Dr. Alexandre Zougman[1], Dr. Stefan Loroch[2,3], Prof. Dr. Darryl J. Pappin[3,4], Dr. John P. Wilson[4]

PROTIFI™
Bringing precision omics to life!™

The Shredder: A New Way to Sequence

[1]University of Leeds, Leeds, England
[2]Medical Proteome-Center, Medica Faculty, Ruhr-University Bochum, Bochum, Germany
[3]ProtiFi LLC, Fairport, NY, United States
[4]Cold Spring Harbor Laboratory, Cold Spring Harbor NY, United States

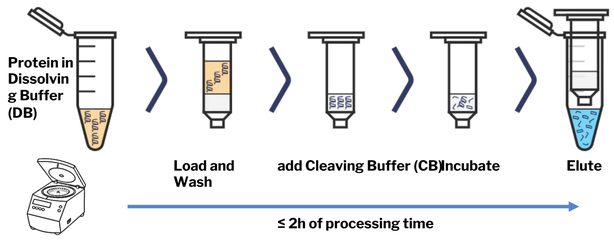
Introduction

Protein sequencing remains at the heart of proteomics, especially in cases of new species, variants, PTMs or FDA-approval of biological drugs. Sequencing has been traditionally been performed via enzymatic and chemical digestion in conjunction with various techniques of gas-phase fragmentation.

Every approaches has limitations: for full sequence coverage, often multiple enzymes and/or chemical fragmentations must be combined due to their cleavage specificities. In particular, a particular sequence might be not cleaved, or might be overly cleaved, in particular regions of interest. For example, the basic tails of histones are typically invisible with tryptic digestions due to the very small resulting peptides. Gas-phase fragmentation is limited not only by the size and length of peptide introduced, but also by fragmentation behavior.

To address these problems, we developed the Shredder, a new approach to bottom-up proteomics sample preparation for in depth protein sequencing. The Shredder randomly cleaves all along the peptide backbone using a combination of low-specificity active sites and activated residues in rapid reaction times.

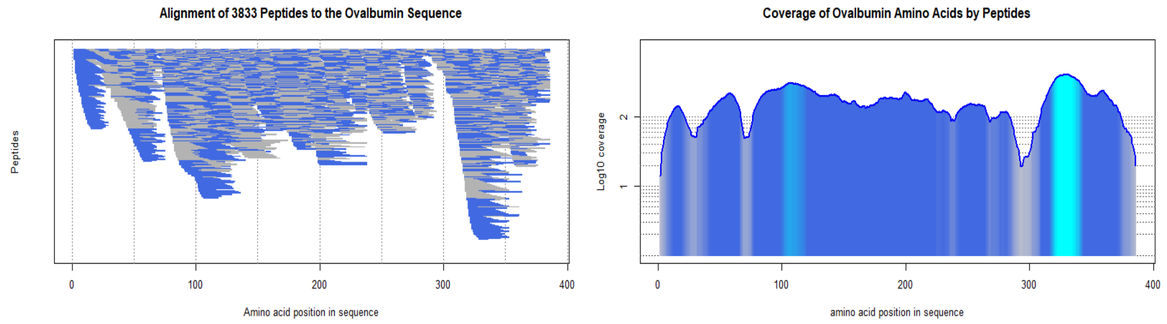
Steps of use



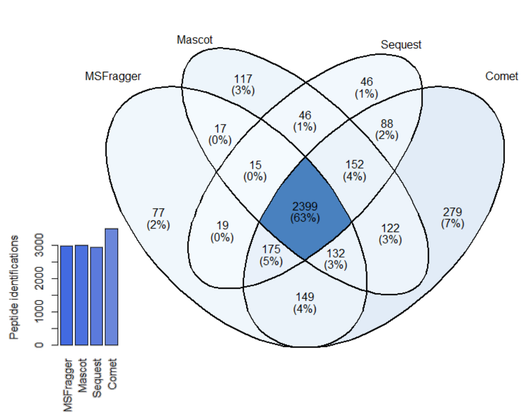
Methods

100 µg of protein (BSA or Chicken Ovalbumin) were digested using the following protocol (specific to the prototype Shredder only). The Shredder unit was placed into a fresh Eppendorf tube and activated with 200 uL of Equilibration Buffer 1 (EB1) then 200 uL of Equilibration Buffer 2 (EB2). Proteins were dissolved in Dissolving Buffer (DB) at concentration of ~0.1 to 1 mg/mL. Samples were loaded onto the cartridge with low-speed centrifugation and washed with 200 uL of Wash Buffer 1 (WB1). The Shredder was moved to a fresh 1.5 mL tube and 200 uL of Cleaving Buffer (CB) was added and brought into the column with low-speed centrifugation. The unit was closed, placed at 90 C for 30 min and the Cleavage Buffer removed via centrifugation. The column was transferred to a fresh tube, washed with 200 uL of Wash Buffer 2 (WB2), then subsequently with 200 uL and 150 uL of Wash Buffer 3 (WB3). 200 uL of Reducing Buffer (RB) was added and loaded into the column with low-speed centrifugation. After 10 min at RT, the RB was removed and 200 uL of Alkylation Buffer (AB) was added and likewise incubated for 10 min at RT; AB was subsequently removed via centrifugation. The column was washed once with 200 uL of Wash Buffer 3 (WB3) and eluted in a fresh tube with 200 uL of Elution Buffer (EB). The resulting peptides were dried down until resuspension and injection to an Exploris480 nanoLC-MS instrument.

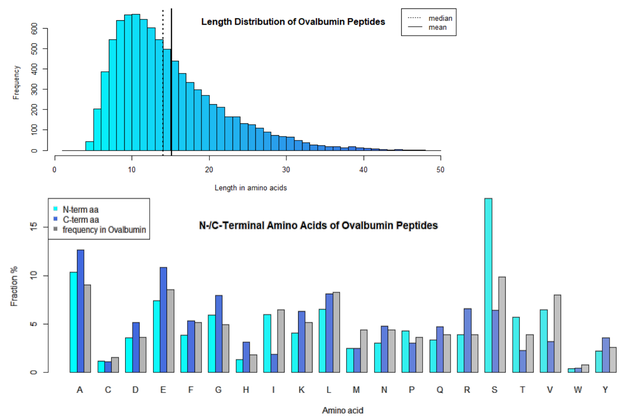
Deep sequencing of chicken ovalbumin



100 µg of chicken ovalbumin was analyzed in 3 LC-MS runs and database search was conducted using 4 different search engines (Comet, Mascot, Sequest, MS Fragger). In total, 3833 Ovalbumin peptide sequences were identified, yielding a remarkable read depth of 167x per amino acid. Single runs yielded between 2300 and 3600 peptides depending on load, instrument and chromatography settings.



Shredder peptides from ovalbumin were identified using Comet, Mascot, Sequest, MS Fragger. All search engines except Comet were run through Proteome Discoverer. Comet was run as an in-house build with similar parameters to those in Proteome Discoverer.

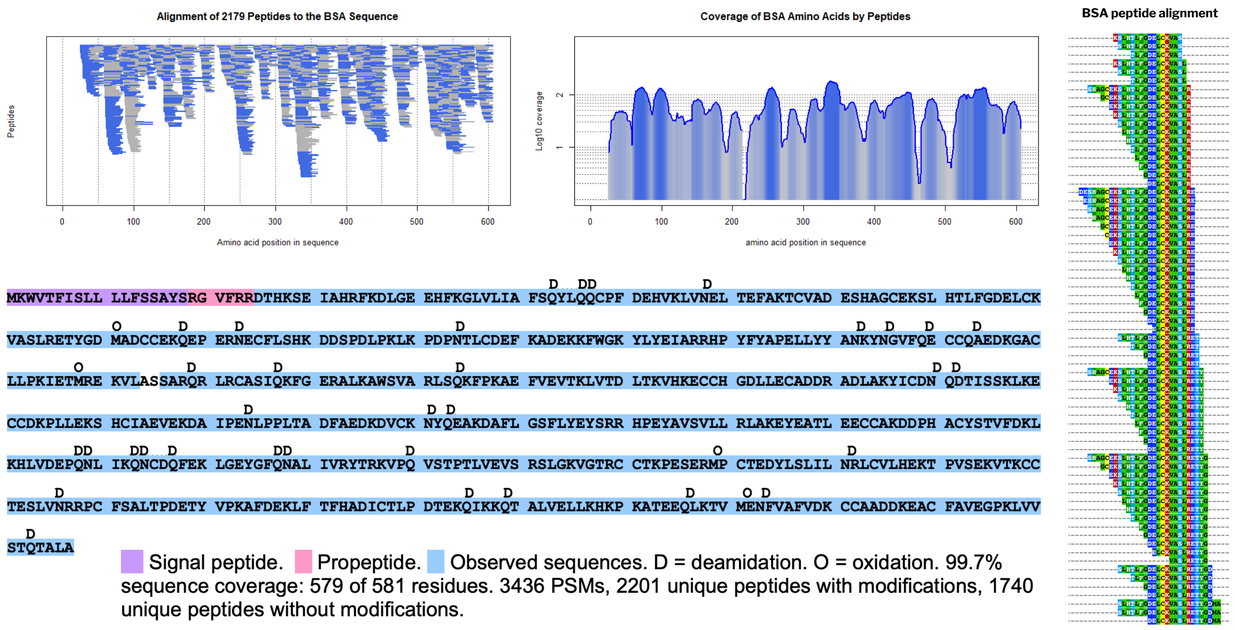


Shredder peptides exhibited a mean/median length of 15/14 amino acids. Cleavage occurs evenly along the protein sequence, reflected by the correlation of the peptides' terminal amino acids occurrence with the amino acid stoichiometry in ovalbumin.

Bovine serum albumin: single shot LC-MS run

The Shredder afforded a 99.7% sequence coverage for Bovine Serum Albumin (BSA). As expected, signal- and propeptide which were, as expected, not detectable). The average read depth ("depth of sequencing") was 62X per amino acid, a concept widespread in genomics but one not yet broadly applied in proteomics. 3436 PSMs were observed containing 2179 unique peptides. Peptide lengths from 7 to 49 amino acids were observed with an average of 16.3 and median of 15. Especially for such a wide range of peptide sizes, the observation of specific peptides is a function of chromatography, mass spec analytical limitations and software search setting; in this case, 7 was the lower limit in the Mascot search settings. Importantly, the resulting peptide with significant overlap holds knowledge of co-occurring PTMs, variants, etc.

This capability is useful in the analysis of biosimilars and other biological drugs, full determination of sequences including isoforms and PTMs, the study of new and even completely unknown proteins, etc.



Shredder analysis with 100 µg of BSA. Peptide alignment displays an extensive laddering allowing for sequence coverage of >99% in a single shot LC-MS run. The read depth was 62X per amino acid on average.

ProtiFi's Shredder

- **Yields a wide range of peptide without need for multiple proteolytic enzymes;**
- **Gives information about the co-occurrence of PTMS, variants, isoforms, etc., especially in longer peptides;**
- **Is designed for relatively refined mixtures of proteins; and**
- **Begins the adaptation of sequence coverage depth, as often reported in genomics, within proteomics. This extended depth of observation brings with it certainty of results, the observation of small changes such as SNPs and detection, by example, of multiple isoforms that have traditionally been hard to study.**
- **Like all techniques, the Shredder has some limitations. First, it is intended to be used on purified or relatively purified proteins or mixtures thereof. Second, PTMs may be liable in the highly acidic cleavage environment. Third, this environment induces deamidation, as expected, which may limit its use in specific settings.**

UNIVERSITY OF LEEDS

RUHR UNIVERSITÄT BOCHUM

PROTIFI™
Bringing precision omics to life!™