Mass Spec Software to Aid Data Interpretation at the Speed of Biotherapeutics Development

Alex Zhai[1], Michael Poltash[1], Ilker Sen[2], St John Skilton[1], Hirsh Nanda[1], Partha Chowdhury[1], Andrew Mahan[1]

[1]Johnson & Johnson Innovative Medicine, Spring House, PA
[2]Protein Metrics, MA

Background & Objective

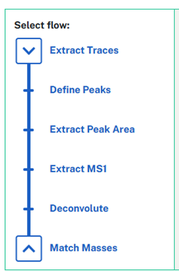
High-Complexity Biotherapeutics Require Faster MS Data Interpretation

The rapid expansion of complex biotherapeutic formats — including multispecific antibodies, antibody-drug conjugates, scFvs, antibody fragments, and other engineered biologics has increased both the complexity and volume of mass spectrometry datasets generated during discovery and development. As pipeline throughput increases, manual review of spectra becomes increasingly difficult for mass spectrometrists and lab managers. Traditional software outputs such as static tables and basic charts provide useful summaries, but they are often insufficient for rapid, scalable decision-making across large datasets.

Objective

To demonstrate how an advanced vendor-neutral mass spectrometry software platform can accelerate interpretation of complex biotherapeutic datasets by combining: 1) Automated raw data processing; 2) Interactive visualization of large datasets; 3) ML-and logic-assisted data triage; 4) Traffic-light quality indicators; 5) Plate-based heatmaps for clone ranking; 6) Automated report generation for remote review.

Vendor-Neutral Automated MS Workflow



Datasets containing multispecific biotherapeutics were processed using a vendor-neutral software platform capable of querying and visualizing large MS datasets.

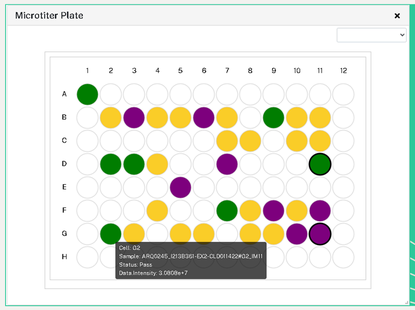
Workflow Steps

- Raw MS data import
- Automated data processing
- Molecular-weight and species identification
- ML-and logic-based triage
- Pass / fail / ambiguous status assignment
- Interactive visualization
- Automated report generation
- Remote review by scientists and lab managers

Intelligent Triage System

Automated Assessment Reduces Manual Spectral Review

An intelligent triage system was used to evaluate the correctness and quality of software-generated identifications. The system combined machine learning with rule-based logic to assign actionable review categories without requiring manual inspection of every spectrum.



Triage Outputs

- •Green: Expected molecular weight and acceptable product quality
- •Yellow: Ambiguous or partially characterized data requiring review
- •Red: Failed identification, poor quality, or undesired species detected
- •Purple: Mixed results from multiplicated samples

Key Benefit

This triage approach allows large datasets to be reviewed rapidly, while preserving human expert attention for the samples most likely to require interpretation.

Case Study: 96-Clone CHO Screening

The automated workflow was applied to a 96-well CHO clone screening dataset. Each clone expressed an anti-RSV scFv-Fc × Fab bispecific antibody. During cell line development, large numbers of clones must be screened to identify candidates with favorable product quality profiles. For multichain biologics, this requires selecting clones that produce the desired therapeutic species while minimizing undesired assembly products.

Clones were evaluated based on:

- Correct molecular weight
- Desired bispecific antibody formation
- Balanced chain expression
- Low abundance of off-target or mispaired species
- Overall data quality
- Software-assigned triage status

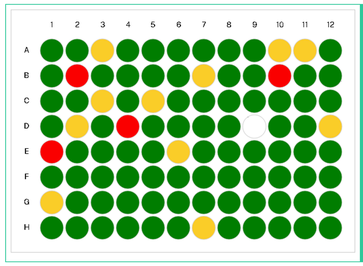
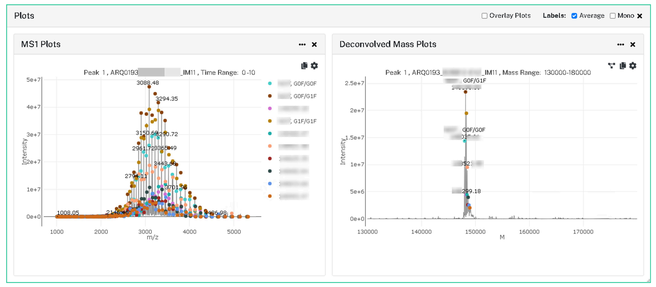


Figure 1. 96-well plate visualization for automated clone ranking.
Green wells indicate high-quality clone candidates, yellow wells indicate ambiguous results requiring review, red wells indicate failed or low-quality candidates, and white indicate the sample is missing in the corresponding location.

Intact Mass Analysis of CHO Antibody



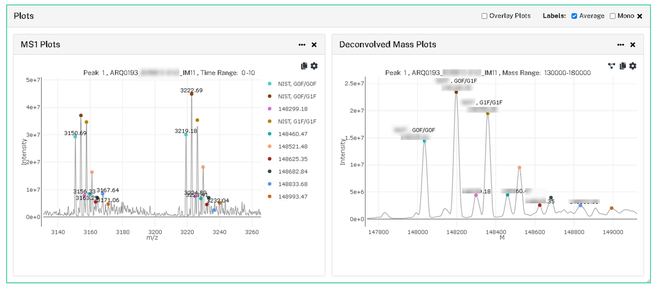


Figure 2. Representative intact MS raw and deconvoluted spectra of one of the 96 samples in the micro titer plate.
Upper panel: zoom out spectra; Lower panel: zoom in spectra of the same sample.

Results and Impact

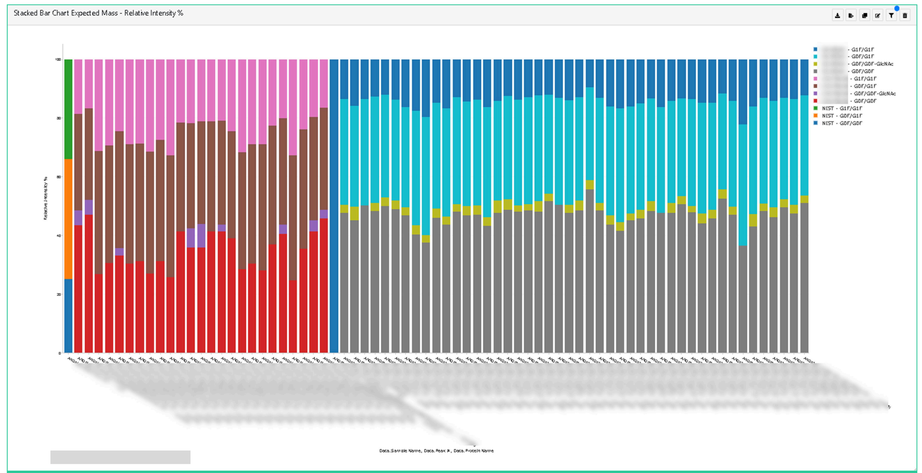


Figure 3. Relative intensity distribution of detected species across clones.

Stacked bar chart showing the relative intensity (%) of expected and related mass species for each analyzed sample. Each bar represents an individual clone, with color-coded segments corresponding to different identified molecular species (e.g., correctly assembled bispecific antibody, glycoforms, and mispaired or off-target species). The left portion of the dataset highlights greater heterogeneity and larger contributions from undesired or mispaired species, while the right portion shows improved profiles dominated by the expected product with reduced impurity signals. This visualization enables rapid comparison of product quality across clones and supports software-assisted ranking and selection of high-quality candidates for further development

Software-Assisted Review Accelerates Decision-Making

The combination of heatmaps, traffic-light indicators, and automated triage enabled rapid interpretation of complex multispecific antibody datasets.

Key Findings

- Large datasets were summarized into intuitive visual outputs.
- Manual review was reduced by focusing expert attention on ambiguous cases.
- High-confidence clone candidates could be rapidly identified.
- Lab managers could assess dataset-level quality without directly reviewing raw spectra.
- Plate-based visualization supported rapid cherry picking of top clones.
- Review time was reduced by hours per dataset compared with traditional manual workflows.

Conclusions

Advanced MS Software Enables Faster Biotherapeutic Decisions

Automated visualization and intelligent triage provide a scalable approach for interpreting complex biotherapeutic MS datasets.

- Complex multispecific datasets can be converted into intuitive visual summaries.
- Traffic-light triage enables rapid identification of high-quality clone candidates.
- Plate heatmaps support efficient 96-well clone screening workflows.
- Automated reports reduce manual interpretation burden.
- Software-assisted interpretation accelerates decision-making in biologics discovery and development.

Final Take-Home Message

Mass spec software can move biotherapeutic characterization from manual data review to rapid, automated, decision-oriented interpretation.

Johnson&Johnson Protein Metrics ThermoFisher SCIENTIFIC Therapeutics Development & Supply