IMSC 2022, POSTER #126

USE OF PASEF FOR ACCELERATED PROTEIN SEQUENCE CONFIRMATION AND *DE NOVO* SEQUENCING WITH HIGH DATA QUALITY

[1]Stuart Pengelley, [1]Waltraud Evers, [1]Eckhard Belau, [2]Ilker Sen, [2]Wilfred Tang, [3]Alain Beck, [1]Detlev Suckau

[1]Bruker Daltonics, Bremen, Germany; [2] Protein Metrics Inc., Cupertino, CA; [3] Centre d'Immunologie Pierre-Fabre, Saint-Julien-en-Genevois, France

Introduction

- Biopharmaceutical sequences can be well confirmed by multiple protease digests - e.g., trypsin, elastase and chymotrypsin – followed by LC-MS/MS data analysis.
- High quality data can be used for *de novo* sequencing as well. PASEF (Parallel Accumulation and Serial Fragmentation) on the timsTOF instrument has been used to accelerate proteome studies and increase sequence coverage concomitantly.
- Here we applied PASEF to generate exhaustive protein sequence coverage maps by combination of results from 2 or 3 enzyme digests using a short LC gradient. The data quality obtained was high and adequate for determining antibody sequences *de novo*.
- Nivolumab and Dulaglutide sequences were confirmed and Nivolumab sequenced *de novo*.

Methods

- Samples Nivolumab (IgG4κ) and Dulaglutide (IgG4 based Fc-fusion peptide) were used as European union available drug products. They were reduced and carbamidomethylated prior to proteolytic digest by trypsin, elastase or chymotrypsin.
- Data acquisition Analytical LC – Acquity 2.1x150 mm C_{18} column (Waters), 45 min gradient, 60 min cycle time; paired with timsTOF Pro (Bruker) using PASEF.
- Data analysis LC-MS/MS datasets were processed directly either in BioPharma Compass 2021 (Bruker) to match spectra with the reference sequence (5 ppm MS tolerance, no enzyme) for confirmation of provided sequences or to Supernovo (ProteinMetrics) to establish sequences *de novo*.

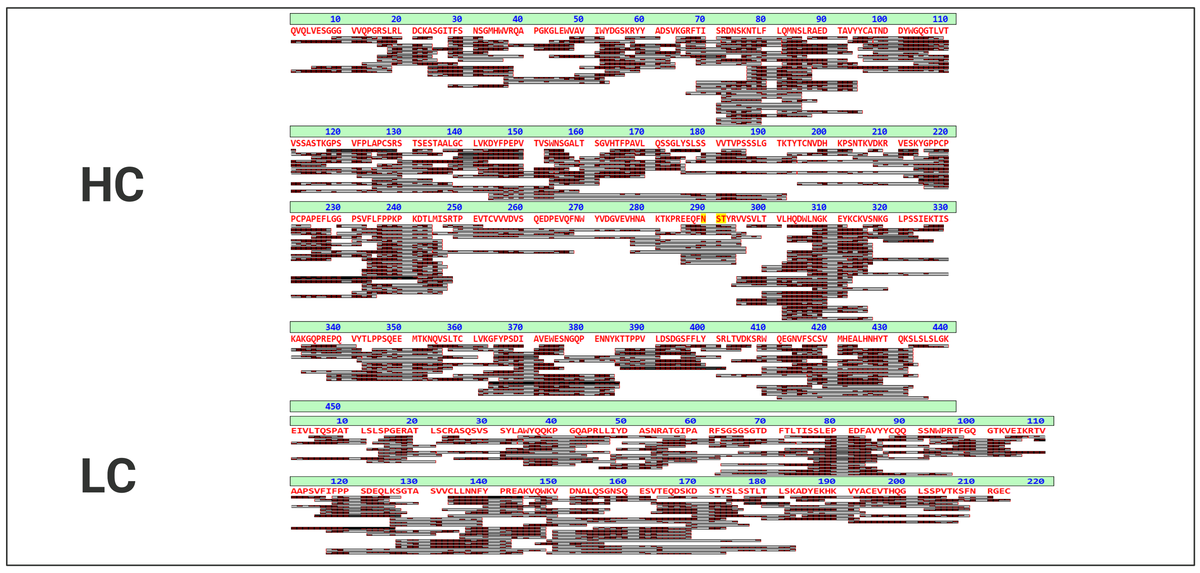


Fig. 3 Sequence coverage maps of the Nivolumab HC (top) and LC (bottom) generated by combination of the results from the individual 3 digests. Red sequence text indicates the presence of identified peptides covering the particular sequence stretch, black: not covered by any peptide - this is not observed here. Grey bars: matching peptides contributing to cover the sequence. Red bricks: matching fragment ions; upper row of bricks: matching N-terminal b-ion fragments, lower row of bricks: matching C-terminal y-ion fragments.

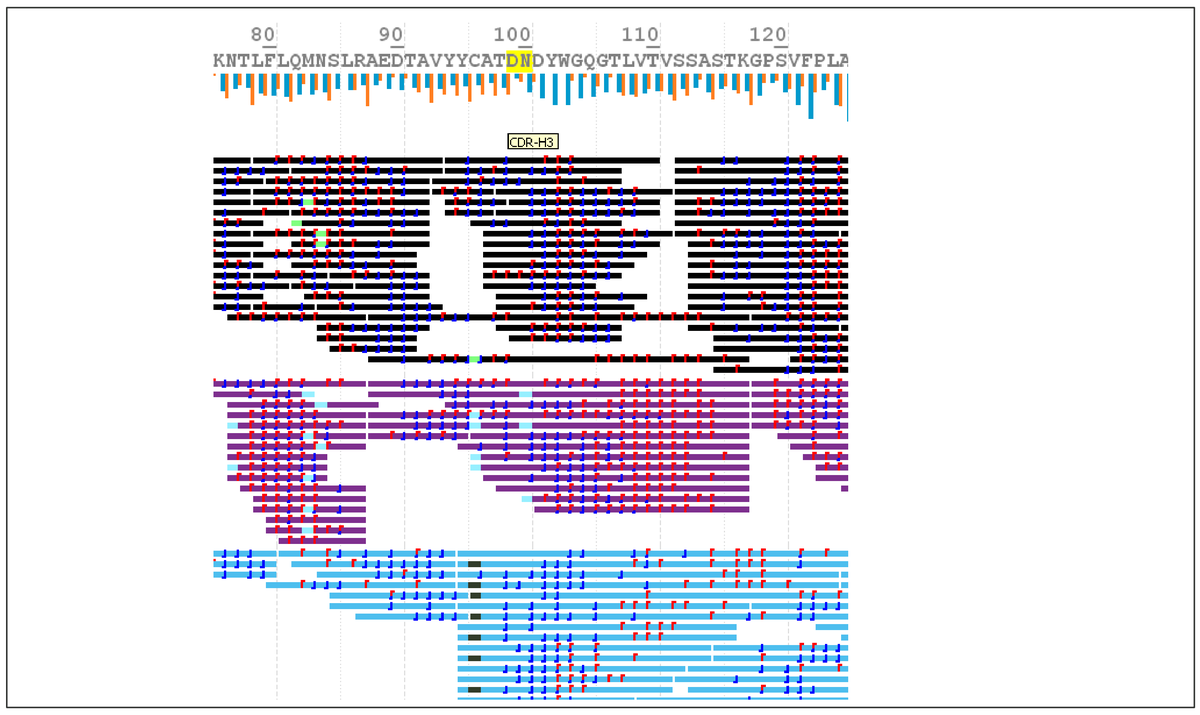


Fig. 4 Sample Supernovo output. Three samples' peptides of Nivolumab are shown. Yellow highlight on the protein sequence indicates low confidence in the determined sequence, and vertical blue bars below the sequence show aggregated fragment coverage in these residues.

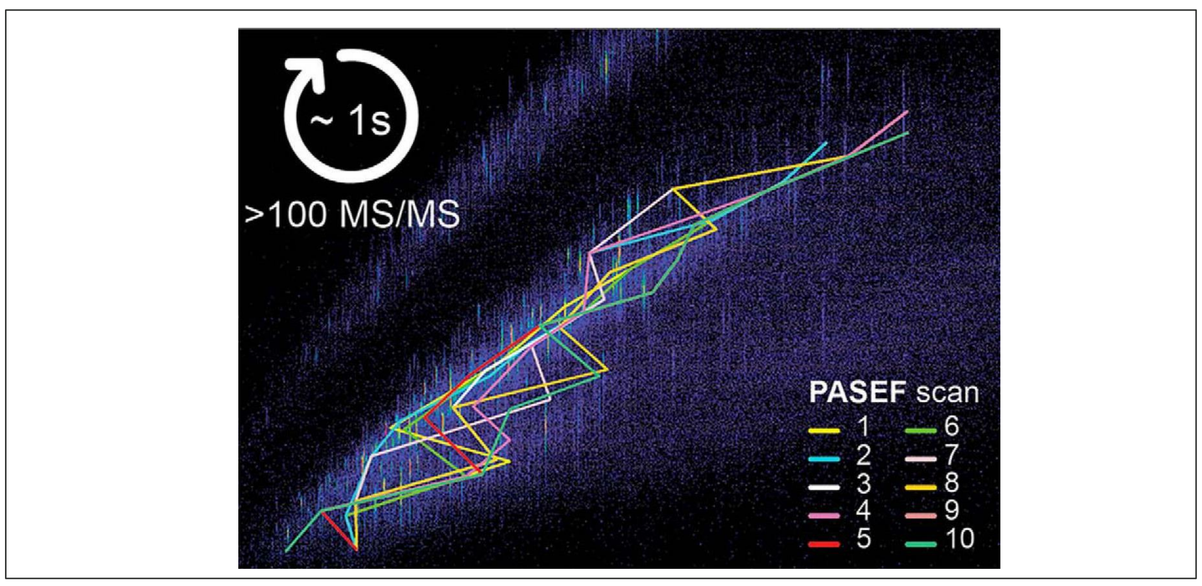


Fig. 1 PASEF spectra were obtained with a standard 1.1 sec PASEF acquisition cycle. In this method, Parallel Accumulation and SErial Fragmentation results in increased sensitivity due to ion mobility focusing of the ions, which are sequentially fragmented at > 100 Hz. This increased sensitivity and analysis depth results in more peptides, fragment ions and a high sequence coverage even in antibody *de novo* sequencing analyses.

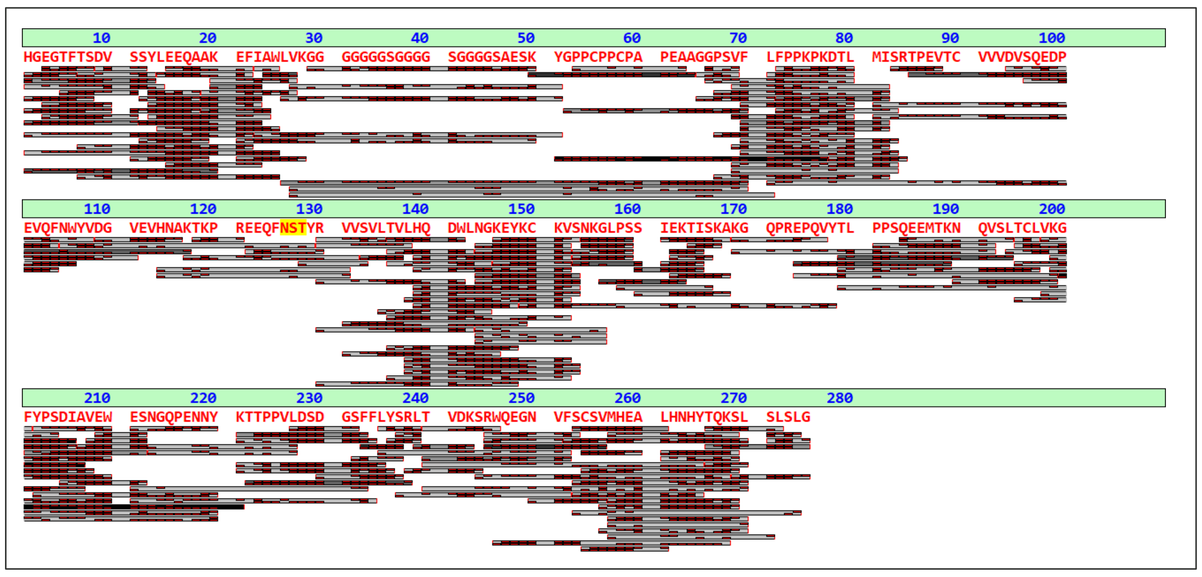
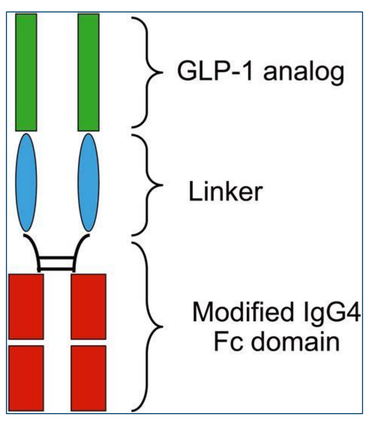


Fig. 2 Sequence coverage map of Dulaglutide generated by combination of the results from the individual 3 digests. For dulaglutide 96/100/90% sequence coverage (SC) were obtained and 92/90/83 % fragment ion coverage (FC); combined, 100 % SC and 100 % FC were obtained.
The merged peptide map in BioPharma Compass derived from the 3 digests resulted in 221 peptides; enough to confidently confirm the full Dulaglutide sequence.

The Scheme of the Dulaglutide structure is shown on the right.



Results

Sequence confirmation of Nivolumab

- For nivolumab 94/94/90% sequence coverage (SC) and 86/84/85% fragment coverage (FC) were obtained from the individual digest analyses with Trypsin/ Chymotrypsin/Elastase, respectively.
- The combined data (*Fig. 3*) resulted in 305 HC and 148 LC peptides and 100% SC and 99% FC confidently confirming the full sequence.

De novo sequencing of Nivolumab

- The combined peptide maps were analyzed using Supernovo yielding a sequence with 99.5% identity to the Drugbank sequence of Nivolumab. The differences were localized at HC residue 27, where the Leucine was determined instead of Isoleucine, and at 99-100 where the determined motif was DN rather than ND.
- The metrics/visualizations, which include the confidence of each deduced residue, an aggregate fragmentation map, differences from germline, and a peptide inspection dashboard, helped to spot and curate the remaining deviations (*Fig.4*).

Summary

- *De novo* analysis and sequence conformation completed with 3 enzyme digests and 60 min LC-MS cycle time on timsTOF Pro 2
- Nivolumab and the fusion protein Dulaglutide sequences were fully confirmed using PASEF
- Nivolumab sequence was fully established *de novo* including the 6 CDR sequences
- PASEF shortened the typical time for sequence analysis of biopharmaceuticals significantly and improved data quality concomitantly
- The novel approach reduces costs for full biologics sequence analysis and makes such analysis more readily available

Biologics