Guidelines for Reporting Proteomic Experiments Using Mass Spectrometry

These guidelines are provided as a reminder for the information required for readers and reviewers to assess the validity of experiments and for experimentalists to reproduce experiments. It is understood that not all of this information is applicable to all papers and that authors will provide the information that is relevant to their studies. This information can be provided as part of supplemental methods.

Experimental design
- Number of replicates
  - Biological
  - Technical
- Labeling strategy (when applicable)

Analytical methods
- Sample preparation
  - Protein isolation
  - Proteolytic digestion (if used)
  - Chemical modification (if used)
    - Labeling scheme can be repeated(?)
  - Off-line chromatography/cleanup
  - Enrichment strategy
- Protein analysis
  - HPLC
    - Instrument vendor/model
    - Column(s), mobile phases, flow rate, gradient, auxiliary detection details
  - Mass spectrometry
    - Instrument vendor/model
    - Instrument parameters/scan strategy
      - Tandem-MS
        - Data-dependent analysis: MS1 mass resolution, MS1 scan range, charge-state screening parameters, mass window for precursor ion isolation, fragmentation mode, collision energy (or other parameter, as appropriate), mass analyzer for MS2, MS2 mass resolution (where appropriate), scan strategy for MS2, dynamic exclusion, analogous details for additional stages of MS (as appropriate)
        - Data-independent analysis: MS1 mass resolution, MS1 scan ranges, collision energy, mass analyzer for MS2, MS2 mass resolution (where appropriate),
      - MALDI-TOF-MS: mass resolution, scan range, laser parameters, data acquisition (*i.e.*, number of shots summed or averaged), flight mode (i.e., linear or reflectron)
      - MALDI-TOF-MS/MS: MS1 mass resolution, MS1 scan range, laser parameters, data acquisition (*i.e.*, number of shots summed or averaged), flight mode (i.e., linear or reflectron), number of ions selected for tandem-MS (including criteria for selection), mass window for precursor ion isolation, fragmentation mode, collision energy

Data processing
- Software/method for peak list generation
- Database searching
  - Software name, vendor or literature citation, version
  - Database(s)
    - Name/source
    - Date/version
    - Taxonomy

- Number of sequences
- Search parameters
  - Precursor and product ion mass tolerances
  - Enzyme specificity
  - Charge states considered
  - Fixed and variable modifications
  - Other search engine-specific settings (*e.g.*, ^{13}C number in Mascot)
- False discovery rate (FDR) determination
  - Method for FDR calculation
  - Decoy database details (if used)
- Criteria for acceptance of peptide assignments and protein identifications

De novo sequencing
- Approach (*e.g.*, manual or computational)
  - Software name, vendor or literature citation, version (if used)
- Validation

Quantitative analysis
- Software name, vendor or literature citation, version
- Quantitation parameters
- Details about inclusion/aggregation of isotopes and/or charge states
- Peak/quality filtering strategy
- Handling of shared peptides
- Data transformation and normalization
- Criteria for acceptance of peptide values
- Criteria for protein inference from peptide assignments

Statistics
- Software/program
  - Name, date/version
  - Literature citation (where appropriate)
- Power analysis (where appropriate)
- Test(s) applied
- Significance levels

Documentation
- Required for proteins that exhibited significant differences in quantity among experimental groups or are of special interest/focus in the study
- Spreadsheet format (e.g., Excel)
  - For long tables, provide legend on a separate worksheet
  - Clear/meaningful, consistent column headings
- Explanation of significance level required for assigned peptides (in legend)
- Tandem-MS
  - Protein group report
    - Protein name, accession number, molecular weight, number of assigned spectra used for identification, number of unique sequences, percent sequence coverage, probability of protein inference (if determined)
  - Peptide report (grouped by protein)

- Peptide sequence, start/stop residue numbers, observed mass, mass error, score/expect value for assignment, modifications, probability for modification site (where appropriate)

MALDI-TOF-MS (PMF) - only suitable for low-complexity samples
- Protein group report
  - Protein name, accession number, molecular weight, number of spectra searched, number of spectra assigned, percent sequence coverage, probability of protein inference

MALDI-TOF/TOF-MS
- Protein group report
  - Protein name, accession number, molecular weight, number of spectra searched for PMF, number of spectra assigned for PMF, number of tandem mass spectra assigned, percent sequence coverage, probability of protein inference

Post-translational modifications (excluding sample preparation artifacts)
- Peptide sequence
- Observed mass, mass error and charge state
- Database search score
- Probability/expect score
- Probability of site localization (with an indication in the legend of how the probability was determined)
- Annotated tandem mass spectra
  - If modified peptide is listed in a table or discussed in the manuscript, spectrum needs to be provided
  - For unusual modifications or surprising findings that are the focus of the manuscript - provided either in the body of the manuscript or as supplemental data
  - Annotations
    - Peptide sequence showing site(s) of modification
    - Observed mass, mass error and charge state
    - Database search score
    - Probability/expect score
    - Assigned ions (designated as corresponding ion series)
    - Neutral loss fragments

Proteins identified on the basis of a single, high-confidence peptide assignment
- Annotated tandem mass spectrum
  - Peptide sequence showing any site(s) of modification
  - Observed mass, mass error and charge state
  - Database search score
  - Probability/expect score
    - Assigned ions (designated as corresponding ion series)
    - Neutral loss fragments

Quantitative analysis
- Supplementary file in spreadsheet format (i.e., Excel)
- Columns added to the identification table described above or provided in a separate table
  - Label-free (peak area)

For each replicate: total number of spectra used for identification, number of unique sequences used for identification, percent sequence coverage, total number of spectra used for relative quantification

Mean peak area (± standard deviation) for each experimental group, the ratio between experimental groups and a measure of significance

Spectral counting

For each replicate: total number of spectra used for identification, number of unique sequences used for identification, percent sequence coverage, total number of spectra used for spectral counting

Mean ratio (± standard deviation) for each experimental group and a measure of significance

MS1-based labeling (e.g., SILAC, reductive methylation)

For each replicate: total number of spectra used for identification, number of unique sequences used for identification, percent sequence coverage, total number of spectra used for relative quantification

Mean (or median) ratio (e.g., H/L or L/H) between experimental groups (± standard deviation), variability of ratios for peptides assigned to each protein, transformed and/or normalized values (as appropriate), measure of significance.

Reporter ion-based methods

For each replicate: total number of spectra used for identification, number of unique sequences used for identification, percent sequence coverage, total number of spectra used for relative quantification

Mean (or median) reporter ion ratio, variability of reporter ion ratios for peptides assigned to each protein, transformed and/or normalized values (as appropriate), measure of significance.