Mass Spectrometry in the Genomes To Life Center for Molecular and Cellular Systems

Gregory B. Hurst[1] (hurstgb@ornl.gov), Robert L. Hettich[1], Nathan C. Verberkmoes[1], Gary J. Van Berkel[1], Frank W. Larimer[1], Trish K. Lankford[1], Steven J. Kennel[1], Dale Pelletier[1], Jane Razumovskaya[1], Richard D. Smith[2], Mary Lipton[2], Michael Giddings[5], Ray Gesteland[4], Malin Young[3], Carol Giometti[6]

[1]Oak Ridge National Laboratory

[2]Pacific Northwest National Laboratory

[3]Sandia National Laboratories

[4]University of Utah

[5]University of North Carolina

[6]Argonne National Laboratory

Overview

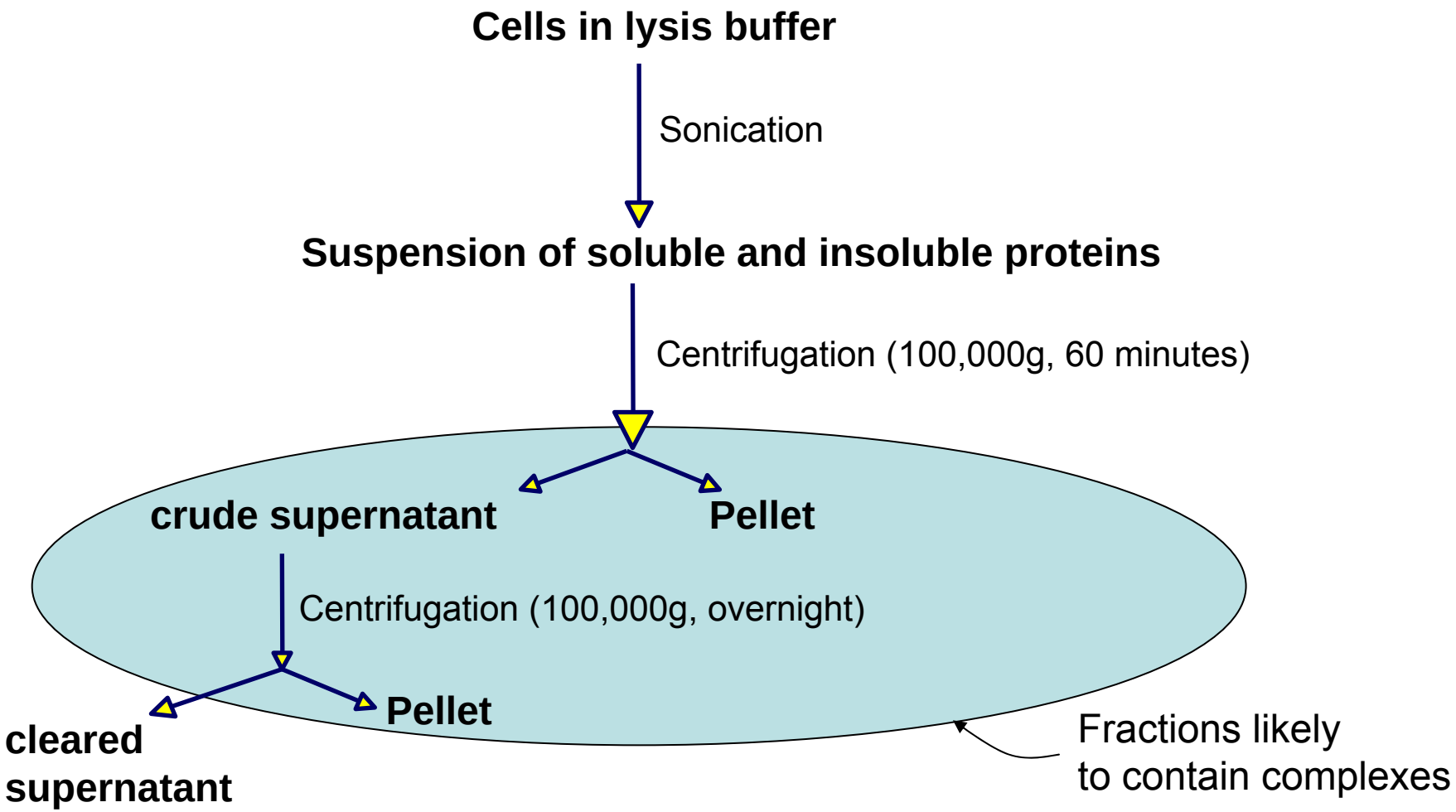
Mass spectrometry is a significant contributor to the Center for Molecular and Cellular Systems due to its capability for high-throughput identification of proteins and, by extension, protein complexes. From the outset of the Genomes To Life (GTL) Program, therefore, mass spectrometry has an important role to play in the pursuit of Goal 1 of the GTL--the identification of the "machines of life." The potential utility of mass spectrometry to GTL, however, extends far beyond current capabilities. In addition to incorporation of state-of-the-art mass spectrometry as a resource, we have also included a mass spectrometry research component as part of the Center for Molecular and Cellular Systems. The aim of this research component is to improve on existing mass spectrometry tools for protein complex characterization, as well as to produce new tools that will further the goals of the GTL program. Key to the success of this research component is close interaction with the protein expression, complex isolation, computational and imaging components of the Center.

Identification of Target Proteins for Expression with Affinity Tags

- Proteomics measurements on *R. palustris* help prioritize proteins for expression with affinity tags:
  - Proteins that are expressed (i.e., observed experimentally)
  - Proteins in pellets after short and overnight centrifugation steps (assumes complexes are “large”)

- Two growth conditions to date:
  - Aerobic, succinate, yeast extract
  - Anaerobic, photoheterotrophic
  - Duplicate bottom-up proteomics measurements complete
  - Top-Down proteomics measurement and data analysis in progress

Bacterial Sample Fractionation for Proteomics



R. palustris Proteome Analysis

"bottom-up*" method

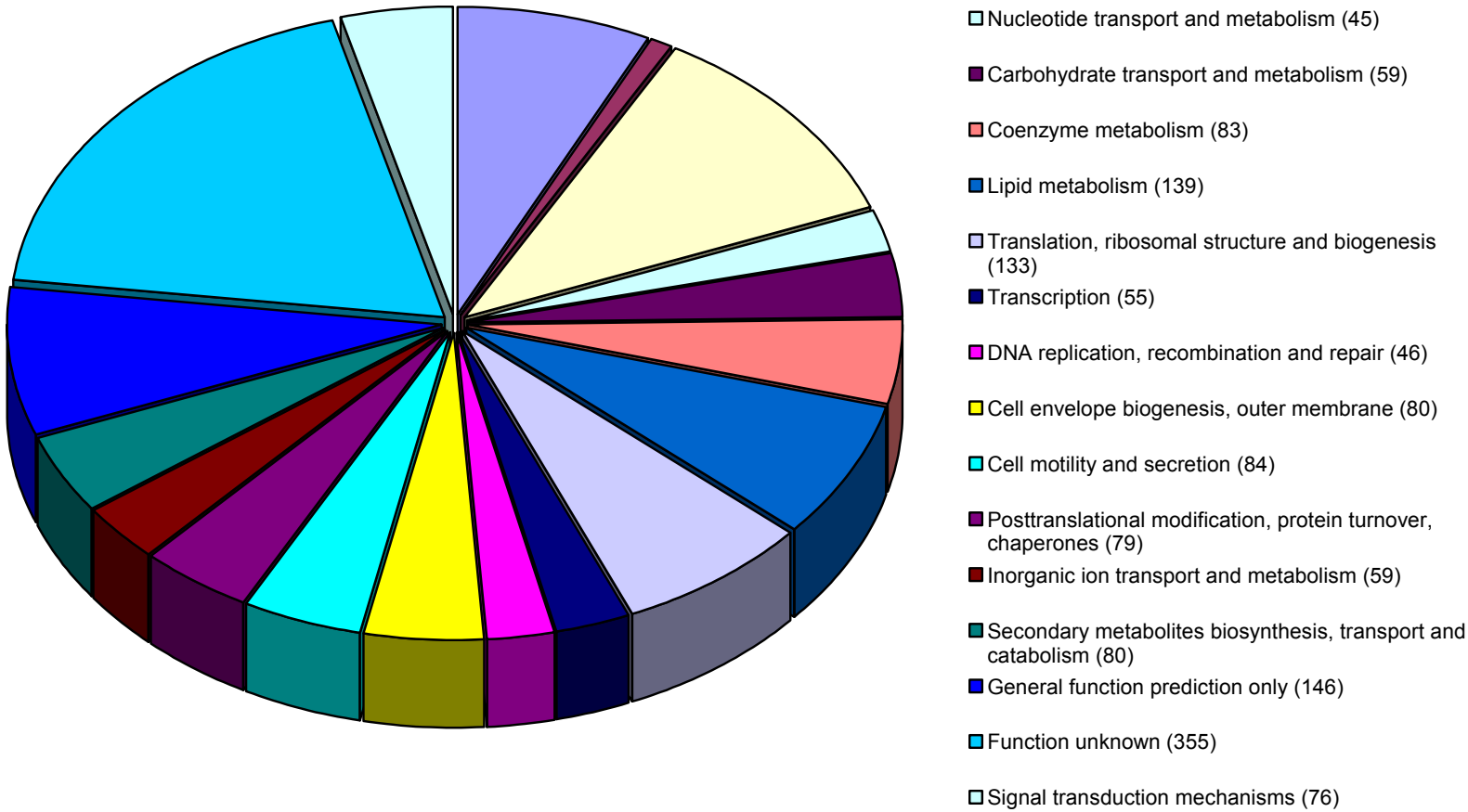
Numbers of proteins identified in fractions:

Fraction	Liberal** run1/run2	Strict‡ run1/run2	Liberal** run1/run2	Strict‡ run1/run2
cleared supernatant	551/598	363/378	641/645	420/447
crude supernatant	842/823	560/548	800/788	542/528
Pellet	661/638	442/418	634/659	433/472
Pellet (membranes +)	494/495	318/305	451/425	277/260
Total	1336/1353	948/939	1265/1243	884/879

Overall Total Proteins Identified: 1882 (liberal), 1283 (strict)

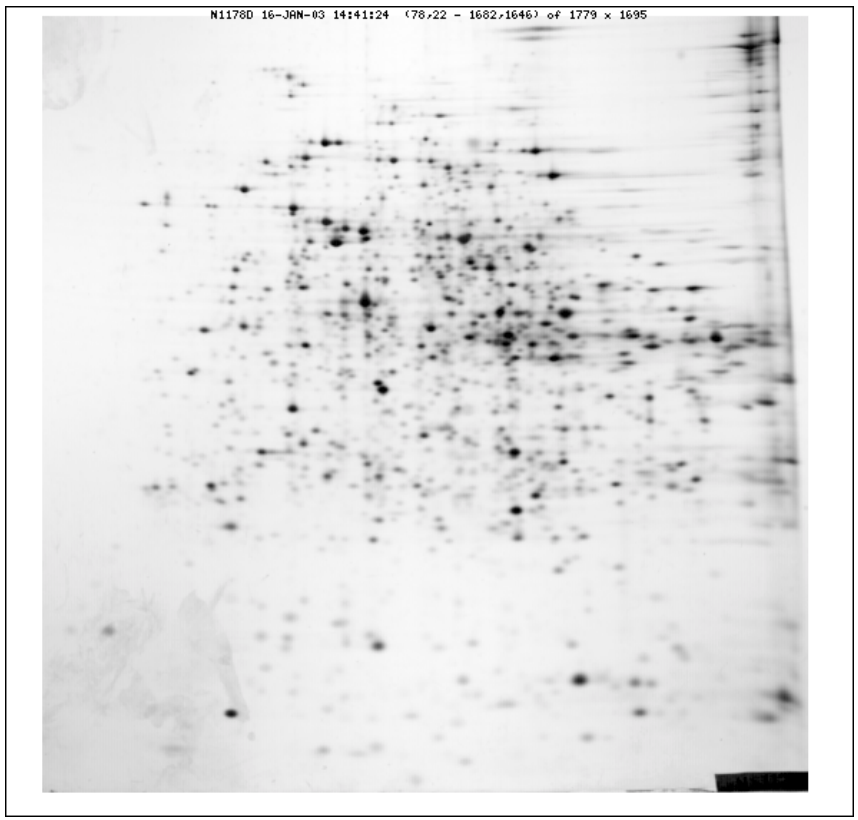
*bottom-up: trypsin-digested proteins analyzed by HPLC/quadrupole ion trap MS-MS
**liberal: ≥1 peptide required for protein ID; minimum Xcorr (charge state): 1.8(+1), 2.5(+2), 3.5(+3)
‡strict: ≥2 peptides required for protein ID; minimum Xcorr (charge state): 1.8(+1), 2.5(+2), 3.5(+3)

R. palustris Wild Type Proteome Analysis: Functional Categories

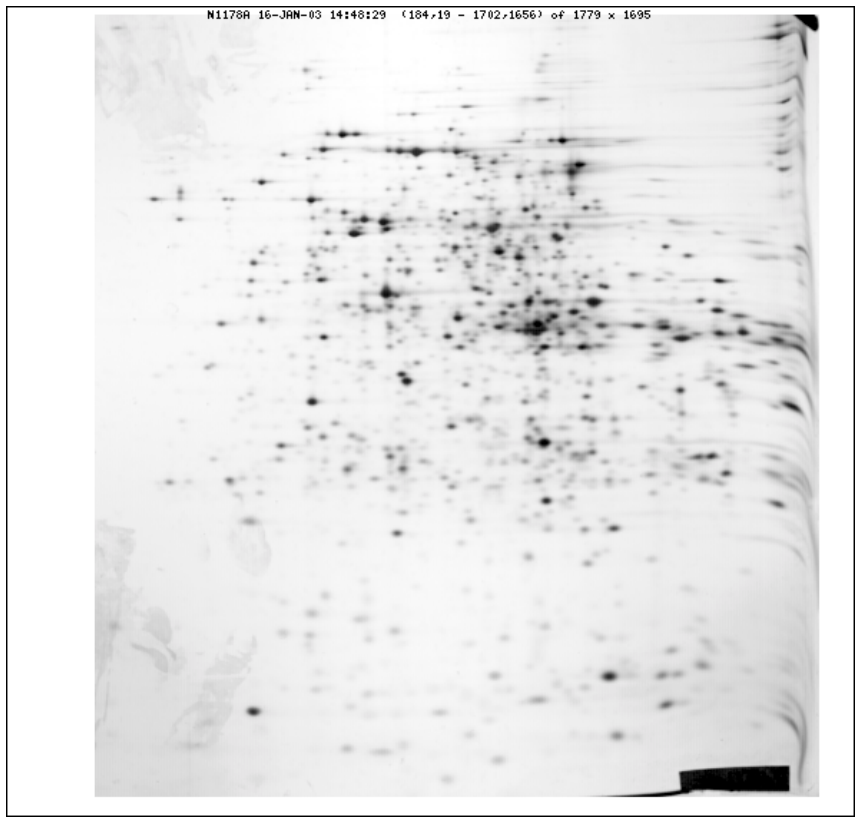


2D Electrophoresis of *R. palustris*

Argonne National Laboratory



Anaerobic growth,
Crude supernatant

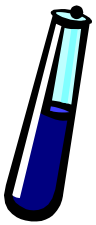


Aerobic growth,
Crude supernatant

R. palustris Proteomics

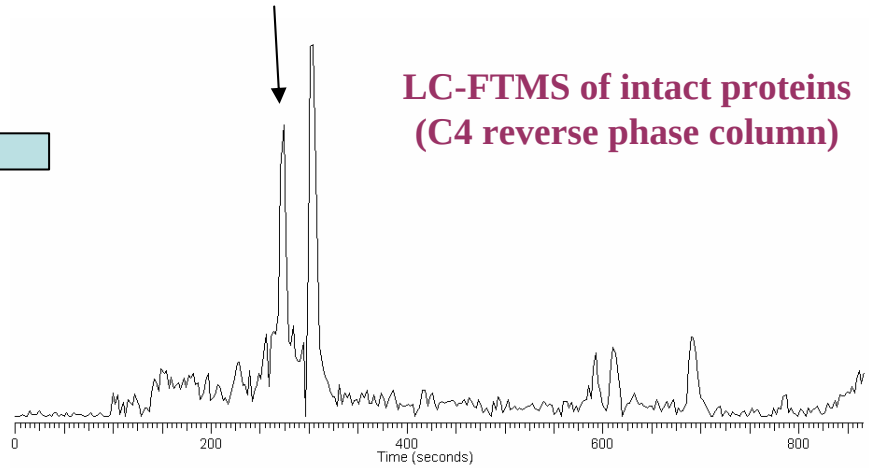
Top-down Approach

Intact protein mixture
(from anion exchange LC fractionation)



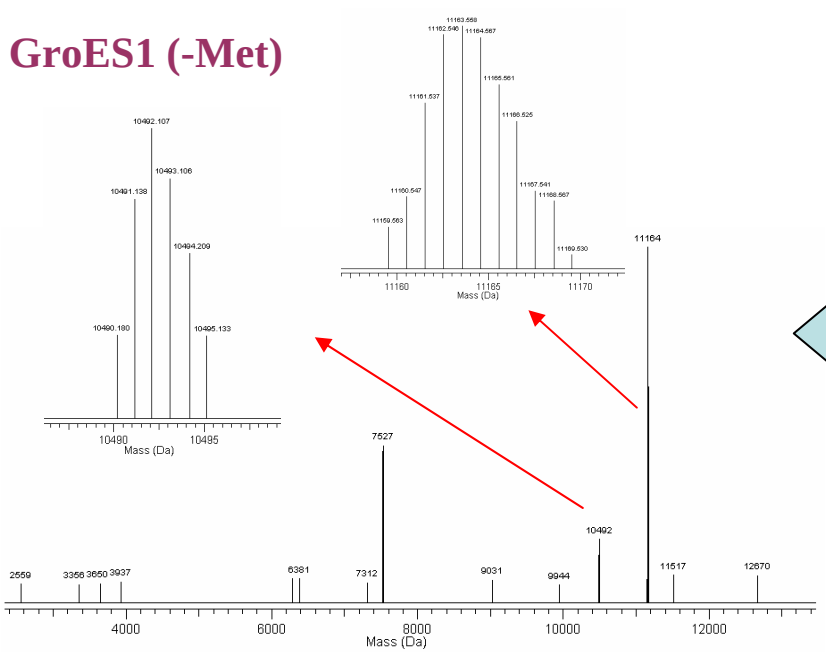
HPLC/FTMS Measurement of Intact Proteins

Scan #111



GroES2

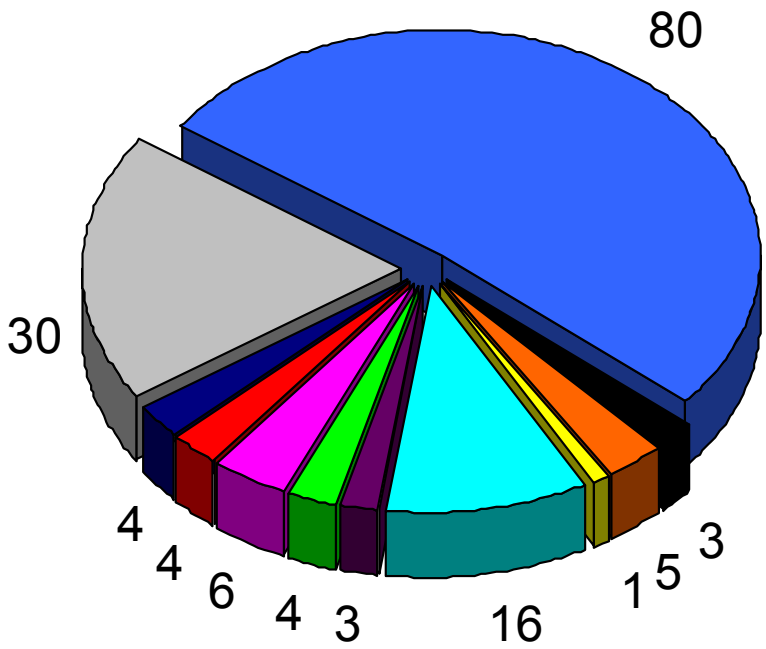
GroES1 (-Met)



Preliminary Top-Down MS Results for *R. palustris*

Intact Protein Identifications, COGS

156 Proteins measured, 76 proteins identified



Energy production and conversion (3)

Amino acid transport and metabolism (5)

Nucleotide transport and metabolism (1)

Translation, ribosomal structure and biogenesis (16)

Transcription (3)

Posttranslational modification, turnover, chaperones (4)

Inorganic ion transport and metabolism (6)

Function unknown (4)

Signal transduction mechanisms (4)

Hypothetical (30)

Proteins measured, but no match in database (80)

Current Status of Top-Down Experiments

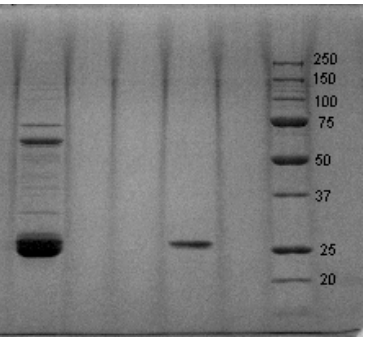
•LC-FTMS is being used to examine the **intact proteins** for *R. palustris* and provides high resolution, accurate mass measurements (< 3 ppm error)

•Preliminary inspection of the crude lysate yielded accurate mass measurements of 265 redundant proteins, including species from at least ten of the COGS functional categories. In total, accurate molecular mass measurements were made for 156 non-redundant proteins, in which 76 could be identified by database querying. The unidentified proteins (~80) could not be matched directly, due to the presence of post-translational processing, protein truncation, or gene annotation errors.

•Notable proteins include both GroES proteins, with GroES2 observed as the expected sequence and GroES1 observed with N-terminal methionine truncation.

•Work is underway to complete a two-dimensional LC-FTMS experiment, in which off-line anion-exchange chromatography is used to fractionate the complex mixture of proteins prior to LC-FTMS detection. This should permit measurement and identification of between 200-500 intact proteins for this organism under a single growth condition.

Verification of *R. palustris* Fusion Proteins Expressed in *E. coli*—Two Approaches



Affinity capture of tagged proteins from lysed cells

1D PAGE

whole eluate digestion

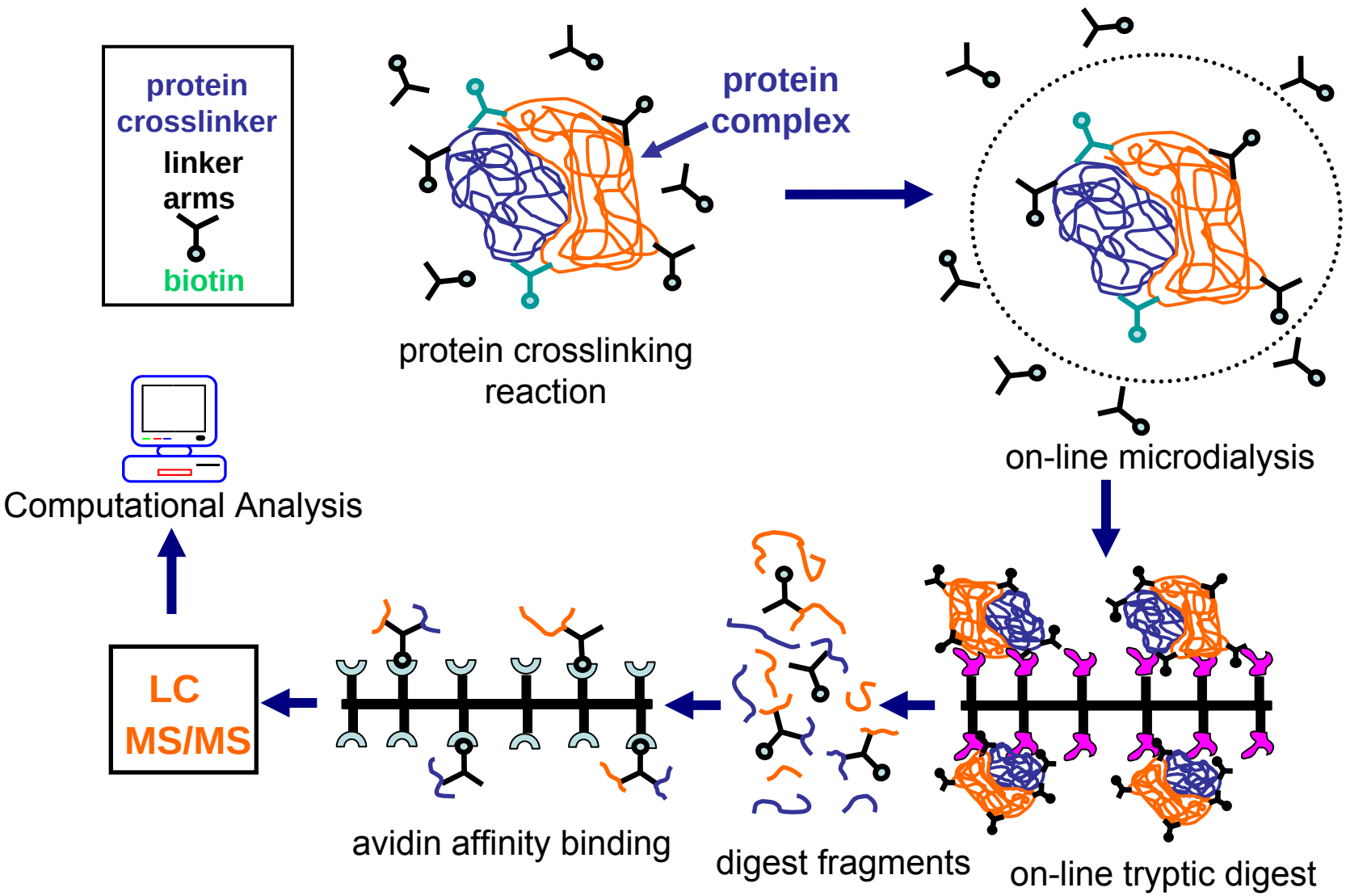
In-gel digestion and mass spectrometric identification of individual gel bands (See poster by Wiley *et al*.)

LC-MS-MS of digest peptides; identification of proteins via SEQUEST

Fusion Protein	No. of peptides identified from: target protein	No. of peptides identified from: affinity tag	Additional proteins identified *E. coli*	Additional proteins identified *R. palustris*
Rpal 4709 + N-terminal GST	45	8	0	2
Rpal 4709 + C-terminal 6-His & V5 epitope	31	3	19	0
Rpal 5426 + C-terminal 6-His & V5 epitope	35	3	7	1

*These are candidate methods for analysis of **protein complexes** isolated via affinity purification*

Protein Complex Analysis: Proposed Affinity Crosslinker Approach



Affinity Purification of Sulfo-SBED Crosslinked Peptide

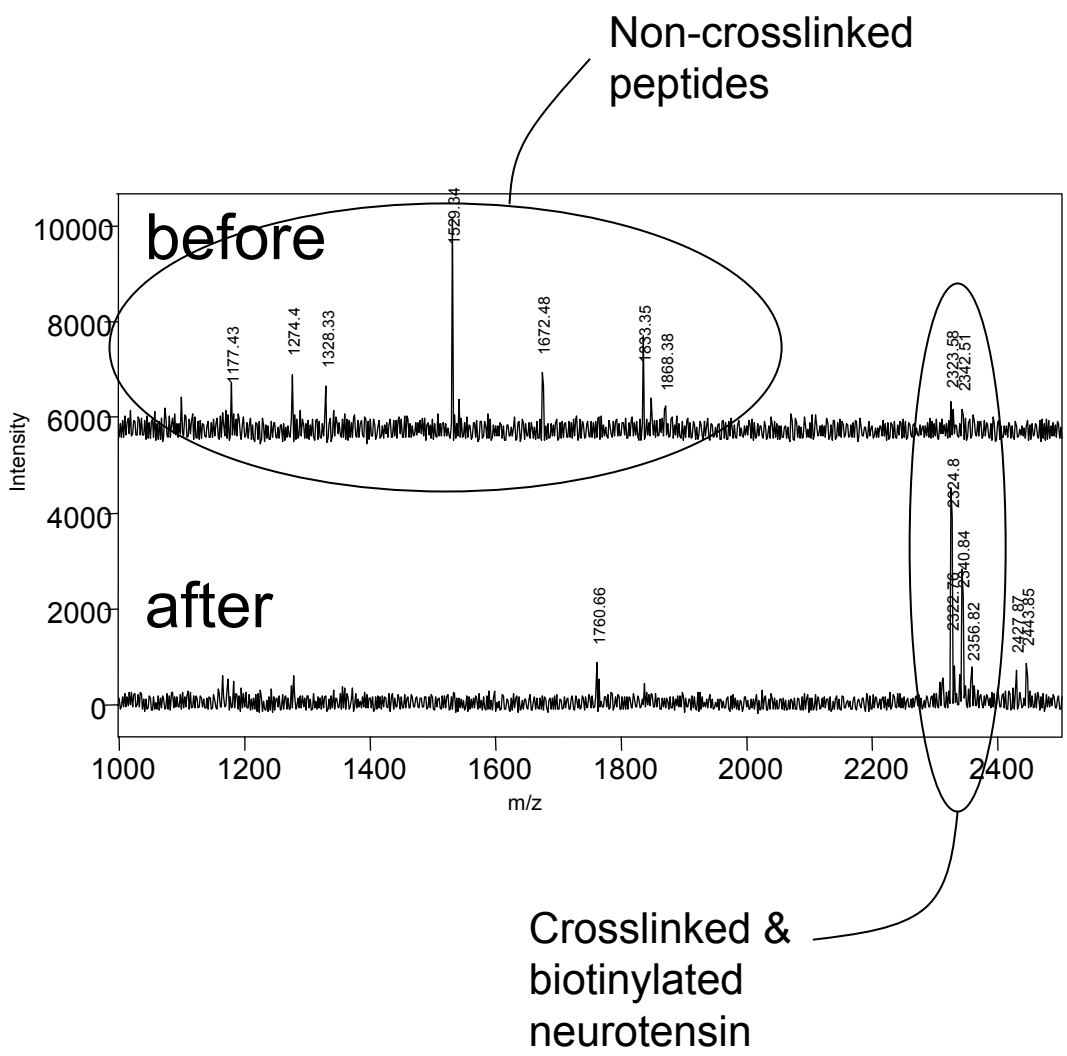
Biotin-labeled crosslinking reagents offer a potential method for “fishing” rare crosslinked peptides from digests of crosslinked protein complexes.

•Crosslink neurotensin internally with biotinylated reagent sulfo-SBED.

•Add crosslinked peptide to tryptic digest of hemoglobin (“interference” peptides)

•MALDI-TOF mass spectra **before** and **after** small-scale avidin affinity separation.

•The crosslinked (and therefore biotinylated) peptide is enriched, while “interference” peptides are greatly reduced.



Acknowledgements

- Research sponsored by Office of Biological and Environmental Research, U.S. Department of Energy.