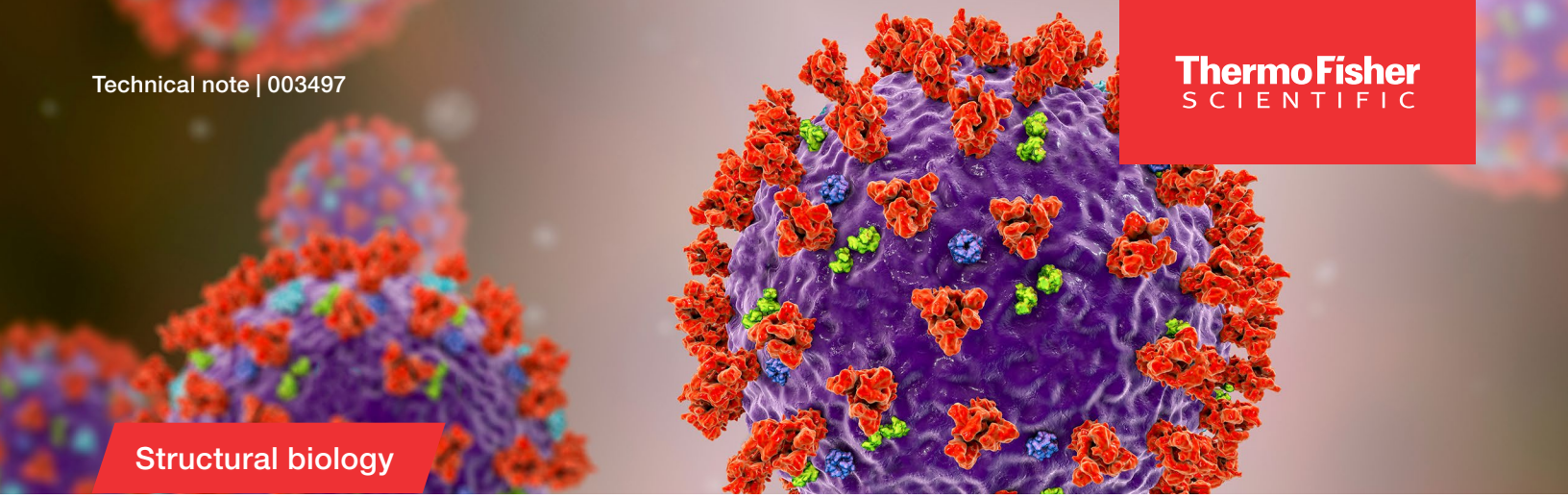

Technical note | 003497

Structural biology

Decipher intricate glycoproteins using data-independent acquisition-proton transfer charge reduction and native top-down mass spectrometry

Authors

Weijing Liu, Yuqi Shi, Christopher Mullen, Julian Saba, Roberto Gamez, Ping Yip, Rosa Viner

Thermo Fisher Scientific, San Jose, CA

Keywords

Glycoprotein, data-independent acquisition, DIA, PTCR, proton transfer charge reduction, native mass spectrometry, native MS, Top-Down, glycosylation, proteoform, ETD, electron transfer dissociation, Orbitrap Ascend Structural Biology Tribrid MS, Orbitrap Eclipse MS

Introduction

The SARS-CoV-2 pandemic underscores the urgent need for rapid viral glycoprotein research. Viral glycoproteins, crucial for host cell attachment, are key targets for neutralizing antibodies produced by vaccines. While RNA sequencing reveals viral mutations, it does not capture post-translational modifications (PTMs) like glycosylation, which influence receptor binding and infection efficiency. A comprehensive understanding of glycoprotein glycosylation is essential for the development of effective vaccines and therapeutic strategies.

Native mass spectrometry (native MS or nMS) is a powerful tool in the mass spectrometry (MS) arsenal for characterizing glycosylation on viral glycoproteins. Unlike other methods, native MS maintains protein structures intact and introduces them into the mass spectrometer in a configuration that closely resembles their natural state in biological conditions. This technique provides an unaveraged snapshot of the solution conditions, allowing for the simultaneous detection of different proteoforms, such as varying glycosylations, which is challenging to achieve with other structural biology techniques. Native MS can identify both the number and types of glycoform compositions present.

However, challenges remain in native MS and native top-down MS characterization of glycoproteins as the heterogeneity leads to complex spectra. This study utilizes the high quadrupole mass filter coupled with proton transfer charge reduction (PTCR) on Thermo Scientific™ Orbitrap™ Tribrid™ platforms to decipher the complexity of glycoproteins, including ones from the COVID family using native MS.[1-2] Top-down analysis employing electron transfer dissociation (ETD) provides structural, sequence, and PTM site information. In summary, the combination of data-independent acquisition-proton transfer charge reduction (DIA-PTCR) (Figure 1) and native top-down MS approaches enables a comprehensive characterization of viral glycoproteins.

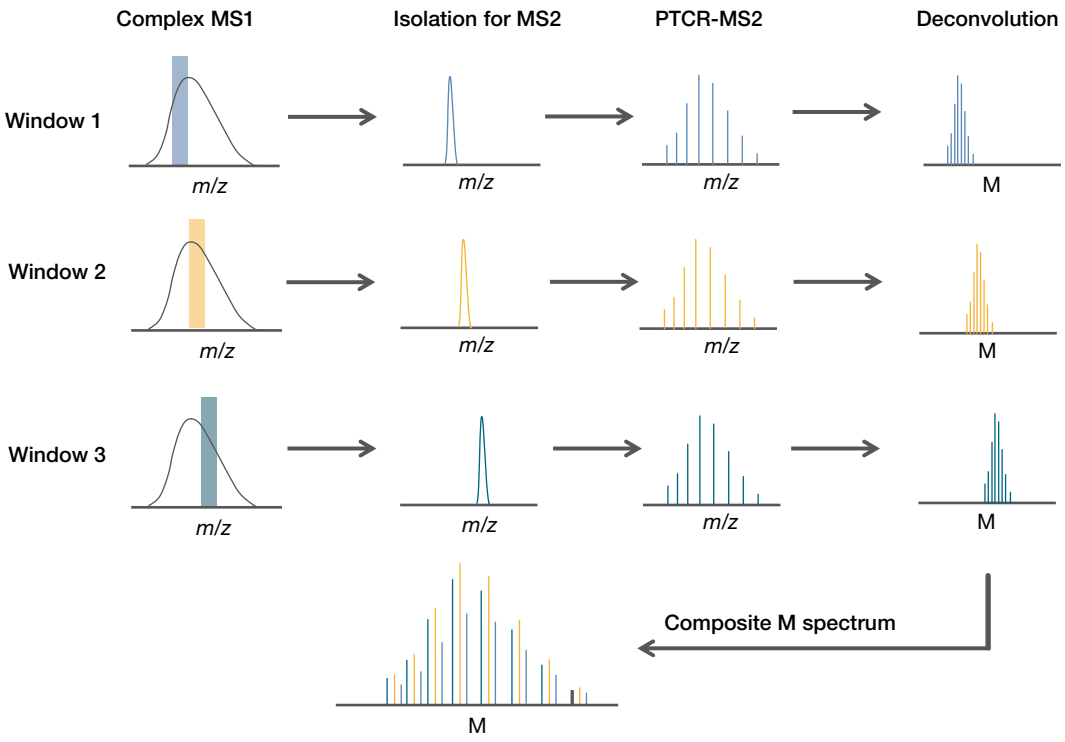

Figure 1. Overlapping windows of DIA-PTCR spectra are acquired and stitched together for deconvolution

Experimental

Materials

- Ammonium acetate, Sigma (P/N 372331-10G)
- Fisher Chemical™ Optima™ LC/MS grade water, (P/N 10505904)

Sample preparation

Human Fetuin (hFet) was purchased from Sigma-Aldrich. Nucleocapsid protein (Nc), spike protein receptor-binding domain (RBD), and full-length spike protein (S-protein) were purchased from Acro Biosystems. Proteins were buffer exchanged into 200 mM ammonium acetate using an Amicon™ Ultra centrifugal filter (Sigma-Millipore). Samples were diluted to 0.1–0.2 mg/mL prior to static nanospray experiments.

MS experiments

Glycoprotein analyses were performed on a Thermo Scientific™ Orbitrap™ Ascend Structural Biology Tribrid™ mass spectrometer. Native top-down analysis of human Fetuin was performed on a Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometer.

Data analysis

The data were analyzed using Thermo Scientific™ BioPharma Finder™ 5.0 software.

Table 1. MS parameters for MS^1 and DIA-PTCR experiments

Source parameters		
Spray voltage (+V)	1,200–1,400	1,200–1,400
Capillary temperature (°C)	275	275
Orbitrap scan parameters		
Method type	Full MS	DIA-PTCR
Scan range (*m/z*)	2,000–16,000	3,000–16,000
Application mode	Intact	Intact
Pressure mode	High (20 mtorr)	High (20 mtorr)
Resolution	7,500 at *m/z* 200	480,000 at *m/z* 200 7,500 at *m/z* 200 for S-protein
RF lens (%)	60	60
AGC target value	200	400
Max injection time (ms)	100	1,000
Isolation mode	-	Quadrupole
Isolation window (MS^2)	-	5–20 Th
Microscans	3	10
Source fragmentation (V)	25–100	25–120
Source CID compensation scaling	0.01–0.02	0.01–0.02
PTCR reaction time (ms)	-	7–15
PTCR reagent target	-	6.00E+05

Results and discussion

1. DIA-PTCR and native top-down MS reveal more proteoforms and structure of hFet

hFet (Uniprot P02765) is a heavily modified glycoprotein with a predicted mass of approximately 37 kDa. It contains two *N*-linked glycosylation sites, suggested seven *O*-linked glycosylation sites, and seven phosphorylation sites. The enriched PTMs lead to partially resolved spectrum of hFet from full scan analysis (Figure 2A). By leveraging the high quadrupole mass filter combined with PTCR on the Orbitrap Ascend Structural Biology MS, we initially compared the spectral quality obtained using the ion trap versus the quadrupole isolation (Figure 2B). Subsequently, we performed gas-phase fractionation and charge reduction using DIA-PTCR analysis to uncover the numerous glycoforms concealed in the full scan data (Figure 3).

When the precursor at *m/z* 3,662 (highlighted in orange in Figure 2A) was isolated in both the ion trap and quadrupole with a 5-Th isolation width, the signal-to-noise ratio (S/N) was nearly 6-fold higher with quadrupole isolation (Figure 2B, left). As a result, after conducting PTCR, the S/N of the charge-reduced envelope was significantly higher with quadrupole isolation compared to ion trap isolation (Figure 2B, right). Similarly, the isolation of the near-baseline low intensity precursor at *m/z* 3,500 resulted in improved isolation accuracy and a higher S/N with quadrupole isolation compared to ion trap isolation (Figure 2C).

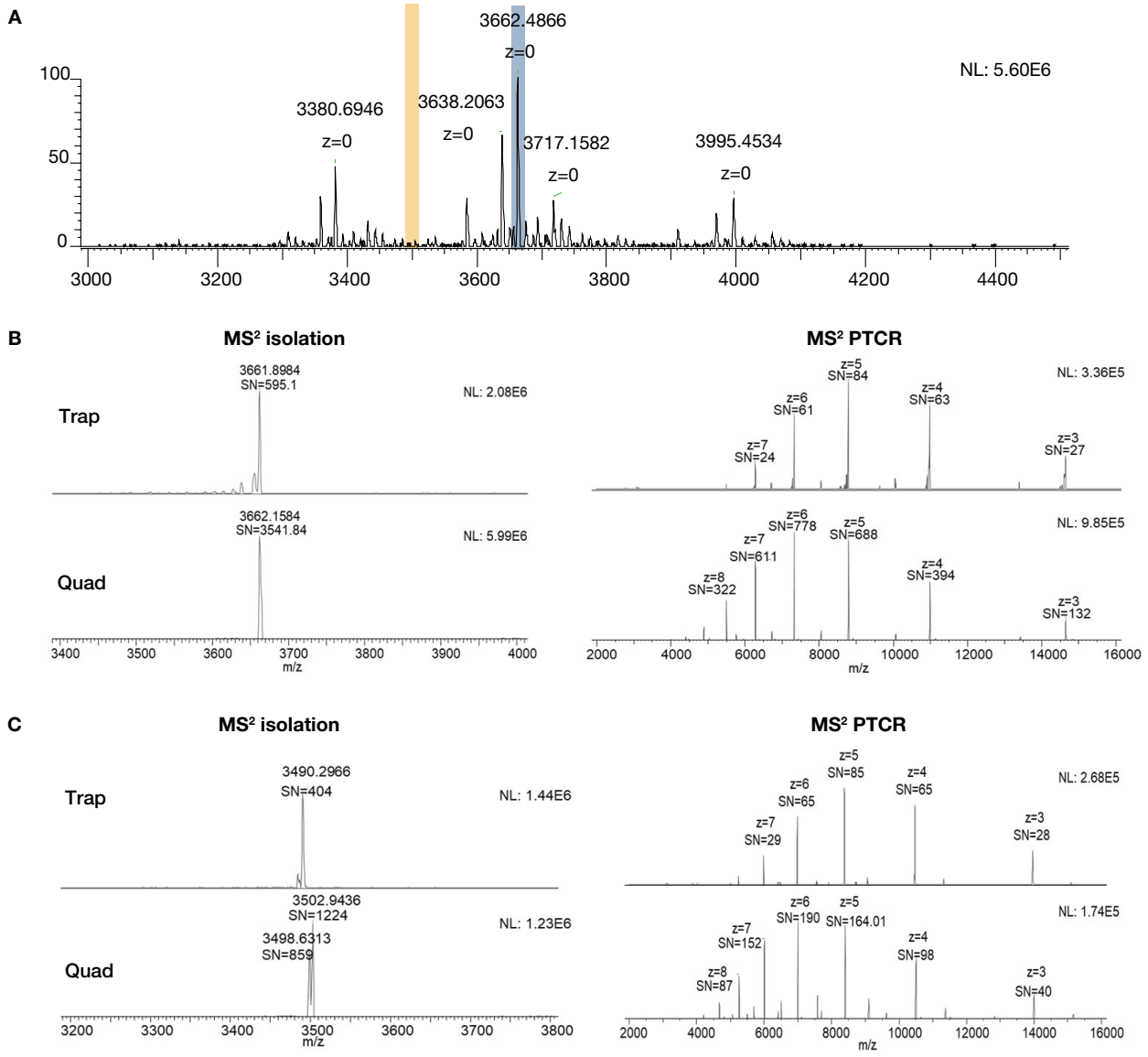

Figure 2. (A) MS¹ spectrum of hFet; (B) ion trap vs. quadrupole isolation of 5-Th followed by PTCR at *m/z* 3,662 for S/N comparison; and (C) *m/z* 3,500 for isolation accuracy comparison

Additionally, performing PTCR revealed new charge envelopes by separating previously overlapping signals in the MS[1] full scan, leading to the discovery of new proteoforms. Comparing MS[1] glycoform assignments to those obtained using DIA-PTCR with Biopharma Finder software, a significantly greater number of proteoforms were assigned from the latter approach (Figure 3).

Structurally, although according to the sequence, the N-terminus of the B-chain in hFet is connected to the C-terminus of the A-chain via a propeptide, native top-down electron-transfer and higher-energy collision dissociation (EThcD) fragmentation shows that the B-chain is disulfide-bonded to the N-terminus of the A-chain through Cys32–Cys358 (Figure 4).

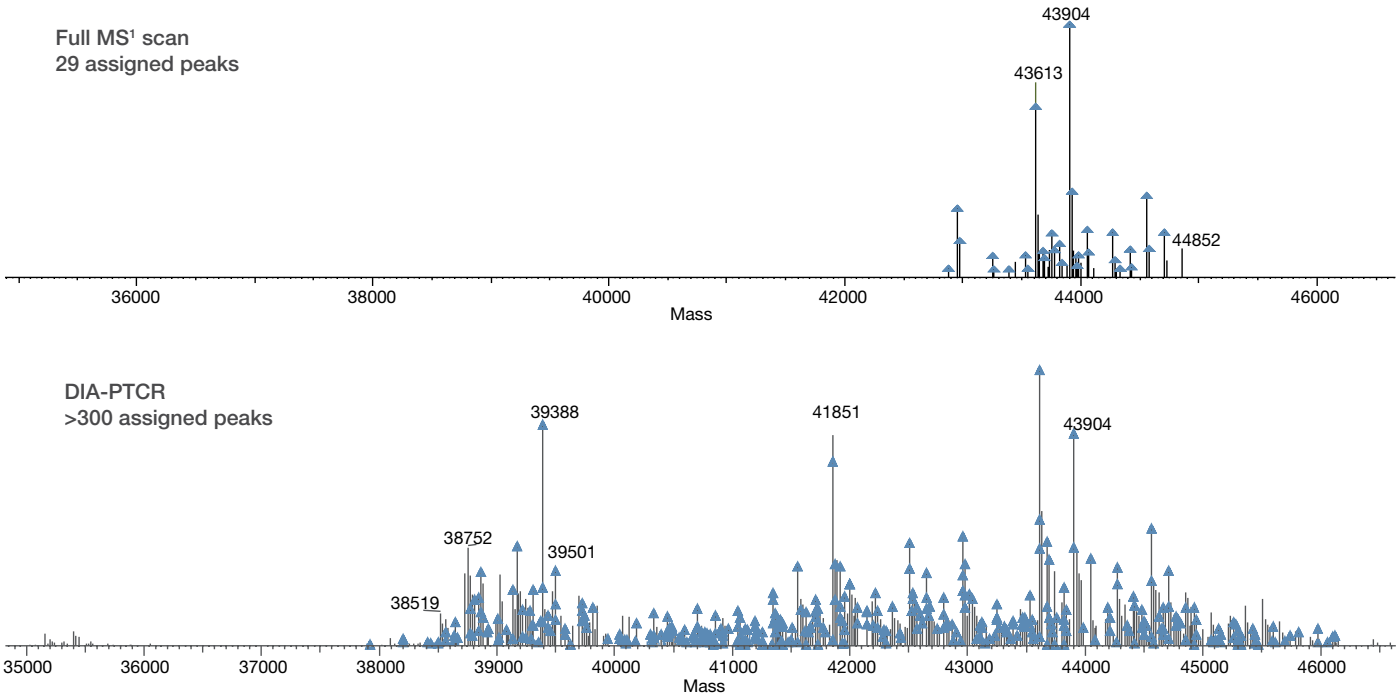

Figure 3. A comparison of hFet glycoforms assigned by Biopharma Finder software using MS[1] and DIA-PTCR

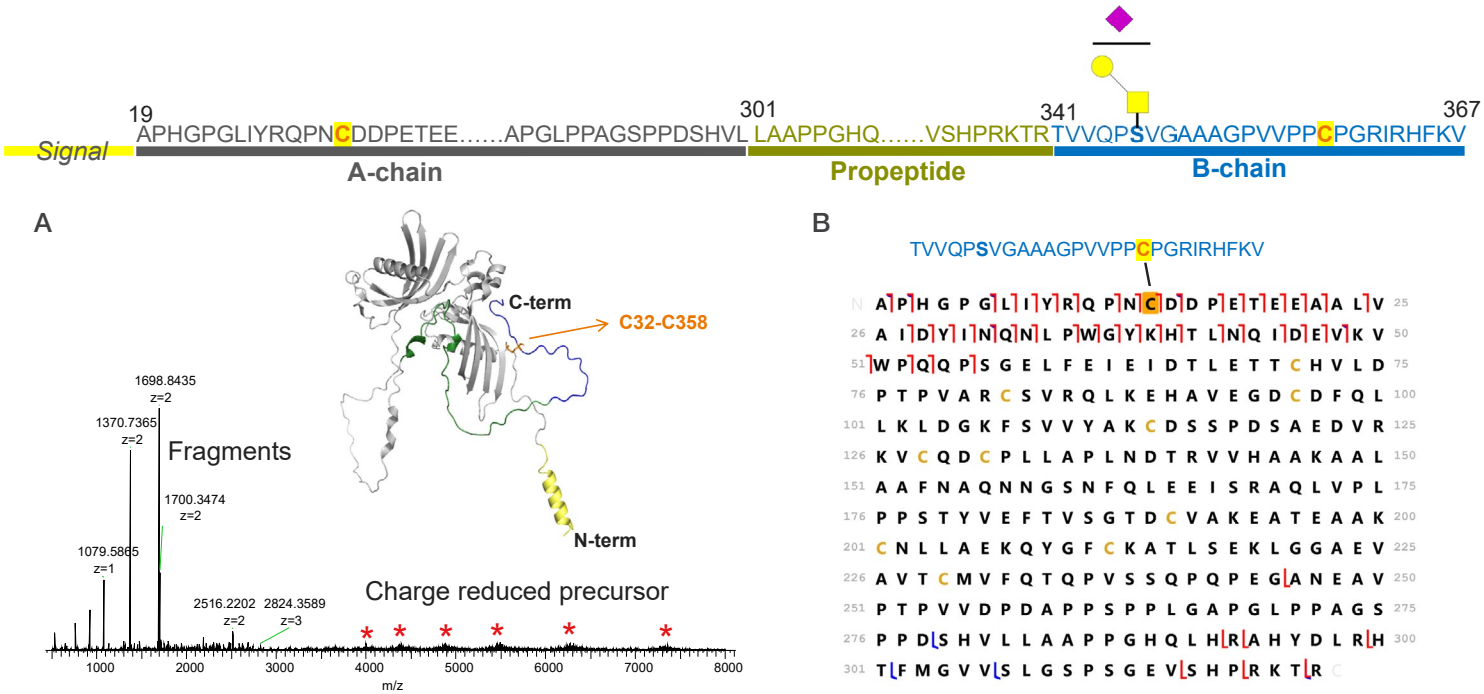

Figure 4. (A) EThcD spectrum and (B) sequence map to resolve chain connection of hFet

2. DIA-PTCR and top-down analyses of Nc

We applied the same methodology to the proteins from the COVID family, including Nc, S-protein RBD, and S-protein. The Nc protein may be glycosylated depending on its leader sequence and expression system. A full MS scan of 49 kDa Nc revealed partially resolved peaks atop the elevated baseline. DIA-PTCR analysis across the entire *m/z* range not only resolved the charge states and identified PTMs on the 49 kDa species but also revealed dimers near 100 kDa (Figure 5A). Some unexpected peaks were observed around 70 kDa and 90 kDa (Figure 5B), which are presumed to be impurities in the sample. DIA-PTCR reveals that the Nc protein is adorned with either covalent modifications or noncovalent interactions. Mild and strong desolvation conditions disclose varying extents of noncovalent (+Na) and covalent interactions (Figure 5B).

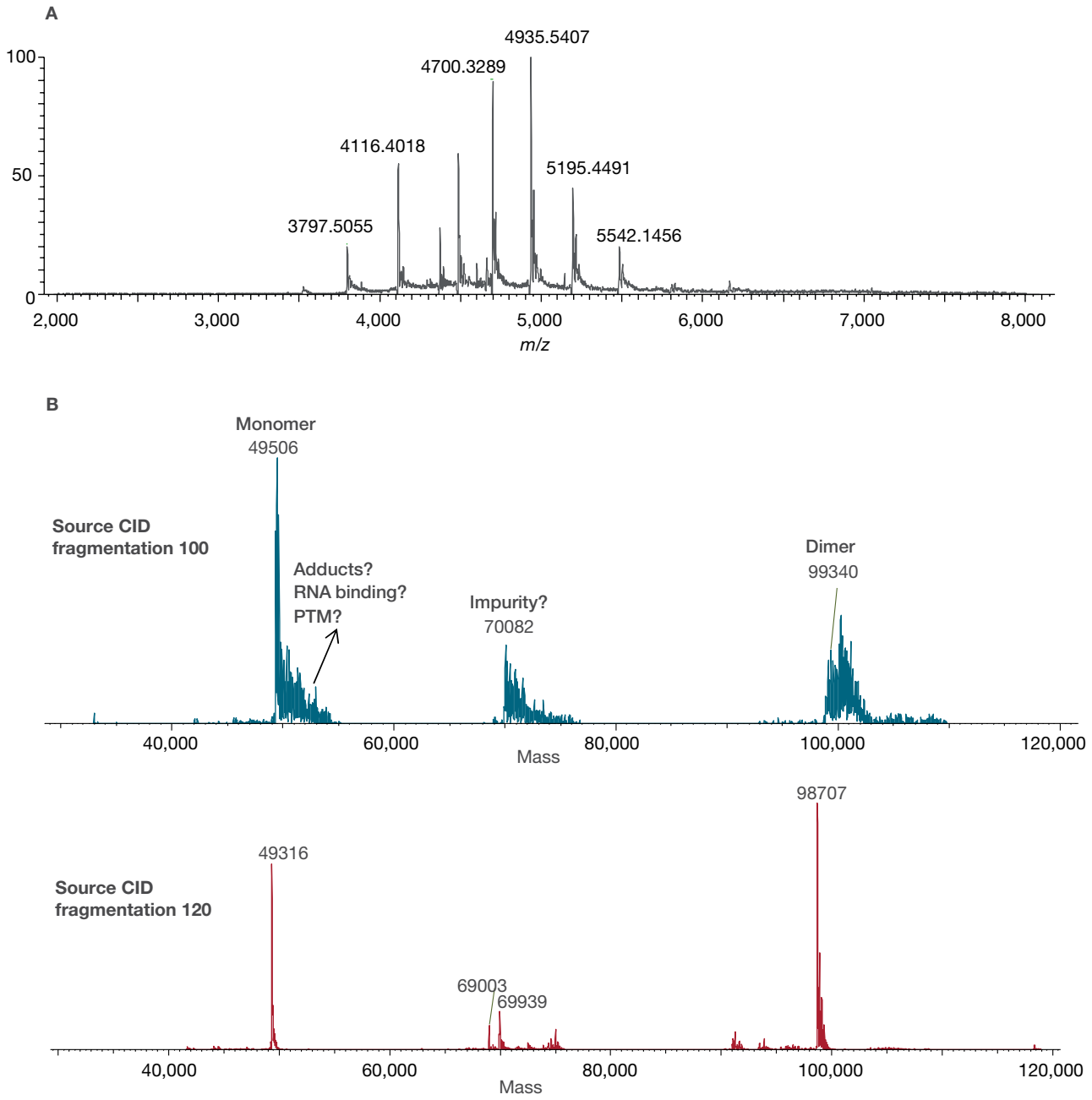

Figure 5. (A) A full MS scan of 49 kDa Nc revealed partially resolved peaks atop the elevated baseline; (B) DIA-PTCR at different desolvation voltages

The Nc sample was received with a confidential N-terminal tag, which hindered the assignment of top-down fragments. To address this, we performed top-down sequencing to obtain the tag mass. Such analysis revealed an N-terminal monoisotopic mass shift on methionine of 3,686.6794 Da (Figure 6).

Deciphering the unknown tag mass benefits and subsequently enhances top-down data analysis for structural elucidation.

Fragments of Nc protein generated from top-down ETD are predominantly c-ions, reflecting an exposed N-terminus. This aligns with the structure featuring a flexible N-terminus and a dimerized C-terminus (Figure 7).

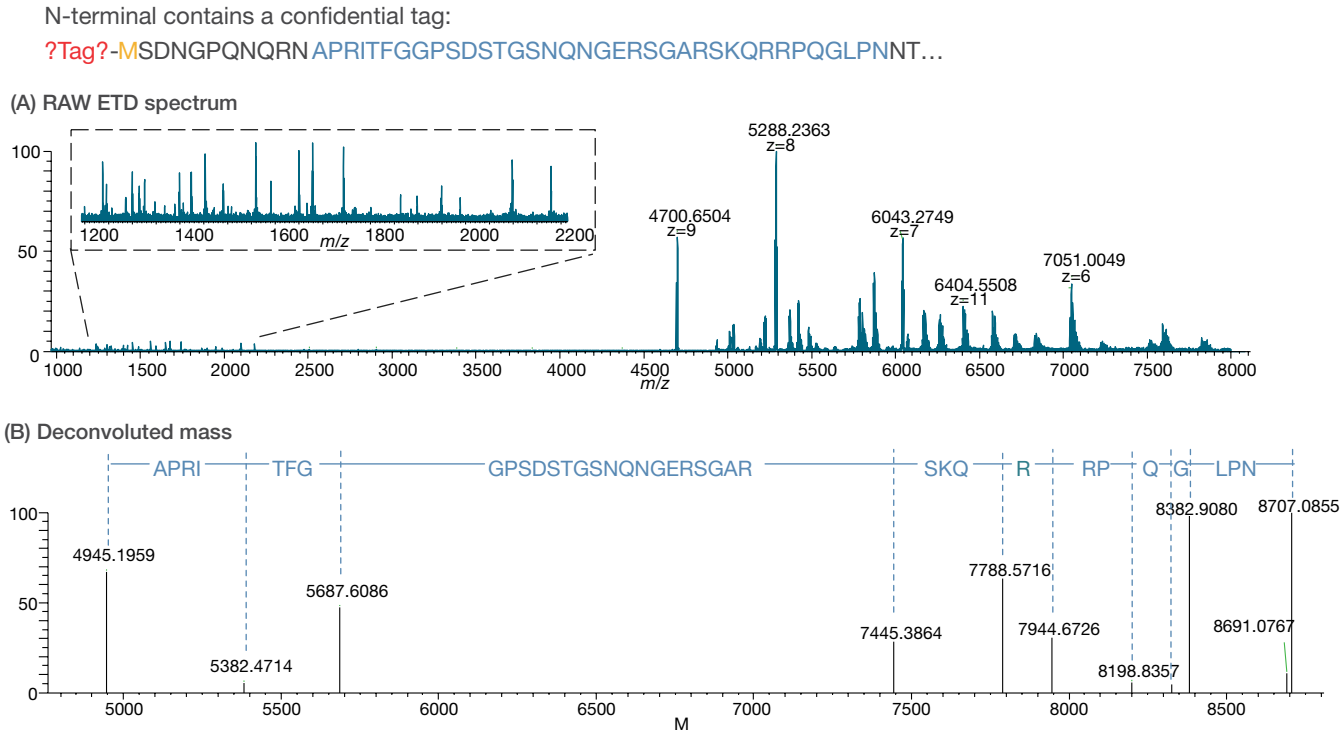

Figure 6. Denovo sequencing was used to identify molecular weight (MW) shift on the N-terminal due to an unknown tag

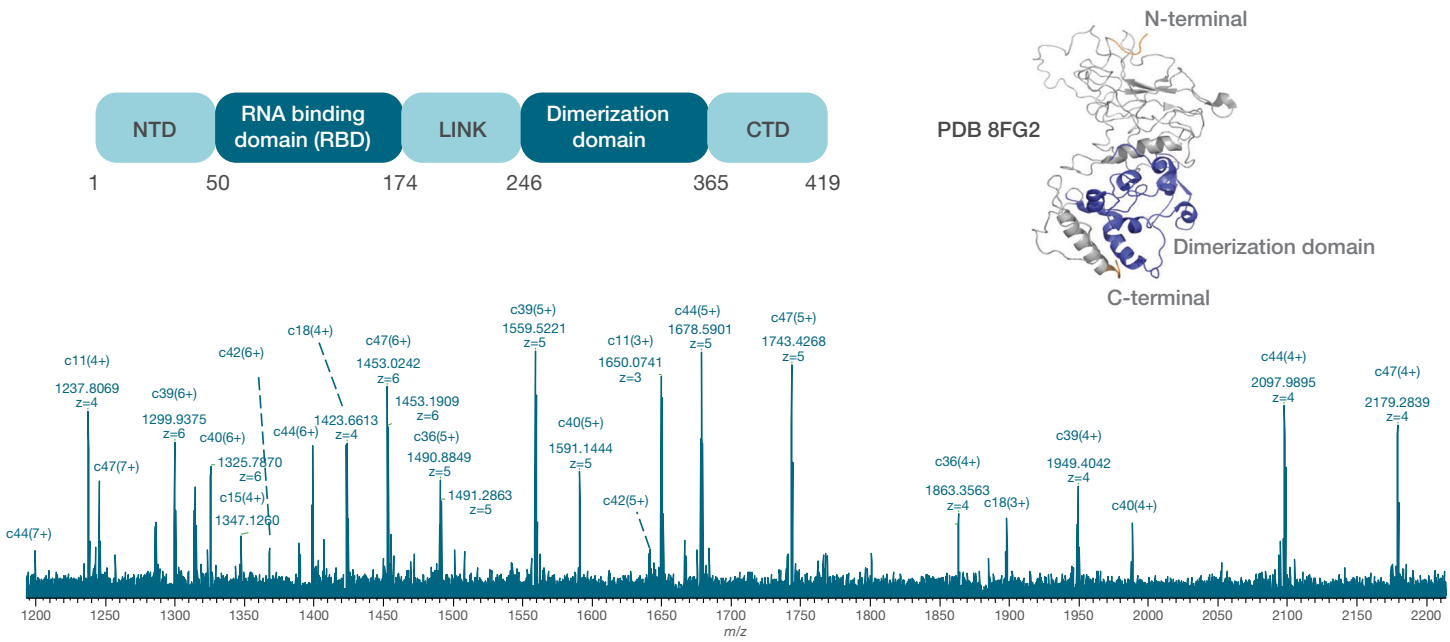

Figure 7. MS²-ETD spectrum of Nc reveals the exposed N-terminus and dimerization domain

C-terminal fragments from the top-down HCD experiment show characteristic patterns of $-H_2O$ and +Na shifts in accurate mass. These shifts reveal Na+ binding sites within the C-terminus, which is rich in basic residues (Figure 8). In summary, top-down analysis provided insights into the previously unknown N-terminal tag and indicated the dimerization domain at the C-terminus, but did not reveal any evidence of glycosylation.

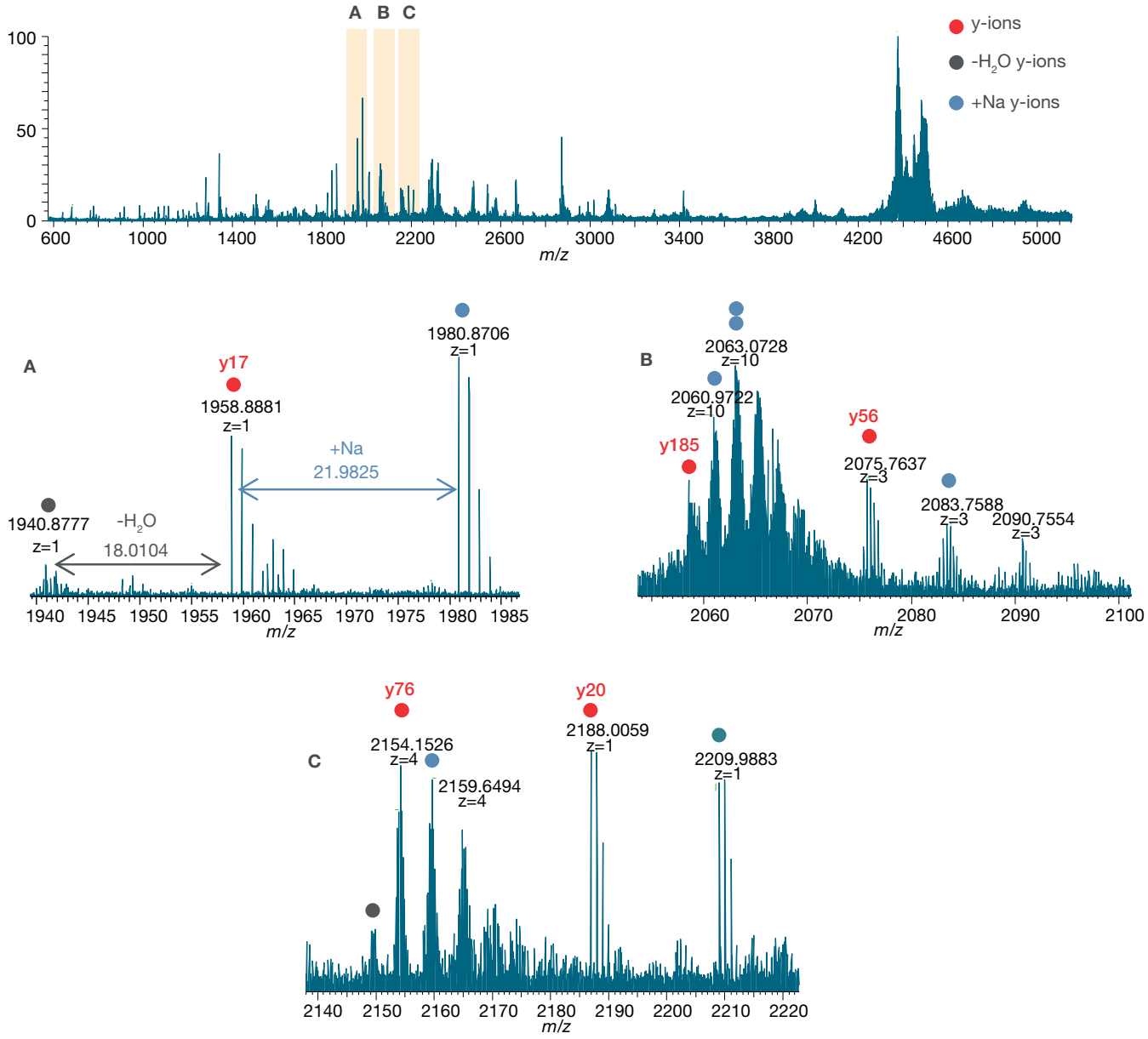

Figure 8. MS²-HCD spectrum of Nc discloses the Na^+ binding region

3. DIA-PTCR of PNGase F treated spike protein RBD to assign *O*-glycosylated proteoforms

The spike protein RBD, featuring a combination of *N*- and *O*-glycosylations, exhibited a broad molecular weight (MW) distribution centered around 31.8 kDa in DIA-PTCR analysis (Figure 9A). Cleavage of *N*-glycans using PNGase F shifted the MW center to 27.6 kDa, indicating the removal of *N*-glycans. A major mass loss of 4,216 Da corresponds to the removal of two *N*-glycans with the composition Hex(9)HexNAc(10)dHex(3) NeuAc(1). After *N*-glycan removal, the MW profile significantly simplified, allowing for the assignment of glycan compositions for major *O*-glycoforms (Figure 9B). Additionally, two proteoforms arising from sequence variants with mass difference of 70 Da (R37T and K21I) were identified. In summary, DIA-PTCR combined with accurate MW measurement at a resolution of 480,000 on *O*-linked RBD, enabled the assignment of *O*-glycan compositions.

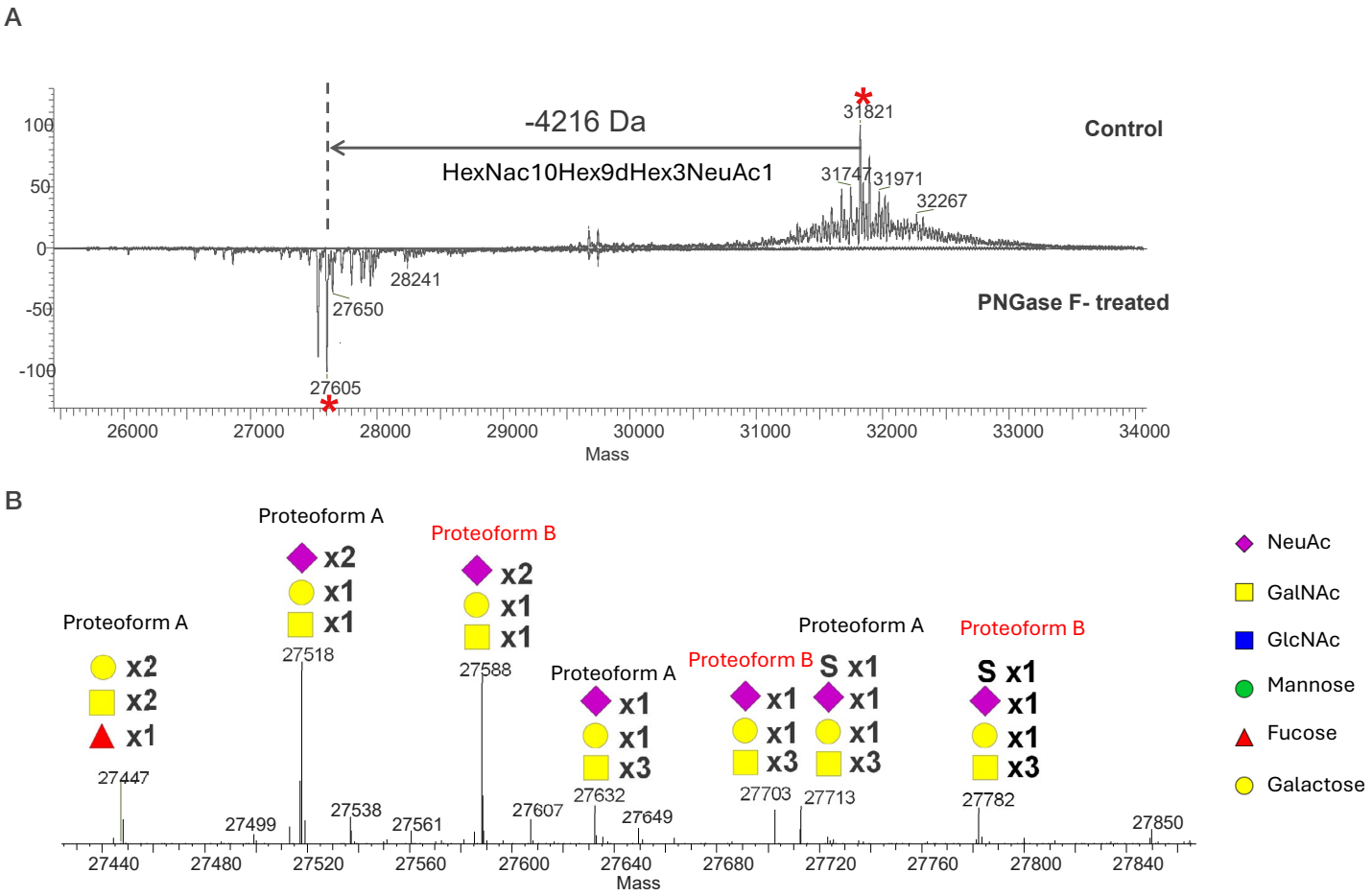

Figure 9. (A) DIA-PTCR shows comparison of RBD MW before and after *N*-deglycosylation; (B) examples of *O*-glycoforms assignment from DIA-PTCR analysis

4. DIA-PTCR analysis of SARS-Cov-2 S-protein

The full-length SARS-CoV-2 S-protein, the most complex glycoprotein in the SARS-CoV-2 family, plays a crucial role in the virus' function, infectivity, and immune evasion. Using the DIA-PTCR approach (5-Th quad isolation followed by 5-ms PTCR) easily resolves the unresolved charge states in full scan (Figures 10A and 10B). The deconvoluted spectrum in Figure 10C clearly shows a pattern of mass shift of 7,000–8,000 Da, indicating complexity of global glycosylation. DIA-PTCR facilitated easy determination of spike protein monomer MW (160–200 kDa) and revealed impurities at 70 kDa and 144 kDa (Figure 10D).

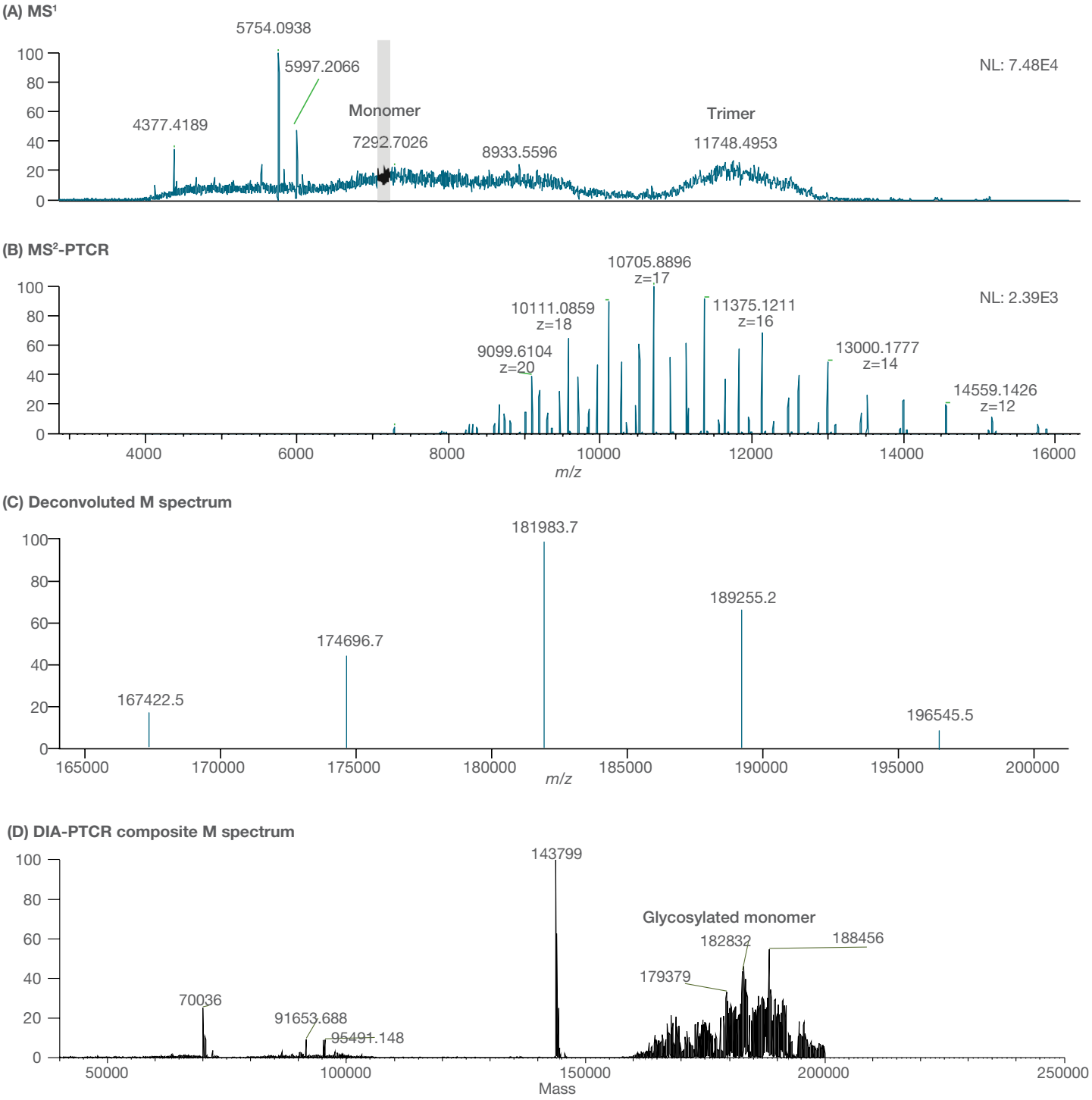

Figure 10. (A) A full MS¹ scan of spike protein only displays unresolved species; (B) MS² spectrum collected at 5-Th quadrupole isolation followed by 5-ms PTCR at *m/z* 7,280; (C) deconvolution of MS²-PTCR at *m/z* 7,280 (D) DIA-PTCR with 10-Th isolation and 5 ms reaction time shows monomer MW ranging from 160 to 200 kDa due to heavy glycosylation, along with some impurities

Conclusions

- The extended quadrupole isolation range, reaching up to *m/z* 8,000, enhances isolation accuracy and S/N compared to ion trap isolation, thereby advancing native omics studies, as it not only works for top-down but for DIA-PTCR as well.
- Combining high quadrupole isolation with gas-phase charge reduction facilitates the rapid assessment of the MW of complex glycoproteins.
- DIA-PTCR streamlines the complete MW measurement process for glycoproteins.
- Native top-down analysis not only identifies the sequence but also unveils structural insights into glycoproteins such as hFet and RBD.

References

1. Huguet, R., et al. Proton transfer charge reduction enables high-throughput top-down analysis of large proteoforms. *Analytical Chemistry* **2019,** *91*(24), 15732–15739.
2. Schachner, L.F., et al. Exposing the molecular heterogeneity of glycosylated biotherapeutics. *Nature Communications* **2024,** *15*(1), 3259.

Learn more at **thermofisher.com/nativems**