A Practical Beginner's Guide to Proteomics

This manuscript (permalink) was automatically generated from jessegmeyerlab/proteomics-tutorial@343a832 on January 15, 2022.

Authors

- **Jesse G. Meyer**
  0000-0003-2753-3926 · jessegmeyerlab · j_my_sci
  Department of Biochemistry, Medical College of Wisconsin · Funded by Grant R21 AG074234; Grant R35 GM142502

Abstract

Proteomics is the large scale study of protein structure and function from biological systems. “Shotgun proteomics” or “bottom-up proteomics” is the prevailing strategy for proteomics, which refers to the fragmentation of proteins into peptide pieces for easier analysis. Mass spectrometry is used to detect peptides from proteome fragmentation. Proteomics is commonly applied to identify and quantity as many proteins as possible from biological systems. Such quantitaive comparisons enable discovery of proteins that change between two or more conditions. More diverse views of the proteome are also possible using mass spectrometry, for example we can measure protein-protein interactions, post-translational modifications, and protein stability. To collect these diverse types of information, there are diverse strategies for proteome analysis, and each step in the workflow is tunable. The complexity and nuance of how proteomic methods differ may be difficulty to understand for new practitioners. Here, we provide a comprehensive tutorial of proteomics methods. Our tutorial starts with protein extraction and covers all steps ending with biological interpretation and study design. We expect that this work will serve as a basic resource for new practitioners of the field of shotgun or bottom-up proteomics.

Introduction

[paragraph about what proteomics means today]

[history of proteomics? how we got here]

[paragraph about what proteomics can do] A wide range of questions are addressable with proteomics experiments, which translates to a wide range of variations of proteomics workflows. Sometimes identifying what proteins are present is desired, and sometimes the quantities of as many proteins as possible are desired.

Protein Extraction

Discussion of methods for protein extraction and solubilizaition.

1. Choice of Lysis buffer

- Urea can cause chemical modifications

2. chemicals to avoid
3. removal of contaminations, Protein Precipitation
4. protein alkylation

- choices of reduction and alkylation reagents, TCEP/DTT/2BME, Chloroacetamide/iodoacetamide, n-ethyl maleimide

Proteolysis

1. discussion of protein sequence coverage is determined by the choice of proteolysis
2. why trypsin is the most common choice (charge and length character)
3. theoretical studies of proteolysis and enzyme [1]
4. Challenges associated with alternative enzyme choices (non-specific and semi-specific enzymes)
5. Alternative enzyme choices (one paragraph each?) - LysC
6. GluC
7. AspN
8. Alpha-lytic protease [2] and how it enables mapping human SUMO sites [3].
9. others?

Peptide and Protein Labeling

Discussion of methods to isotopically label peptides or proteins that enable quantification

1. SILAC/SILAM
2. iTRAQ
3. TMT
4. dimethyl labeling

Peptide or Protein Enrichment

Protein enrichment (e.g. for protein protein interactions)

- coIP
- APEX
- bioID
- bioplex

Peptide enrichment

- antibody enrichments of modifications, e.g. lysine acetylation [4].
- TiO2 and Fe enrichment of phosphorylation
- Glycosylation
- SISCAPA

Methods for Peptide Purification

1. Reverse phase including tips and cartridges
2. stage tips
3. in stage tip (iST)
4. SP2, SP3
5. s traps

Types of Mass Spectrometers used for Proteomics

1. QQQ
2. Q-TOF
3. Q-Orbitrap
4. LTQ-Orbitrap
5. TOF/TOF
6. FT-ICR
7. types of ion mobility

- SLIM
- FAIMS
- traveling wave
- tims

Peptide Ionization

The 2002 Nobel Prize in Chemistry was awarded to partially to John Fenn and Koichi Tanaka “for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules” [5/].

MALDI

Electrospray Ionization

Data Acquisition

Data acquisition strategies for proteomics fall generally within targeted or untargeted, and they can depend on the data (data dependent acquisition or DDA) or be data independent (data-independent acquisition or DIA).

DDA

Targeted DDA

Untargeted DIA

DIA

Targeted DIA

Untargeted DIA

Analysis of Raw Data

The goal of basic data analysis is to convert raw spectral data into identities and quantities of peptides and proteins that can be used for biologically-focused analysis.

Analysis of DDA data

Strategies for analysis of DIA data

Biological Interpretation

1. term enrichment analysis (KEGG, GO)
2. network analysis methods
3. structure analysis
4. isoform analysis
5. follow-up experiments

Experiment Design

This section should discuss trade offs and balancing them to design an experiment. 1. constraints: Each experiment will have different constraints, which may include the number of samples needed for analysis, or desire to quantify a specific subset of proteins within a sample. 2. sample size 3. statistics 4. costs

References

1. **In Silico Proteome Cleavage Reveals Iterative Digestion Strategy for High Sequence Coverage**
   Jesse G Meyer
   ISRN Computational Biology (2014-04-22) https://doi.org/gb6s2r
   DOI: https://doi.org/10.1155/2014/960902

2. **Expanding Proteome Coverage with Orthogonal-specificity α-Lytic Proteases**
   Jesse G Meyer, Sangtae Kim, David A Maltby, Majid Ghassemian, Nuno Bandeira, Elizabeth A Komives
   Molecular & Cellular Proteomics (2014-03) https://doi.org/f5vgcg
   DOI: https://doi.org/10.1074/mcp.m113.034710

3. **Site-specific identification and quantitation of endogenous SUMO modifications under native conditions.**
   Ryan J Lumpkin, Hongbo Gu, Yiying Zhu, Marilyn Leonard, Alla S Ahmad, Karl R Clauser, Jesse G Meyer, Eric J Bennett, Elizabeth A Komives
   Nature communications (2017-10-27) https://www.ncbi.nlm.nih.gov/pubmed/29079793
   DOI: 10.1038/s41467-017-01271-3 · PMID: 29079793 · PMCID: PMC5660086

4. **Simultaneous Quantification of the Acetylome and Succinylome by 'One-Pot' Affinity Enrichment**
   Nathan Basisty, Jesse G Meyer, Lei Wei, Bradford W Gibson, Birgit Schilling
   PROTEOMICS (2018-08-19) https://doi.org/gn4cmb
   DOI: 10.1002/pmic.201800123 · PMID: 30035354 · PMCID: PMC6175148

5. **The Nobel Prize in Chemistry 2002**
   NobelPrize.org
   https://www.nobelprize.org/prizes/chemistry/2002/summary/