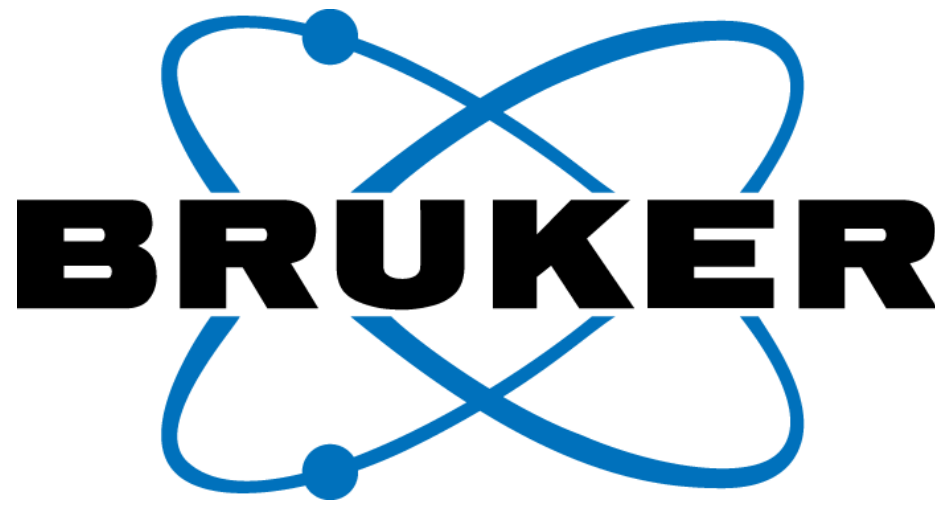


In-Depth Protein Sequence Verification by TIMS-enabled MALDI Top-Down Sequencing



Detlev Suckau¹, George Alevizos², Georgia Orfanoudaki², Guillaume Tremintin³, Mariangela Kosmopoulou², Eckhard Belau¹, Waltraud Evers¹, Arndt Asperger¹

¹Bruker Daltonics GmbH & Co. KG, Bremen, Germany

²Fasmatech Science & Technology, Chalandri, Greece

³Bruker, San Jose, CA

Introduction

- TIMS-enabled MALDI Top-Down Sequencing (MALDI-TDS) using in-source decay (ISD) provides a powerful method for protein *de novo* sequencing and sequence confirmation.
- TIMS-enabled MALDI-TDS and Bruker's Top-Down sequence analysis software OmniScape™ simplify data analysis and improve sequencing accuracy and proteoform assignment.
- Our approach demonstrates the powerful combination of MALDI-MSD and TIMS for *de novo* sequencing and advanced protein characterization, providing critical insights into protein structure and function.

Methods

Bovine carbonic anhydrase II and NISTmAb 8671 (Merck) were prepared and analyzed using MALDI-MSD, MALDI-TIMS-MS, and T³-Sequencing (1).

NISTmAb was IgGZERO and IdeS digested (both: Genovis) and reduced to obtain the Fd, Fc/2 and LC subunits, which were LC-separated and dried. MALDI samples were prepared in sDHB.

MALDI-MSD TIMS data were acquired on a timsTOF flex (Bruker), using red phosphorus as mass calibrant. Data preprocessing involved TIMS extraction, smoothing, baseline correction using DA 6.1 (Bruker). TIMS separated 1+ and 2+ charged ion series were analyzed individually.

Bruker's OmniScape software was used for sequence confirmation and *de novo* sequence analysis followed by homology searches using online MS-BLAST (2) or prototypic local NovorTag (Rapid Novor Inc., Canada) searches in a SWISSPROT database with added NISTmAb sequence. The overall results were compiled across multiple analyses using OmniScape's Result Combination functionality.

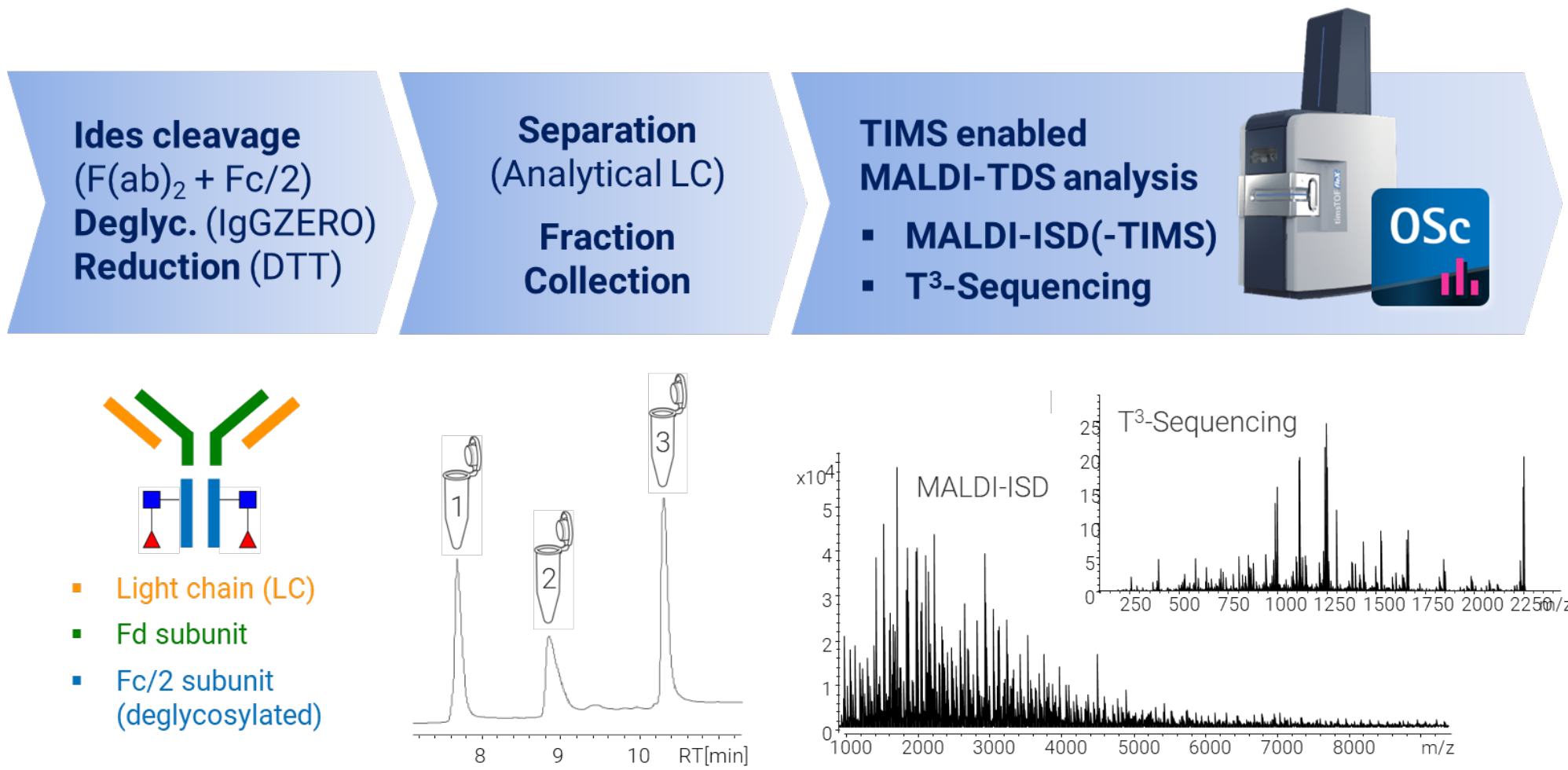


Fig 1: Offline LC-MALDI-MSD analysis workflow of NISTmAb subunits

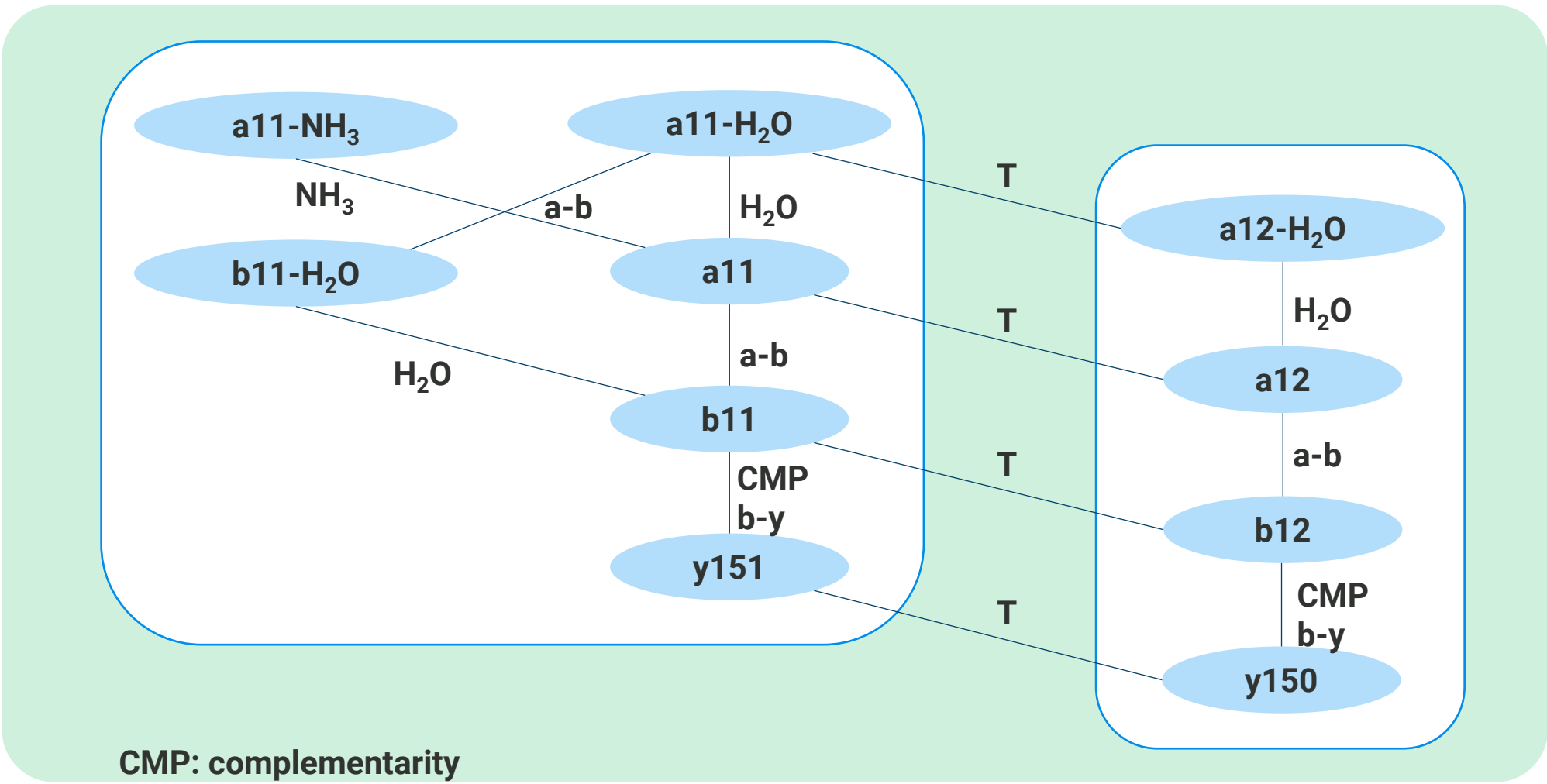


Fig 2: De novo sequencing algorithm in OmniScape is based on a graph with defined relationships

Results

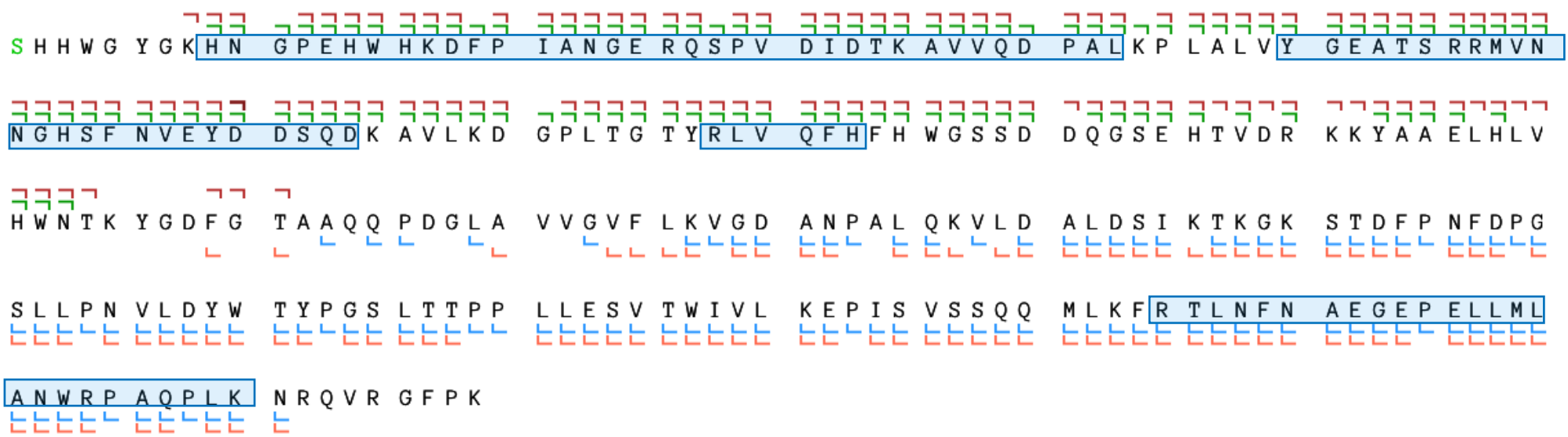


Fig 3: De novo sequencing result based on sequence tags and MS-BLAST search. The matching tags are overlaid on the confirmation result for bovine CAII.

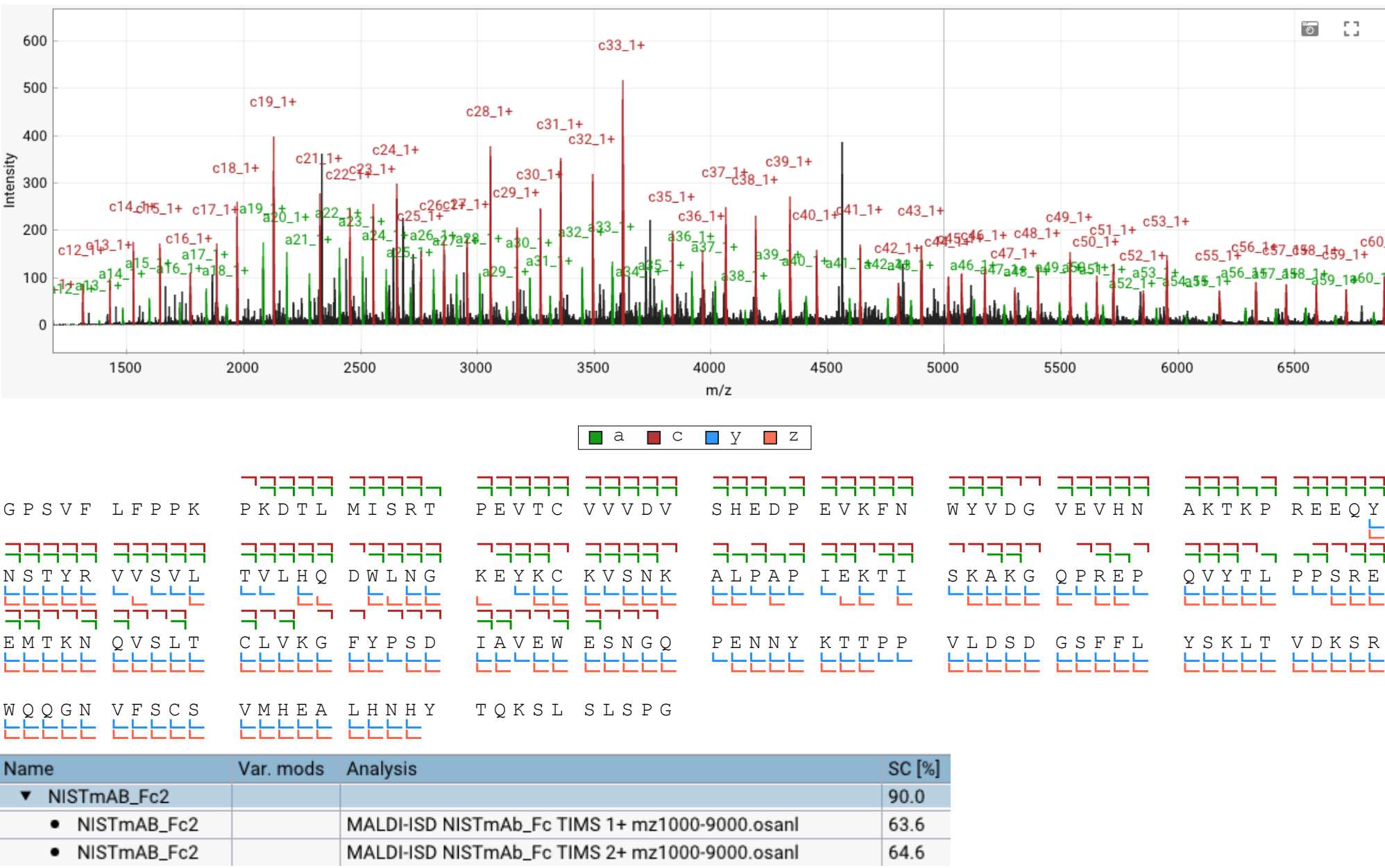


Fig 4: Top: MALDI-MSD spectrum of NISTmAb Fc/2. Bottom: Combined sequence map from TIMS-MSD of 1+ and 2+; fragments cover the entire sequence – except the termini.

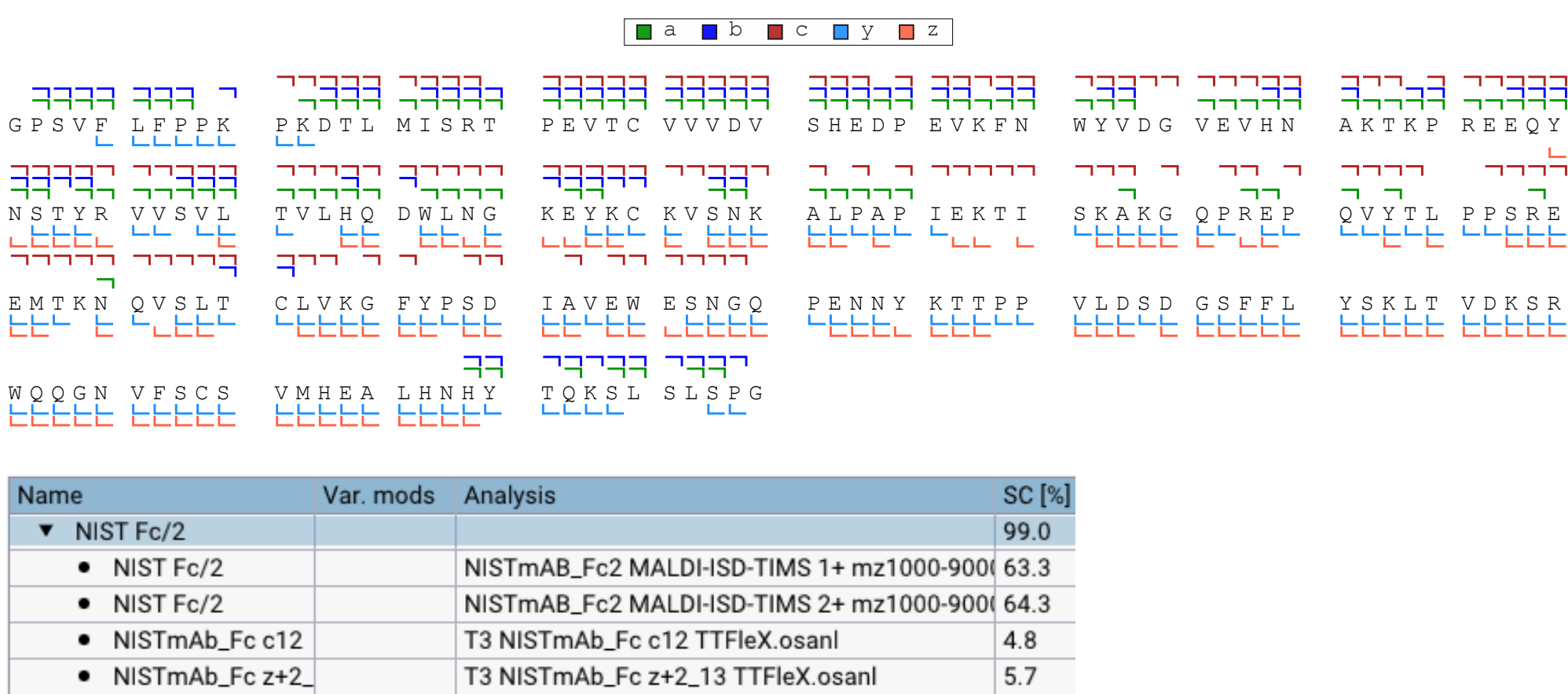


Fig 5: Combined MALDI-MSD-TIMS 1+/2+ sequence map of NIST-Fc/2 (99% SC), including terminal fragments from T³-Sequencing.

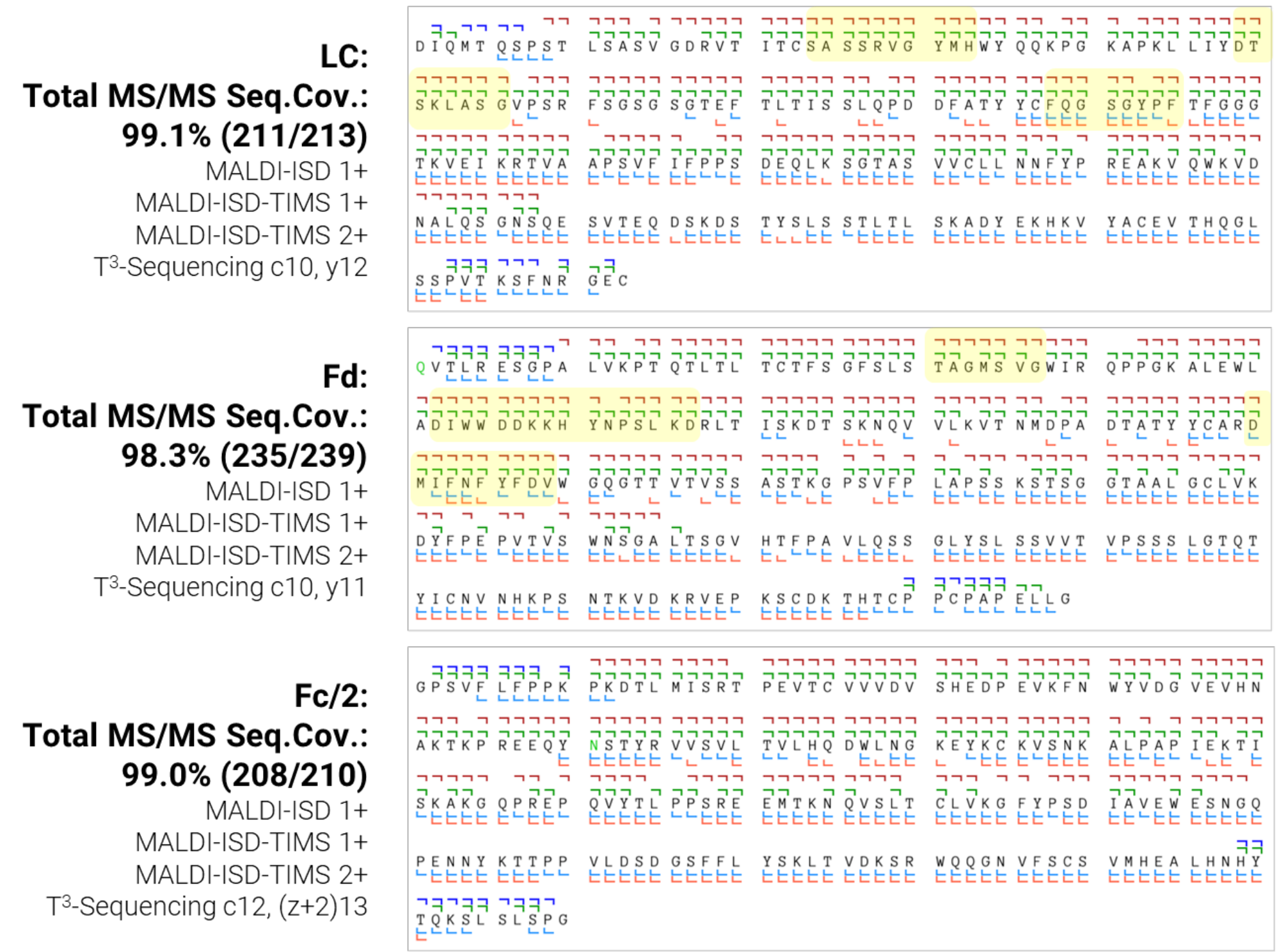


Fig 6: NISTmAb Fc/2: Matching *de novo* sequencing tags in the MALDI-MSD-TIMS 1+ spectrum.

The Fc/2 sequence was largely covered, suggesting *de novo* sequencing was a promising approach to identify the NIST Fc in a local database of the SWISSPROT with the added NISTmAb subunit sequences.

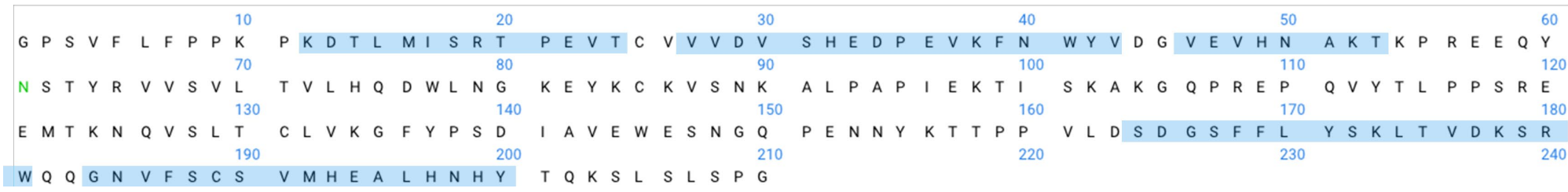


Fig 7: NISTmAb Fd: Matching *de novo* sequencing tags in MALDI-MSD-TIMS 1+ (blue) and 2+ (pink) spectra. CDRs 1-3 (orange).

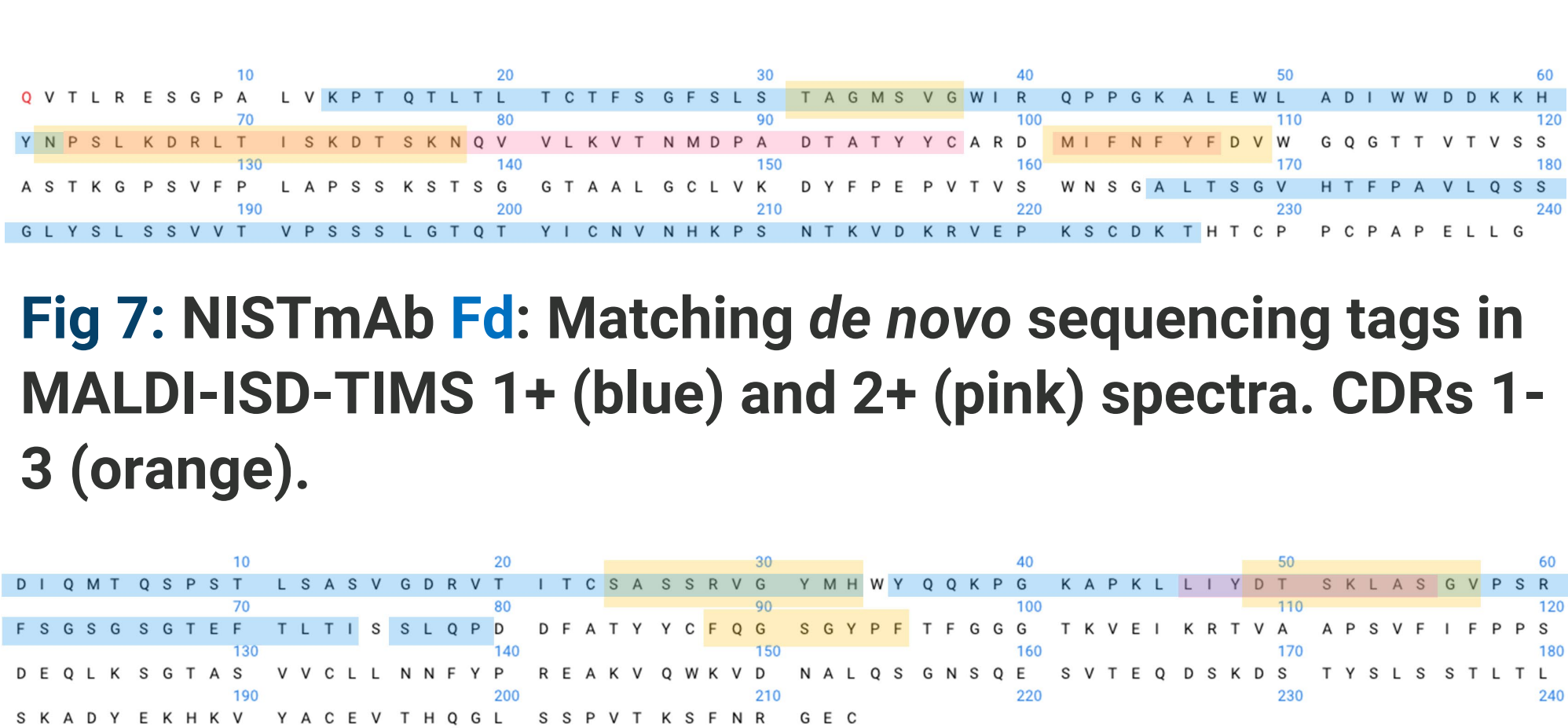


Fig 8: NISTmAb LC: Matching *de novo* sequencing tags in MALDI-MSD-TIMS 1+ (blue) and 2+ (pink) spectra. CDRs 1-2 were covered, CDR3 was not (orange).

While the *Fc/2* sequence tag-based homology searches in SWISSPROT were quite successful due the conserved sequence, the identification for the variable domain sequences of NISTmAb LC and Fd provided much better results in local searches of a NISTmAb+SWISSPROT database (Figs. 7 and 8).

Fd CDRs 1 and 2 were fully covered and CDR3 by 80%. Particularly the CDR3 *de novo* sequencing result depended fully on the 2+ charge state spectrum analysis of the TIMS isolated fragment ion series.

LC CDRs 1 and 2 were fully covered via 1+ fragment-derived sequence tags, whereas CDR3 was not *de novo* sequenced successfully and additional experiments such as MS4 on a timsOMNI instrument (WOA pm 04:10, MP 713, MP 714, TP 023, WP 582) or else could be added.

Summary

TIMS provided great benefit to sequence confirmation and *de novo* sequence analysis by yielding „charge-purified“ Top-Down MS/MS spectra as a result of mobility separation of fragment charge states 1+ and 2+.

Near 100% sequence coverage (SC) was obtained in the confirmation analysis of NISTmAb subunits (Fc/2, LC and Fd) by combining TIMS separated MALDI-MSD 1+ and 2+ ion series and pseudo-MS³ fragmentation of near-terminal c, y and z+2 fragment ions (T³-Sequencing).

TIMS separation of charge states 1+ and 2+, for the first time, enabled sequence readout based on low-abundant 2+ MALDI-MSD fragments. TIMS separated 2+ fragments provided added sequence readout for the core regions of the subunits, which had been covered incompletely when utilizing 1+ MALDI-MSD fragments solely.

For *de novo* sequencing, only the TIMS-separated 1+ and 2+ ions were used from a single MALDI-MSD-TIMS-MS/MS spectrum per subunit. For the Fd, CDR1 and CDR2 were fully sequenced and CDR3 by ~80%, For the LC, CDR1 and CDR2 were fully sequenced.

Offline LC-MALDI-MSD has the added advantage that separation and acquisition are decoupled. Thus, the acquisition time is not restricted by the chromatographic peak width and leads to high quality spectra with high resolution, mass accuracy and isotopic pattern fidelity.

References

1. D. Suckau, A. Resemann (2003), T³-Sequencing: targeted characterization of the N- and C-termini of undigested proteins by mass spectrometry, Anal Chem 75(21).
2. Shevchenko A, et al. (2001), Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching, Anal Chem 73(9).

Acknowledgement

Qixin Liu, Mingjie Xie, Bin Ma
Rapid Novor, Kitchener, Ontario

Conclusion

- ~99% Sequence Coverage for all NISTmAb subunits were obtained from 1+, 2+ TIMS separated MALDI-MSD ion series and T³-Sequencing.
- *De novo* sequence analysis and homology searches in a database containing target sequences allowed to ID large part of the CDR 1-3 sequences.
- Ability to charge-separate 1+ and 2+ ion series by TIMS enhanced the sequence coverage in a single dataset significantly.

Top-Down Sequencing