Robust Optimization of SWATH-MS workflow for human blood serum proteome analysis using a Quality by Design approach

Edith Serrano-Blesa
Newcastle University
Andrew Porter
Newcastle University Protein and Proteome Facility
Dennis William Lendrem
Newcastle University
Costantino Pitzalis
Queen Mary University of London
Anne Barton
The University of Manchester
Achim Treumann
Newcastle University Protein and Proteome Facility
John D. Isaacs (john.isaacs@ncl.ac.uk)
Newcastle University

Research

Keywords: SWATH-MS, proteomic quantification, design of experiments, quality by design, screening, optimization, robustness

Posted Date: March 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-347514/v1

Version of Record: A version of this preprint was published at Clinical Proteomics on August 12th, 2021. See the published version at https://doi.org/10.1186/s12014-021-09323-z.

Abstract

Background: It is not enough to optimize proteomics assays. It is critical those assays are robust to operating conditions. Without robust assays, proteomic biomarkers are unlikely to translate readily into the clinic. This study outlines a structured approach to the identification of a robust operating window for proteomics assays and applies that method to Sequential Window Acquisition of all Theoretical Spectra Mass Spectroscopy (SWATH-MS).

Methods: We used a sequential Quality by Design approach exploiting a fractional screening design to first identify critical SWATH-MS parameters, then using response surface methods to identify a robust operating window with good reproducibility, before validating those settings in a separate validation study.

Results: The screening experiment identified two critical SWATH-MS parameters. We modelled the number of proteins and reproducibility as a function of those parameters identifying an operating window permitting robust maximization of the number of proteins quantified in human serum. In a separate validation study, these settings were shown to give good proteome-wide coverage and high quantification reproducibility.

Conclusions: Using design of experiments permits identification of a robust operating window for SWATH-MS. The method gives a good understanding of proteomics assays and greater data-driven confidence in SWATH-MS performance.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures

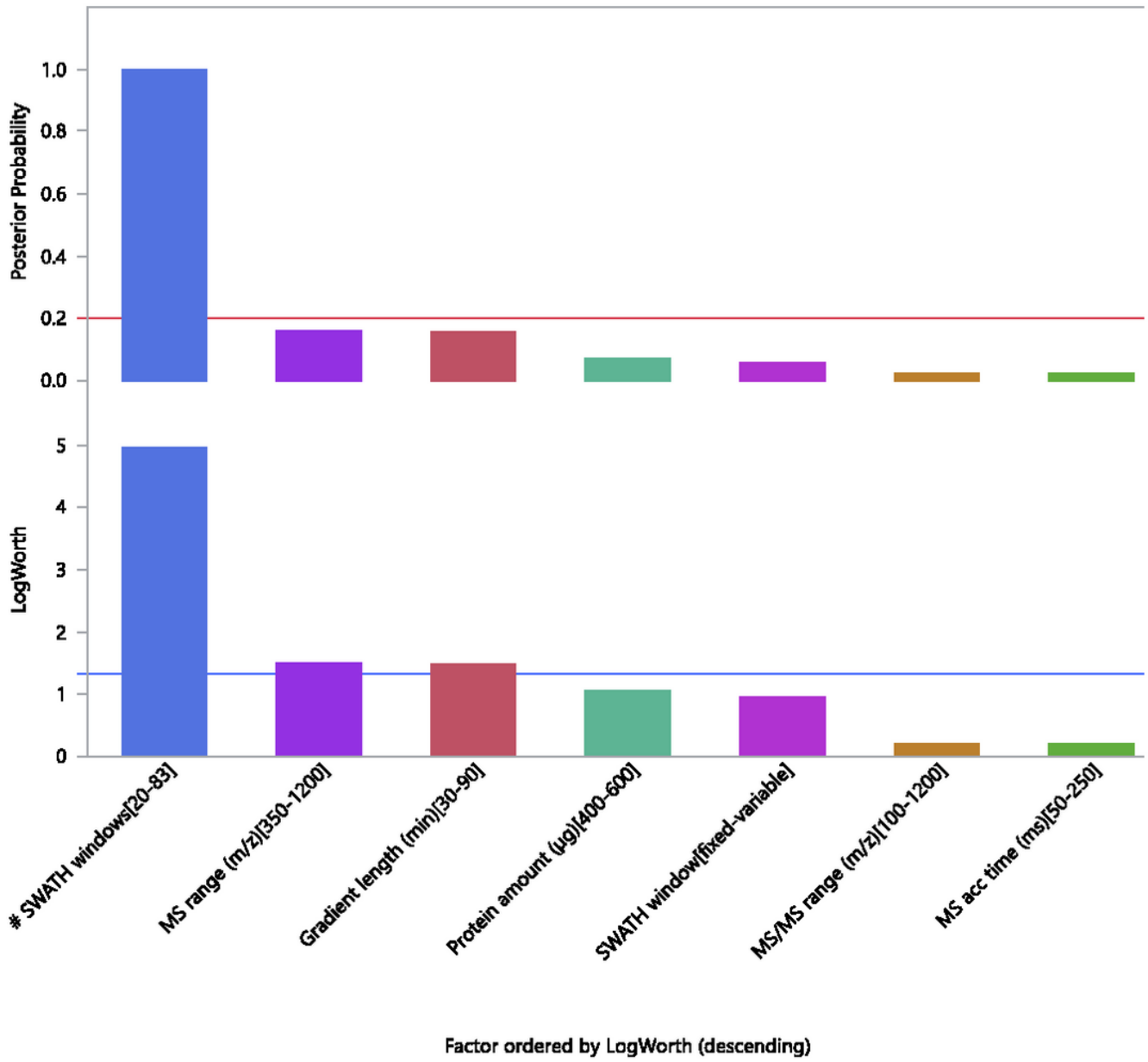

Figure 1

Bayes plot showing the posterior probabilities revised in light of the screening study data. The prior probabilities (p=0.20) are indicated by the red horizontal line. Note the large increase in probability for the number of SWATH windows. The LogWorth value corresponding to a statistically significant effect at the p<0.05 level of significance is indicated by the blue horizontal line. While the Number of SWATH windows, MS range and Gradient length are all statistically significant, the posterior probability increases only for the Number of SWATH windows. Increasing the number of precursor windows from 20 to 83 increases the number of proteins recovered. Note that within the experimental ranges used the method is relatively robust to variation in protein amount, choice of fixed or variable window, MS/MS range and MS accumulation times.

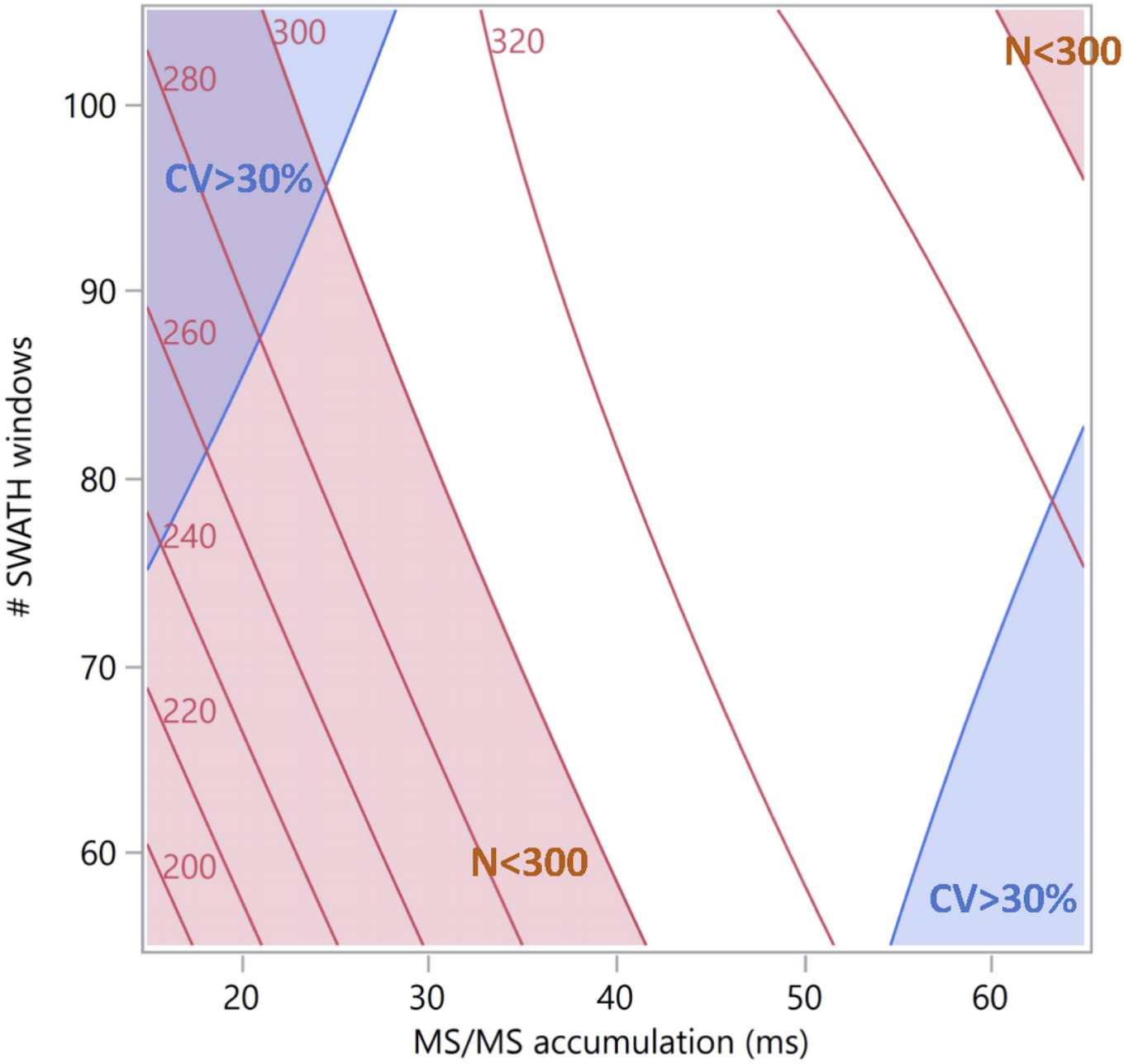

Figure 2

Contour plots showing the mean number of proteins (in red) as a function of the Number of SWATH Windows and the MS/MS accumulation time. The blue regions indicate areas of high variability (coefficient of variation > 30%). The pink regions indicate areas of low yield (less than 300 proteins recovered). The remaining space (in white) marks a safe operating region where we are likely to meet both constraints recovering higher numbers of proteins with reduced variability. The goal of robust optimization is to identify such regions.

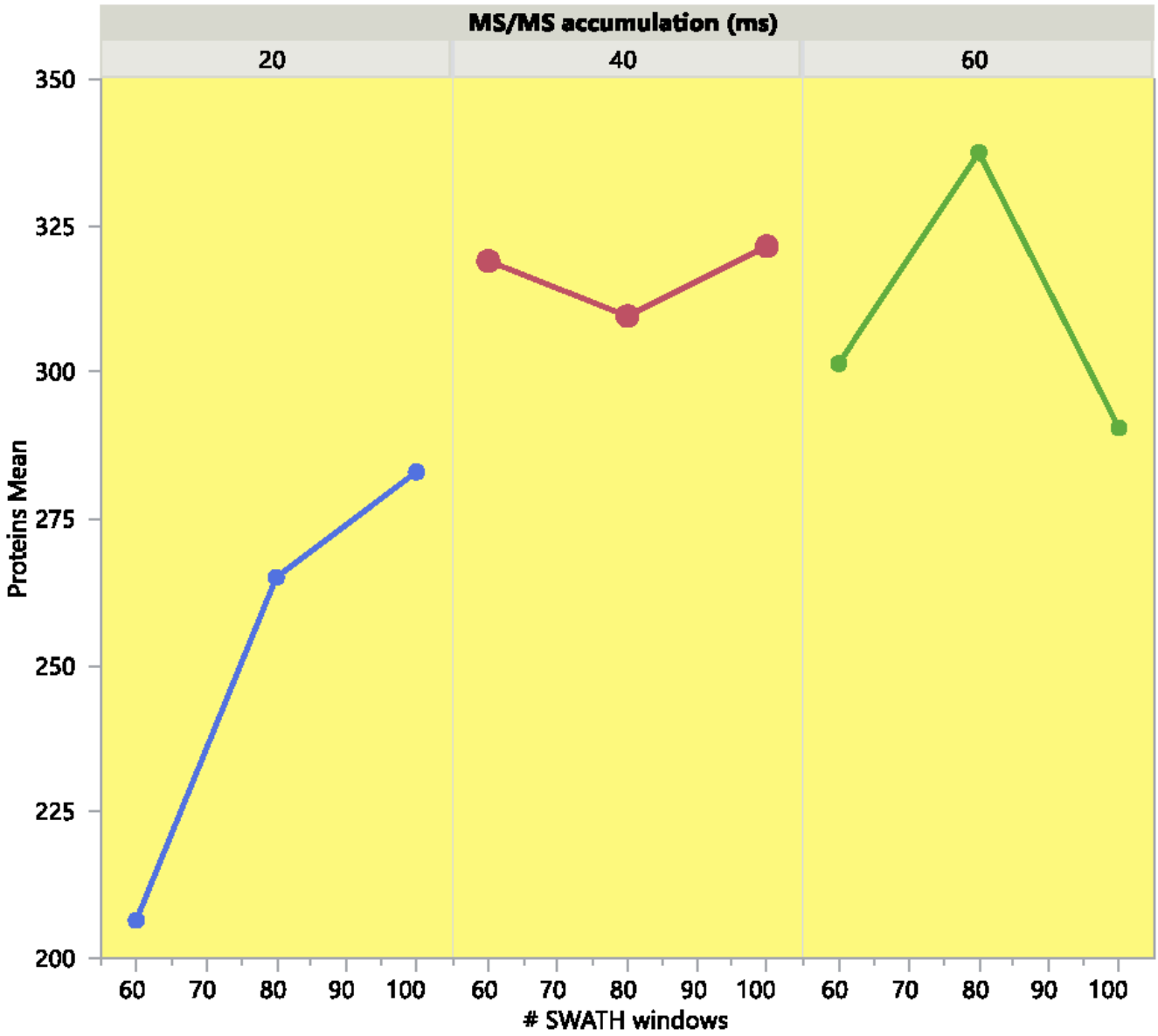

Figure 3

Mean number of proteins recovered as a function of MS/MS accumulation time and number of SWATH windows. More than 300 proteins were recovered at settings of both: 80 SWATH windows and an Accumulation time of 60ms; 100 SWATH windows and an Accumulation time of 40ms. While the number of proteins was maximized at 80 SWATH windows and an Accumulation time of 60ms, the choice of 100 SWATH windows and 40ms for the Accumulation time was more robust - giving consistently over 300 proteins - with a cycle time of 4s permitting satisfactory peak integration.

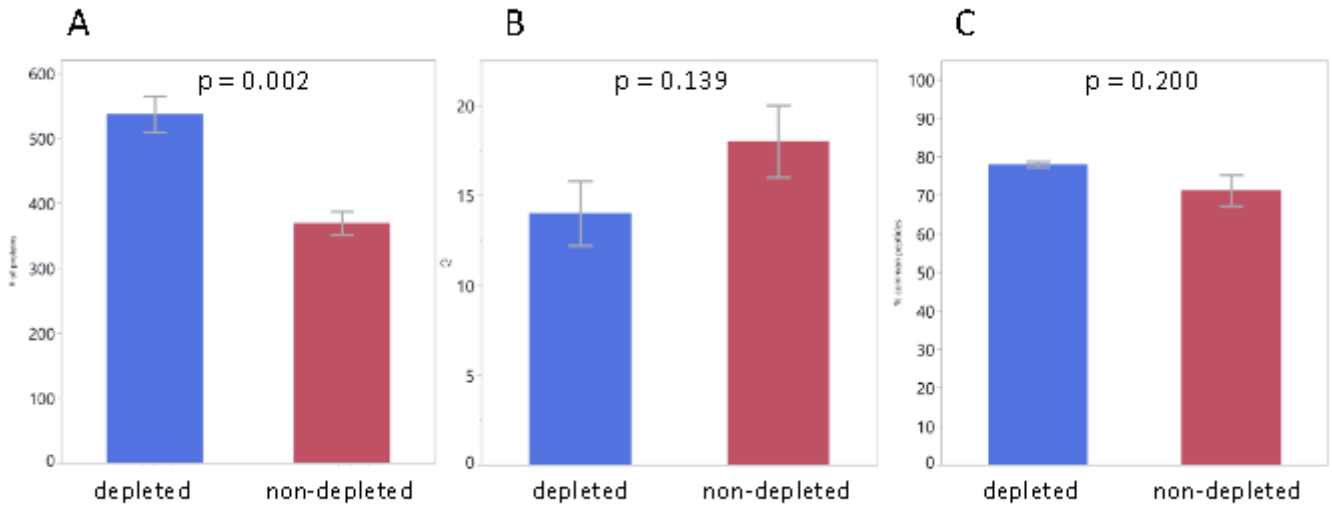

Figure 4

Means and standard errors for depleted (blue) and non-depleted (red) samples analysed using the optimized SWATH-MS method. Depletion has a significant effect on the number of proteins quantified (A) but not on the reproducibility between replicates as measured by the mean CV of peptide intensities (B) or the percentage of common peptides – those recovered in all replicates (C).