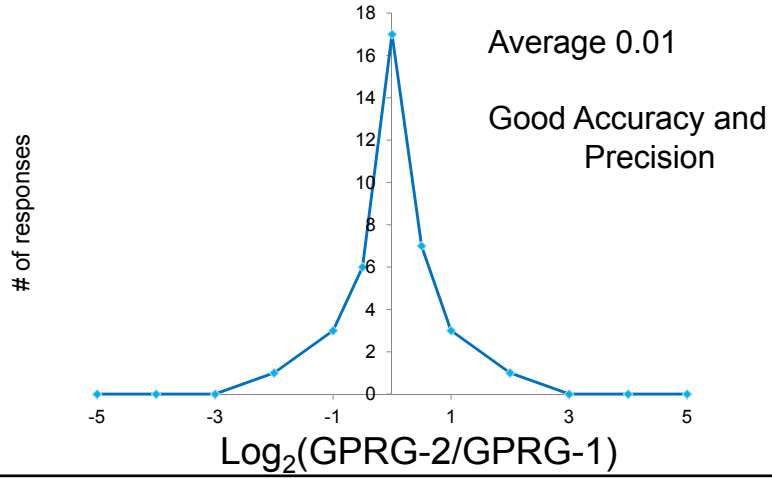
All labs analyzed with MS
- 10 MALDI – 7 LC/MS – 2 ESI

Double blind study

All labs used a label free approach.

Quantification – Glycans that do not change

Each sample contained the same amount of Ov.
Responses for Man_{5+6} for GPRG-2+3 vs GPRG-1

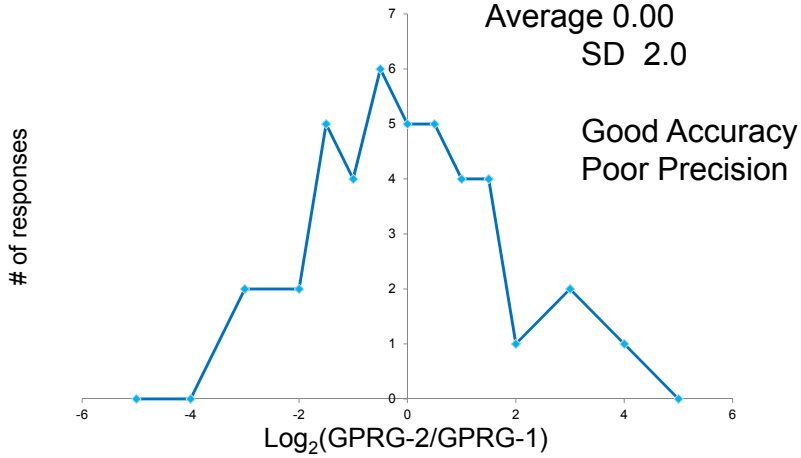


Average 0.01

Good Accuracy and Precision

Quantification – Glycans that do not change

Each sample contained the same amount of Ov.
Responses for Man_{8+9} (lower abundance) for GPRG-2+3 vs GPRG-1

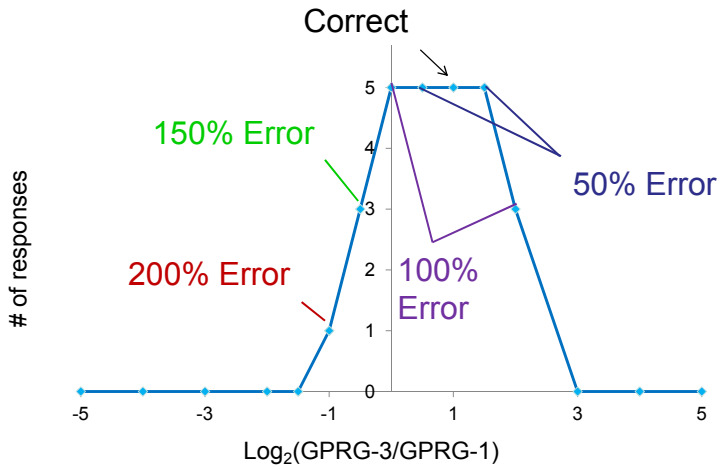


Average 0.00
SD 2.0

Good Accuracy
Poor Precision

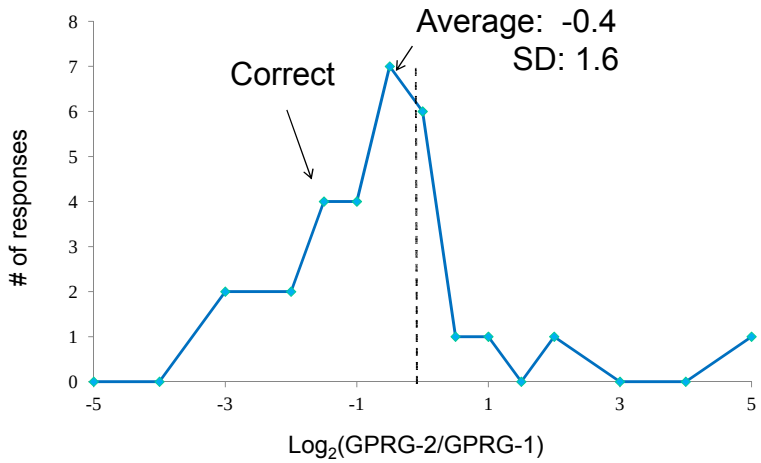
Quantification – Results Submitted “2x Up regulated”

Exp. Average 1.6x increase, Standard Dev. 110%



Quantification – II

GPRG-2 vs GPRG-1, “down regulated”



Label Free Summary

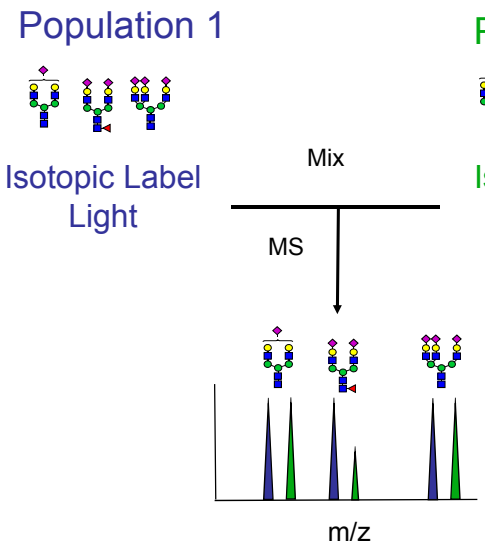
- Advantages
  - Easy
  - Cheap
  - Simple, no need for special reagents and no disruption of your sample flow
- Disadvantages
  - Provides less certainty and lower confidence than other methods

Its free, but you get what you pay for

Outline

- Introduction
  - Background/Terminology
    - Absolute vs. Relative Quantification
    - Calibration curve, Sensitivity, LOD, LOQ, LOL, etc.
    - Statistics
  - Sources of error
    - Matrix affects, instrument response, sample preparation, etc.
- Quantification
  - Label free
    - Direct MS analysis
    - Normalization
    - Results from multilaboratory studies
  - **Isotopic Labeling**
    - ***in vitro* labeling**
    - *in vivo* tagging
- Conclusions

Relative Quantification using Isotopic Labeling

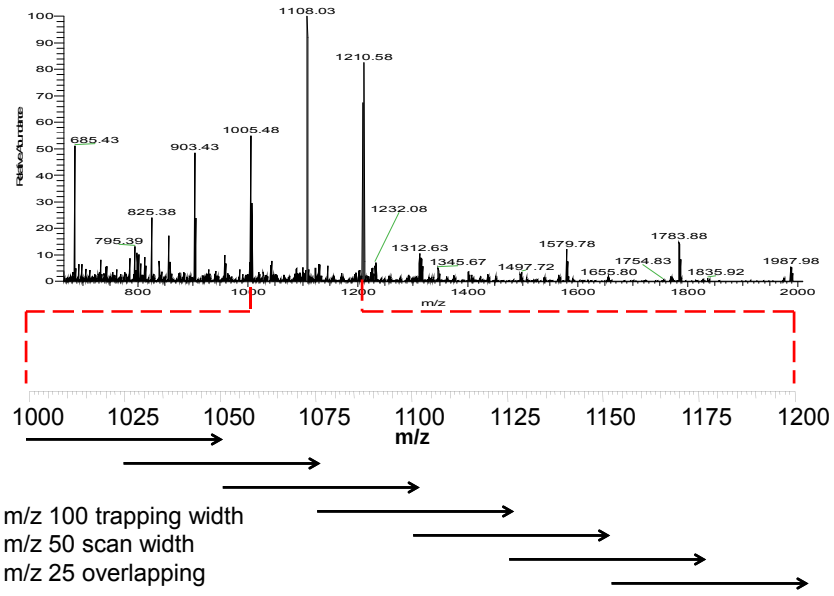


Isotope labeling compensates for errors resulting from:

Matrix effects
Variable instrument response
Variations between instruments

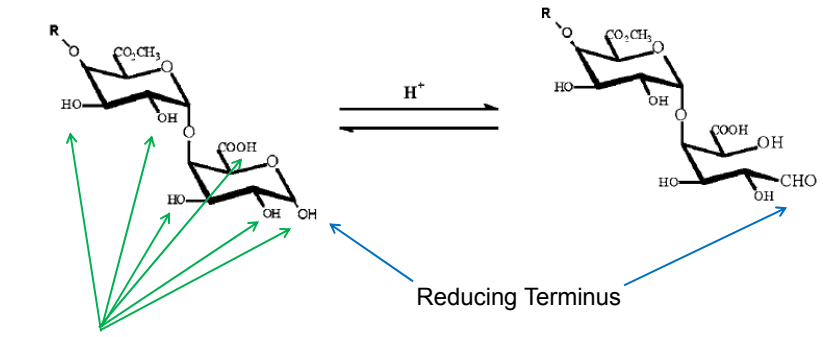
Improving dynamic range with rolling trapping scans

(no collision energy)



m/z 100 trapping width
m/z 50 scan width
m/z 25 overlapping

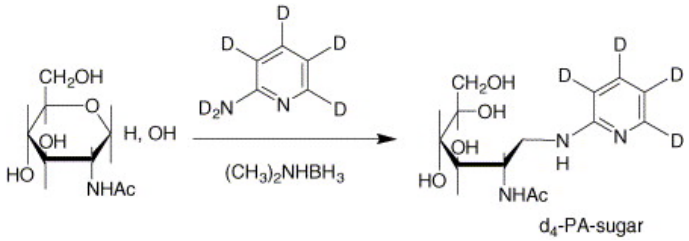
Sites for Isotopic Labeling



Hydroxyl Groups

Isotopic Labeling of Reducing Terminus

2-Aminopyridine
(2-AP)



d4-PA-sugar

Isotope effect introduced by hydrogen/deuterium substitution may be noticed as chromatographic shifts during LC separations.

J. Yuan, N. Hashii, N. Kawasaki, S. Itoh, T. Kawanishi, T. Hayakawa, *J. Chromatog. A*, 2005, *1067*, 145-152.

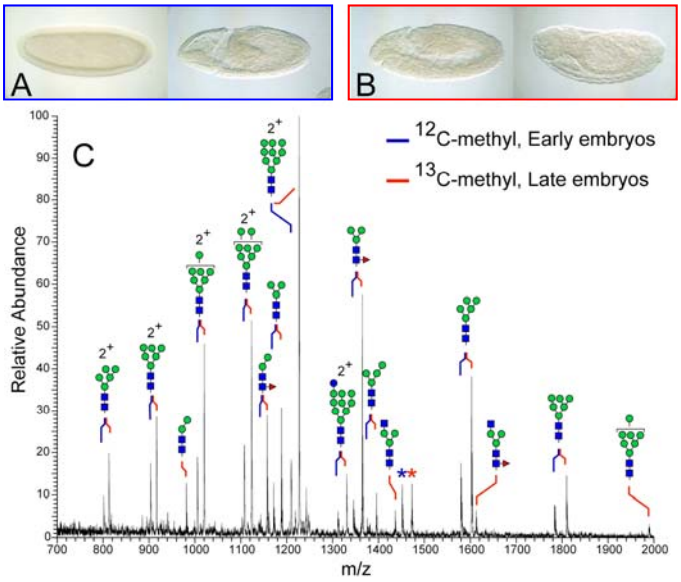
Permethylation of Glycans is Common

During permethylation of oligosaccharides the OH groups are converted to O-methyl groups. . (I. Ciucanu and F. Kerek. 1984. *Carbohydr. Res*., 131: 209-217)

It increase the sensitivity of oligosaccharides for subsequent MS analysis.

The mass increase is not too much to shift the mass to higher mass range and decrease sensitivity.

It allows for diagnostic molecular ions which are easier to interpret than the native oligosaccharides.



K. Aoki, M. Perlman, J. Lim, R. Cantu, L. Wells, M. Tiemeyer, Journal of Biological Chemistry, 2007, 282, 9127-9142.

Problem: Quantification of Changes in Glycoforms

Isobars

Isobar (definition): One of two or more atoms which have a common mass number (A) but which differ in atomic number (Z). Thus, isobars possess approximately equal masses, …

Molecular Isobars

$^{13}CH_3$ I= 142.931299 $^{12}CH_2D$ I = 142.9342217

ΔM = 0.0029227

Isobaric Labeling: the concept

$-^{13}CH_3$ = 16.026829 $^{12}CH_2D$ = 16.0297517

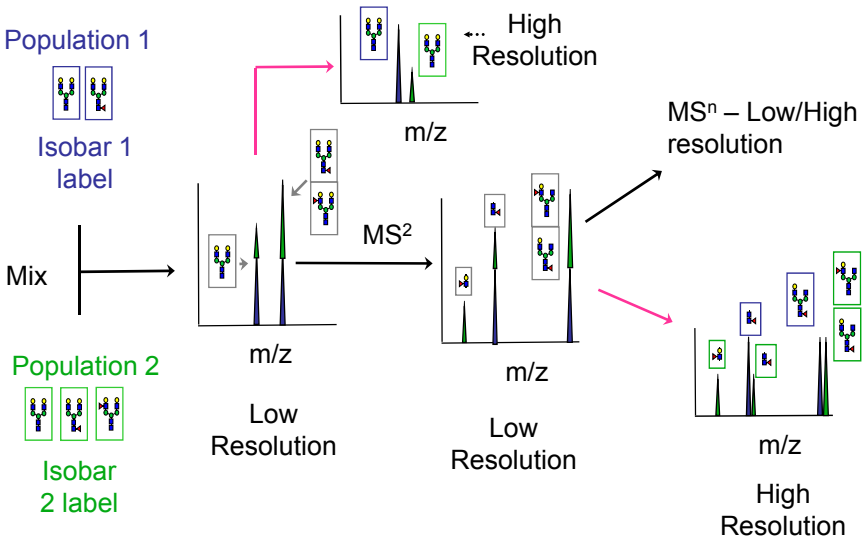
$\Delta M = 0.0029227$

Hex: 6 sites of methylation, MW = 233.X
DM = 6 x 0.0029227 = 0.0175362
Resolution required (M/ΔM) = 233/0.0175362 = **13,286**

$Hex_5HexNAc_2$: 26 sites of methylation, MW = 1572.X
DM = 23 x 0.0029227 = 0.0672221
Resolution required (M/ΔM) = 1572/0.0672221 = **23,385**

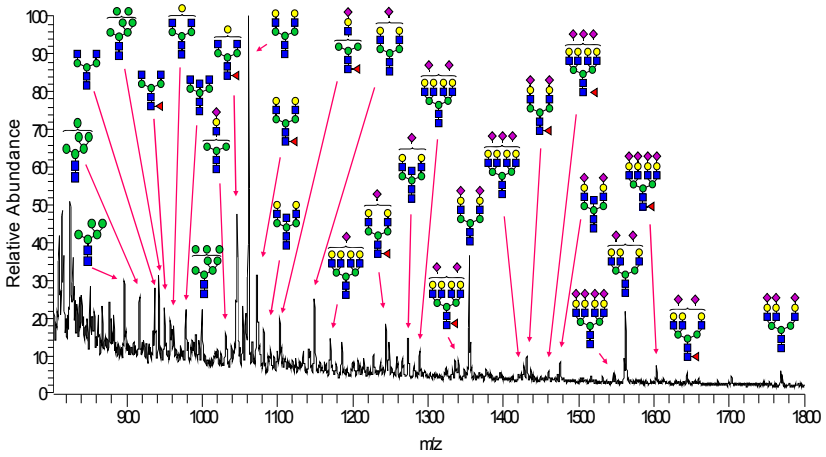
$Hex_{100}HexNAc_2$: 285 sites of methylation, MW = 21237.X
DM = 308 x 0.0029227 = 0.9001916
Resolution required (M/ΔM) = 21237/0.9001916 = **23,591**

Quantitation by Isobaric Labeling (QUIBL)[1]

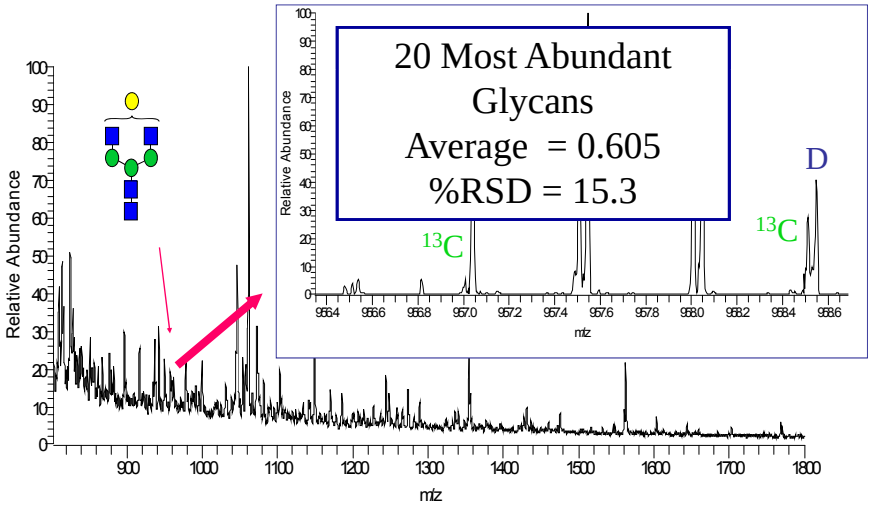


[1]J.A. Atwood III, L. Cheng, G. Alvarez-Manilla, W. York, R. Orlando, "Quantitation by Isobaric Labelling: application to Glycomics," *J Prot Res*., **2008**, *7*, 367–374.

Isobaric Labeling: Complex Mixture
Glycans from Human Serum, 6:10 (^{13}C:D) mix

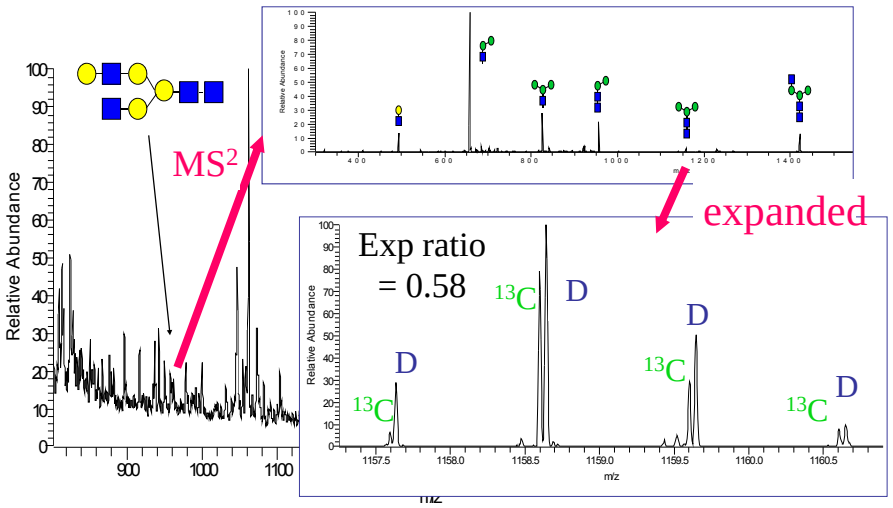


Isobaric Labeling: Complex Mixture
Glycans from Serum, 6:10 (^{13}C:D) mix

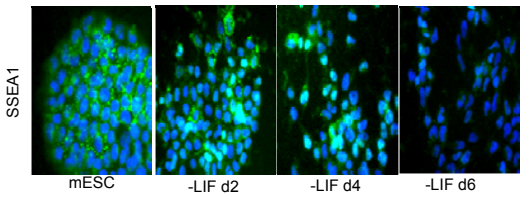


Quantitation from Fragment Ions

MS^2, N-glycans from Serum, 6:10 (^{13}C:D) mix



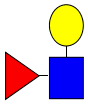
QUIBL applied to the Stem Cell Differentiation



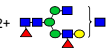
Immuno staining with anti SSEA 1 antibody

Demonstrates a decrease in SSEA-1 anitgen when mESC's are allowed to spontaneously differentiate

SSEA-1 or Lewis X antigen



QUIBL comparison of fucosylated glycans from mESCs and mEBs

	m/z	Charge state	Glycan structure	Log_2^* {EB (^{13}C)/ES (D)} Run 1	Run 2	Run 3	Ave	SDEV
Core Fucose	942.51	$[M+2Na]^{2+}$		0.16	0.05	0.08	0.10	0.07
	1046.06	$[M+2Na]^{2+}$		0.13	0.11	0.12	0.12	0.01
	1170.13	$[M+2Na]^{2+}$		0.14	0.17	0.08	0.13	0.02
SSEA-1	1046.06	$[M+2Na]^{2+}$		-1.13	-1.24	-1.19	-1.19	0.07
	1170.13	$[M+2Na]^{2+}$		-1.34	-1.27	-1.18	-1.23	0.05
	1258.68	$[M+2Na]^{2+}$		-1.37	-1.25	-1.43	-1.35	0.08
	1362.24	$[M+2Na]^{2+}$		-2.08	-2.06	-2.19	-2.11	0.02

Results are consistent with Immuo staining

*EB/ES ratios are shown as their Log_2 values to normalize spacing between up and down expression, i.e., a 4 fold increase/decrease is **2/-2** with Log_2, as opposed to **4.0/0.25** for the ratio. This also sets zero as the value for no change.

QUIBL Analysis (gPRG-3 vs gPRG-1)

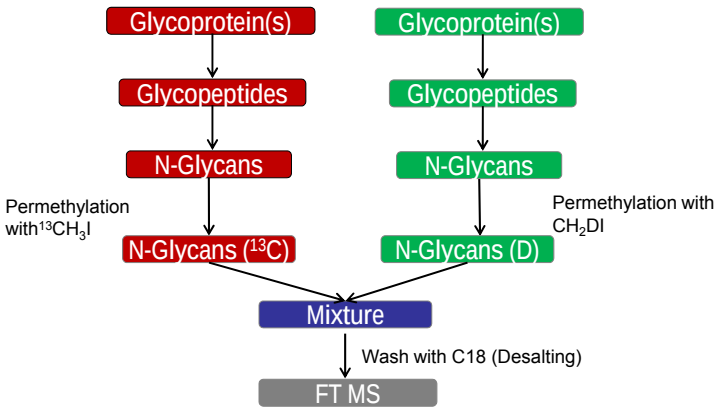
Glycan Composition	Glycoprotein	z	Monoisotopic Mass gPRG 1 (^{13}C)	Monoisotopic Mass gPRG 3 (D)	Theoretical Ratio	Experimental Ratio
$HexNAc_3Hex_3$	OV	1	1436.78	1436.84	1	0.963±0.011
$HexNAc_3Hex_3$	OV	2	729.88	729.91	1	1.020±0.018
$HexNAc_2Hex_5$	OV	1	1602.86	1602.93	1	1.008±0.077
$HexNAc_2Hex_5$	OV	2	812.92	812.96	1	0.986±0.018
$HexNAc_3Hex_4$	OV	1	1643.89	1643.95	1	0.753±0.016
$HexNAc_3Hex_4$	OV	2	833.44	833.47	1	0.856±0.018
$HexNAc_4Hex_3$	OV				1	0.932±0.025
$HexNAc_4Hex_3$	OV				1	1.052±0.033
$HexNAc_2Hex_6$	OV				1	0.912±0.031
$HexNAc_2Hex_6$	OV	2	916.48	916.52	1	1.010±0.019
$HexNAc_3Hex_5$	OV	2	936.99	937.03	1	0.923±0.050
$HexNAc_2Hex_7$	OV	2	1102.09	1102.13	1	1.337±0.018
$HexNAc_5Hex_5$	OV	2	1185.13	1185.18	1	1.103±0.012
$HexNAc_4Hex_5$	AF	2	1288.68	1288.73	2	2.124±0.010
$HexNAc_5Hex_6$	AF	3	866.79	866.82	2	2.066±0.011

Accurate Results

But, Large Systematic Variability

Glycan Composition	Glycoprotein	z	Monoisotopic Mass		Theoretical Ratio	Experimental Ratio
			gPRG 1 (^{13}C)	gPRG 3 (D)		
$HexNAc_3Hex_3$	OV	1	1436.78	1436.84	1	0.863±0.011
$HexNAc_3Hex_3$	OV	2	729.88	729.91	1	1.029±0.018
$HexNAc_2Hex_5$	OV	1	1602.86	1602.93	1	1.138±0.077
$HexNAc_2Hex_5$	OV	2	812.92	812.96	1	0.946±0.018
$HexNAc_3Hex_4$	OV	1	1643.89	1643.95	1	0.723±0.016
$HexNAc_3Hex_4$	OV	2	833.44	833.47	1	0.856±0.018
$HexNAc_4Hex_3$	OV	1	1684.91	1684.98	1	0.912±0.025
$HexNAc_4Hex_3$	OV	2	853.95	853.98	1	1.152±0.033
$HexNAc_2Hex_6$	OV	1	1809.97	1810.05	1	0.842±0.031
$HexNAc_2Hex_6$	OV	2	916.48	916.52	1	1.117±0.019
$HexNAc_3Hex_5$	OV	2	936.99	937.03	1	0.803±0.050
$HexNAc_2Hex_7$	OV	2	1102.09	1102.13	1	1.337±0.018
$HexNAc_5Hex_5$	OV	2	1185.13	1185.18	1	1.253±0.012
					Ave deviation	16.5%
					Dev. Range	27.7% to -33.7%

Source of Systematic Variability

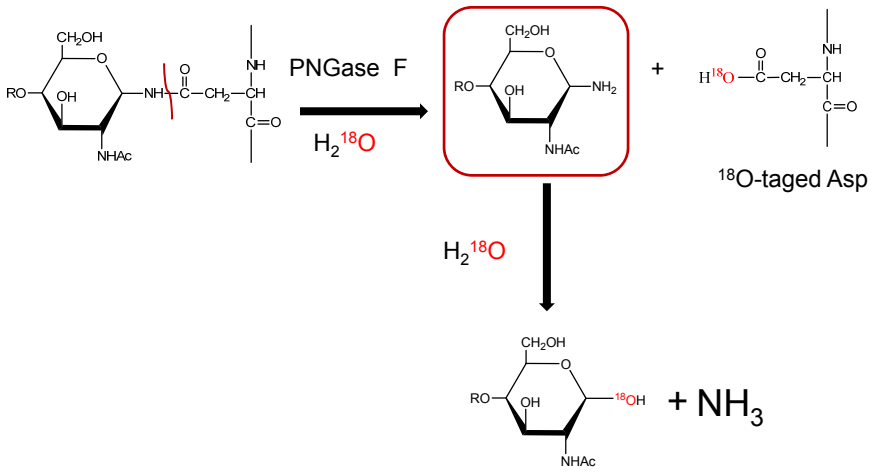


i) Multiple sample preparation steps: differential sample losses

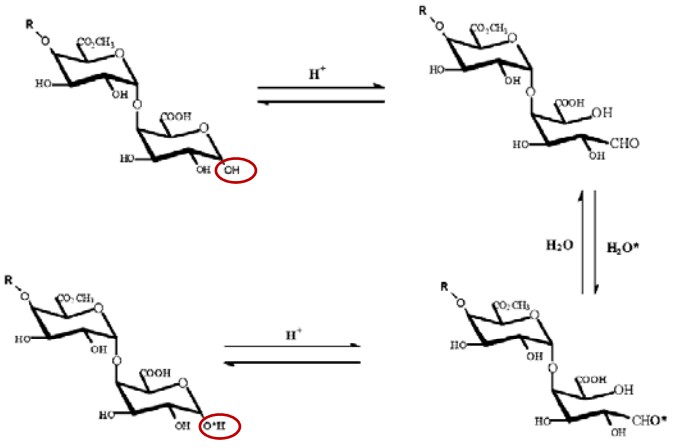
ii) Permethylation efficiency

For example, a 0.1% change in permethylation efficiency leads to a 5% error when there are 50 sites of permethylation

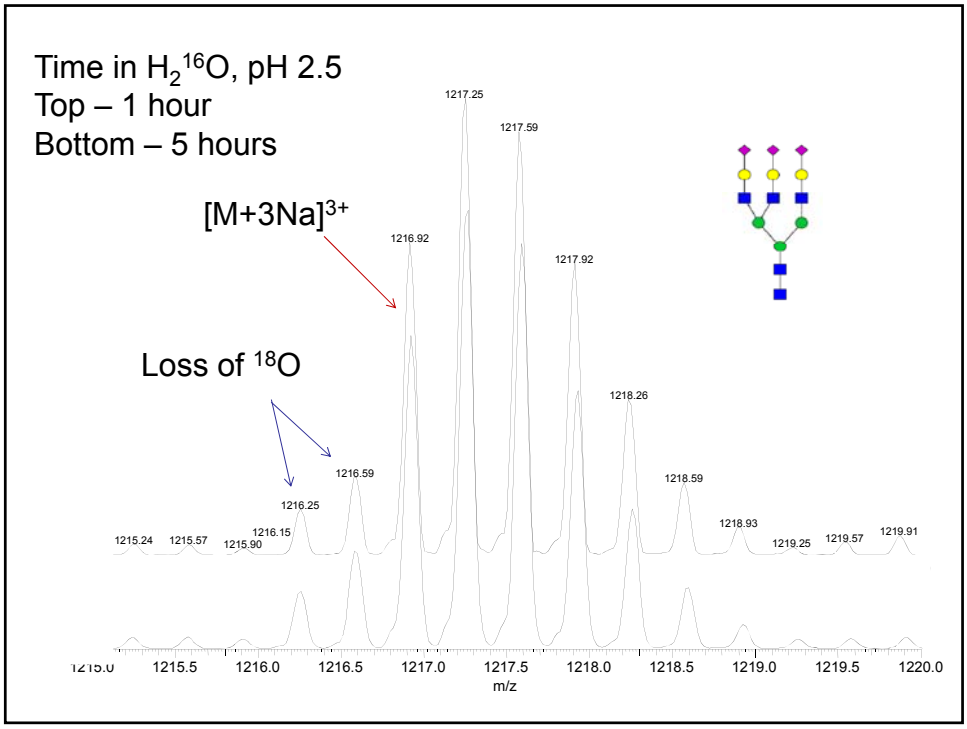
New Labeling Method ($H_2{}^{18}O$)

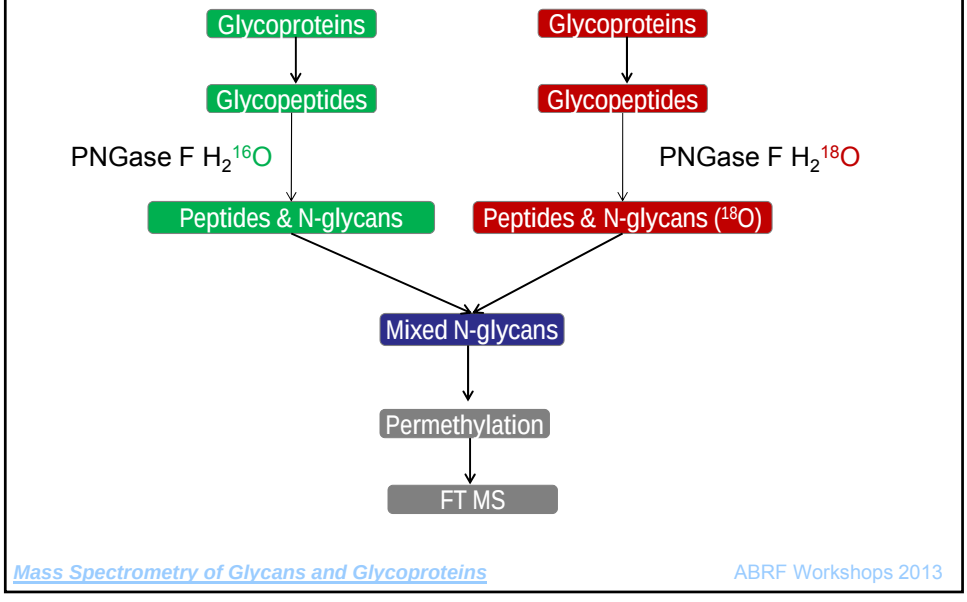


Loss of ^{18}O resulting from acidic hydrolytic exchange



Methods in Molecular Biology, 2006, 367, 289-301


Time in $H_2{}^{16}O$, pH 2.5
Top – 1 hour
Bottom – 5 hours
$[M+3Na]^{3+}$
Loss of ^{18}O
1217.25
1217.59
1216.92
1217.92
1218.26
1216.59
1218.59
1216.25
1218.93
1216.15
1215.24
1215.57
1215.90
1219.25
1219.57
1219.91
1215.5
1216.0
1216.5
1217.0
1217.5
1218.0
1218.5
1219.0
1219.5
1220.0
m/z


Glycoproteins
Glycopeptides
PNGase F $H_2{}^{16}O$
Peptides & N-glycans
Glycoproteins
Glycopeptides
PNGase F $H_2{}^{18}O$
Peptides & N-glycans (^{18}O)
Mixed N-glycans
Permethylation
FT MS

QUIBL vs ^{18}O

Glycan Composition	Theoretical Ratio	QUIBL	^{18}O
$HexNAc_3Hex_3$	1	0.863	0.972
$HexNAc_3Hex_3$	1	1.029	1.02
$HexNAc_2Hex_5$	1	1.138	0.986
$HexNAc_2Hex_5$	1	0.946	1.043
$HexNAc_3Hex_4$	1	0.723	0.956
$HexNAc_3Hex_4$	1	0.856	0.932
$HexNAc_4Hex_3$	1	0.912	1.052
$HexNAc_4Hex_3$	1	1.152	1.011
$HexNAc_2Hex_6$	1	0.842	1.01
$HexNAc_2Hex_6$	1	1.117	0.923
$HexNAc_3Hex_5$	1	0.803	1.057
$HexNAc_2Hex_7$	1	1.337	1.049
$HexNAc_5Hex_5$	1	1.253	1.014
	Ave deviation	16.50%	3.90%
	Dev. Range	27.7% to -33.7%	5.7% to -7.7%

Variability reduced by 4x

Conclusions

The incorporation of isotopic labels significantly improves the accuracy and reproducibility of glycan quantification.

^{18}O labeling decreases the variability compared to permethylated approaches to incorporate isotopic labels, presumably because this method labels a single site and decreases the number of steps done in parallel

Neither of these approaches alters normal glycomic workflows.

Isotopic Labeling Summary

- Advantages
  - Higher level of accuracy
  - Can reduce number of analyses performed
- Disadvantages
  - Reagents can be expensive
  - May require special (expensive) instrumentation

Outline

- Introduction
  - Background/Terminology
    - Absolute vs. Relative Quantification
    - Calibration curve, Sensitivity, LOD, LOQ, LOL, etc.
    - Statistics
  - Sources of error
    - Matrix affects, instrument response, sample preparation, etc.
- Quantification
  - Label free
    - Direct MS analysis
    - Normalization
    - Results from multilaboratory studies
  - **Isotopic Labeling**
    - *in vitro* labeling
    - ***in vivo* tagging**
- Conclusions

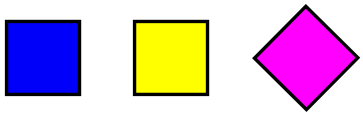
in vivo labeling of glycans

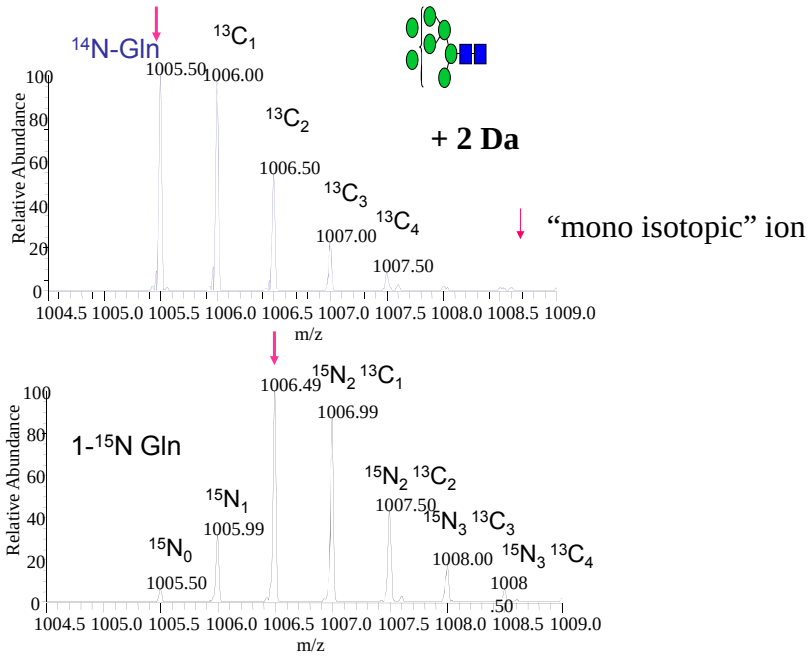
Isotopic Detection of Aminosugars with Glutamine
I-DAWG

A Novel Quantitative Method for Glycomics

I-DAWG:
Isotopic Detection of Aminosugars With Glutamine

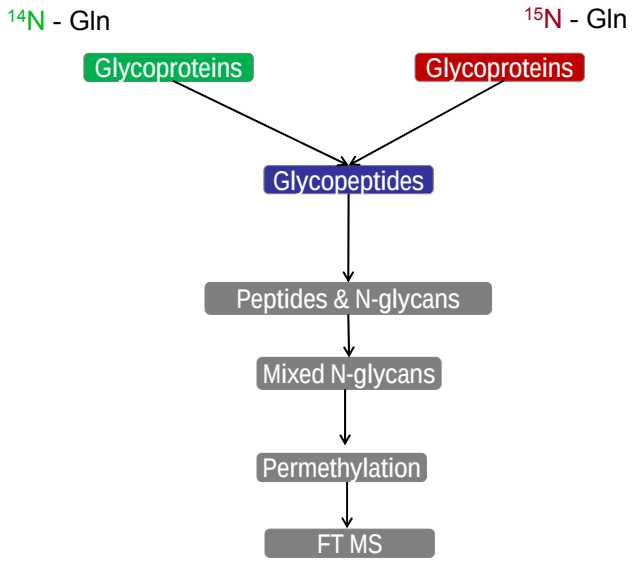
UDP-GlcNAc
↓
GLYCOCONJUGATE

+1 Dalton → 

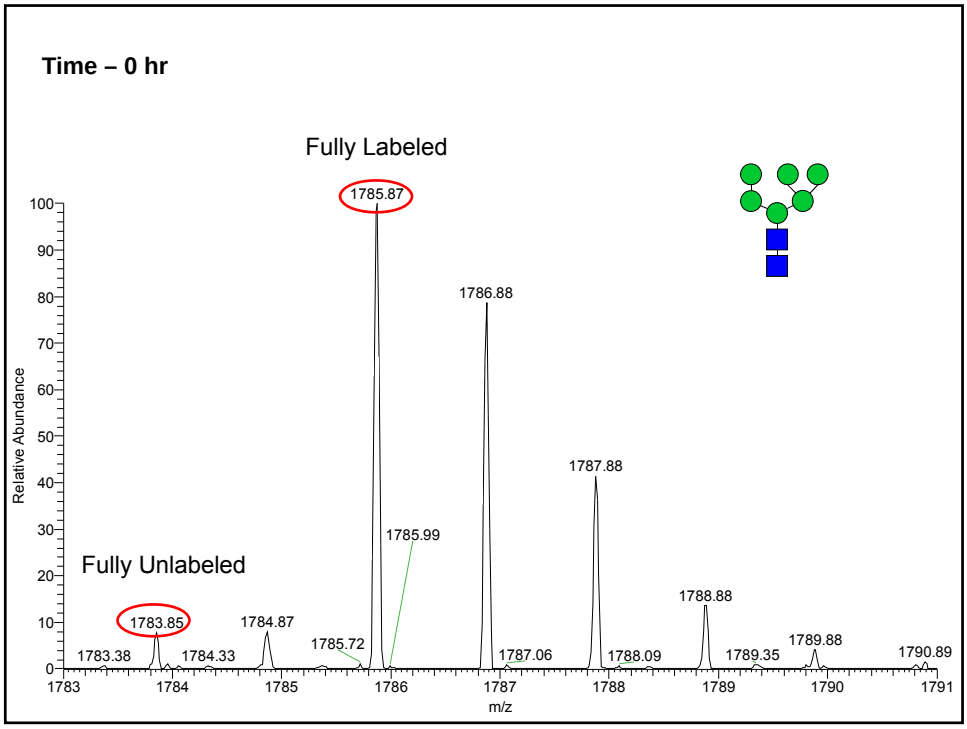


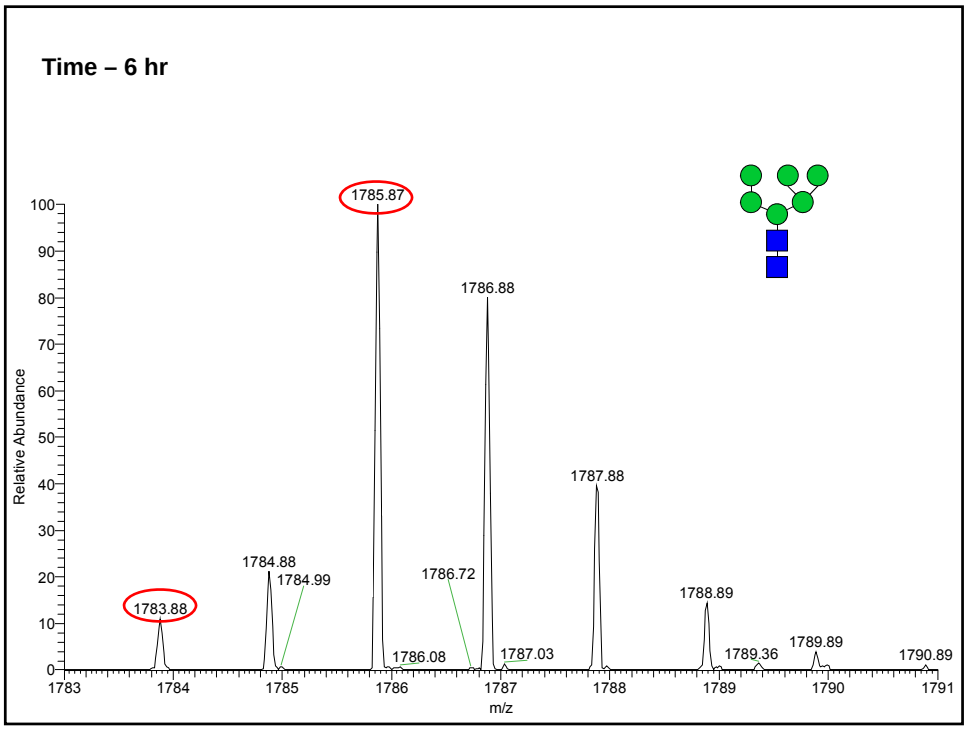
Labeling Efficiencies for a Variety of N-linked Glycans

# of Nitrogens	Composition	% ^{15}N Incorporation
2	$(Hex)_4$ - $(Man)_3$ $(GlcNAc)_2$	98.1
	$(Hex)_5$ - $(Man)_3$ $(GlcNAc)_2$	98.9
	$(Hex)_6$ - $(Man)_3$ $(GlcNAc)_2$	98.7
4	$(Hex)_2$ $(HexNAc)_2$ $(Deoxyhexose)_1$ - $(Man)_3$ $(GlcNAc)_2$	98.5
	$(Hex)_2$ $(HexNAc)_2$ $(Deoxyhexose)_2$ - $(Man)_3$ $(GlcNAc)_2$	98.0
	$(Hex)_3$ $(HexNAc)_2$ $(Deoxyhexose)_1$ - $(Man)_3$ $(GlcNAc)_2$	98.8
5	$(Hex)_1$ $(HexNAc)_3$ $(Deoxyhexose)_1$ - $(Man)_3$ $(GlcNAc)_2$	100.3
	$(Hex)_2$ $(HexNAc)_3$ $(Deoxyhexose)_3$ - $(Man)_3$ $(GlcNAc)_2$	100.6
	$(Hex)_1$ $(HexNAc)_3$ $(Deoxyhexose)_2$ - $(Man)_3$ $(GlcNAc)_2$	100.3
Average		**99.1**

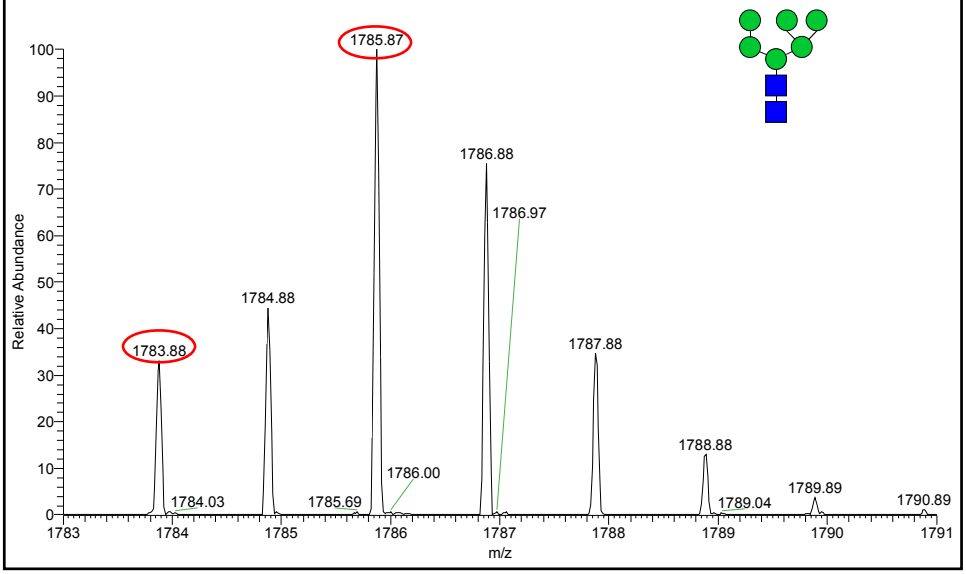


Assessing the dynamics of turnover and synthesis for individual glycans using pulse-chase I-DAWG

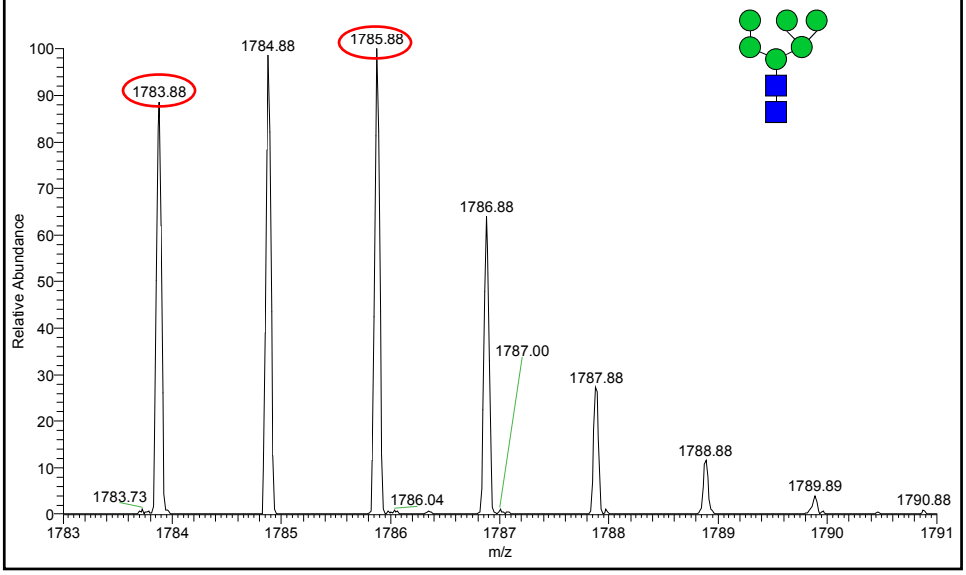

Time – 0 hr
Fully Labeled
Fully Unlabeled
Relative Abundance
m/z
1785.87
1786.88
1787.88
1785.99
1783.85
1784.87
1785.72
1783.38
1784.33
1787.06
1788.09
1788.88
1789.35
1789.88
1790.89


Time – 6 hr
Relative Abundance
m/z
1785.87
1786.88
1787.88
1784.88
1784.99
1783.88
1786.72
1788.89
1786.08
1787.03
1789.36
1789.89
1790.89

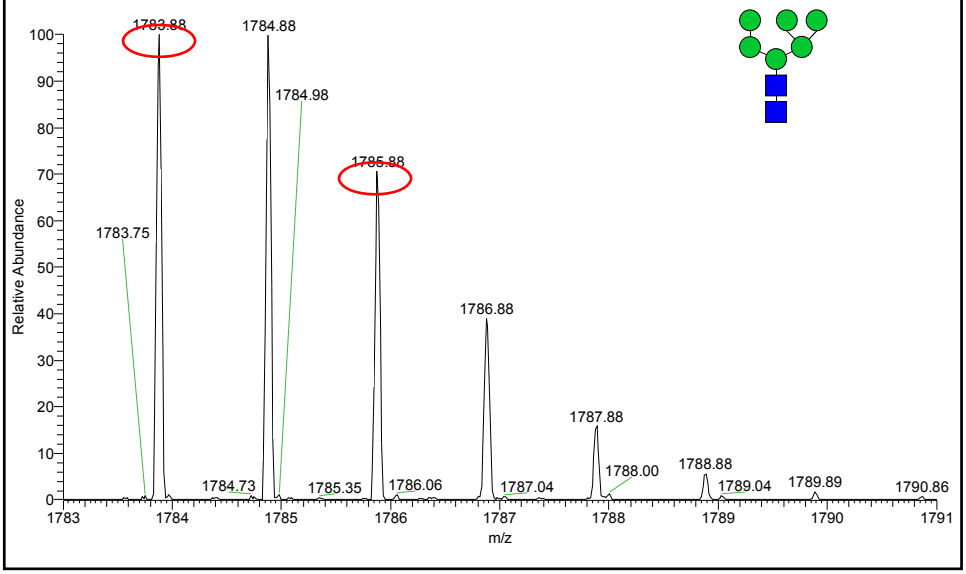
Time – 12 hr



Time – 24 hr

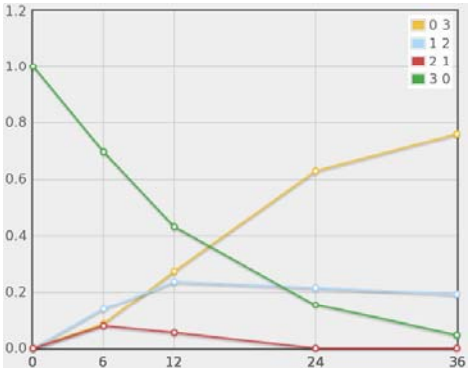


Time – 36 hr



Kinetic turnover of an O-linked glycan

0hr	1.0	0	0	0
6hr	0.644	0.125	0.141	0.089
12hr	0.390	0.062	0.219	0.328
24hr	0.147	0	0.215	0.638
36hr	0.044	0	0.191	0.765



Conclusions

- I-DAWG is introduced as an *in vivo* isotopic labeling strategy for quantitative - comparative glycomics/glycoproteomics.
- I-DAWG is expected to be applicable to a variety of glycans
- Currently, the use of I-DAWG in other cell lines, including those that are post mitotic, is being explored to optimize labeling times.
- I-DAWG enables other studies, kinetics, pulse-chase, etc

The I-DAWG approach is an easily applied and powerful new tool in the glycomics toolbox.

I-Dawg Summary

- Advantages
  - Higher level of accuracy than label free and presumably isotope labeling approaches
  - Can be used for glycan turnover-kinetic experiments
- Disadvantages
  - Performed on cell cultures
  - Additional cost
  - Best performed on a high resolution MS

Methods for Quantitation – Proteomics/Glycomics

Label Free	AMT Spectral Counts	Total Ion Mapping
in vitro labeling	I-CAT, ^{18}O, etc. I-TRAQ	$^{12}CH_3$ - $^{13}CH_3$ QUIBL
in vivo labeling	SILAC	I-DAWG