Characterization of Sequence Variants in Protein Therapeutics by Liquid Chromatography Tandem Mass Spectrometry and Automated Data Processing

Scott A. Ugrin[1], Michelle English[2], and Colin Barry[1]
[1]*Alliance Pharma, Inc., 17 Lee Blvd, Malvern, PA 19355*
[2]*Protein Metrics, 20863 Stevens Creek Blvd #450, Cupertino, CA 95014*

Introduction

- As biologics continue to gain prominence in the pharmaceutical industry, stringent characterization of biosynthetic products is essential for ensuring the clinical safety and efficacy of drug products. By its nature, biosynthesis can result in variations to the amino acid sequence that can have deleterious effects on a protein's structure and function. Therefore, characterization of sequence variants is crucial to help direct adaptations to upstream processes during early drug development.
- Sequence variant characterization workflows typically involve time-intensive data interpretation to differentiate false positives from true positives.
- A semi-automated data analysis workflow is presented using a vendor neutral software package, Protein Metrics Byos®, to filter out false positive results and dramatically reduce the time required for data interpretation.

Methods

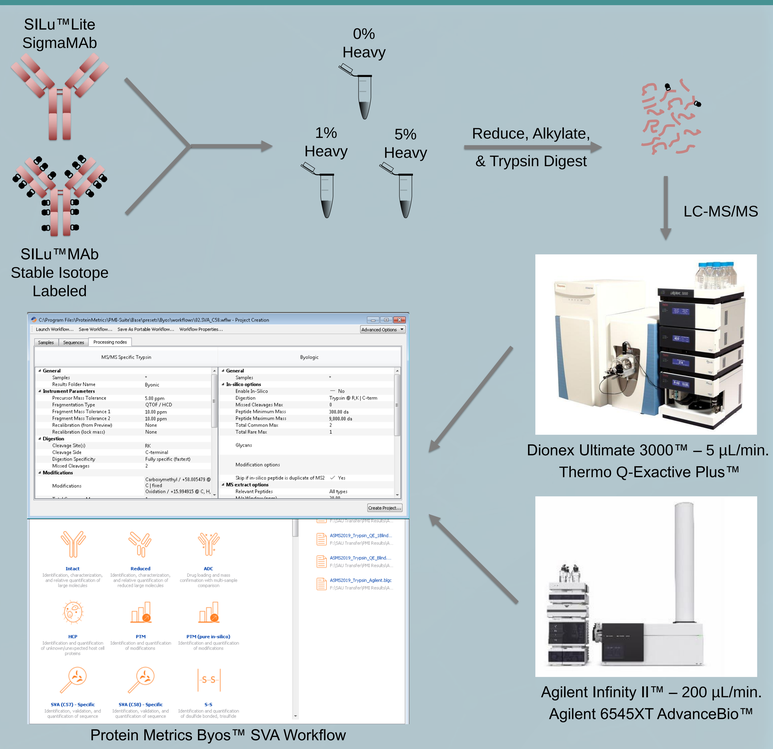


Figure 1. Trypsin peptide mapping workflow for sequence variant analysis (SVA).

Methods (Cont.)

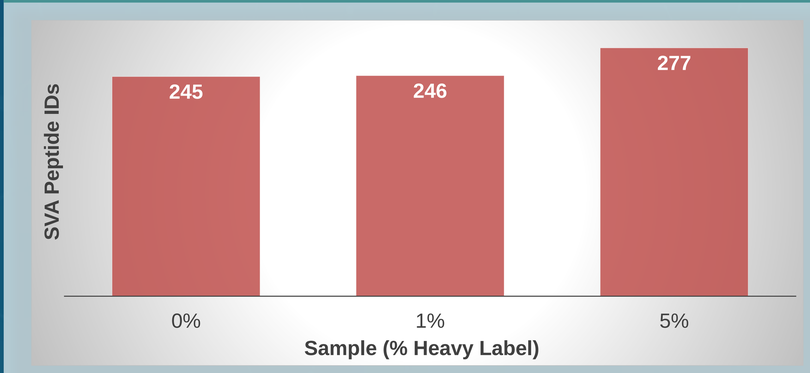
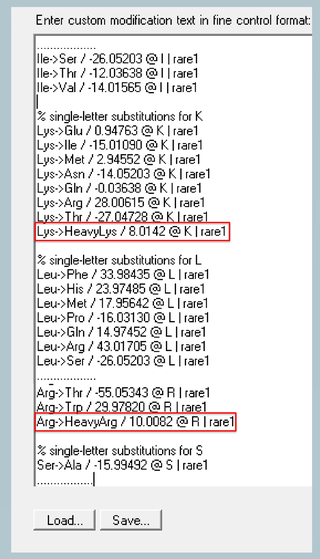


Figure 2. Initial sequence variant search results for each sample.

- Search results contain a significant number of false positives.
- Manual interpretation of search results is time-consuming and labor-intensive.
- Heavy-labelled Lys [$^{13}C_6$, $^{15}N_2$] & Arg [$^{13}C_6$, $^{15}N_4$] added to preloaded list of single-base substitutions included in Byos SVA Workflow.
- Allowed monitoring of true positive versus "false positive" results, while applying preset filters.
- 37 true positives expected.



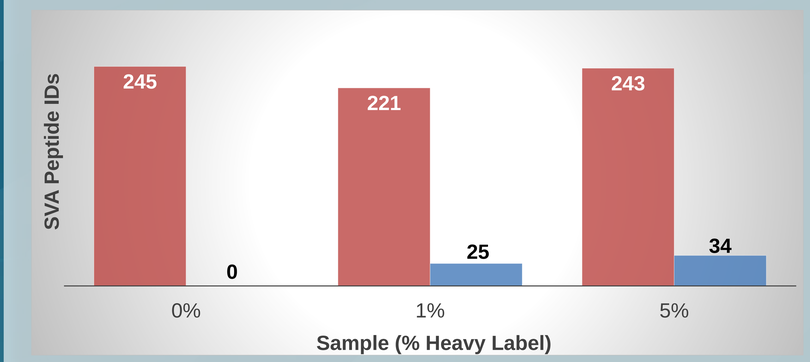


Figure 3. Initial true positive (Blue) vs. "false positive" (Red) sequence variant results.

Results

Filter #	Description
1	Modified peptides with no corresponding wildtype peptide
2	Sequence variant mass shift aligns with a common PTM (oxidation & deamidation)
3	Modified peptides that have a low MS/MS score (50)
4	Assignment based on an incorrect precursor mass
5	Peptides with more than one modification
6	Excessive retention time (RT) shift compared to calculated RT shift
7	XIC ratio higher than 10% or lower than 0.2%

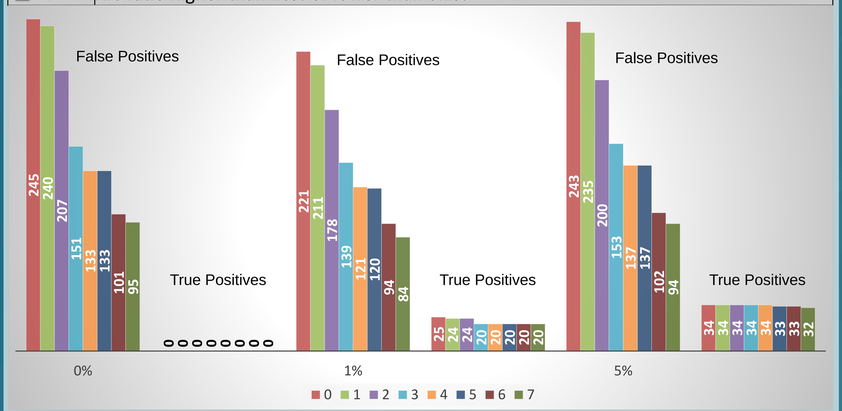


Figure 4. Remaining "false positive" and true positive sequence variant identifications with the application of each preset filter. Application of all filters results in the removal of approximately 61% of false positive results and only 20% and 6% of true positive results from the 1% and 5% samples, respectively.

Conclusions

- Protein Metrics Byos™ software used in conjunction with preset filters for removing likely false positive results dramatically reduced the number of sequence variant identification results and reduced the time necessary for manual interpretation of results.
- By introducing known true positive mass shifts and treating them as sequence variants, we were able to assess the accuracy of this methodology.
- A majority of false positive results were removed by the preset filters, cutting data interpretation time in half.

Acknowledgements

The authors would like to thank Ilker Sen, St. John Skilton, and James Moore of Protein Metrics, Inc.™ for providing the preset filters and invaluable software support.

Reference

1. Protein Metrics Application Note – Sequence Variant Analysis, February 2019.