

Efficient Protein Characterization with timsOmni and ProSight Native

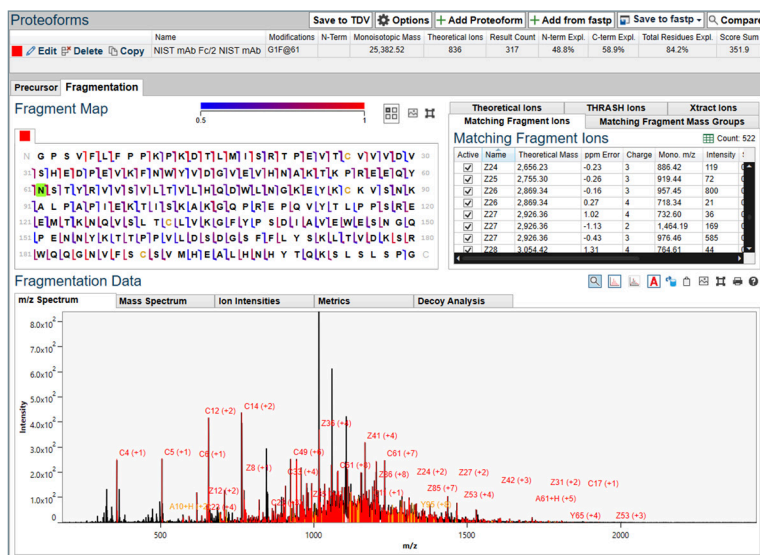
In the rapidly evolving field of top-down proteomics and biological drug development, the ability to accurately and efficiently analyze proteins is paramount. ProSight Native offers a comprehensive suite of workflows designed to process proteoform data acquired with the timsOmni™ system, ensuring precise and reliable results.

A complex data interpretation challenge

Researchers in top-down proteomics face significant challenges, including the need for comprehensive characterization of protein isoforms, post-translational modifications, and complex protein structures. Additionally, top-down spectra are more congested and overall more complex than those of smaller molecules. These challenges often result in time-consuming data interpretation and potential inaccuracies.

Faster top-down data annotation

ProSight Native addresses these challenges by offering increased throughput with a range of workflows to analyze LC-MS or infusion data. The data processing time is significantly reduced by removing tedious steps. The timsOmni open data format is read natively and user-friendly tools simplify the task of extracting and deconvoluting the spectra of interest either manually or with an automated routine.



Top-down spectra from the timsOmni, which offers multiple fragmentation modes such as C_{CID} , R_{CID} , ECD or EID, can be efficiently annotated by ProSight Native. This software supports the flexible definition of fragments and target proteins (UniProt XML databases, FASTP proteoform definitions, or FASTA). Finally, the TDValidator tool saves users time when reviewing results by compiling sequence map details, quality metrics and decoy validation tools in a single window.

Display of sequence map, matched fragments, and annotated spectrum provides a comprehensive overview of results.

Data acquisition

A mixture of reduced NISTmAb subunits was directly infused in the timsOmni with the NEOS source. The light chain (23+), Fd' (22+), Fc2-G0F (25+) and Fc/2-G1F (25+) were isolated and fragmented with Trapped eXd (50 ms reaction time, 0 eV electrons) in the Omnitrap Q5 segment.

Data import

Data files were directly imported into ProSight Native using the targeted top-down workflow to produce an average spectrum for each MS² file.

Spectrum annotation

Proteoform sequences for NISTmAb LC, Fd', and Fc/2 with N61 glycosylation were specified as FASTP entries such that the exact proteoform composition could be used to generate theoretical isotope distributions for every possible fragment ion.

Results validation

All selected subunits were confirmed with >75% sequence coverage, including the region of the Fc/2 modified with large glycans (1444.534 Da and 1607.474 Da respectively) which were covered by many fragments, including a 21-residue sequence tag for the G1F glycosylation (Figure 1). Additionally, the Fc/2-G1F had 522 fragments assigned while the same sequence length shuffled and randomized yielded an average of only 49 and 7 matches, respectively (Figure 2).

Conclusion

This experiment enabled the rapid optimization of the acquisition time required to comprehensively characterize these subunits (Figure 3). Here most of the sequence coverage could be achieved in under 100 scans, which represents ~1 minute of data collection.

N G P S V[F][L][F P P][K][P][K][D][T][L][M][I][S][R][T P][E][V][T][C V][V][D][V 30
31 [S][H][E][D][P][E][V][K][F][N][W][Y][V][D][G][V][E][V][H][N][A][K][T][K P][R][E][E][Q][Y 60
61 [N][S][T][Y][R][V][V][S][V][L][T][V][L][H][Q][D][W][L][N][G][K][E][Y][K][C K V][S][N][K 90
91 [A L P][A][P][I][E][K][T][I][S][K][A][K][G][Q P][R][E P][Q V][Y][T L P P][S][R][E 120
121 [E][M][T][K][N][Q][V][S][L T][C][L][V][K][G][F][Y P S][D][I A][V][E][W][E][S][N G][Q 150
151 [P E][N][Y][K][T][T][P][V][L][D][S][D][G][S F][L Y S][K][L][T][V][D][K][S][R 180
181 [W][Q][G][Q][N][V][F][S C][S][V M][H][E][A][L][H][N][H Y T][Q][K][S L S L S P][G C

Figure 1

Sequence map for Fc/2-G1F with a long sequence tag (highlighted in light blue) confirming the presence of the N61 G1F glycan.

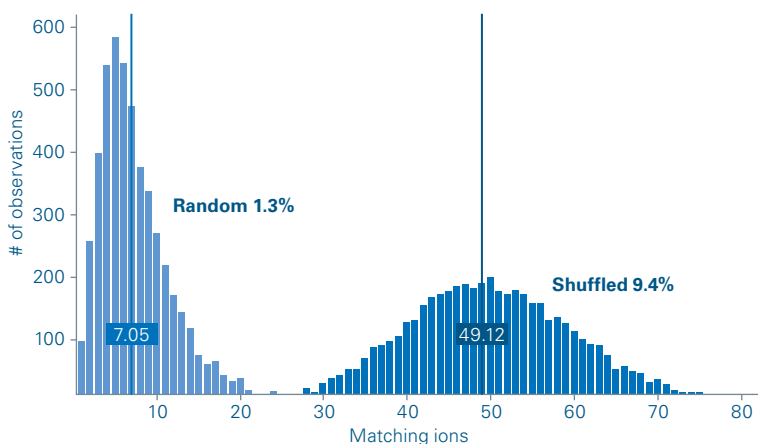


Figure 2

Analysis of shuffled and random decoy sequences of the same length as Fc/2-G1F show low numbers of matching ions compared to the identified proteoform.

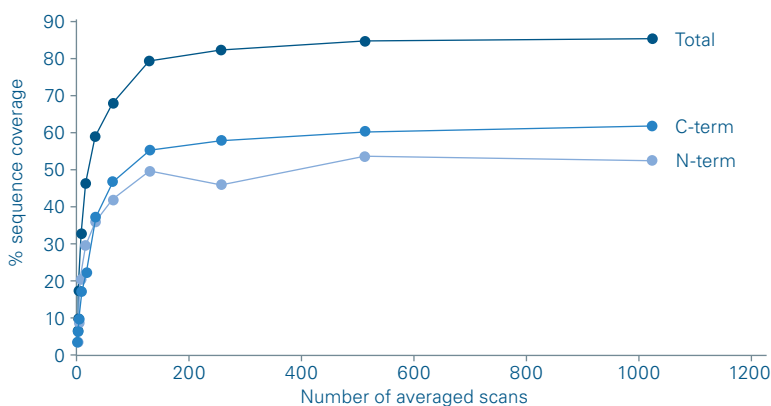


Figure 3

Sequence coverage for Fc/2-G1F as a function of number of averaged spectra.

For Research Use Only. Not for use in clinical diagnostic procedures.

Bruker Switzerland AG

Fällanden · Switzerland
Phone +41 44 825 91 11

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

