Sample preparation
FT-ICR MS sample analysis

Top down of proteins
Data interpretation

Petr Novák, **Zdeněk Kukačka**, Petr Man

IMIC

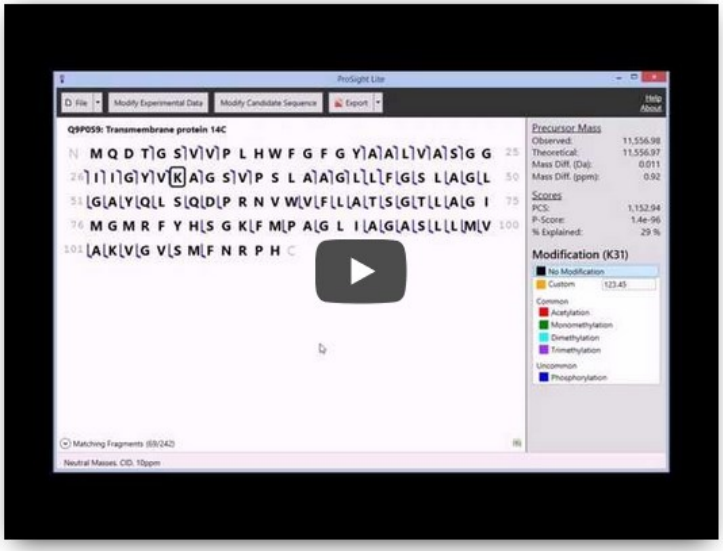
ProSight Lite – freely available

http://prosightlite.northwestern.edu/

- Groups of Neil L. Kelleher and Paul M. Thoms
- Windows operating systems

PROTEOMICS CENTER OF EXCELLENCE
NORTHWESTERN UNIVERSITY

work with us / our team / technologies / publications



ProSight Lite

ProSight Lite is a free Windows application for matching a single candidate protein sequence and its modifications against a set of mass spectrometric observations.

Download

View changelog

End User License Agreement

ProSight Lite – Manual and software description

ProSight Lite: Graphical Software to Analyze Top-Down Mass Spectrometry Data

Ryan T. Fellers, Joseph B. Greer, Bryan P. Early, Xiang Yu, Richard D. LeDuc, Neil L. Kelleher*, and Paul M. Thomas*
Departments of Chemistry, Molecular Biosciences and the Proteomics Center of Excellence Northwestern University, 2145 N. Sheridan Road, Evanston, IL 60208

Abstract

Many top-down proteomics experiments focus on identifying and localizing post-translational modifications and other potential sources of "mass shift" on a known protein sequence. A simple application to match ion masses and facilitate the iterative hypothesis testing of PTM presence and location would assist with the data analysis in these experiments. ProSight Lite is a free software tool for matching a single candidate sequence against a set of mass spectrometric observations. Fixed or variable modifications, including both post-translational modifications and a select number of glycosylations, can be applied to the amino acid sequence. The application reports multiple scores and a matching fragment list. Fragmentation maps can be exported for publication in either PNG or SVG format. ProSight Lite can be freely downloaded from http://prosightlite.northwestern.edu, installs and updates from the web, and requires Windows 7 or higher.

Keywords

Top-down proteomics; proteomics software; MS Analysis; proteoform characterization

Top-down proteomics describes the study of intact proteins with mass spectrometry [1, 2]. Traditional bottom-up proteomics experiments are marked by the use of an enzyme, typically trypsin, to proteolyze intact proteins into more analytically manageable peptides (0.5-3 kDa)[3]. This proteolysis effects a loss of information between the ribosomally-expressed pro-protein, which in eukaryotes often contains RNA splice variants, and the post-translationally modified intact protein (termed a "proteoform")[4]. If a modification has been identified on two separate peptides, a typical bottom-up proteomics experiment cannot know whether those modification existed singly on two separate proteoforms or in tandem on a single proteoform. While more analytically challenging, top-down proteomics provides that full information[5].

Bioinformatics Analysis of Top-Down Mass Spectrometry Data with ProSight Lite

Caroline J. DeHart, Ryan T. Fellers, Luca Fornelli, Neil L. Kelleher, and Paul M. Thomas

Abstract

Traditional bottom-up mass spectrometry-based proteomics relies on the use of an enzyme, often trypsin, to generate small peptides (typically < 25 amino acids long). In top-down proteomics, proteins remain intact and are directly measured within the mass spectrometer. This technique, while inherently simpler than bottom-up proteomics, generates data which must be processed and analyzed using software tools "purpose-built" for the job. In this chapter, we will show the analysis of intact protein spectra through deconvolution, deisotoping, and searching with ProSight Lite, a free, vendor-agnostic tool for the analysis of top-down mass spectrometry data. We will illustrate with two examples of intact protein fragmentation spectra and discuss the iterative use of the software to characterize proteoforms and discover the sites of post-translational modifications.

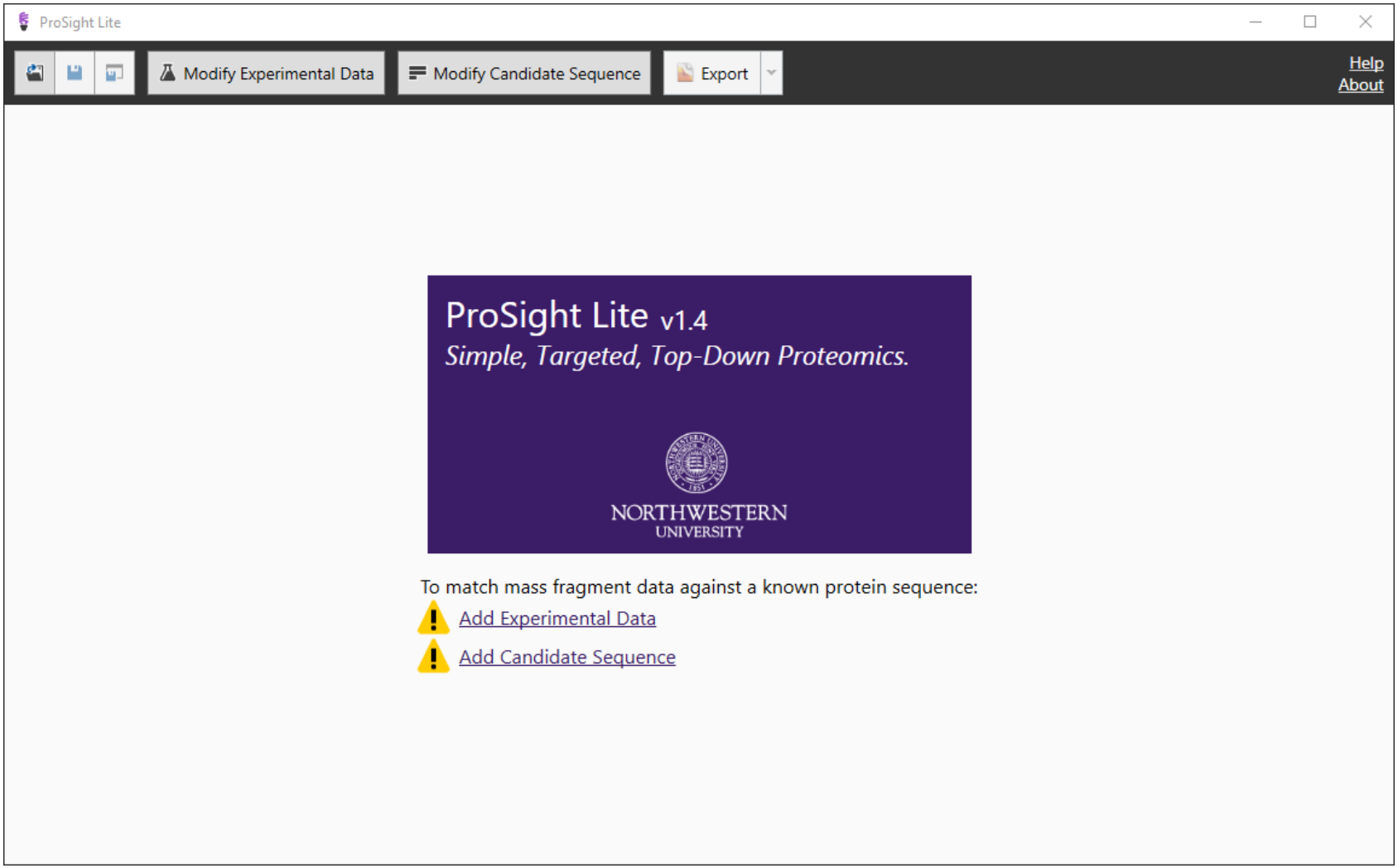
Key words Top-down, Mass spectrometry, Proteomics, ProSight Lite, Intact protein, Bioinformatics

1 Introduction

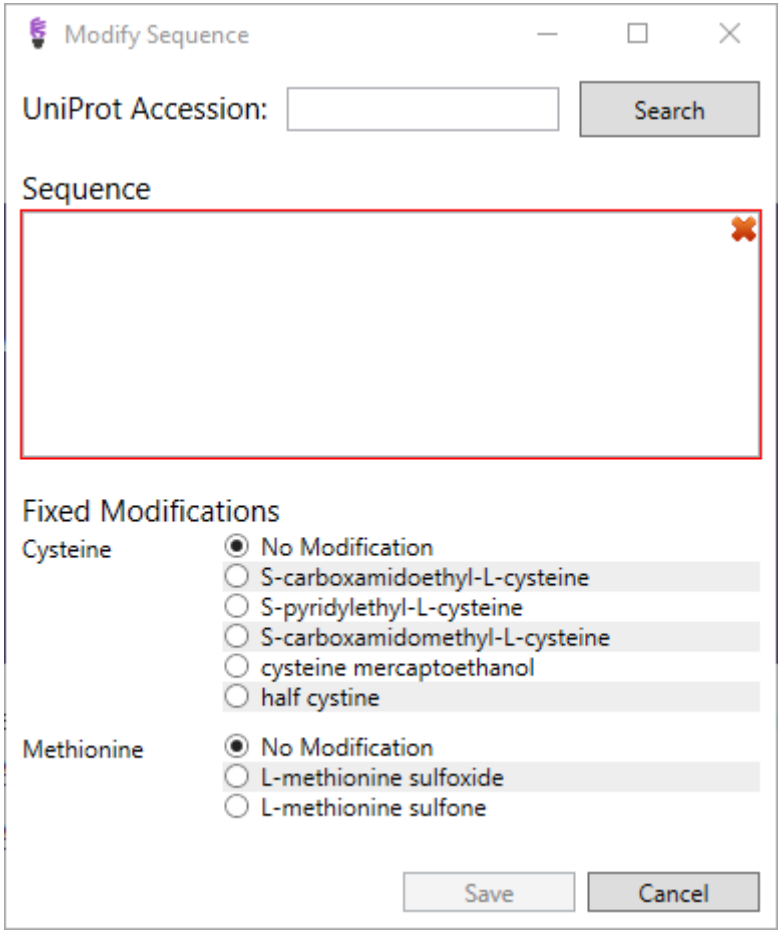
Complementing the speed and sensitivity of bottom-up proteomics [1, 2], top-down proteomics [3] offers a comprehensive view of proteoforms present in the sample (for a more detailed description of a proteoform, *see* **Note 1**) [4]. It is important to distinguish between top-down mass spectrometry and top-down proteomics. In top-down mass spectrometry, one or a few proteoforms are isolated and studied. This technique is typified by studies of histones [5–7] and immunoglobulins [8–10], among others, including the routine analysis of small (< 40 kDa) proteins. In top-down proteomics, a whole, unknown proteome is typically isolated, fractionated, and analyzed by mass spectrometry [11–13]. The tools used for the study of top-down mass spectrometry and top-down proteomics are related, but fundamentally different. Top-down proteomics requires searching tandem MS data against a database of known proteins and proteoforms [14, 15], while top-down mass

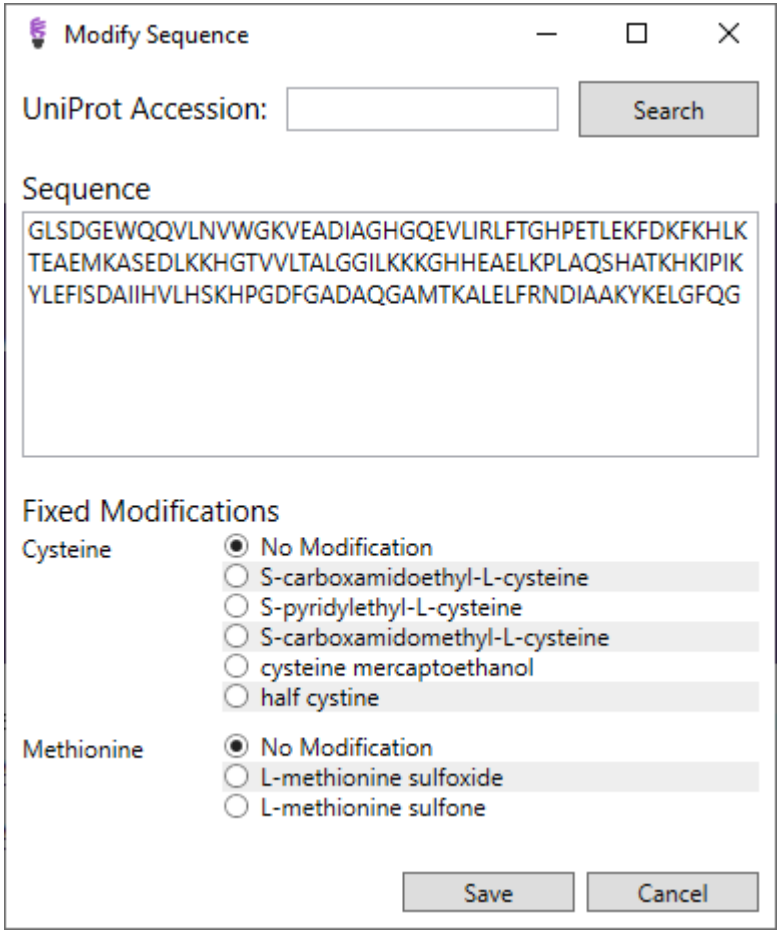
Cathy H. Wu et al. (eds.), *Protein Bioinformatics: From Protein Modifications and Networks to Proteomics*, Methods in Molecular Biology, vol. 1558, DOI 10.1007/978-1-4939-6783-4_18, © Springer Science+Business Media LLC 2017

ProSight Lite – User interface

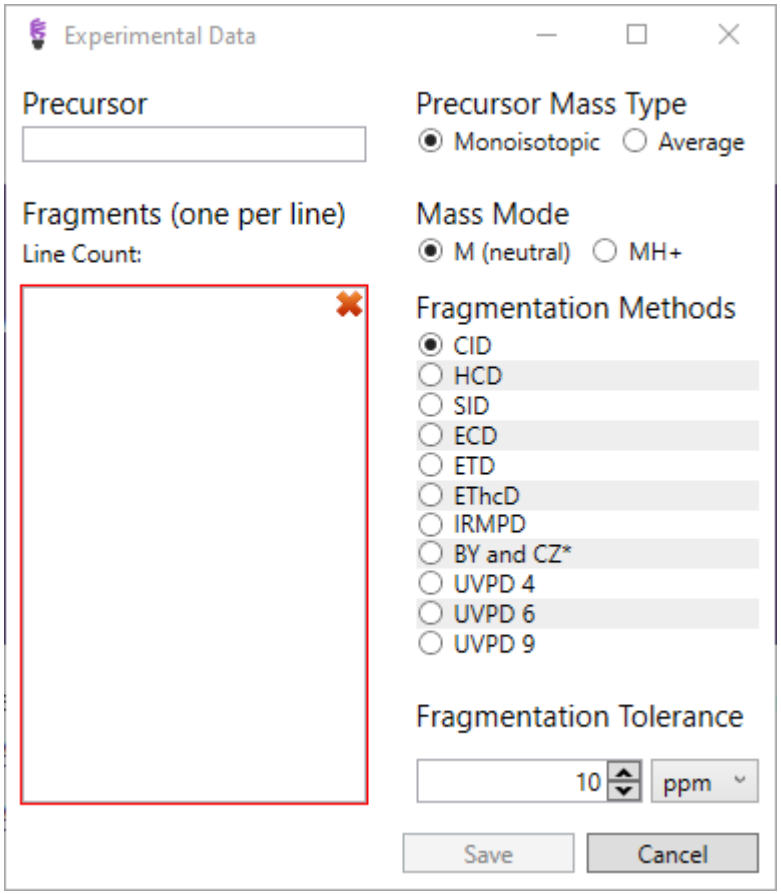


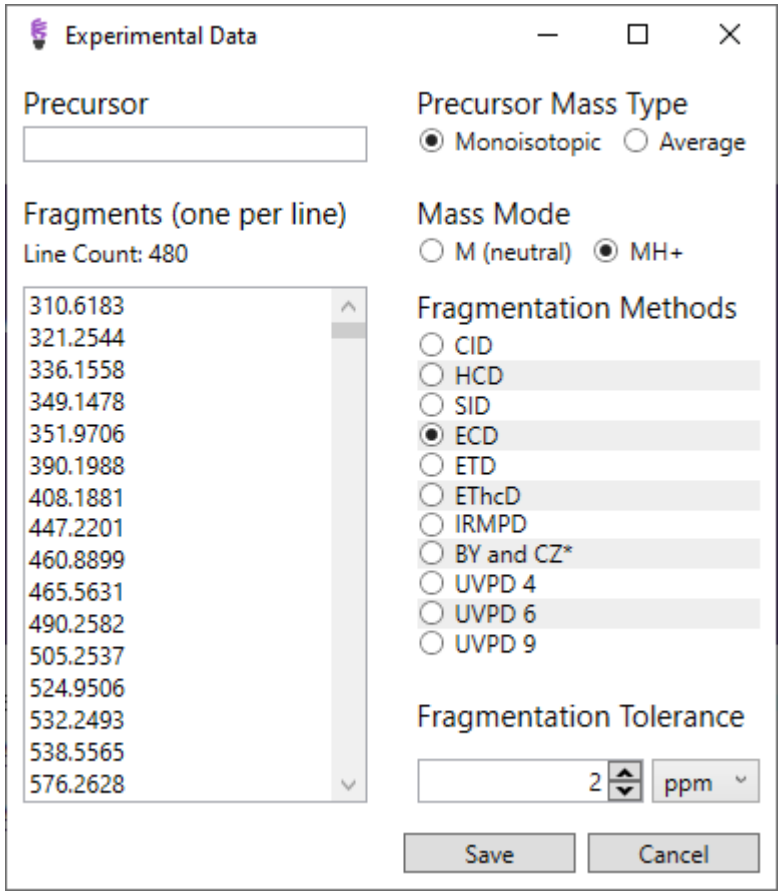
ProSight Lite – Sequence Uniprot/fasta





ProSight Lite – Experimental data





First Example

Myoglobin – modification by NHS-propionate (mass shift 56.026)

ECD – fragmentation of singly modified protein

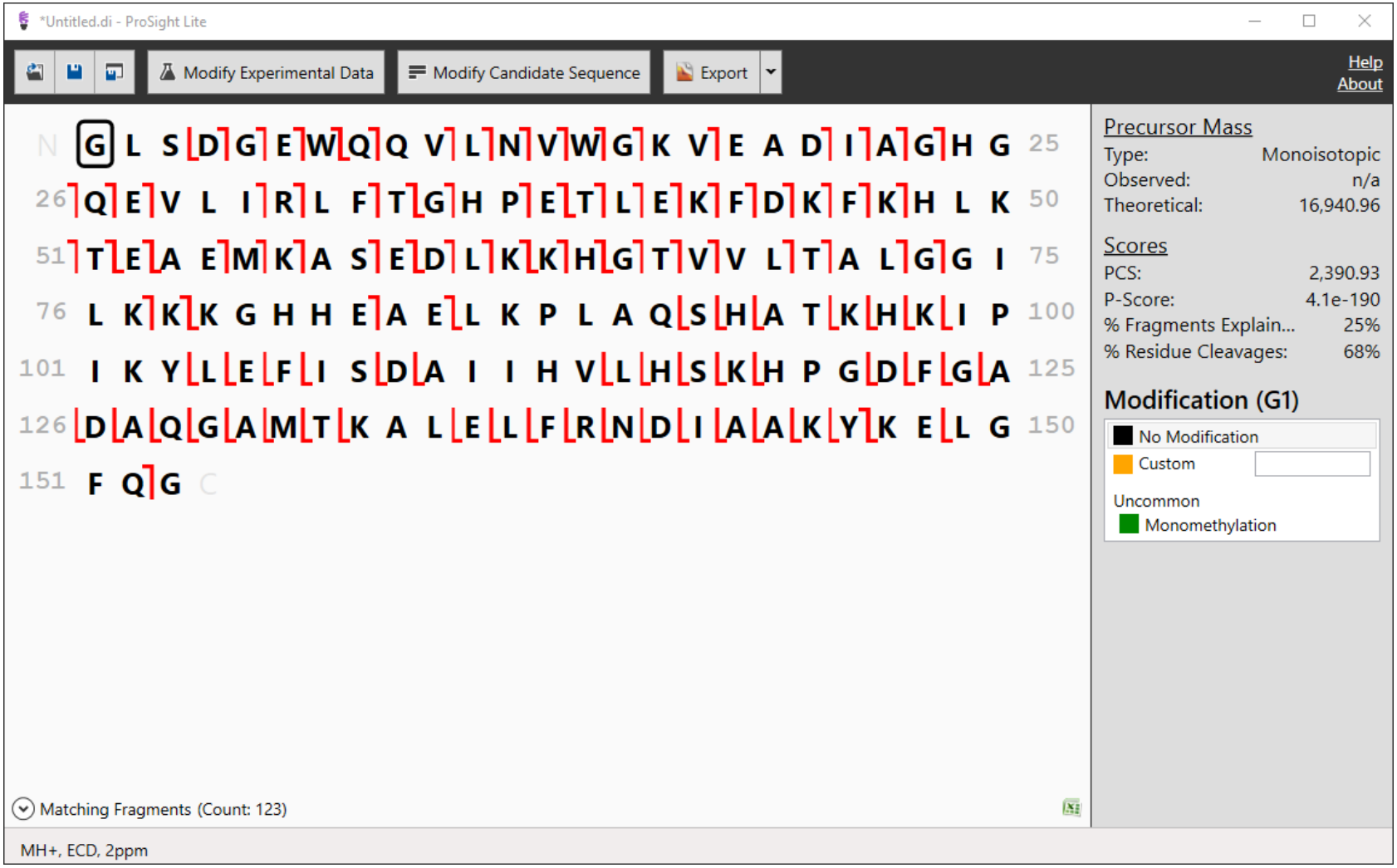
ProSight Lite – Experimental data from mgf file

Myo_ECD_NHSP.mgf - Notepad

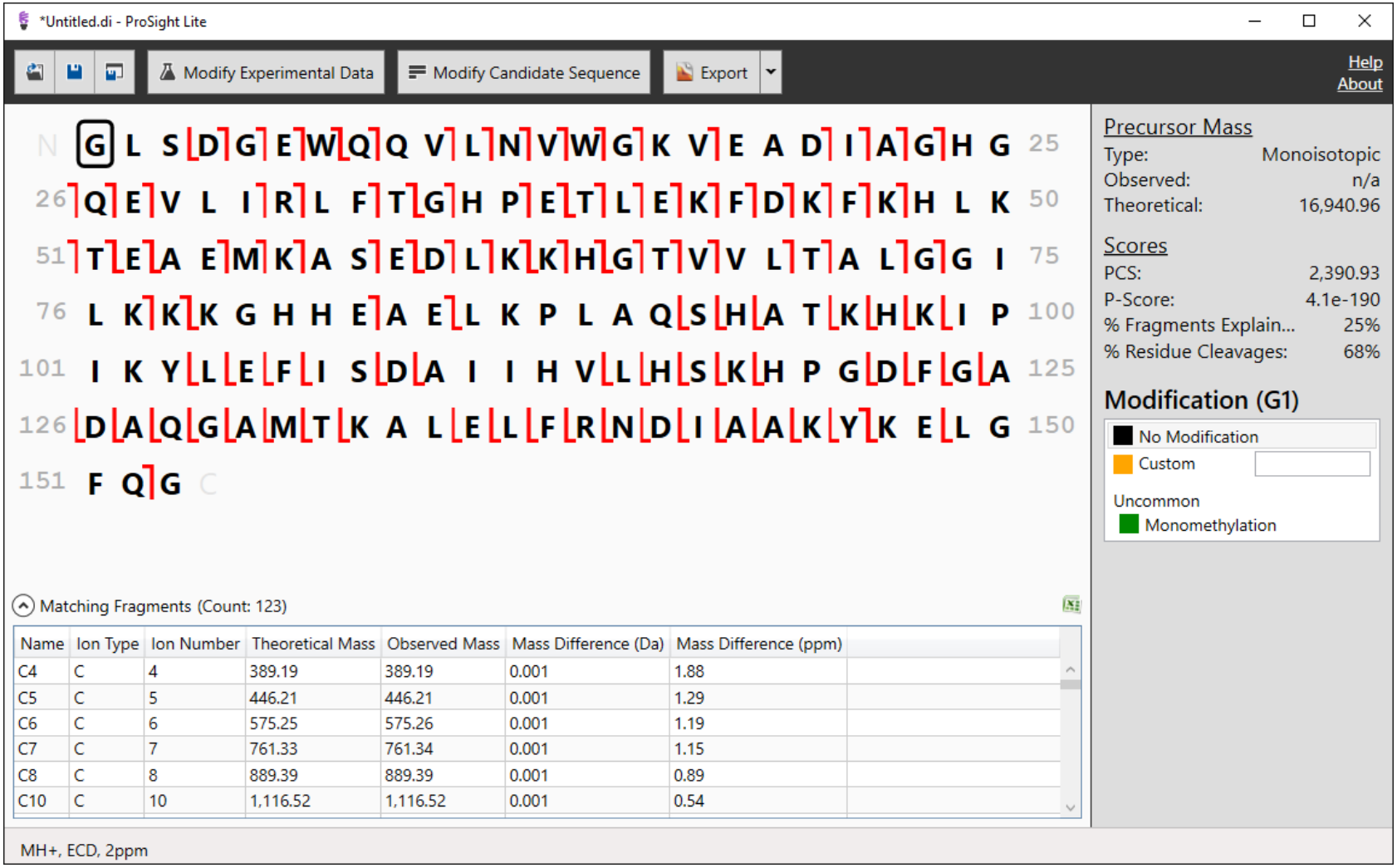
File Edit Format View Help

```
###     C:\Users\Administrator\Desktop\Myo_ECD_NHSP.mgf
###     D:\Biocev\Data\MultiCASI_ECD_NHSP_apo_MYO.d\analysis.baf
###     MultiCASI_ECD_NHSP_apo_MYO.d
###     Instrument: solariX XR
###    apo_MYO NHSP
###    MultiCASI_ECD
###    MultiCASI_ECD_NHSP_+19/+18/+17/+16/+15
###    Quality 0.50, S/N 0, BB C 4.9384 N 1.3577 O 1.4773 S 0.0417 H 7.7583, BB2 , RelIntens 0, MinIntens 0, MaxCharge 15, SNAP2 1
###    Retain Residuals no, Create neutrally no,
###    FullScan Parameters: AdductIon (+)+H (-)-H, LowMass 250, HighMass 4000, MaxCharge charge decon 4, MinPeaks 3
###    Perf. isotope decon yes, Max Charge Istopic decon 10, MW agreement 0.00, Abund. cutoff 10.00, Envelope cutoff 75.00
###    MaxRes   Parameters: AdductIon (+)+H (-)-H, LowMass 250, HighMass 4000
###    Perf. isotope decon yes, Max Charge Istopic decon 10, Abund. cutoff 2.00
###    MSn      Parameters: AdductIon (+)+H (-)-H, LowMass 50, HighMass 2000
###    Perf. isotope decon yes, Max Charge Istopic decon 8, Abund. cutoff 2.00
###    Global Charge Limit: no 3, Prefer FullScan Result: yes, Export what: 1 (0=non deconv/1=deconv only/2=mixed), NoMSMSSignals: 2000, Intensity threshold: 2000,
NoMSSignals (deconv): 2000, NoMSSignals (non-deconv): 2000, NormalizeData: no, SingleCharged: yes
###     +MS(CASI 895.58580)

###FS:    #m/z: 895.58580 #charge  19
###MaxRes:#m/z: 0.000000 #charge-128
###MS:
###MSMS:
BEGIN IONS
TITLE=+MS2(qCID 895.58580)
RTINSECONDS=0
SCANS=MS:  MSMS:
PEPMASS=895.58580       148506160

310.6183        1032778 1+
321.2544        625904  1+
336.1558        12251523        1+
349.1478        767920  1+
351.9706        1008276 1+
390.1988        30597142        1+
408.1881        4508194 1+
447.2201        137446480       1+
460.8899        812333  1+
465.5631        890910  1+
490.2582        930141  1+
505.2537        8062103 1+
524.9506        859948  1+
532.2493        4148949 1+
```

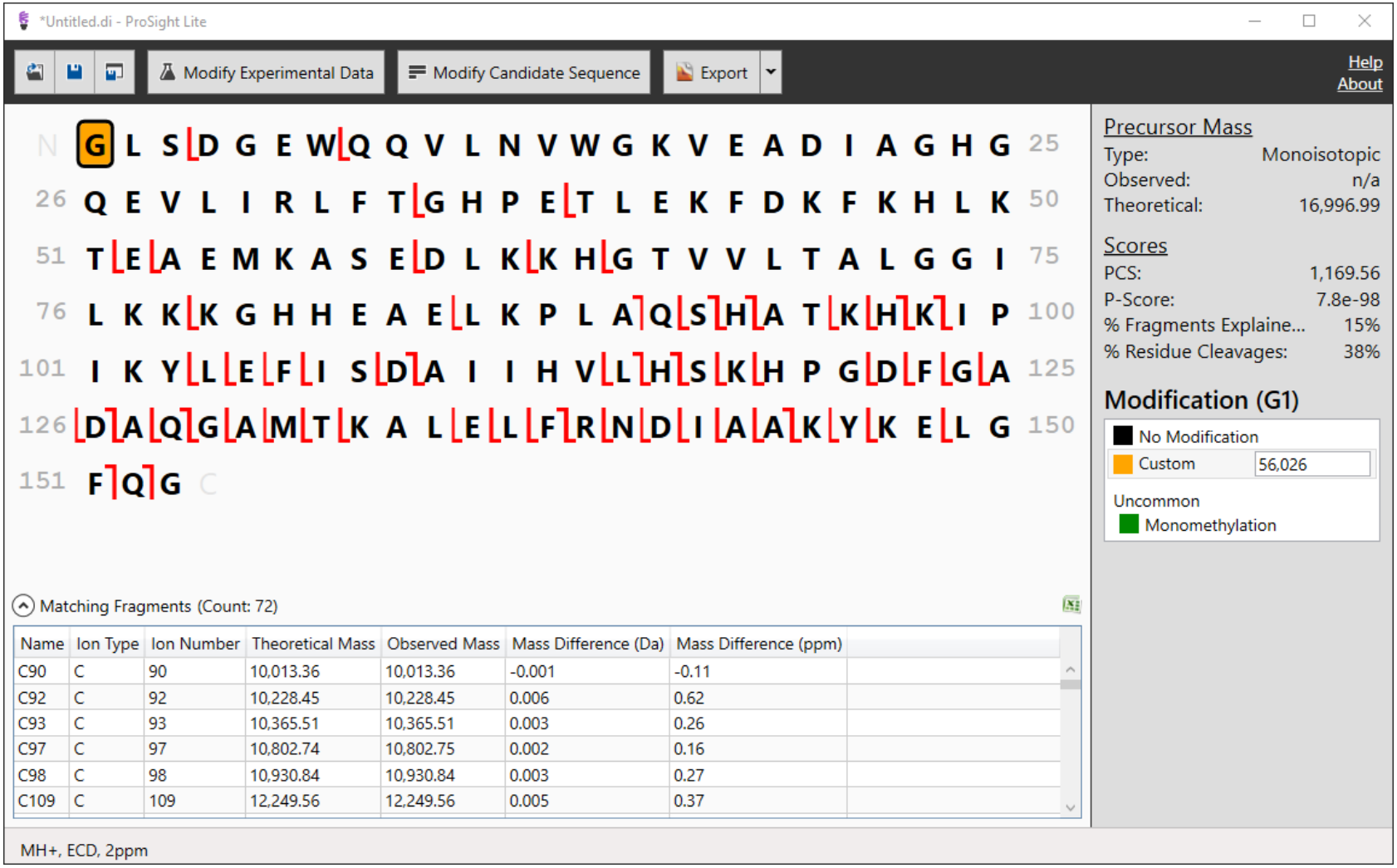
ProSight Lite – Results window



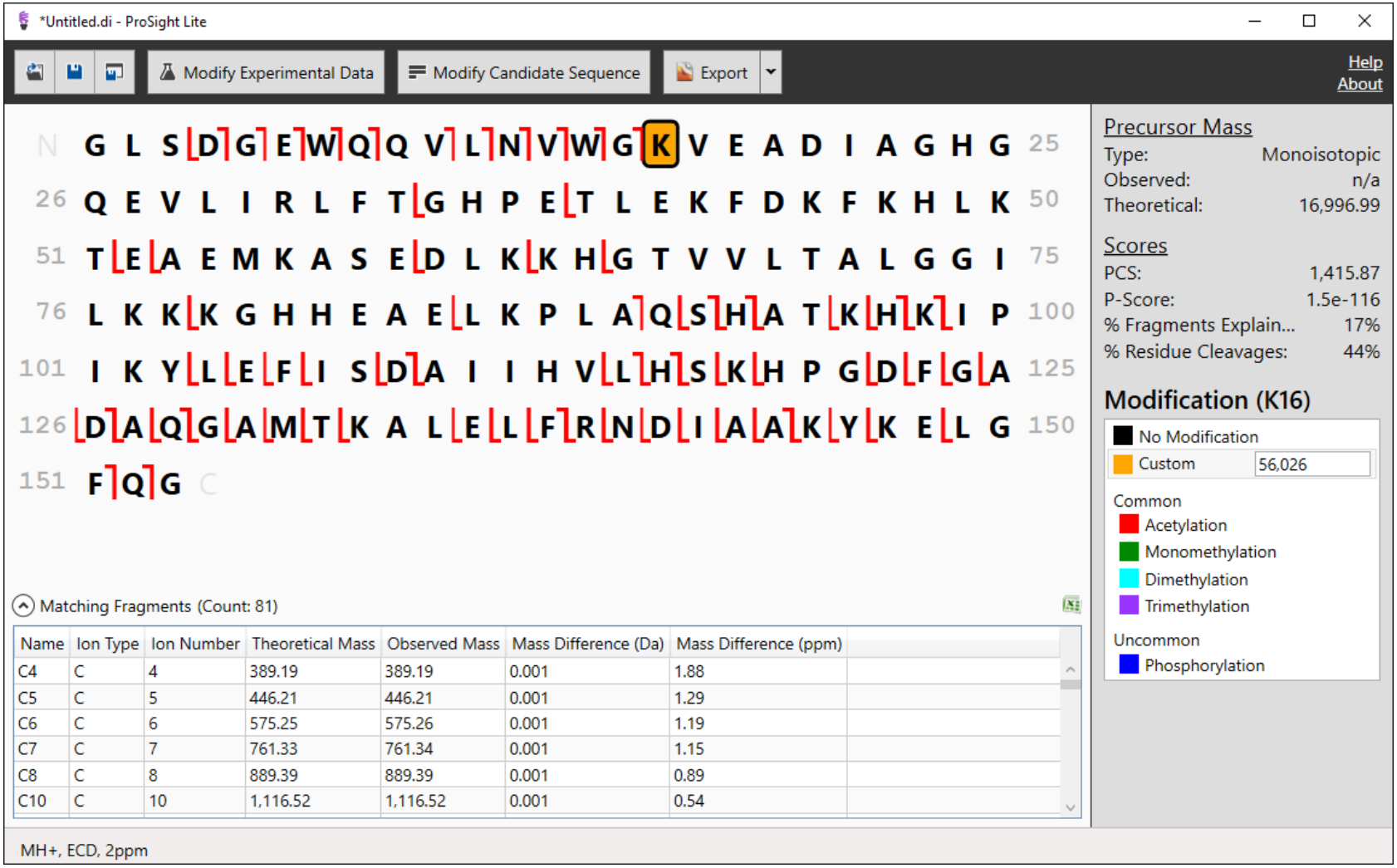
ProSight Lite – List of matching fragments



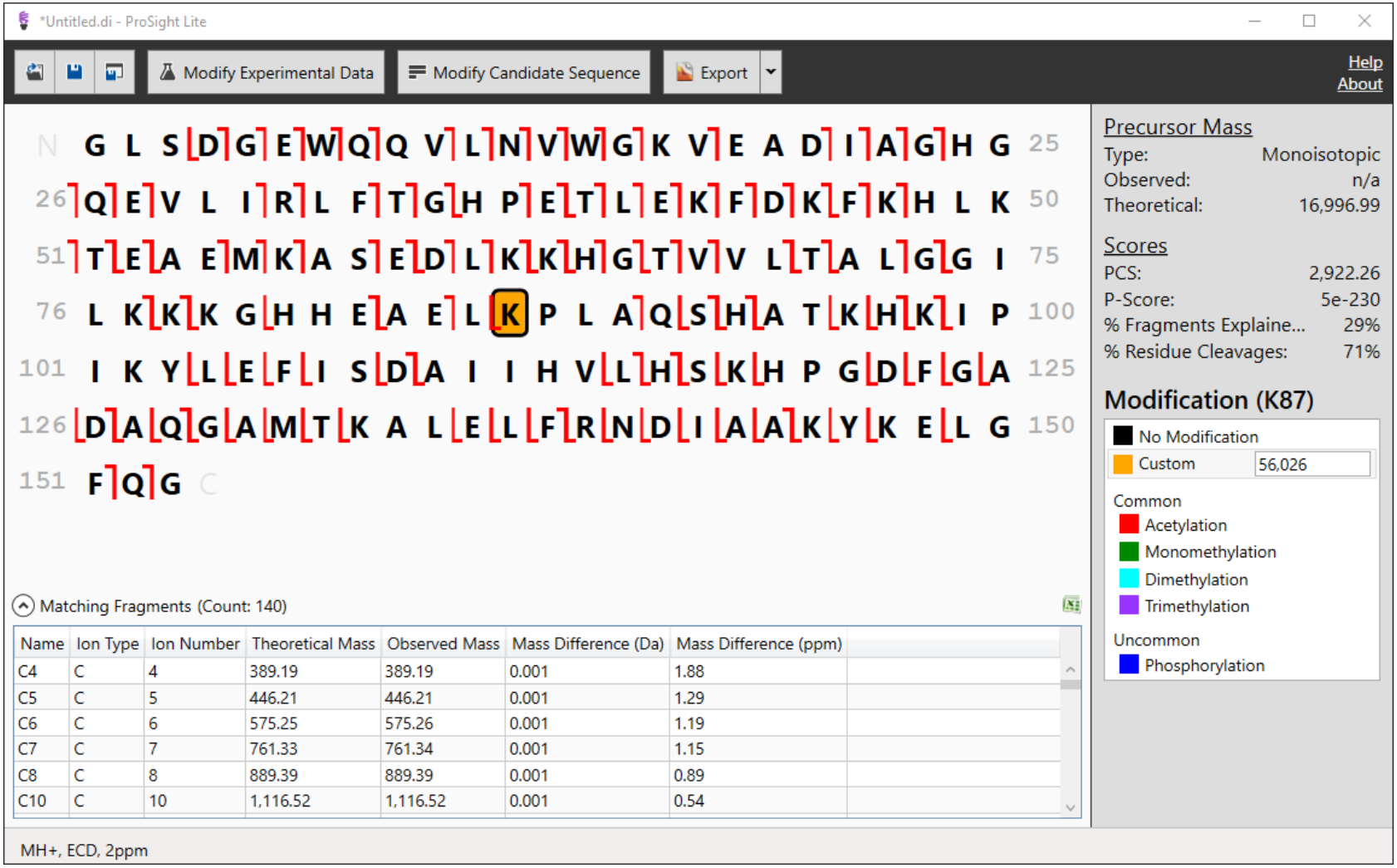
ProSight Lite – Modification of N-terminus



ProSight Lite – Modification of K16



ProSight Lite – Modification of K87

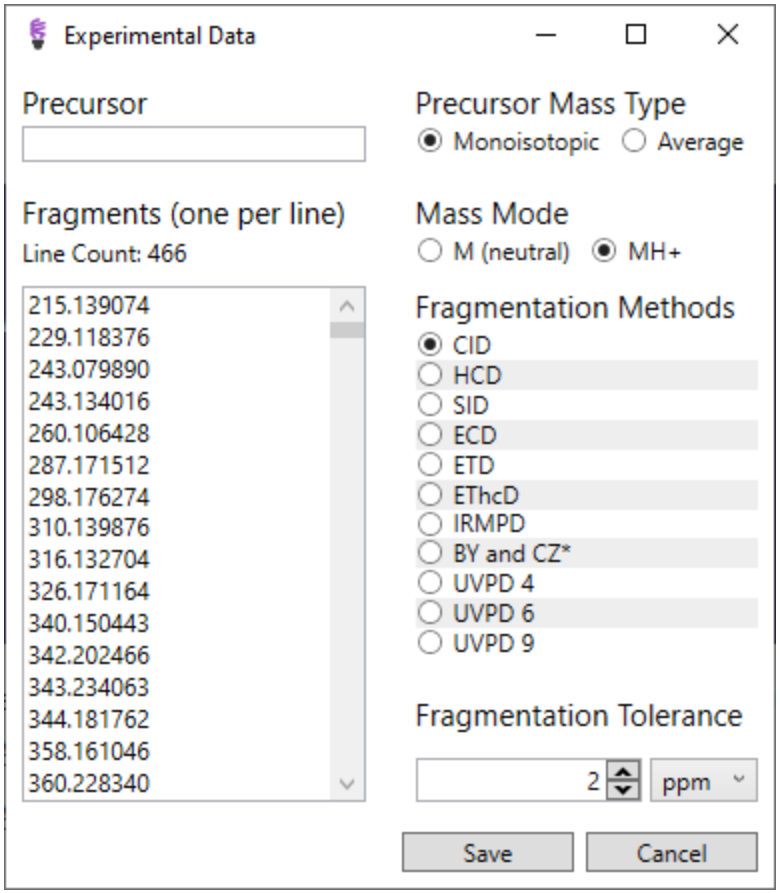


Second Example

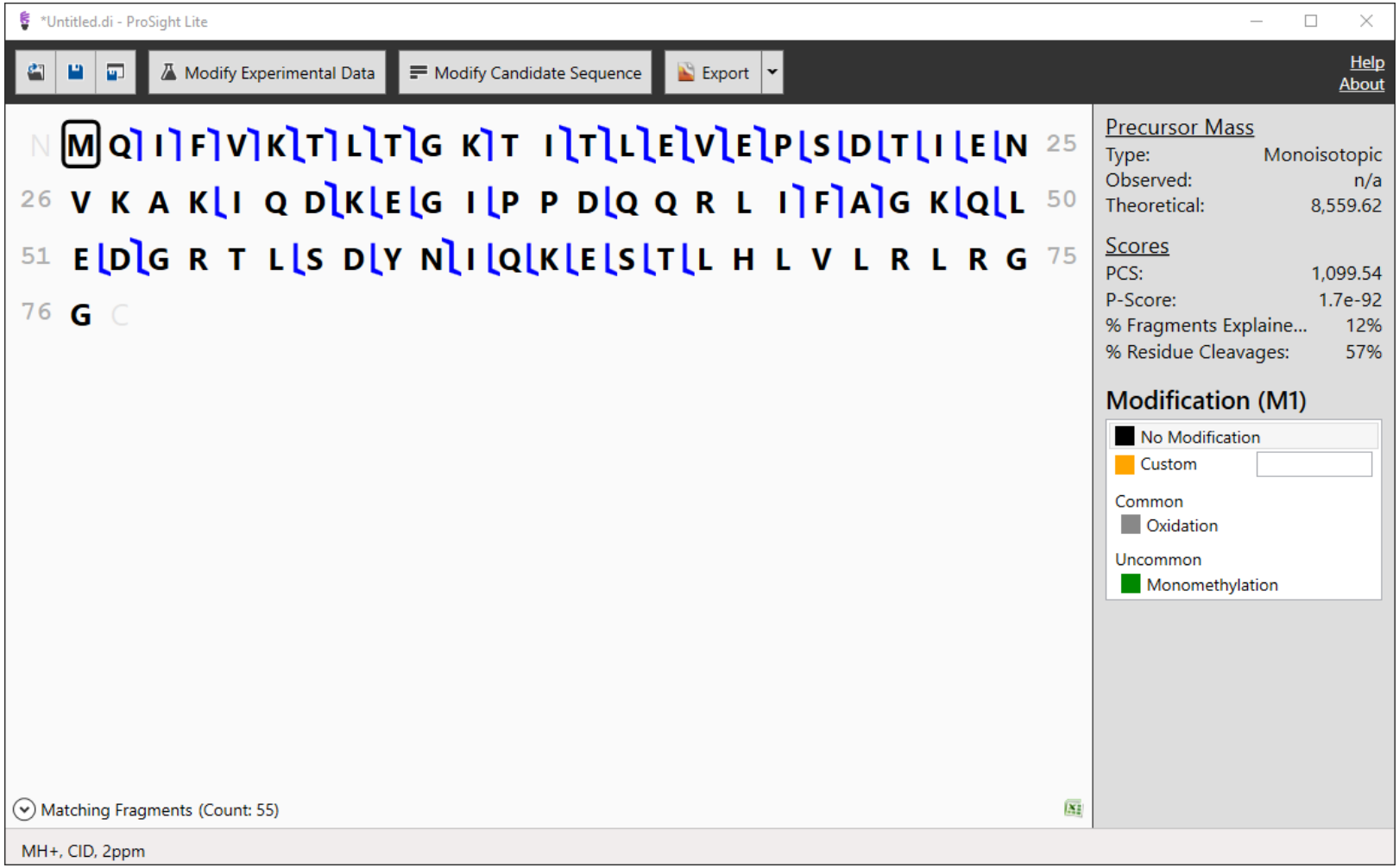
Ubiquitin – modification by NHS-propionate (mass shift 56.026)

CID – fragmentation of singly modified protein

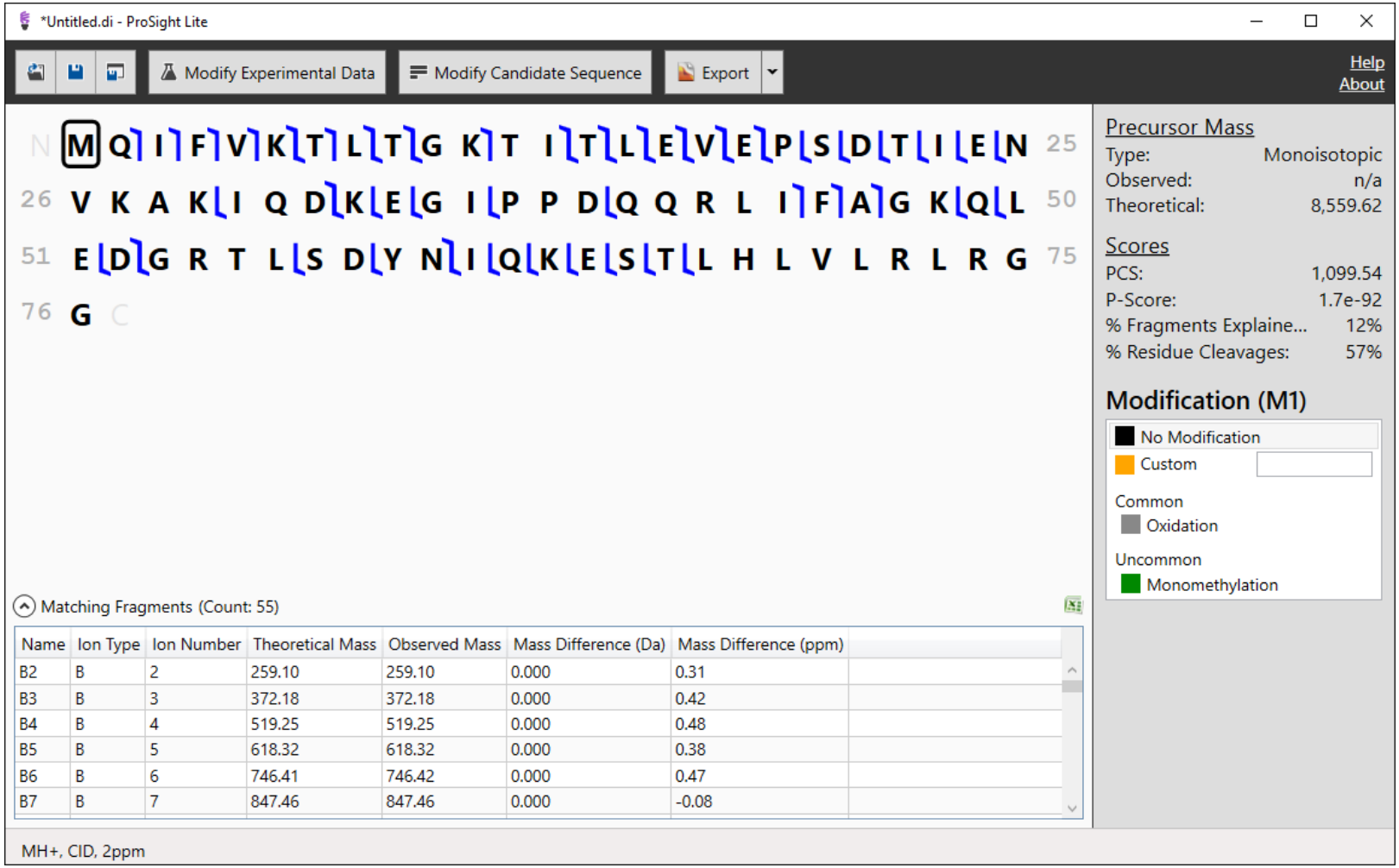
ProSight Lite – Experimental data



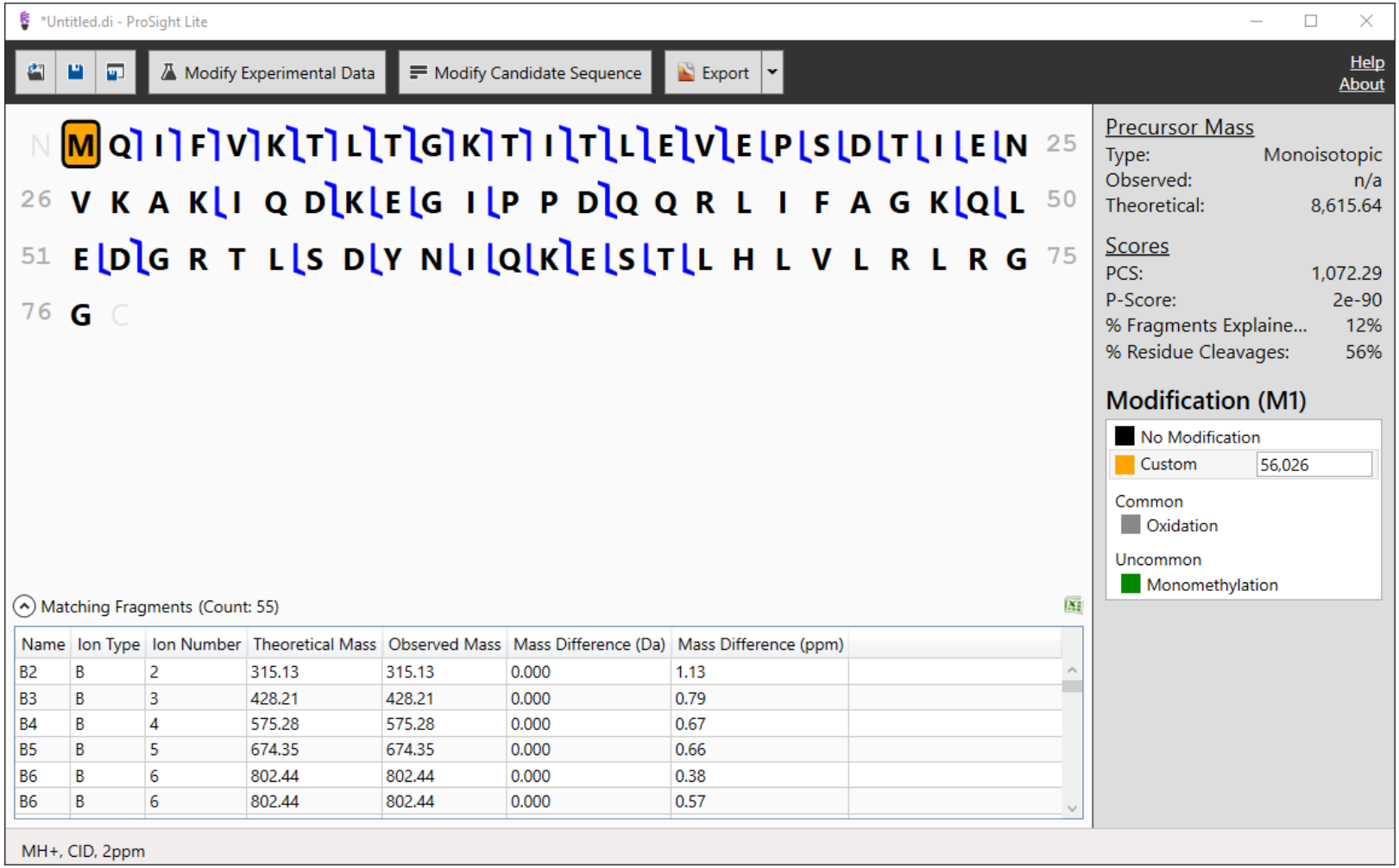
ProSight Lite – Result window



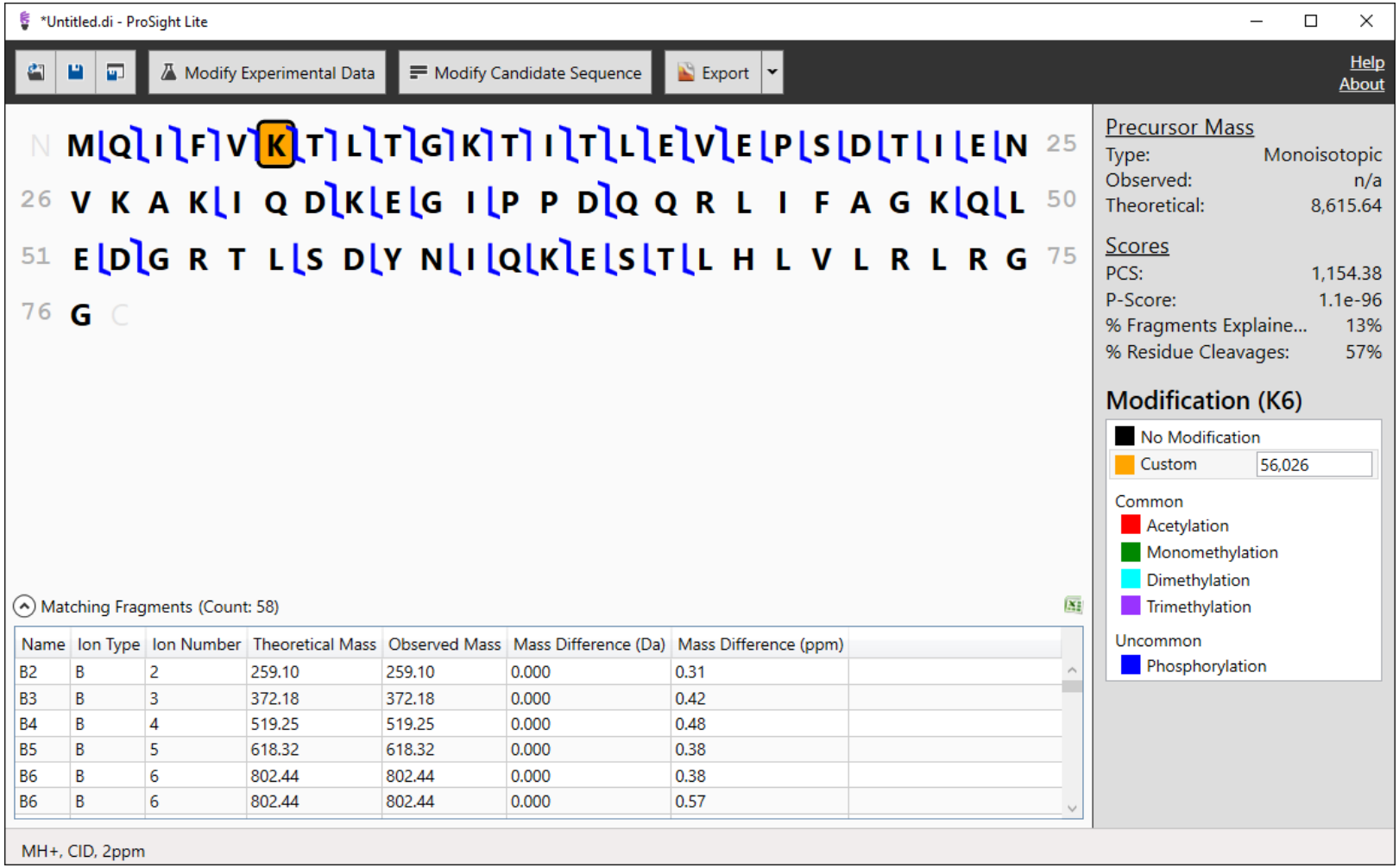
ProSight Lite – Matching fragments



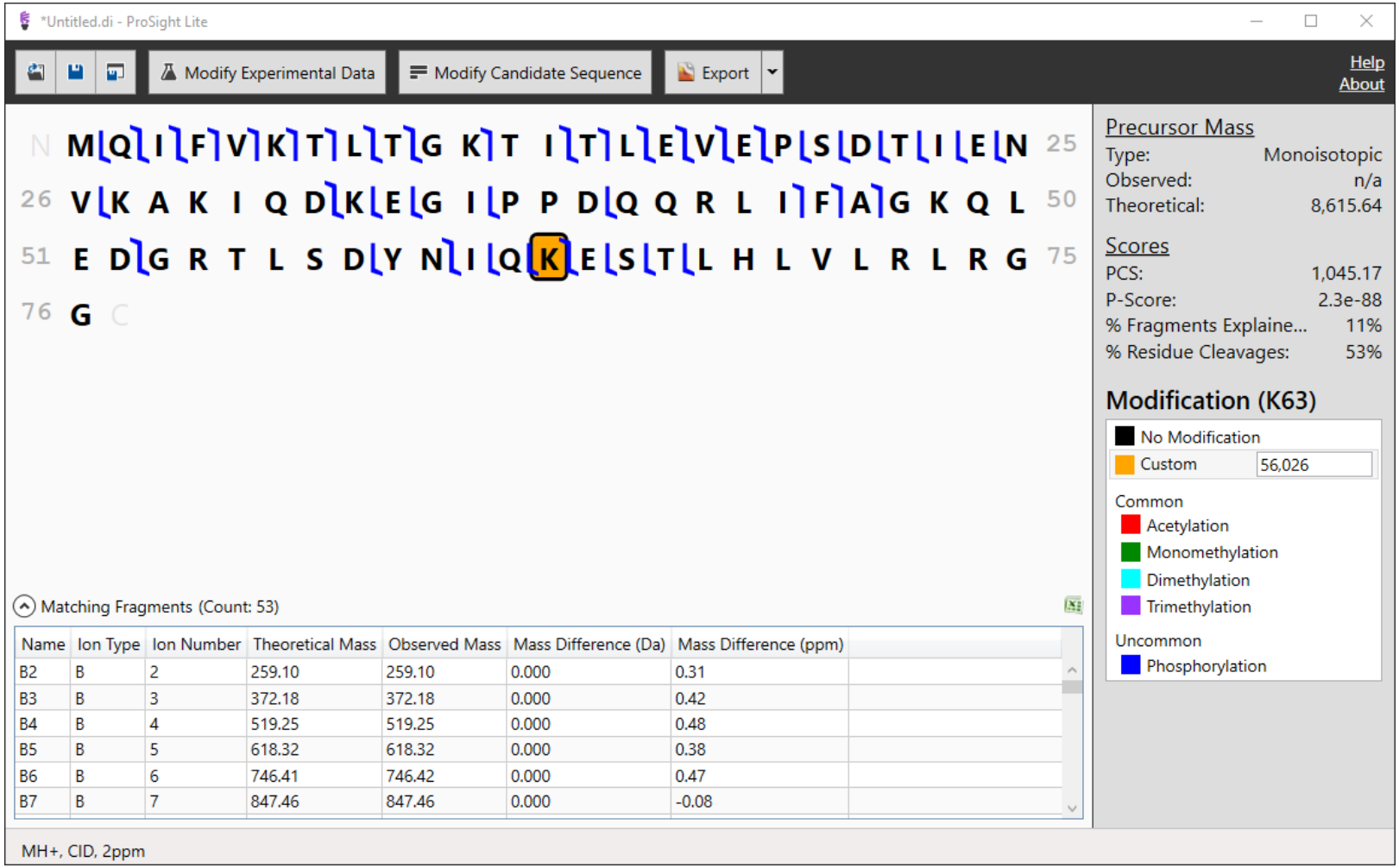
ProSight Lite – Modification of N-terminus

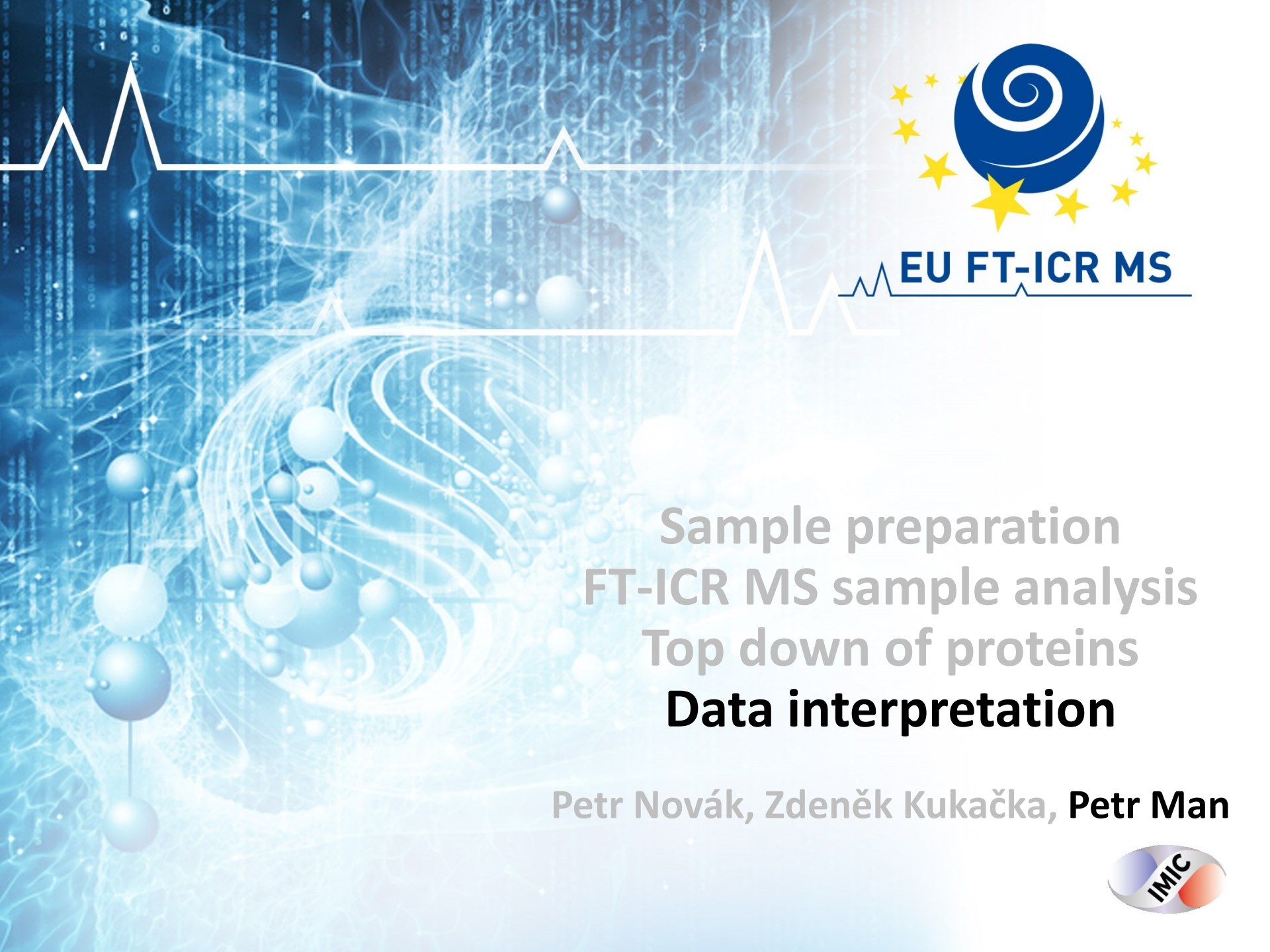


ProSight Lite – Modification of K6



ProSight Lite – Modification of K63





Sample preparation
FT-ICR MS sample analysis
Top down of proteins

Data interpretation

Petr Novák, Zdeněk Kukačka, **Petr Man**

IMIC