Analytica Chimica Acta 1297 (2024) 342349

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Mass spectrometry study on SARS-CoV-2 recombinant vaccine with comprehensive separation techniques to characterize complex heterogeneity

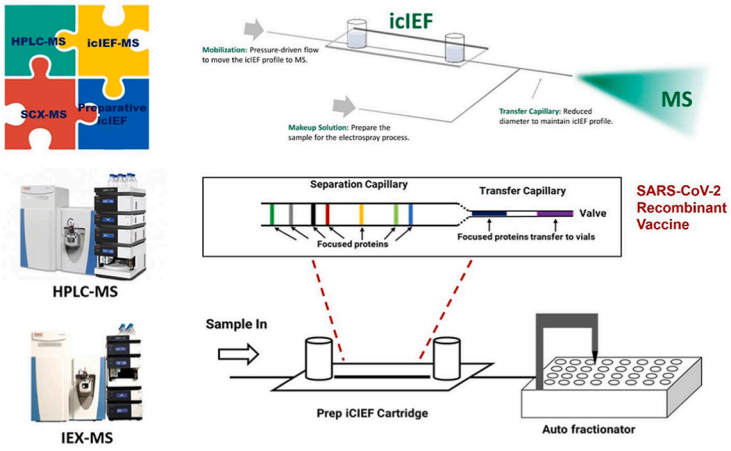
Gang Wu [a], Jialiang Du [c], Chuanfei Yu [c], Zhihao Fu [c], Xiaoxi Zhang [b], Lan Wang [c], Junzhi Wang [a,*]

[a] School of Life Science and Biopharmaceutics, Shenyang Pharmaceutical University, No. 103 Wenhua Road, Shenyang, Liaoning, 110016, China
[b] Thermo Fisher Scientific, A Building, Henggu1976, No.1976 Middle Gaoke Road, Pudong District, 201203, Shanghai, China
[c] Key Laboratory of the Ministry of Health for Research on Quality and Standardization of Biotech Products, National Institutes for Food and Drug Control, Daxing District, Beijing, 102629, China

HIGHLIGHTS

- Comprehensive separation techniques coupled to high-resolution mass spectrometry have been utilized for in-depth heterogeneity characterization of SARS-CoV-2 recombinant vaccine.
- As The cutting-edge technologies, icIEF-MS online and icIEF fractionation integrating IEX-MS were successfully explored to investigate their applicability in in-depth characterization of SARS-CoV-2 recombinant vaccine compositions.
- The constructed workflow underlined the outstanding importance of comprehensive platforms based on MS technology for the extremely complex heterogeneity of SARS-CoV-2 recombinant vaccine.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Dr. L. Liang

Keywords:
Imaged capillary isoelectric focusing
Ion exchange chromatography
High performance liquid chromatography
Mass spectrometry

ABSTRACT

SARS-CoV-2, the causative agent of COVID-19, has imposed a major public health threat, which needs effective therapeutics and vaccination strategies. Several potential candidate vaccines being rapidly developed are in clinical evaluation and recombinant vaccine has gained much attention thanks to its potential for greater response predictability, improved efficacy, rapid development and reduced side effects. Recombinant vaccines are designed and manufactured using bacterial, yeast cells or mammalian cells. A small piece of DNA is taken from the virus or bacterium against which we want to protect and inserted into the manufacturing cells. Due to

Abbreviations: icIEF, Imaged capillary isoelectric focusing; SCX, Strong cation exchange; IEX, Ion exchange chromatography; MS, Mass spectrometry; mAbs, Monoclonal antibodies; UHPLC, Ultra-high performance liquid chromatography; pI, Isoelectric point; TIC, Total ion current; AD, Acrylamide; MW, Molecular weight; PTM, Post-transitional modification.

* Corresponding author.
E-mail address: wangjz_nifdc2014@163.com (J. Wang).

https://doi.org/10.1016/j.aca.2024.342349
Received 8 December 2023; Received in revised form 1 February 2024; Accepted 4 February 2024
Available online 12 February 2024

Fractionation
SARS-CoV-2
Recombinant vaccine
Heterogeneity

the extremely complex heterogeneity of SARS-CoV-2 recombinant vaccine, single technology platform cannot achieve thorough and accurate characterization of such difficult proteins so integrating comprehensive technologies is essential. This study illustrates an innovative workflow employing multiple separation techniques tandem high-resolution mass spectrometry for comprehensive and in-depth characterization of SARS-CoV-2 recombinant vaccine, including ultra-high performance liquid chromatography (UHPLC), ion exchange chromatography (IEX) and imaged capillary isoelectric focusing (icIEF). The integrated methodology focuses on the importance of cutting-edge icIEF-MS online coupling and icIEF fractionation applied to revealing the heterogeneity secret of SARS-CoV-2 recombinant vaccine.

1. Introduction

A novel coronavirus (CoV) identified at the end of 2019 has been rapidly spread and therefore was declared as a public health emergency of international concern by the World Health Organization (WHO) on January 30, 2020 [1–3]. To date, the COVID-19 pandemic has resulted in a huge of confirmed infections, over several million deaths worldwide, and big numbers of cases increasing rapidly. Despite having a fatality rate lower than that recorded for SARS-CoV during the epidemic in 2003, SARS-CoV-2 is associated with severe respiratory malfunction and has a high mortality rate among the elderly and in individuals with chronic health issues, including diabetes and hypertension [4,5]. SARS-CoV has been found to share high levels of genomic similarity (spike glycoprotein sequence similarity 87.2%) to SARS-CoV-2 [6,7], providing a basis for the development of a vaccine. Efforts are currently underway to obtain further information on SARS-CoV-2 and its mechanism of rapid transmission to promote the development of effective control measures. Both the current pandemic and the threat of future epidemics underline a prerequisite for the development of precautionary strategies to fight coronaviruses (CoVs). In this context, vaccines represent an effective measure for the control of widespread viral infections associated with high morbidity and mortality. Research and development (R&D) organizations and institutes worldwide are currently using various vaccine platforms in attempts to develop a treatment against SARS-CoV-2 infection [8]. In the most simplified terms, a vaccine is a biological preparation that provides active acquired immunity against a certain disease. Usually, a vaccine consists of a biological agent that represents the disease-causing microorganism. It is often made from a weakened or killed form of the microorganism, its toxins or one of its surface protein antigens [9–11]. An individual that has been vaccinated produces antibodies against the protein antigen that protect him/her from contracting the disease upon attack from the pathogenic microorganism. In addition to live vectored and inactivated viruses, novel recombinant technologies are being used in the development of COVID-19 vaccine [12,13]. A recombinant vaccine is a vaccine produced through recombinant DNA technology. This involves inserting the DNA encoding an antigen (such as a bacterial surface protein) that stimulates an immune response into bacterial or mammalian cells, expressing the antigen in these cells and then purifying it from them. The advantage of recombinant vaccines is their greater response predictability and improved efficacy [14]. Thus, recombinant *S*-protein mimicking SARS-Cov-2 S protein can be used as an effective vaccine. SARS-CoV-2 enters human target cells via its viral transmembrane spike (S) protein. The spike protein is a trimeric class I fusion protein consisted of two subunits, namely S1 and S2, and the receptor-binding domain (RBD) is part of subunit S1 [15]. S protein plays the role of binding and membrane fusion in host cell receptor through the receptor binding domain (RBD), so a SARS-CoV-2 vaccine based on RBD of recombinant *S*-protein may induce the production of antibodies and virus neutralization efficiently. That's why the S protein is considered as a good candidate for vaccine development [16,17].

The development of a successful manufacturing process of recombinant vaccines requires their characterization in both upstream and downstream. Due to the extreme complexity of recombinant vaccines from the presence of numerous sialic acids, high diversities occur in terms of isoelectric points and complex glycosylation profiles including multiple *N*- and O-linked glycosylation sites. Hence, single analytical platform cannot achieve accurate characterization of such complex heterogeneities. Employing a portfolio containing comprehensive technologies is essential to address numerous properties (i.e., quality at tributes) for recombinant vaccines as well as for biosimilars with a special focus on so-called critical quality attributes (CQAs), which affect product safety and drug efficacy [18,19]. A recent publication reviewed the analytical strategies of complex protein characterization including mass spectrometry and chromatography [20]. Ultra-high Performance Liquid Chromatography (UHPLC),high-resolution tandem mass spectrometry (HRMS) and ion exchange chromatography coupled to mass spectrometry (IEX-MS) has been widely utilized for in-depth characterization of complex protein drugs [21–25]. Capillary electrophoresis (CE) is an essential electromigration technology for the characterization of complex proteins [26,29], which has important applications in characterizing recombinant proteins. Imaged capillary isoelectric focusing (icIEF) has become an indispensable tool in therapeutic protein development and manufacturing because of its high analytical throughput, ease of use, fast method development, and excellent reproducibility [27,28]. Recently, icIEF-MS has attracted much attention to utilize for protein charge variant analysis [30–39]. Our recent study demonstrated a straightforward and sensitive icIEF-MS platform for charge heterogeneity characterization of nine therapeutic mAbs [40]. A preparative icIEF workflow was illustrated to fractionate protein charge variants that were characterized by UHPLC-MS at the levels of intact and peptide mapping [21,41], which greatly simplified the sample complexity with more accurate in-depth characterization of protein heterogeneity.

There are already multiple publications on recombinant SARS-COV-2 S protein characterization [15,16,42–46]. However, these proteins were all expressed in human cell lines and the sample used in this study was Chinese Hamster Ovary (CHO) expressed trimeric RBDs. Different expression systems and production process conditions will bring heterogeneity to recombinant proteins, especially for this protein containing 657 amino acids and multiple N and O glycosylation sites [15,16,42–46].

In this study, an MS-based workflow integrating comprehensive separation techniques was developed for the complex heterogeneity characterization of CHO cell expressed recombinant trimeric SARS-CoV-2 RBD vaccine for the first time. UHPLC-MS analysis at the levels of native intact MS was carried out and UHPLC-MS/MS technology was utilized for in-depth characterization of SARS-CoV-2 recombinant vaccine including peptide mapping and post-translational modification (PTM) identification. IcIEF and icIEF-MS were utilizing for fingerprinting the subtle heterogeneity of SARS-CoV-2 recombinant vaccine, and high-resolution carrier ampholytes were used for high-efficiently separating its extremely complex compositions. Meanwhile, the results were consistently verified by the PTMs from UHPLC-MS characterization. Based on our latest study of the charge variant of therapeutic mAbs [40], in this study icIEF-MS was further developed for extremely complex protein characterization such as SARS-CoV-2 recombinant vaccine with high resolution and sensitive detection. To simplify the complexity of SARS-CoV-2 recombinant vaccine, preparative icIEF recently developed in our study [21,41], was firstly employed to fractionate the charge

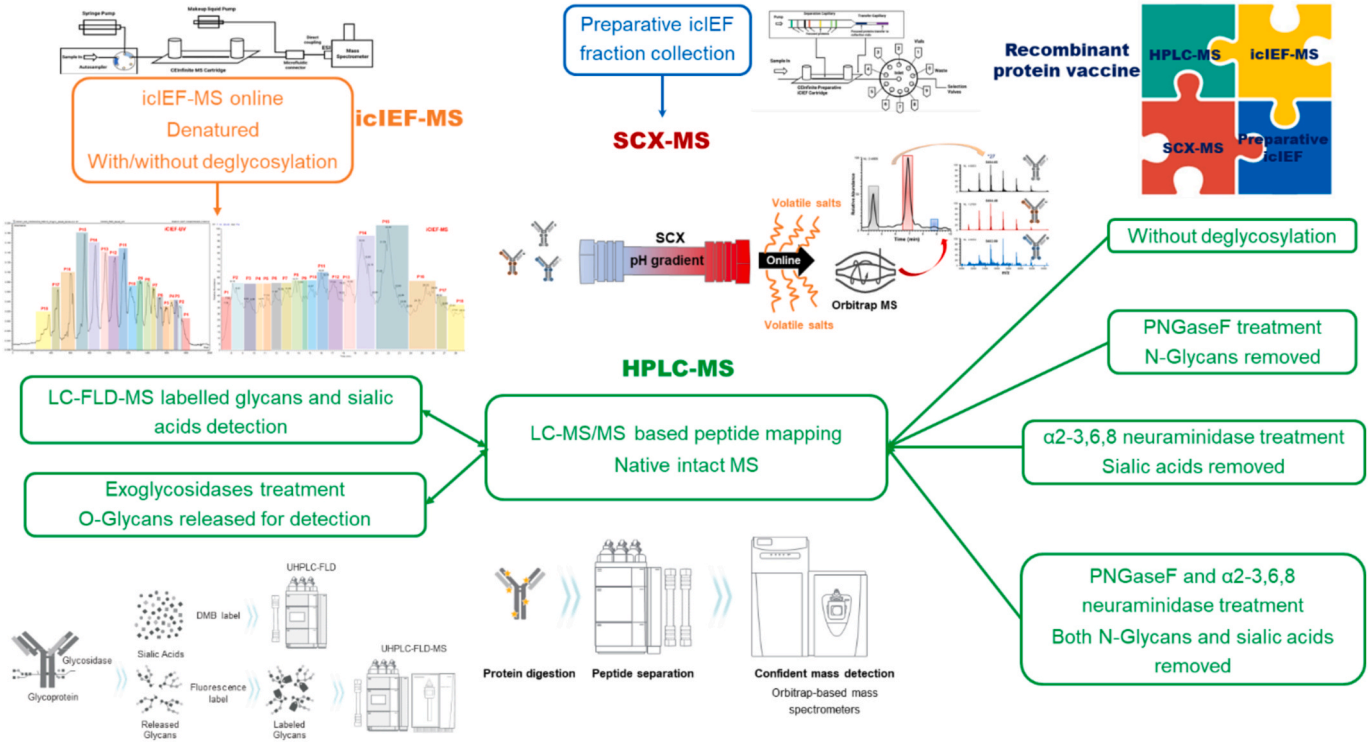


Fig. 1. MS based workflow integrating comprehensively chromatographic techniques with various separation mechanisms for characterizing recombinant vaccine heterogeneity.

variants of such a complex vaccine and then IEX-MS was used for intact-level characterization of collected fractions. During the whole workflow, different sample preparation approaches were optimized for the native and denatured MS identification. The established workflow was comprehensive and innovatively integrated multiple separation techniques coupled to high resolution MS for extremely complex SARS-CoV-2 recombinant vaccine. In addition, as new technologies, icIEF-MS and icIEF-based fractionation further contributed to dissecting extremely complex protein system with deeper understanding of complex charge heterogeneity.

2. Experimental

2.1. Chemicals and enzymes

All ampholytes (AESlytes) were obtained from Advanced Electrophoresis Solutions Ltd (AES, Cambridge, Ontario, Canada). The SARS-CoV-2 recombinant vaccine studied in this work was produced in the CHO cell expression system [42,43]. Mass spectrometry grade acetonitrile (ACN), formic acid, acetic acid and 1 M Tris-HCl solution were purchased from Fisher Scientific. Ammonium bicarbonate, 1 M ammonium hydroxide solution, 8 M guanidine hydrochloride solution, bio-ultra grade DL-Dithiothreitol (DTT) and sodium iodoacetate were purchased from Sigma-Aldrich. Sequence grade trypsin was purchased from Promega. PNGaseF and α2-3,6,8 neuraminidase were purchased from New England Biolabs.

2.2. Parameters for icIEF - MS

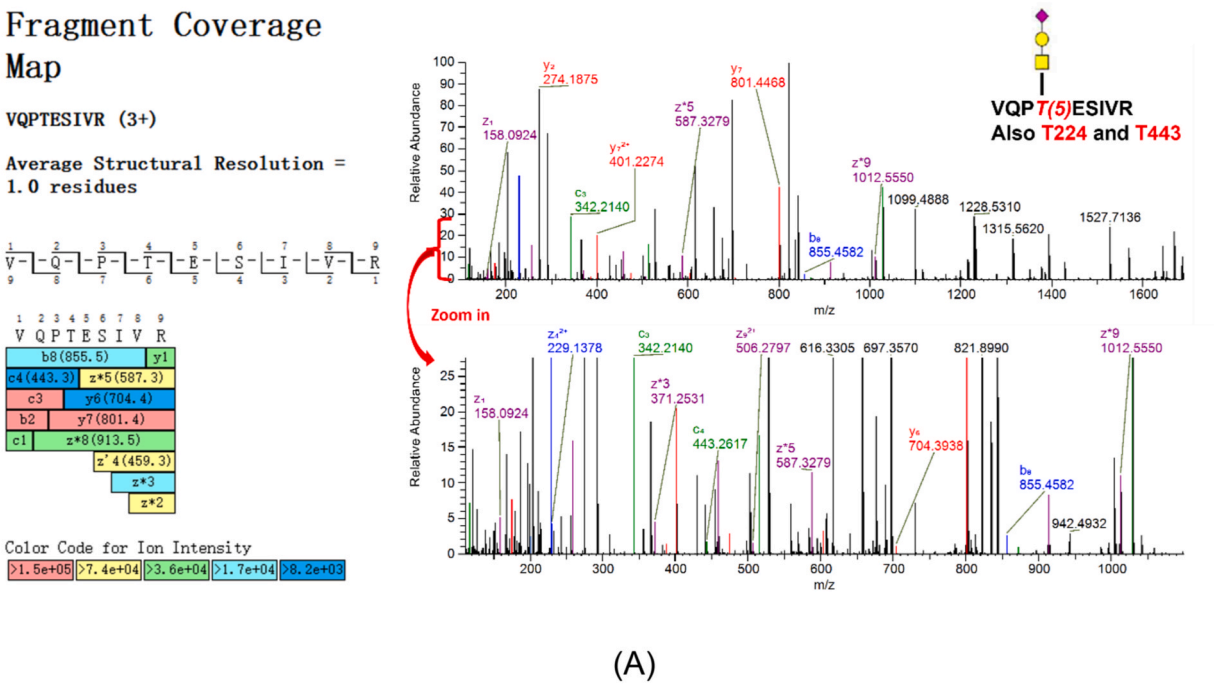
For icIEF separation, the CEInfinite icIEF (Advanced Electrophoresis Solutions Ltd, Cambridge, Canada) was utilized with on-column UV detection at 280 nm. The sample was desalted and concentrated to ~10 mg/mL in ddH_2O using 3 k ultrafiltration Spin Column (Milipore). Final protein concentration is 2 mg/mL in ddH_2O with 0.5%HR3-10, 2.5% HR8.5–9.5 and 1.0%HR9-12 carrier ampholytes (AES). 200 μm ID acrylamide derivative coated (AD) capillary cartridge with 5 cm separation length and micro-tee integrated (AES, cat. No. CP00303 M) were used for icIEF-MS. For icIEF separation, the focusing was performed using 1 min at 1000 V, 1 min at 2000 V and 10 min at 3000 V. For icIEF-MS coupling, 3000 V was applied during mobilisation of focused protein bands; the mobilisation speed was 80 nL/min with water containing 0.1% (*v/v*) formic acid, across the separation capillary, and 6 μL/min make-up solution (water: ACN = 1:1, *v/v*, containing 0.1% formic acid, v/v) added through a micro tee. Mobilisation time was 60 min.

As for icIEF fractionation, 320 μm ID AD coated cartridges (AES, cat. No. CP00307) are used for preparative icIEF with a 5 cm long separation capillary and 50 μm ID transfer capillary assembled. Final protein concentration is 3 mg/mL with 0.5%HR3-10, 2.5%HR8-10.5 and 1.0%HR9-12 carrier ampholytes (AES). After the focusing phase (1 min at 1000 V, 1 min at 2000 V and 10 min at 3000 V), 3000 V was applied during mobilisation; the mobilisation speed was 80 nL/min with water containing 0.1% (v/v) formic acid, across the separation capillary, and 1.6 μL/min make-up solution (ddH_2O) added through a micro tee. Mobilisation time was 50 min.

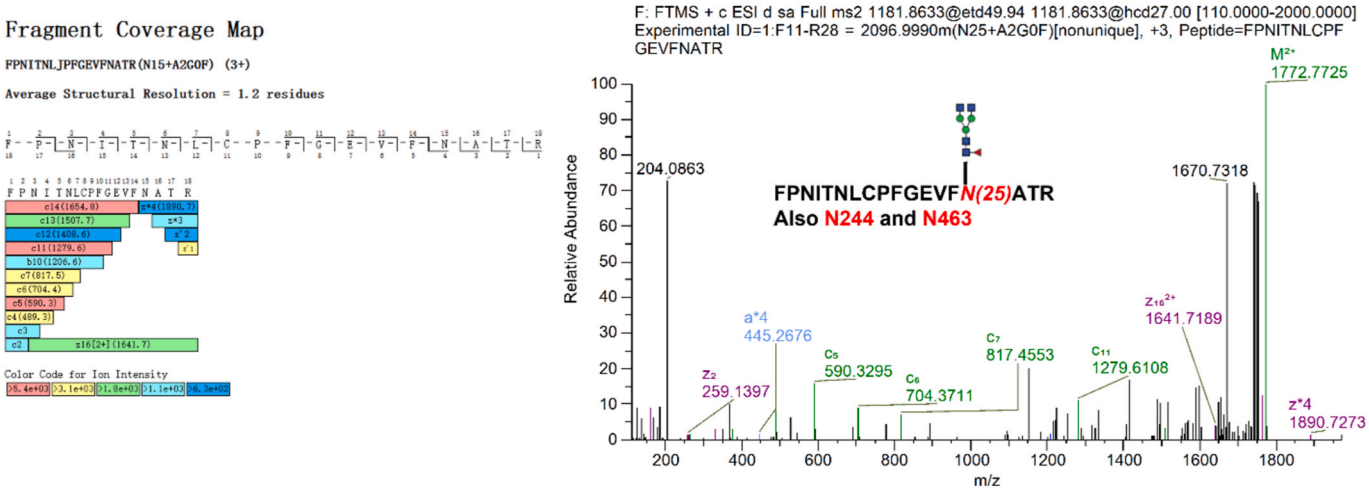
A Thermo Q Exactive™ Plus mass spectrometer equipped with Biopharma option, an Ion Max ESI Ion Source and a 34-gauge needle (Thermo Fisher Scientific, Bremen, Germany) was used for mass measurement. The spray voltage: 3.8 kV, sheath gas: 8 L/min, auxiliary gas 0 L/min, *S*-lens RF 60, capillary temperature 275 °C, vaporizer temperature 100 °C, resolution 35,000@*m/z* 200, scan range of precursor ion 1500–5000 *m/z*, in source CID value 50 eV and maximum injection time 200 m.

2.3. Parameters for SCX – MS

For SCX separation, a Thermo Scientific™ Vanquish™ UHPLC system was employed with mobile phases composed of relative combinations of A (25 mM ammonium bicarbonate and pH 5.3 adjusted by acetic acid) and B (10 mM ammonia, pH 10.9). A Thermo Scientific™ MAb-Pac™ SCX-10 R S column (5 μm, 2.1 × 50mm) was used for the



(A)

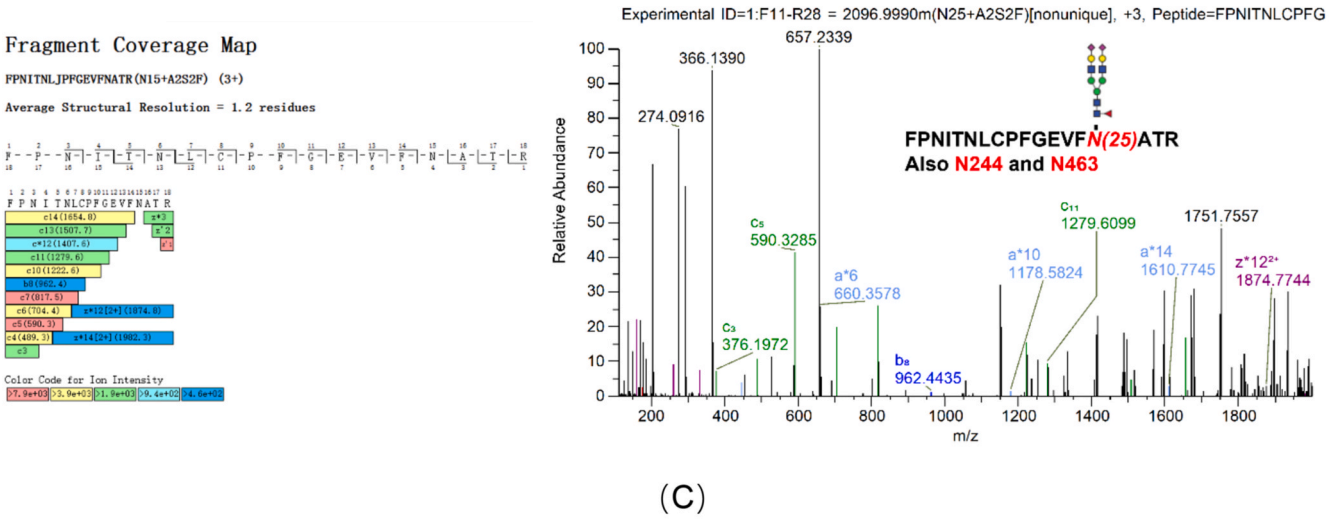


(B)

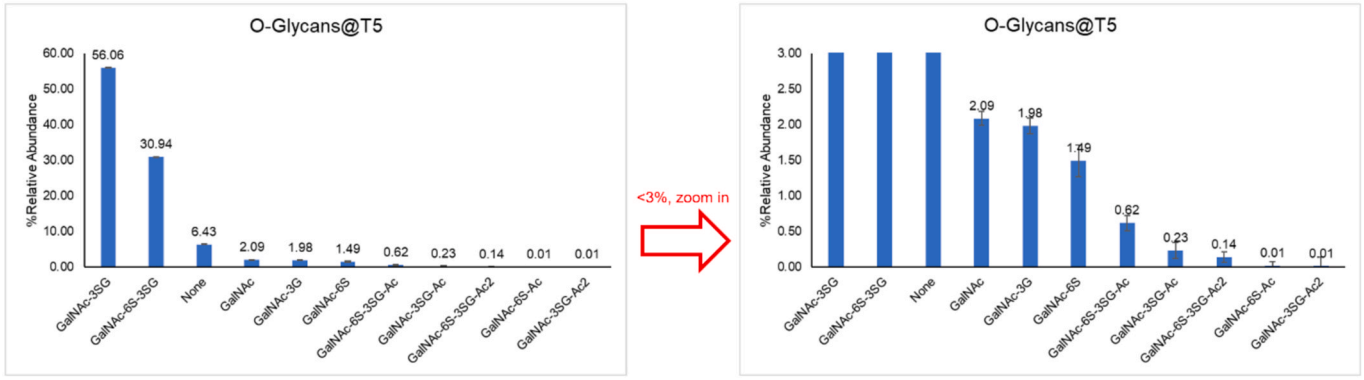
Fig. 2. HPLC-MS for in-depth characterization of SARS-CoV-2 recombinant vaccine heterogeneity. (A) EThcD MS2 spectra of peptide VQPTESIVR, with an O-glycan GalNAc-3SG at T5; (B) EThcD MS2 spectra of peptide FPNITNLCPFGEVFNATR, with an *N*-glycan FA2 at N25; (C) EThcD MS2 spectra of peptide FPNITNLCPF-GEVFNATR, with an *N*-glycan FA2G2S2 at N25; (D) EThcD MS2 identified O-glycans and abundance% of peptide VQPT (5)ESIVR; (E) EThcD MS2 identified *N*-glycans and abundance% of peptide FPN(13)ITNLCPFGEVFN(25)ATR; (F) The comparison between EThcD peptide mapping and released glycans approach. MS PM, trypsin digested sample followed by LC-MS/MS based peptide EThcD MS2 mapping at Orbitrap Fusion Lumos; HPLC: O-glycans were released and labelled using commercial kit, followed by UHPLC-FLD-MS.

separation. The gradient elution programs are 0–10min, 40–100%B, 10–18min, 100%B, 18.1–20min, 0%B, 20–20.1min, 0–40%B. Flowing rate was set at 0.3 mL/min with 30 °C column temperature, 280 nm UV wavelength and 20 μg sample injection. A Thermo Orbitrap Exploris™ 240 Mass Spectrometer equipped with Biopharma option, an Ion Max ESI Ion Source and a 34-gauge needle (Thermo Fisher Scientific, Bremen, Germany) was used for mass measurement. The spray voltage: 3.6 kV, sheath gas: 45 L/min, auxiliary gas 15 L/min, *S*-lens RF 100, capillary temperature 250 °C, vaporizer temperature 150 °C, resolution 30,000@*m/z* 200, scan range of precursor ion 1000–8000 *m/z*, in source CID value 75eV and maximum injection time 100 m.

Sample digestion conditions are indicated in Supplementary data 1. All UHPLC-MS based native intact MS and UHPLC-MS/MS peptide mapping conditions are indicated in Supplementary data 2. Released N/O linked glycans and sialic acid labelling and experimental conditions are indicated in Supplementary data 3.



(C)



(D)

Fig. 2. (*continued*).

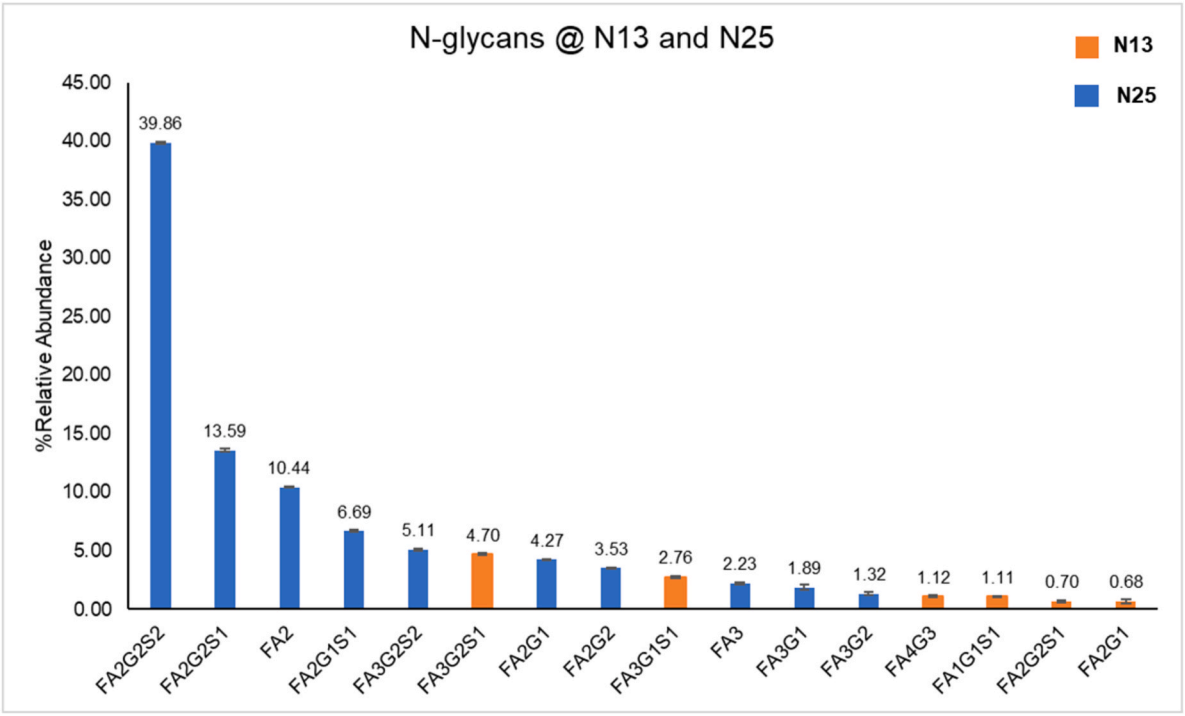
Biopharma Finder software (version 5.1) from Thermo Fisher was used for data analysis. The glycan database used for identifying glycome and glycoforms is the software built-in version with a CHO host cell-line type.

3. Results and discussion

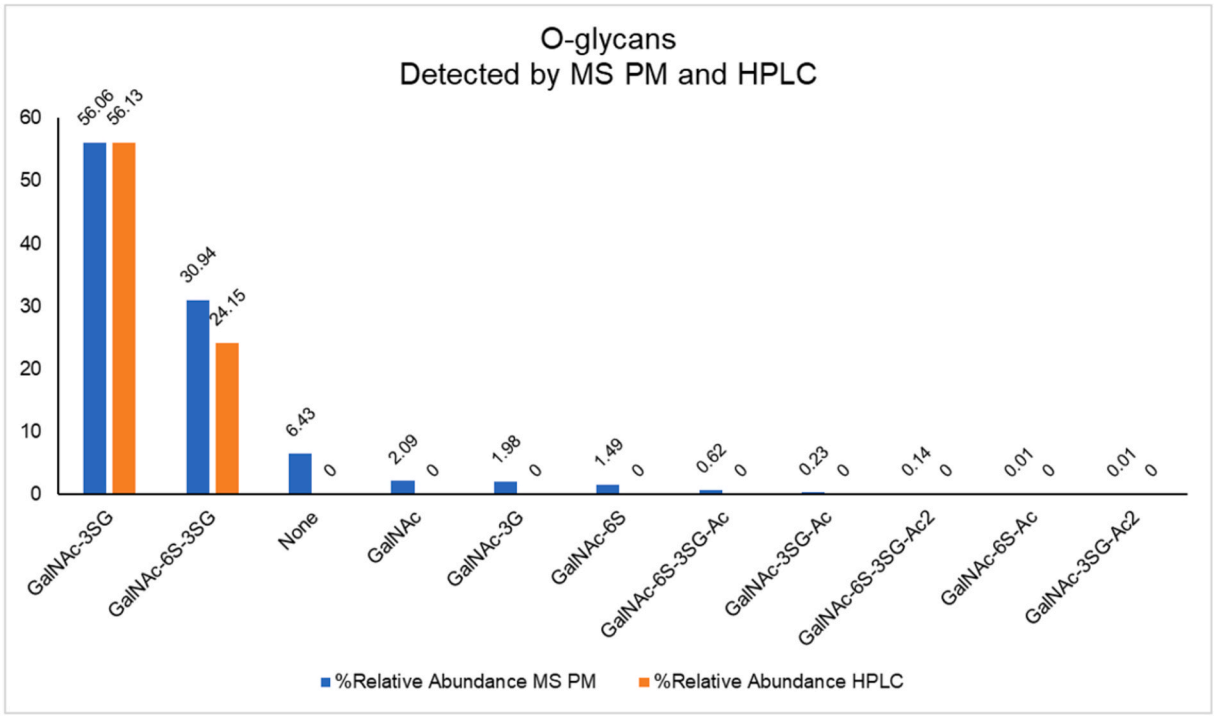
Fig. 1 illustrates a MS-based workflow integrating comprehensively chromatographic techniques with various separation mechanisms for characterizing recombinant vaccine heterogeneity, by which obtained results can be compared and cross validated for more comprehensive and accurate insights into the complexity of recombinant vaccine heterogeneity.

3.1. HPLC-MS for in-depth characterization of SARS-cov-2 recombinant vaccine heterogeneity

As discussed in previous section, the studied protein has multiple potential PTMs, such as *N*- and O- glycosylation, deamidation and oxidation [15,16,42–46]. Deamidation is a common degradation of proteins and it can significantly impact protein structure and function. At neutral and basic pH, deamidation proceeds via the formation of a five-membered ring intermediate succinimide [47]. Oxidation is a common protein modification, mainly occurring on methionine, cysteine, histidine, and tryptophan residues under the treatment of *t*-butyl hydroperoxide (tBHP), hydrogen peroxide, ultraviolet light, and high-temperature conditions. After the oxidation of these sites, the hydrophobicity of the product decreased [47]. Since these PTMs may induce quality or safety issues of the vaccine [16], they are usually defined as product quality attributes (PQAs). To ensure the desired product quality, identification and monitoring of PQAs are essential for biotherapeutic product development, manufacturing, and quality control [48]. Therefore, LC-MS based in-depth characterization of this complex glycoprotein is necessary. First, both reduced and non-reduced trypsin digested peptide mapping data were acquired at a Thermo Scientific Vanquish UHPLC system tandem with a Thermo Scientific Q Exactive Plus mass spectrometer. The sequence coverage of reduced digestion sample is 89.5% since high glycan heterogeneity made it very hard for the identification using HCD fragmentation. However, PTMs other than glycosylation, such as deamidation and oxidation can be identified and relatively quantified. This recombinant protein is consisted of three repeated sequences so each modification appears three



(E)



(F)

Fig. 2. (*continued*).

times in the protein. At W118/337/556, both oxidation (0.42%) and double oxidation (0.32%) were identified. Four deamidation sites were identified based on peptide mapping results. For N130/319/568, its abundance of deamidation is 0.57%. The %deamidation of N214/433/652 is 1.49%, a little bit higher than 1%. However, 10.74% deamidation was observed for N163/382/601, while the %deamidation of N218/437/656 is 66.5%. As described above, these PQAs should be monitored during the whole life cycle of the biotherapeutic due to their

Table 1
O-glycan identification and quantification results comparison: peptide mapping *vs*. UHPLC-FLD-MS.

Mod	%Relative Abundance MS peptide mapping	%Relative Abundance HPLC-FLD-MS	Monoisotopic mass (Da)
GalNAc-3SG	56.06	56.13	656.228
GalNAc-6S-3 S G	30.94	24.15	947.323
None	6.43	–	0
GalNAc	2.09	–	203.079
GalNAc-3G	1.98	–	365.132
GalNAc-6S	1.49	–	494.175
GalNAc-6S-3 S G-Ac	0.62	–	989.333
GalNAc-3SG-Ac	0.23	–	698.238
GalNAc-6S-3 S G-Ac2	0.14	–	1031.343
GalNAc-6S–Ac	0.01	–	536.185
GalNAc-3SG-Ac2	0.01	–	740.248

MS peptide mapping: identified and quantified by EThcD MS2 peptide mapping at Fusion Lumos.
UHPLC-FLD-MS: released glycans assay at UHPLC-FLD-MS.

potential efforts on product quality, efficacy and safety [48]. In our study, higher %deamidation of N163/382/601 and N218/437/656 should be noticed since deamidation can affect protein structure and function [47,48]. The HCD fragmentation-based peptide mapping result provides important information for the following intact protein analysis, especially the glycoforms determination. Also, the combination of reduced and non-reduced trypsin digestion results contribute to identifying three disulfide bonds (C61–C114, C73–C207, C162–C170) with high confidence MS/MS spectra, which is essential to stabilize right protein with higher order structure [17].

From previous publications [15,16,42–46], two *N*-glycosylation and two potential O-glycosylation sites at RBD region were reported. However, the samples analyzed in these publications were human cell line expressed. The trimeric RBDs used here was expressed in CHO cell. *N*-glycosylation can affect the immunogenicity, potency, antibody-dependent cell-mediated cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), serum clearance, and pharmacokinetics of the biotherapeutics [47]. Therefore, the *N*-glycosylation must be closely monitored and controlled to ensure product efficacy and safety. It is well known that the glycoforms can be quite different in different express cell line, and it is necessary to use multiple techniques for glycan and glycosylation sites identification in our samples. A Thermo Fisher Scientific Orbitrap Fusion Lumos mass spectrometer was employed for glycan and glycosylation sites identification at peptide mapping level, using the HCD trigger EThcD MS2 method. It can provide high quality EThcD MS2 spectra for both glycan type identification and modification sites determination. Samples treated with or without PNGaseF followed by denatured trypsin digestion were used for the characterization of O-linked glycans and *N*-linked glycans, respectively. Using this approach, due to sufficient fragmentation MS information of glycopeptides, sequence coverage of trypsin digested only sample achieved 97.7% (without PNGaseF treatment) and 99.5% (with PNGaseF treatment) for O-glycopeptides contained samples. Uncovered regions are those with continuous K and R that generated short peptides by trypsin digestion, which are not suitable for MS2 identification. However, it can be solved by using other enzymes. Fig. 2A–C shows EThcD MS2 spectra of three glycopeptides, respectively. Fig. 2D and E displays UHPLC-MS based N/O-linked glycans identification and quantification results. EThcD fragmentation can keep the glycans at side chain while breaking the peptide bonds so the peptide sequence, glycan type and modification sites can be identified at the same time. One O-glycosylation site T5 was identified and it is repeated two more times (T224 and T443) in the whole sequence. From the peptide mapping quantification results, it is easy to find out that the most abundant O-glycan is GalNAc-3SG (+656.228Da, monoisotopic mass) with 56.06% calculated with the use of peak areas of modified/unmodified peptides (Supplementary data 4). The second is GalNAc-6S-3 S G (+947.323Da, monoisotopic mass) and relative abundance is 30.94%. None modified form ranks the third (6.43%). This is unique information the released glycan approach cannot provide. For other glycoforms, they are all below 3%. We also compared LC-MS based peptide mapping results with released N/O-glycans identification and quantification results by UHPLC-FLD-MS. For O-glycans, it is clearly that HCD trigger EThcD MS2 peptide mapping at Orbitrap Fusion Lumos can identify and quantify more glycans than released glycans method, and peptide mapping method can provide glycosylation sites information which released glycans method cannot. The comparison results in Fig. 2F and Table 1 shows high consistency for top glycoforms. For example, relative abundance of GalNAc-3SG is 56.06% in peptide mapping results and 56.13% in release glycan results. Two *N*-glycosylation sites were identified in peptide mapping result, one is N13 (also N232 and N451) and the other is N25 (also N244 and 463). Therefore, 6 *N*-glycosylation sites were identified in total. The most abundant *N*-glycan is FA2G2S2 at N25 (39.86%) and top 5 *N*-glycans are all at N25. Except FA2 (third place, 10.44%), others are all highly sialic acid contained type (FA2G2S1, 13.59%; FA2G1S1, 6.69% and FA3G2S2, 5.11%). The top *N*-glycan at N13 is FA3G2S1 (4.70%). The *N*-glycan profile of this sample is quite different from previous publication [15,16,42–46]. This may be caused by difference in expressed cell lines, cell passaging, medium selection and protein purification process, etc. For released *N*-glycans, it is hard to quantify because of non-baseline separation at chromatography level (Supplementary data 4) but the most abundant glycoforms are consistent with the peptide mapping results. Released sialic acids quantification was achieved by 1,2-diamino-4,5-methylenedioxybenzene-2HCl (DMB) dye labelling following UHPLC-FLD. On average, every molecule carries ~9 Neu5Ac and 0.09 Neu5Gc (Supplementary data 4). As reported [15], O-glycans are involved in protein stability and function. Some viruses can use host-sialylated structures for binding to their target host cell [15,16] and that is why it is essential to watch highly sialylated glycans in this sample. All these detailed *N*- and O-glycan identification and relative quantification results provide sufficient information for product efficacy and safety assessment.

All of these characterization results, especially the glycosylation parts, contribute to essential information for intact MS data elucidation.

Multiple N/O glycosylation sites and dozens of glycoforms make intact protein analysis of this glycoprotein quite challenging. Intact MS was studied at first but the spectra was extremely complex only providing rather limited information (data not shown). Then glycosidases were utilized to remove glycans to reduce complexity. PNGaseF was employed to remove *N*-linked glycans. Fig. 3A and B shows the native intact MS and deconvoluted spectrum of PNGaseF treated protein. The components distributed from 75,000Da to 77,000Da show combination of different O-linked glycans at three O-glycosylation sites in the sequence. The complexity of deconvoluted spectrum is significantly reduced and it helps us to better understand O-glycan distributions at all three sites at the same time, which can only be observed at intact level. For components greater than 78,000Da, the *N*-linked glycans weren't completely removed. Fig. 3C and D shows native intact MS and deconvoluted spectra of α2-3,6,8 neuraminidase treated proteins. Part of sialic acids were removed but due to *N*-glycosylation, the spectrum is still complex. PNGaseF followed by α2-3,6,8 neuraminidase treatment was used to remove both *N*-linked glycans and sialic acids. Fig. 3E and F shows the native intact MS and deconvoluted spectrum of protein after this treatment, and for the most abundant components, only O-glycan cores were left. The percentage of O-glycans at intact level are consistent with peptide mapping and released glycan results (Fig. 3G and Table 2).

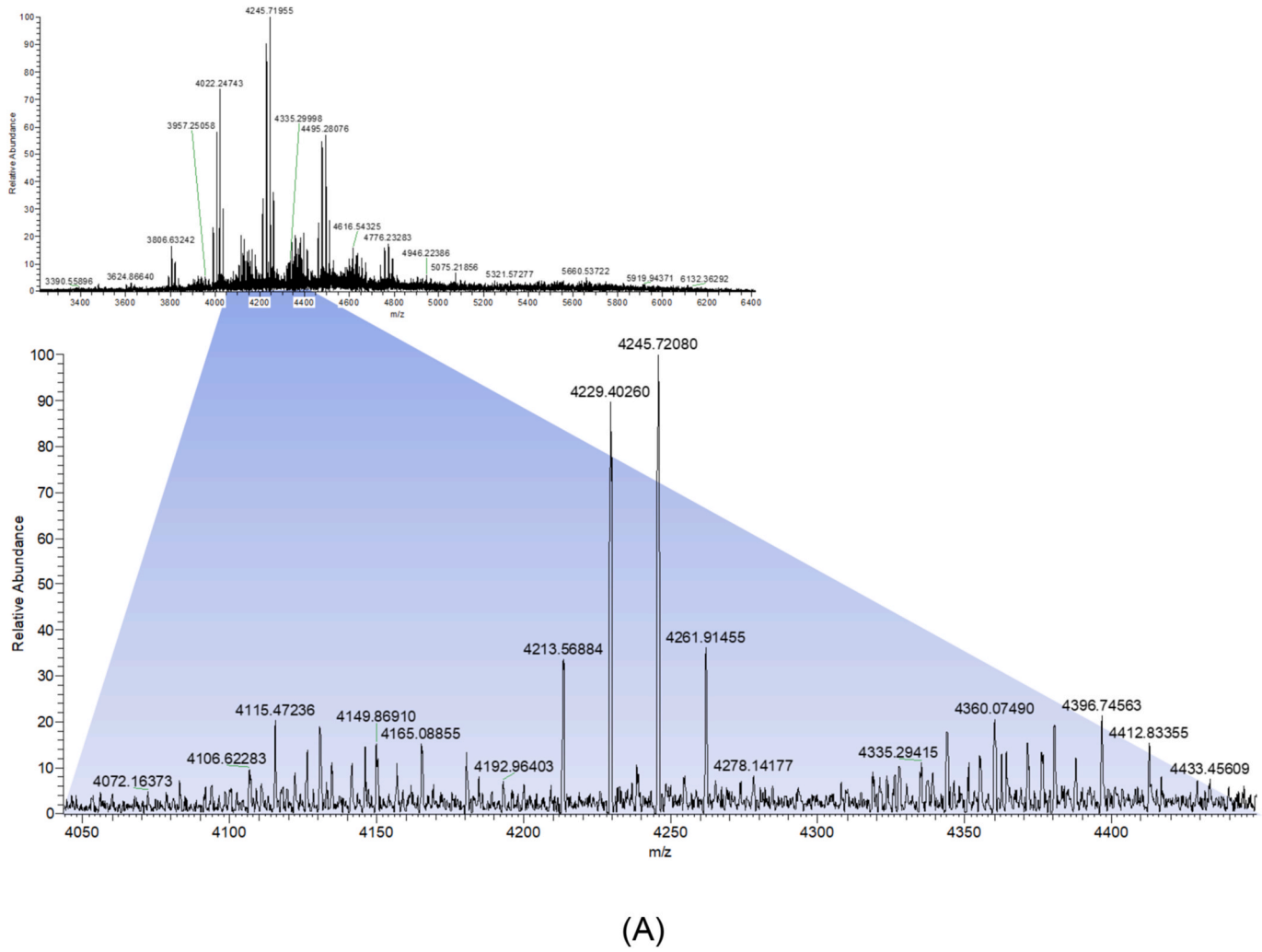


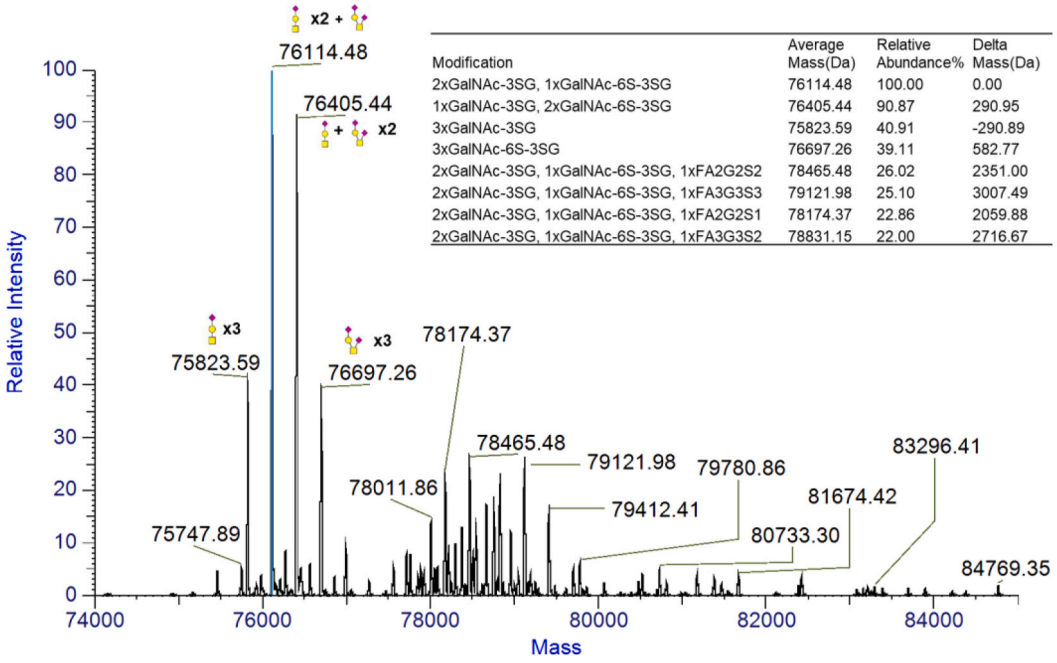
Fig. 3. Native intact MS identification of glycoforms with glycosidases treated sample. (A) Native intact MS spectrum of PNGaseF treated sample (Resolution = 70,000); (B) deconvoluted spectrum of PNGaseF treated sample. Insert table, glycoforms of components>20%; (C) Native intact MS spectrum of α2-3,6,8 neuraminidase treated sample (Resolution = 70,000); (D) Deconvoluted spectrum of α2-3,6,8 neuraminidase treated sample. Insert table, glycoforms of top 5 components; (E)Native intact MS spectrum of sample treated with PNGaseF followed by α2-3,6,8 neuraminidase (Resolution = 70,000); (F) Deconvoluted spectrum of sample treated with PNGaseF followed by α2-3,6,8 neuraminidase; (G) Comparison of O-glycans identification and quantification results using different methods. Intact level is based on native intact MS results, peptide mapping level is based on EThcD MS2 results and released glycan level is from UHPLC-FLD-MS results.

3.2. icIEF and icIEF-MS for characterization of SARS-cov-2 recombinant vaccine heterogeneity

Although glycosidase treatment can help to reduce MS spectra complexity, considering the effect of glycans on protein stability, bioactivity, and immunogenicity, it is necessary to use multiple analytical platforms to achieve accurate characterization of such complex heterogeneities. In this study, icIEF- high resolution MS tandem platform was employed to intact protein analysis.

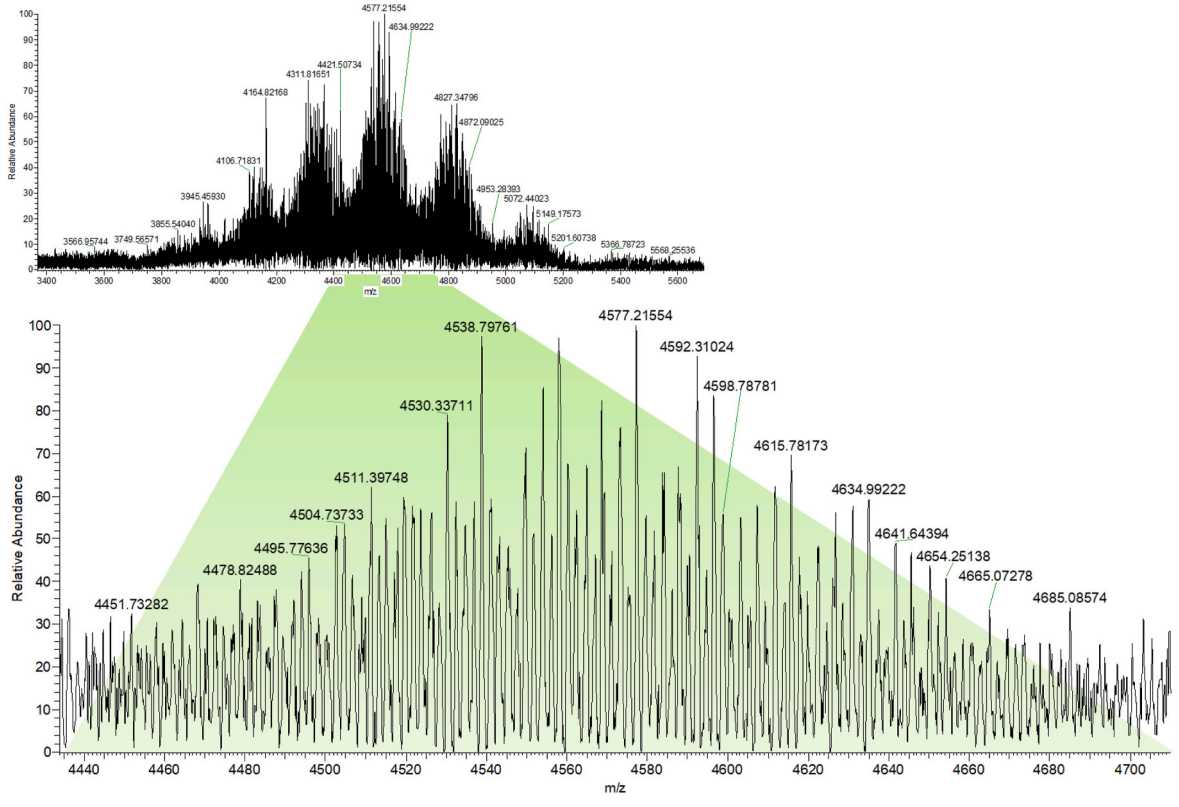
Protein sample was desalted and prepared for icIEF-MS following conditions described in Section 2.2. Since the molecular weight, pI and post-translational modifications of this protein is quite different from mAbs, we optimized the composition of carrier ampholytes by mixing three high-resolution carrier ampholytes together. These three different carrier ampholytes cover different pI ranges respectively, which can provide better separation of charge variants at icIEF level compares to using single carrier ampholytes. Also, the parameters of MS, such as scan range and dissolvent fragmentation energy were optimized for full MS data acquisition. Fig. 4A and B shows the icIEF- UV and icIEF-MS profile, respectively. Due to high glycosylation heterogeneity, 18 peaks were identified using icIEF platform. Corresponding MS peaks were labelled in Fig. 4B. The icIEF separation in front of mass spectrometry can significantly reduce sample heterogeneity and MS spectra were also simplified. Fig. 4C shows raw MS spectra of peak 1 and peak 2 with the significant difference observed. Fig. 4D illustrates deconvoluted spectra mirror plot of these two peaks and it is clear that the molecular distribution is different. From basic peak to acidic peak (P1 to P18), the molecular weight of detected components and sialic acid% demonstrate increasing trend. In peak 1 to 4, for most components, 5 of 6 *N*-glycosylation sites are occupied, the dominant O-glycans are 2xGalNAc-3SG+1xGalNAc-6S-3 S G, and 3xGalNAc-3SG starts to appear in peak 3. As for the peak 5 to 8, 5 of 6 *N*-glycosylation sites are occupied but fully occupied forms cannot be observed. The dominant O-glycans in peak 5 to 8 are still 2xGalNAc-3SG+1xGalNAc-6S-3 S G but the percentage of 3xGalNAc-3SG grows up when moving to acidic side. As for peak 9 to 18, all 6 *N*-glycosylation sites fully occupied by the highly sialic acid, and the typical and dominant O-glycans are 3xGalNAc-3SG. Top 5 components in each peak are shown in Supplementary data 5.

Compared to mAb, the glycosylation heterogeneity of studied protein is much higher, and one way to reduce analytical difficulty is to remove



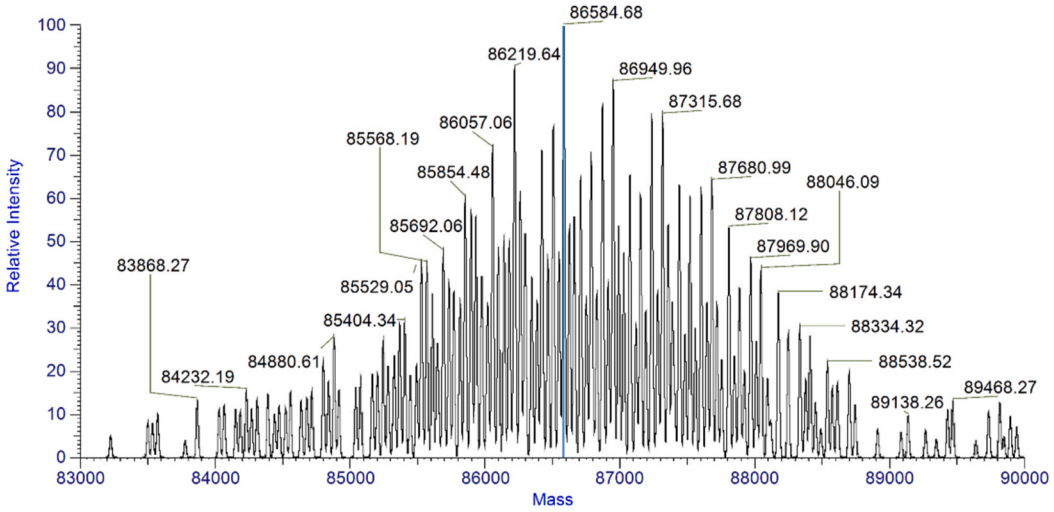
Modification	Average Mass(Da)	Relative Abundance%	Delta Mass(Da)
2xGalNAc-3SG, 1xGalNAc-6S-3SG	76114.48	100.00	0.00
1xGalNAc-3SG, 2xGalNAc-6S-3SG	76405.44	90.87	290.95
3xGalNAc-3SG	75823.59	40.91	-290.89
3xGalNAc-6S-3SG	76697.26	39.11	582.77
2xGalNAc-3SG, 1xGalNAc-6S-3SG, 1xFA2G2S2	78465.48	26.02	2351.00
2xGalNAc-3SG, 1xGalNAc-6S-3SG, 1xFA3G3S3	79121.98	25.10	3007.49
2xGalNAc-3SG, 1xGalNAc-6S-3SG, 1xFA2G2S1	78174.37	22.86	2059.88
2xGalNAc-3SG, 1xGalNAc-6S-3SG, 1xFA3G3S2	78831.15	22.00	2716.67

(B)



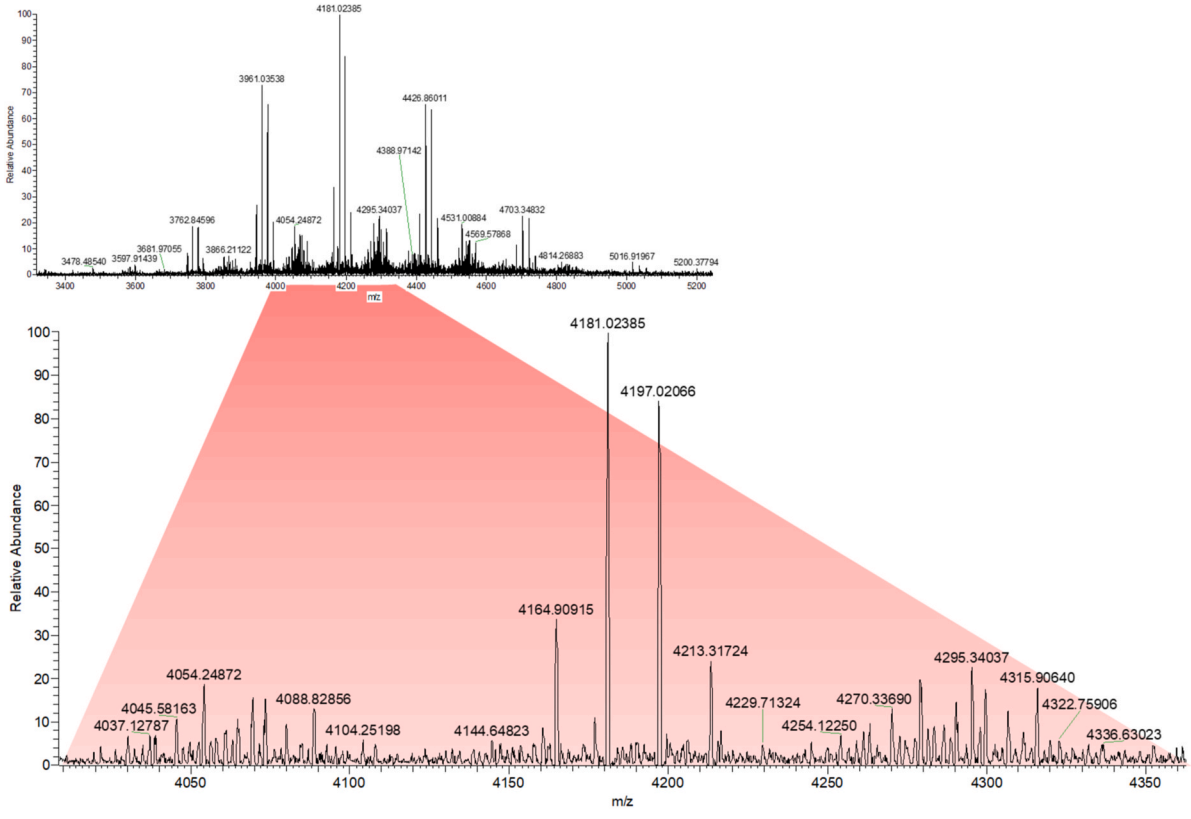
(C)

Fig. 3. (*continued*).



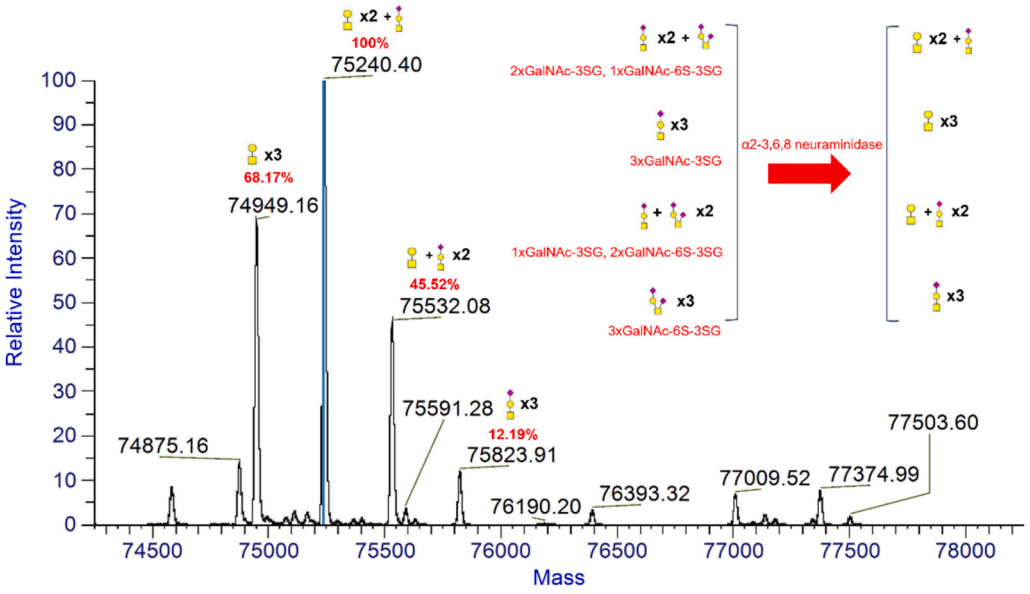
Modification	Average Mass(Da)	Relative Abundance(%)	Delta Mass(Da)
3xGalNAc-3SG,1xA2S1G0F, 2xA2S1G1F, 1xA3S2G1F, 2xA1G0F	86584.68	100.00	0.00
3xGalNAc-3SG,1xA2S1G0F, 1xA2S1G1F, 1xA3S2G1F, 3xA1G0F	86219.64	90.21	-365.04
3xGalNAc-3SG,1xA2S1G0F, 1xA2S1G1F, 1xA3S2G1F, 2xA2G1F, 1xA1G0F	86949.96	86.93	365.28
2xGalNAc-3SG, 1xGalNAc-6S-3SG, 3xA2S1G0F, 1xA2S1G1F, 1xA3S3F	86874.30	81.80	289.62
3xGalNAc-3SG,1xA2S1G0F, 2xA2S1G1F, 1xA3S2G1F, 2xA1G0F	87315.68	79.31	731.00

(D)

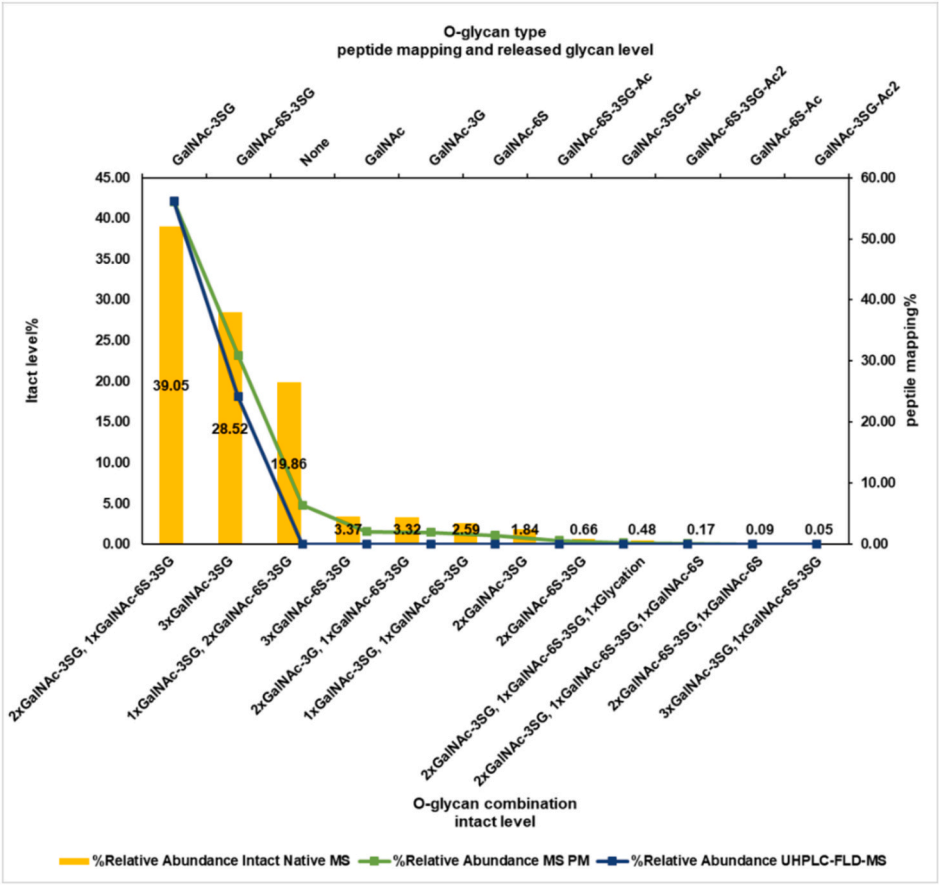


(E)

Fig. 3. (*continued*).



(F)



(G)

Fig. 3. (*continued*).

Table 2
O-glycan identification and quantification results comparison using different methods.

Modification	%Relative Abundance MS PM	%Relative Abundance UHPLC-FLD-MS	O-glycan combination	%Relative Abundance Intact Native MS
GalNAc-3SG	56.06	56.13	2xGalNAc-3SG, 1xGalNAc-6S-3 S G	39.05
GalNAc-6S-3 S G	30.94	24.15	3xGalNAc-3SG	28.52
None	6.43	0	1xGalNAc-3SG, 2xGalNAc-6S-3 S G	19.86
GalNAc	2.09	0	3xGalNAc-6S-3 S G	3.37
GalNAc-3G	1.98	0	2xGalNAc-3G, 1xGalNAc-6S-3 S G	3.32
GalNAc-6S	1.49	0	1xGalNAc-3SG, 1xGalNAc-6S-3 S G	2.59
GalNAc-6S-3 S G-Ac	0.62	0	2xGalNAc-3SG	1.84
GalNAc-3SG-Ac	0.23	0	2xGalNAc-6S-3 S G	0.66
GalNAc-6S-3 S G-Ac2	0.14	0	2xGalNAc-3SG, 1xGalNAc-6S-3 S G,1xGlycation	0.48
GalNAc-6S–Ac	0.01	0	2xGalNAc-3SG, 1xGalNAc-6S-3 S G,1xGalNAc-6S	0.17
GalNAc-3SG-Ac2	0.01	0	2xGalNAc-6S-3 S G,1xGalNAc-6S	0.09
			3xGalNAc-3SG,1xGalNAc-6S-3 S G	0.05

MS PM: identified and quantified by EThcD MS2 peptide mapping at Fusion Lumos.
UHPLC-FLD-MS: released glycans assay at UHPLC-FLD-MS.
Intact Native MS: PNGaseF treated sample analyzed using SEC-MS.

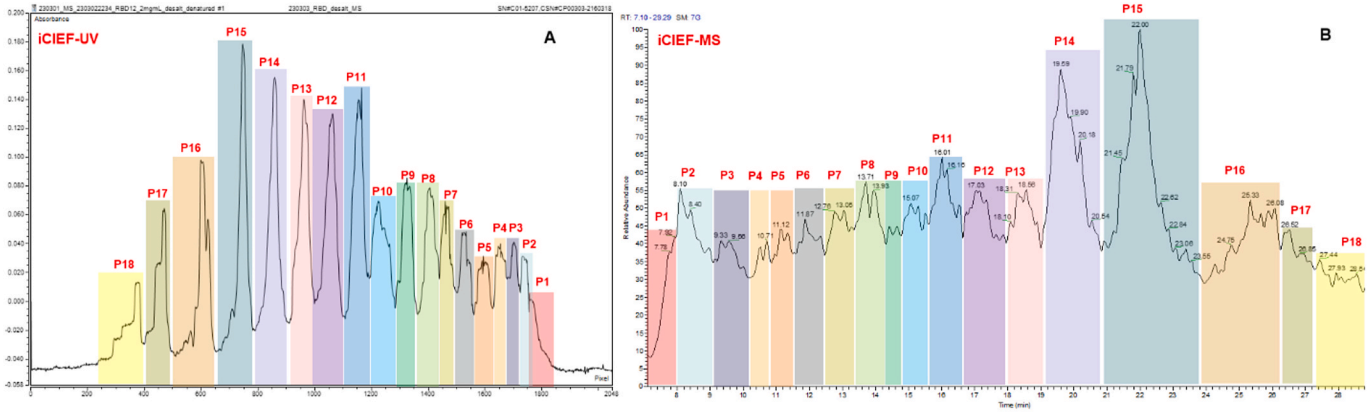
glycans. First, the sample was treated by PNGaseF using the established icIEF-MS workflow. Fig. 4E and F demonstrate the icIEF- UV and icIEF-MS profiles of PNGaseF treated sample, respectively. Although *N*-glycans were removed, due to high O-glycosylation heterogeneity, 14 peaks were identified using icIEF platform. Fig. 4G illustrates raw MS spectra of peak 5, peak 7 and peak 10 as examples. Fig. 4H is deconvoluted spectra of these three peaks. Fig. 4I displays O-glycans distribution trend across all peaks in this PNGaseF treated sample. The *N*-glycans were not completely removed from major components in peak1-3. From peak 4 to 10, *N*-glycans removed components are dominated. In peak 4 and 5, two/three two/three *O*-glycosylated species are co-existed. From peak 6, three O-glycosylation sites are fully occupied. It is worth noting that multiple deamidation can be detected in sample before/after PNGaseF treatment. This is consistent with our peptide mapping results in Session 3.1. One benefit of using high-resolution carrier ampholytes is that even for such a complex sample, the charge variants can be separated based on their pI value, which makes the following MS data interpretation much easier. These differences in molecular weight distribution are in agreement with glycan distribution. Also, more sialic acids can be observed in acidic peaks than basic peaks.

3.3. *icIEF fractionation integrating IEX-MS for SARS-cov-2 recombinant vaccine heterogeneity*

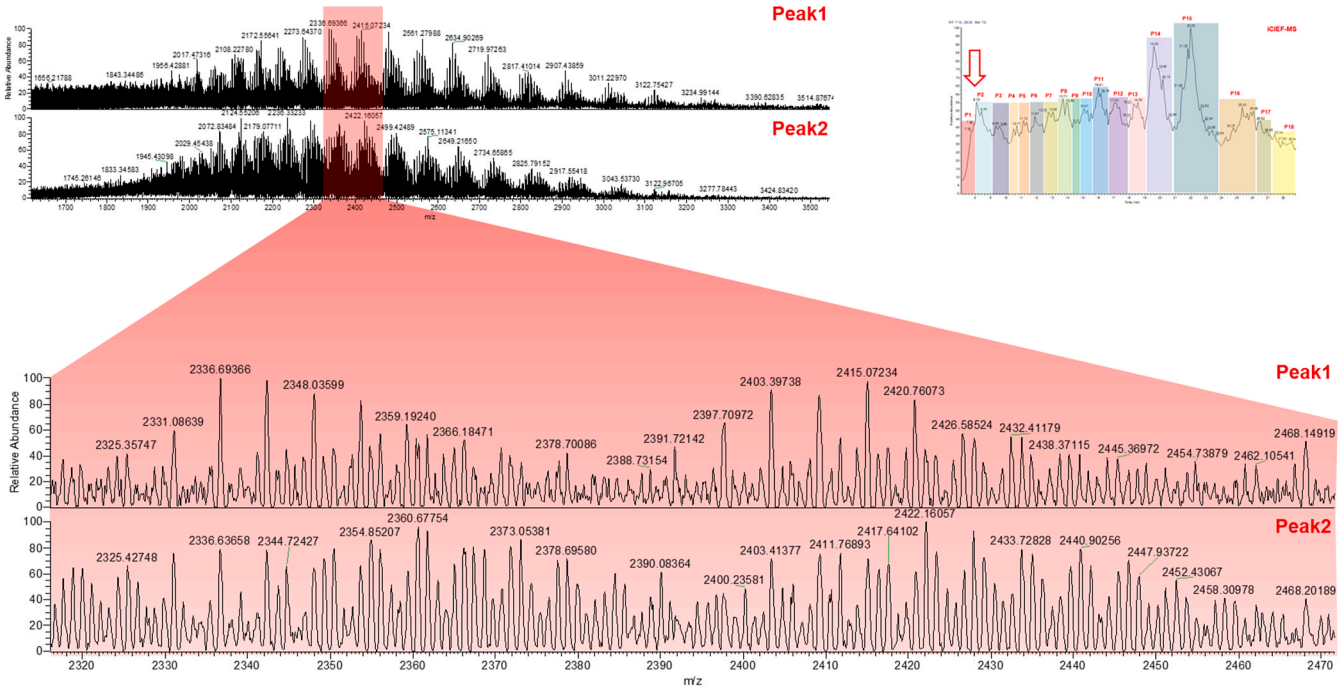
The CE infinite icIEF platform enables fractionation [21,41]. In this study, all 18 peaks were fractionated into 5 groups based on their pI values and collected fractions from each group were for offline SCX-MS characterization. Fig. 5A and B shows icIEF-UV profile of trimeric RBD and 5 fractions. Differences among fractions can be observed. Fig. 5C and D are BPC profiles and full MS spectra of 5 fractions, analyzed by SCX-MS. The MS spectra as illustrated in Fig. 5E shows significant difference of various fractions collected by preparative icIEF and indicates the separation effects of icIEF fractionation on simplifying sample complexity. Fig. 5E demonstrates the deconvolution results of these 5 fractions and corresponding icIEF-MS online peaks are labelled for each fraction in Fig. 5E. Supplementary data 6 lists top 5 components in each fraction analyzed by icIEF-MS and SCX-MS, respectively. Identified peaks from 1 to 4 using icIEF-MS online workflow are missed in the SCX-MS identification, and a possible explanation is icIEF-MS online has better sensitivity than SCX-MS due to icIEF sample precontraction and nano-flow rate mobilisation in icIEF-MS, which is consistent with our study before [40]. The separation mechanism of icIEF and IEX is different from each other. Although these two methods are both broadly used in biotherapeutics charge variant analysis, by combining them together, more complementary information can be collected.

3.4. *Prospect of icIEF-MS and icIEF fractionation in complex protein heterogeneity*

IcIEF technology is a regulated method in biopharma development, quality control and manufacture release in biopharma industry. Although imaged capillary isoelectric focusing (icIEF) has got wide recognition for protein heterogeneity, biopharma industry is in an era of great change in biopharmaceutical analysis with a thirst for new solutions to address big challenges in protein characterization, including complex protein characterization, new mass spectrometry (MS) tandem technology and preparation of protein charge variants. However, for a long term, icIEF-UV has been the single mode and what to frustrate biopharma industry is how to extend icIEF capability to MS online coupling and fractionation purpose with more straightforward strategy. Otherwise, icIEF analysis cannot build an effective bridge with traditional HPLC-MS and IEX-MS, especially for extremely complex proteins with complicated PTMs including recombinant vaccine and fusion protein. In addition, routine carrier ampholytes cannot solve complex protein sample with satisfactory resolution in icIEF separation. Recently, four latest review articles highlight and evaluate the important developments of icIEF technology including high-resolution ampholytes, MS online and fractionation of protein charge variants in biopharma industry [49–52]. The reviews introduce the solutions and workflows we have recently developed in terms of icIEF-MS and fractionation. Especially, Prof. Karger in his review insights into the perspectives on capillary electrophoresis in the Trends in Analytical Chemistry by citing 37 latest publications with important influences [51], and highlights that “Isoelectric focusing, as we have noted, is an important protein characterization method and is widely used in CE. However, the coupling of CE and MS has been limited in part because of interference by the carrier ampholytes. There is a clear need for on-line coupling of CIEF and MS as isolation of specific CIEF peaks for MS analysis can take days to complete. There is active work going on direct cIEF-MS coupling and even MS coupling to imaged CIEF, generally using low levels of ampholytes. We can anticipate increasing use of cIEF-MS systems.“. Promisingly, icIEF-MS technology makes the protein heterogeneity

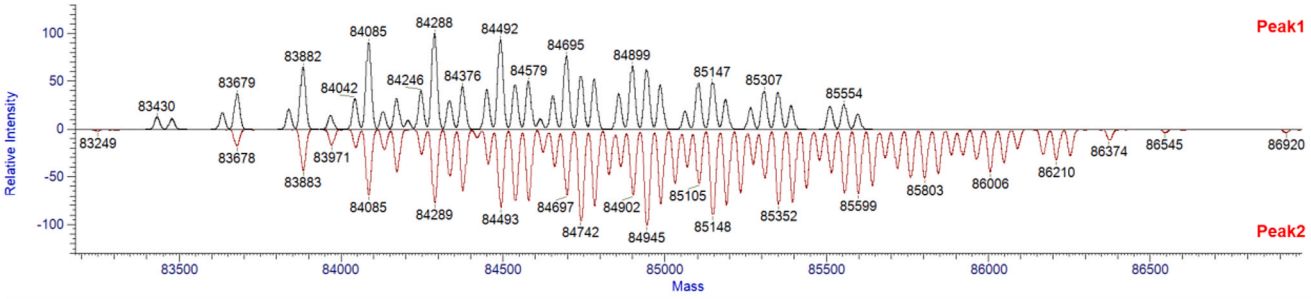


(A) and (B)



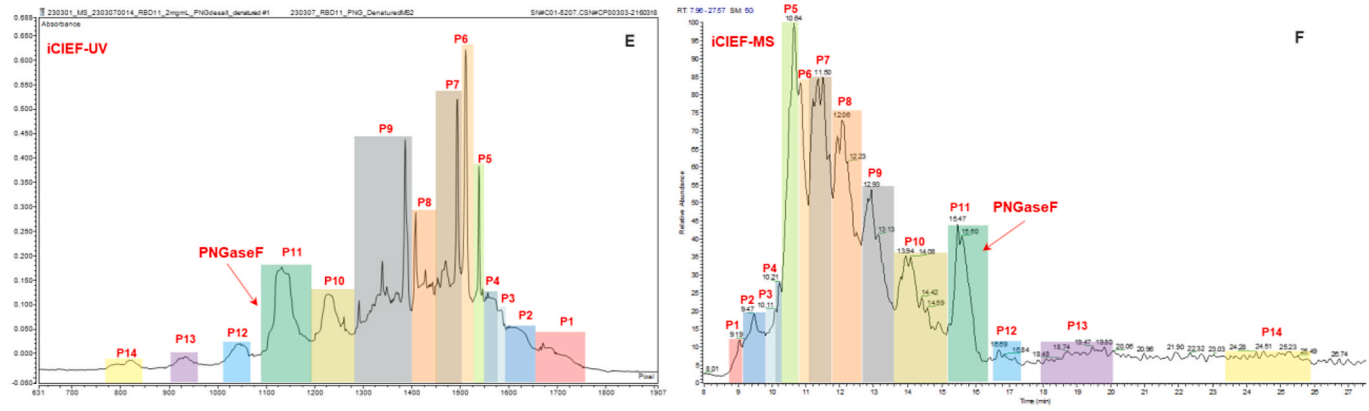
(C)

Fig. 4. IcIEF and icIEF-MS results. (A) IcIEF-UV profile of trimeric RBD; (B) icIEF-MS profile of trimeric RBD; (C) Raw spectra of peak1 and peak2; (D) Deconvoluted spectra mirror plot of peak1 and peak2. Insert table, glycoforms of top 5 components in each peak; (E) icIEF-UV profile of PNGaseF treated trimeric RBD; (F) icIEF-MS profile of PNGaseF treated trimeric RBD; (G) Raw spectra of peak5, peak7 and peak10 from PNGaseF treated trimeric RBD. (H) Deconvoluted spectra of peak5, peak7 and peak10 from PNGaseF treated trimeric RBD; (I) O-glycan distribution trend in icIEF – MS results of PNGaseF treated trimeric RBD.



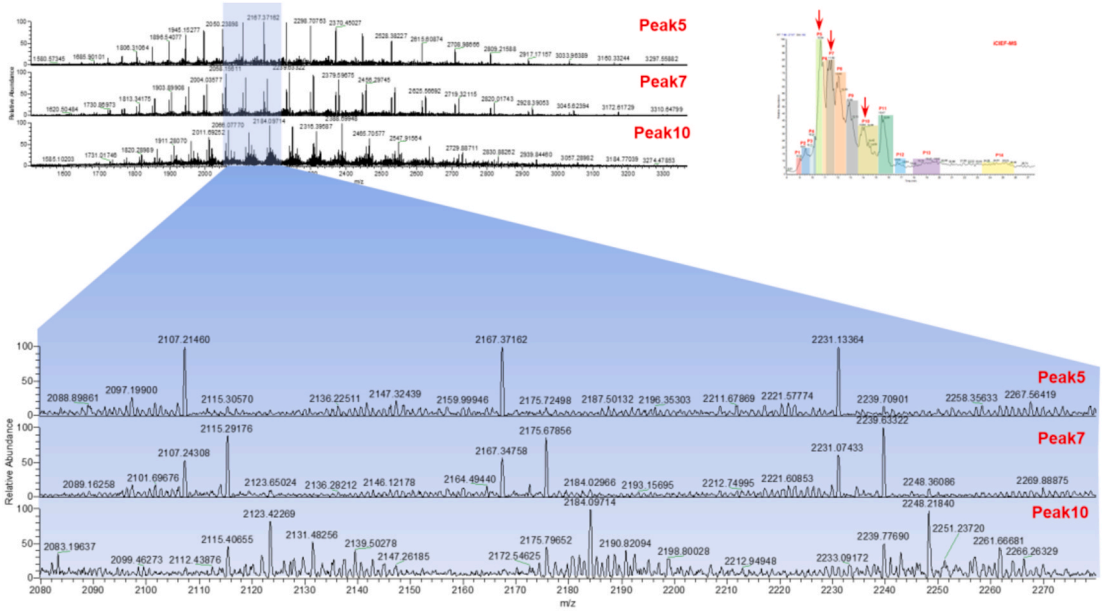
Peak1			Peak2		
Avg.mass (Da)	Modification	Relative abundance(%)	Avg.mass (Da)	Modification	Relative abundance(%)
84288.02	2xGalNAc-3SG, 1xGalNAc-6S-3SG, 1xFA2,1xFA2G1S1,1xFA2G2S2,2xFA1	100.00	84945.04	3xGalNAc-3SG, 1xFA2,1xFA2G2S2,1xA2G1F,3xFA1	100.00
84491.91	2xGalNAc-3SG, 1xGalNAc-6S-3SG, 2xFA2,1xFA2G1S1,1xFA2G2S2,1xFA1	93.96	84741.70	3xGalNAc-3SG, 1xFA2G2S2,1xFA2G1,4xFA1	95.31
84084.92	2xGalNAc-3SG, 1xGalNAc-6S-3SG, 1xFA2G1S1,1xFA2G2S2,3xFA1	90.44	85148.28	3xGalNAc-3SG, 2xFA2,1xFA2G2S2,1xFA2G1,2xFA1	89.47
84695.30	2xGalNAc-3SG, 1xGalNAc-6S-3SG, 3xFA2,1xFA2G1S1,1xFA2G2S2	77.02	84493.49	2xGalNAc-3SG, 1xGalNAc-6S-3SG 2xFA2,1xFA2G1S1,1xFA2G2S2,1xFA1	81.77
84899.02	2xGalNAc-3SG, 1xGalNAc-6S-3SG, 4xFA2,1xFA3G3S3	66.25	84783.88	3xGalNAc-3SG, 3xFA2,2xFA1G1S1,1xFA1	80.17

(D)

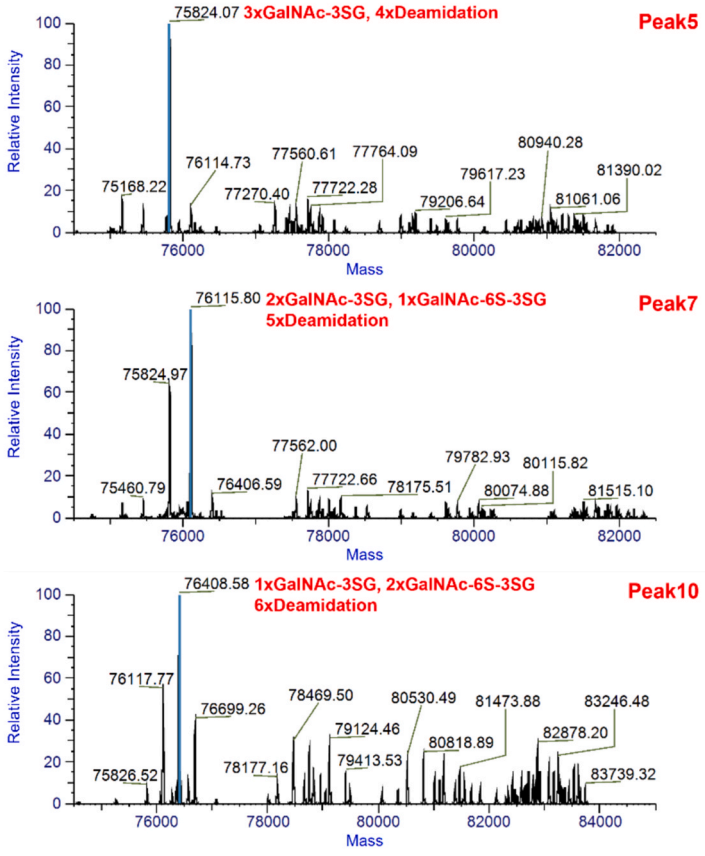


(E) and (F)

Fig. 4. (*continued*).

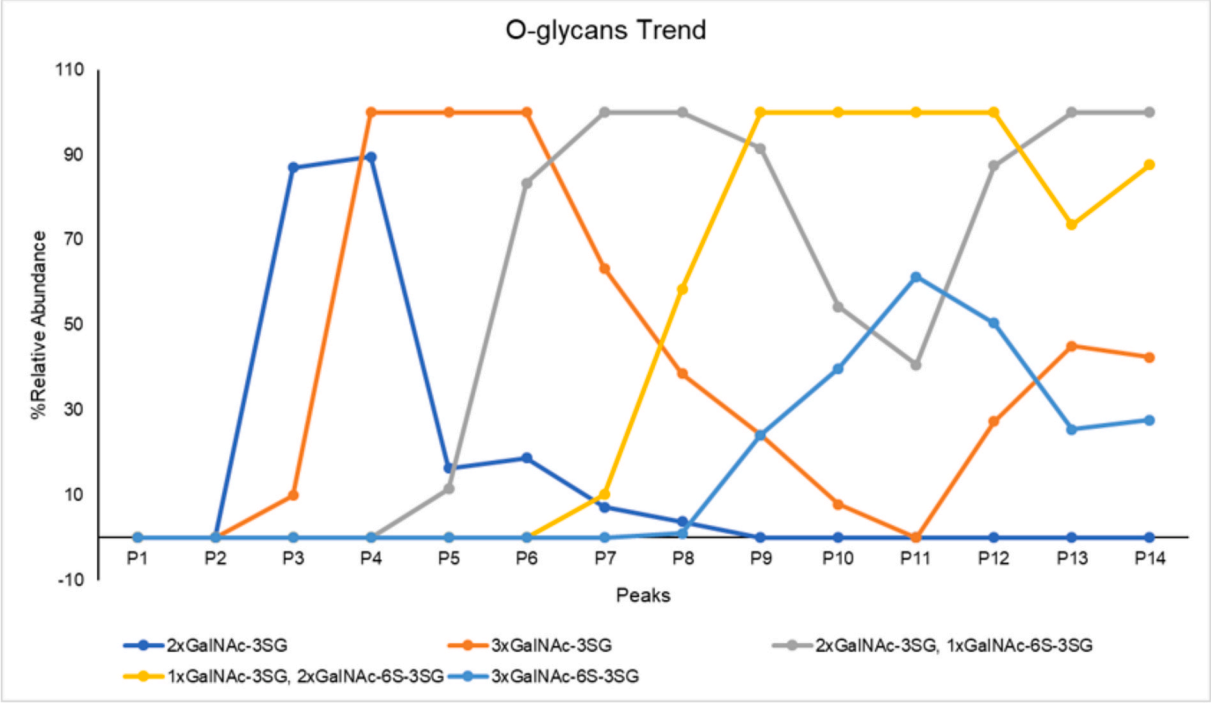


(G)



(H)

Fig. 4. (*continued*).



(I)

Fig. 4. (*continued*).

characterization more "elegant"- straightforward, sensitive, high-throughput and high data consistent. The whole workflows of icIEF-MS and fractionation still need to be further improved in terms of higher resolution, higher automation and developing critical reagents for higher performance icIEF. In this work, we demonstrated an essential role of high-resolution carrier ampholytes in high-efficient icIEF separation of extremely complex protein system with the use of SARS-CoV-2 recombinant vaccine and exhibited an important contribution of icIEF-MS online to rapidly fingerprinting the complex heterogeneity of recombinant vaccine. Moreover, icIEF fractionation greatly simplified the complexity of protein samples, which allowed it to carry out in-depth and more accurate characterization of collected fractions by HPLC-MS and IEX-MS. icIEF-MS online and icIEF fractionation can give fresh vitalities to the heterogeneity characterization of complex protein samples.

4. Conclusions

Due to extreme complexity of SARS-CoV-2 recombinant vaccine, a workflow containing comprehensive separation techniques coupled to high-resolution MS was developed for dissecting its complicated heterogeneity. Traditional HPLC-MS and HPLC-MSMS were utilized for characterization of SARS-CoV-2 recombinant vaccine at the levels of intact protein analysis and peptide mapping with PTMs identification. The cutting-edge technologies including icIEF-MS online and icIEF fractionation integrating IEX-MS were explored to investigate their applicability in in-depth characterization of complex protein compositions with high-throughput, straightforward methodology and excellent sensitivity. Achieved results on the multiple approaches were compared and cross validated. The workflow constructed underlined the outstanding importance of comprehensive platforms based on MS technology and the essential contributions from innovative icIEF-MS and preparative icIEF in the development of SARS-CoV-2 recombinant vaccine and similar complex protein systems.

Ethics declarations

The sample was provided voluntarily and with informed consent.

CRediT authorship contribution statement

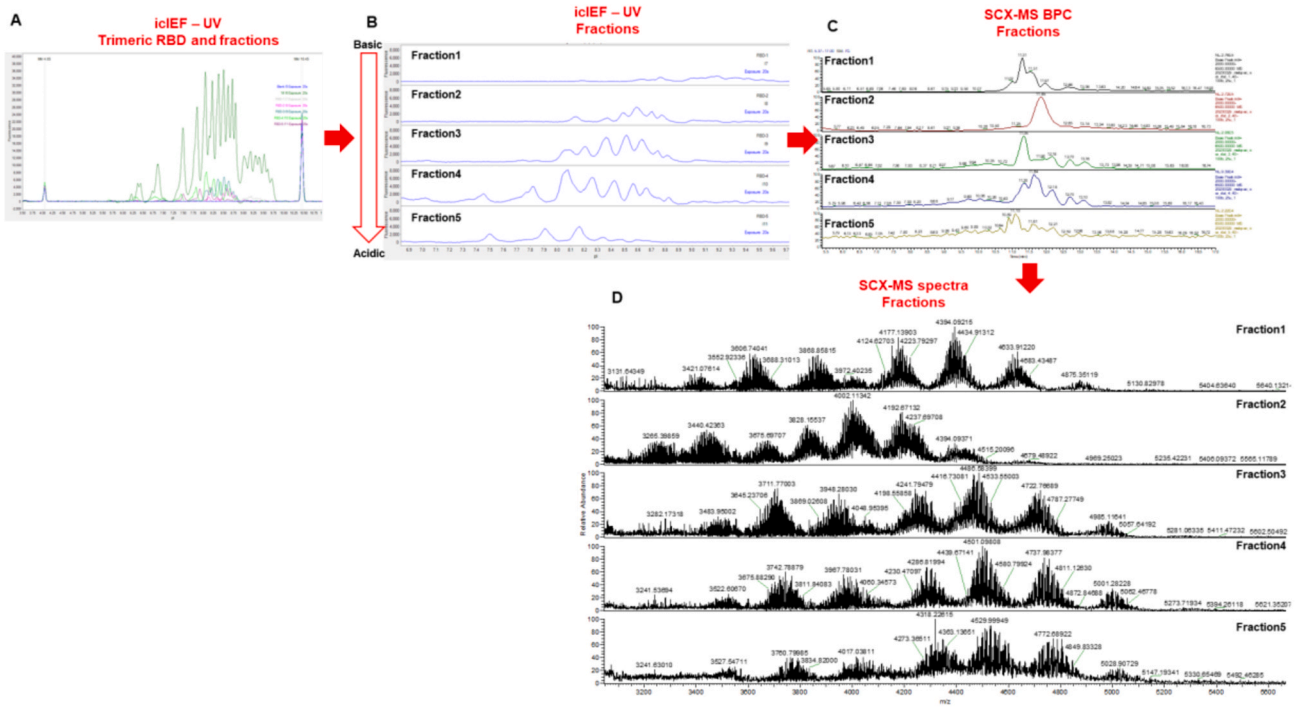
Gang Wu: Writing – original draft, Validation, Investigation, Formal analysis. **Jialiang Du:** Visualization, Data curation. **Chuanfei Yu:** Methodology, Conceptualization. **Zhihao Fu:** Validation, Data curation. **Xiaoxi Zhang:** Validation, Investigation. **Lan Wang:** Supervision, Methodology. **Junzhi Wang:** Supervision, Funding acquisition.

Declaration of competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.



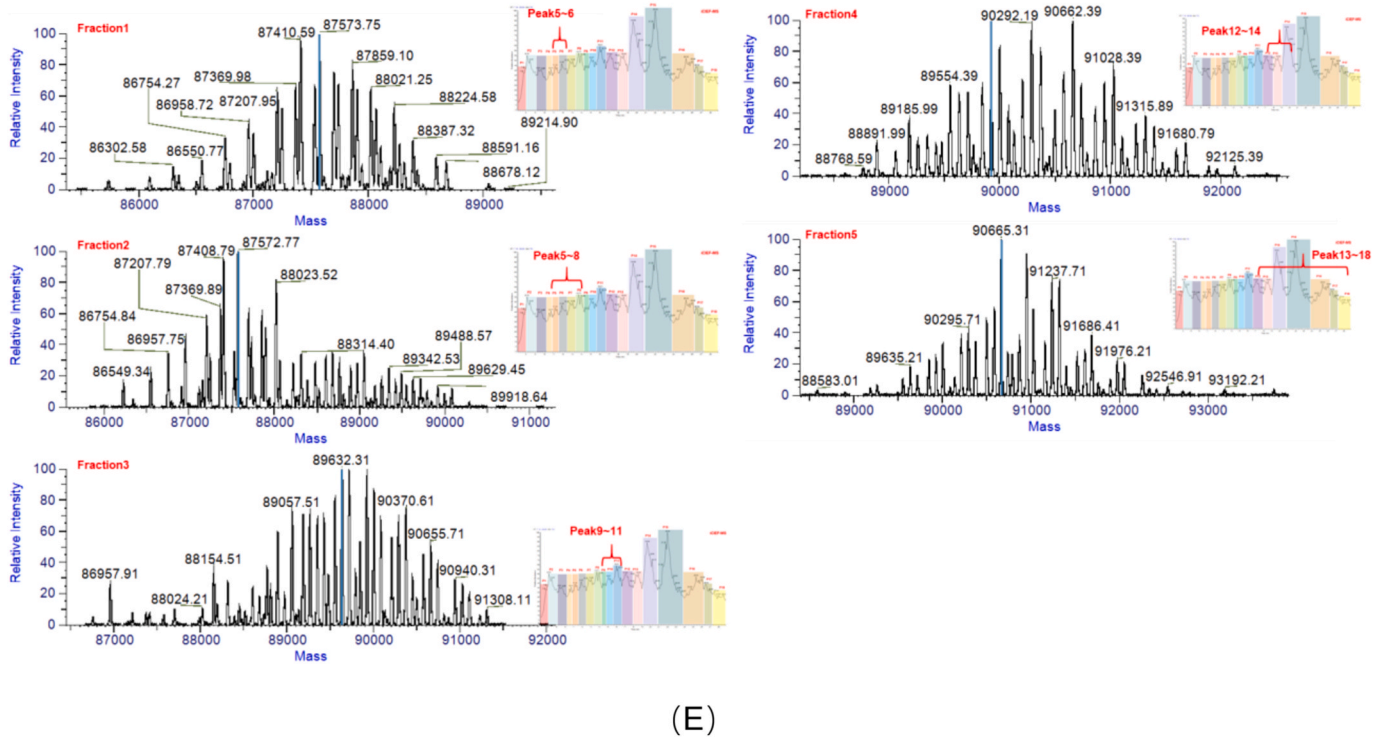


Fig. 5. IcIEF fractionation of trimeric RBD and integrated SCX-MS. (A) IcIEF – UV Trimeric RBD and fractions; (B) icIEF – UV fractions; (C) SCX-MS BPC fractions; (D) SCX-MS spectra fractions and (E) Deconvoluted spectra of icIEF fractionated trimeric RBD, followed by SCX-MS. Corresponding icIEF-MS online peaks are labelled for each fraction.

Acknowledges

This work was supported by the Research on precise quantitative technology of protein biological products under the National Key R&D Program, under grant of No. 2021YFF0600804–C02-05.

The authors gratefully acknowledge Dr. Tao Bo, Dr. Tong Chen, Dr. Tiemin Huang from the Advanced Electrophoresis Solutions Ltd (AES), who provided technical support in this work. The authors also thank AES for the provision of icIEF instrument access to support this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2024.342349.

References

[1] A.E. Gorbalenya, S.C. Baker, R.S. Baric, R.J. de Groot, C. Drosten, A.A. Gulyaeva, B. L. Haagmans, C. Lauber, A.M. Leontovich, B.W. Neuman, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nat Microbiol 5 (2020) 536–544.

[2] A.L. Phelan, R. Katz, L.O. Gostin, The novel coronavirus originating in wuhan, China challenges for global health governance, JAMA 323 (2020) 709–710.

[3] Who, World-health-organization coronavirus disease (COVID-19) outbreak. https://www.who.int/emergencies/diseases/novel-coronavirus-2019, 2020. (Accessed 31 January 2020).

[4] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K.S.M. Leung, E.H.Y. Lau, J.Y. Wong, Early transmission dynamics in wuhan, China, of novel coronavirus–infected pneumonia, N. Engl. J. Med. 382 (13) (2020) 1199–1207.

[5] H.A. Rothan, S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak, J. Autoimmun. 109 (2020) 102433.

[6] S. Kumar, V.K. Maurya, A.K. Prasad, M.L.B. Bhatt, S.K. Saxena, Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV), Virus Disease 31 (2020) 13–21.

[7] J. Xu, S. Zhao, T. Teng, A. Abdalla, W. Zhu, L. Xie, Y. Wang, X. Guo, Systematic comparison of two animal-to-human transmitted human coronaviruses: SARS-CoV-2 and SARS-CoV, Viruses 12 (2020) 244.

[8] W. Shang, Y. Yang, Y. Rao, X. Rao, The outbreak of SARS-CoV-2 pneumonia calls for viral vaccines, NPJ Vaccines 5 (2020) 18.

[9] V. Aida, C. Pliasas, P.J. Neasham, J.F. North, K.L. McWhorter, S.R. Glover, C. S. Kyriakis, Novel vaccine technologies in veterinary medicine: a herald to human medicine vaccines, Front. Vet. Sci. 8 (2021) 654289.

[10] T. Mebatsion, Introduction to veterinary vaccines, in: T. Vanniasinkam, S.K. Tikoo, S.K. Samal (Eds.), Viral Vectors in Veterinary Vaccine Development: A Textbook, Springer, Cham, 2021, pp. 3–12.

[11] E.N. Meeusen, J. Walker, A. Peters, P.P. Pastoret, G. Jungersen, Current status of veterinary vaccines, Clin. Microbiol. Rev. 20 (2007) 489–510.

[12] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082.

[13] A.V. Aid, V.C. Pliasas, P.J. Neasham, J.F. North, K.L. McWhorter, S.R. Glover, C. S. Kyriakis, Novel vaccine technologies in veterinary medicine: a herald to human medicine vaccines, Front. Vet. Sci. 8 (2021) 654289.

[14] M. Teixeira de Pinho Favaro, J. Atienza-Garriga, C. Martínez-Torró, E. Parladé, E. Vázquez, J.L. Corchero, N. Ferrer-Miralles, A. Villaverde, M. Teixeira de Pinho Favaro, J. Atienza-Garriga, C. Martínez-Torró, E. Parladé, E. Vázquez, J. L. Corchero, N. Ferrer-Miralles, A. Villaverde, A. Microbial cell factories, Microb. Cell Factories 21 (2022) 203–219.

[15] A. Shajahan, N.T. Supekar, A.S. Anne S Gleinich, P. Parastoo Azadi, Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2, Glycobiology 30 (12) (2020) 981–988.

[16] T. Yadav, N. Srivastava, G. Mishra, K. Dhama, S. Kumar, B. Puri, S.K. Saxena, Recombinant vaccines for COVID-19, Hum. Vaccines Immunother. 16 (12) (2020) 2905–2912.

[17] H. Yao, Y.T. Song, Y. Chen, et al., Molecular architecture of the SARS-CoV-2 virus, Cell 183 (2020) 730–738.

[18] Y. Jin, Z. Lin, Q. Xu, C. Fu, Z. Zhang, Q. Zhang, W.A. Pritts, Y. Ge, Comprehensive characterization of monoclonal antibody by Fourier transform ion cyclotron resonance mass spectrometry, mAbs 11 (1) (2019) 106–115.

[19] F. Higel, A. Seidl, F. Sörgel, W. Friess, N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins, Eur. J. Pharm. Biopharm. 100 (2016) 94–100.

[20] L.D. Bastiaan, A. Murisier, J. Camperi, S. Fekete, A. Beck, D. Guillarme, V. D'Atri, Therapeutic Fc-fusion proteins: current analytical strategies, J. Separ. Sci. 1 (2020) 35–62.

[21] X.Z. Zhang, T. Chen, V. Li, T. Bo, M. Du, T.M. Huang, Cutting-edge mass spectrometry strategy based on imaged capillary isoelectric focusing (icIEF) technology for characterizing charge heterogeneity of monoclonal antibody, Anal. Biochem. 660 (2023) 114961.

[22] Y.T. Yan, T. Xing, S.H. Wang, N. Li, Versatile, sensitive, and robust native LC-MS platform for intact mass analysis of protein drugs, J. Am. Soc. Mass Spectrom. 31 (2020) 2171–2179.

[23] S. Tamara, M. A. den Boer, A. J. R. Heck, High-resolution native mass spectrometry, Chem. Rev. (Washington, DC, U. S.) 122 (8) 7269–7326.

[24] M. Zarei, P. Wang, J. Jonveaux, F.M. Haller, B.N. Gu, A.V. Akoulov, M. Jahn, A novel protocol for in-depth analysis of recombinant adeno-associated virus capsid proteins using UHPLC–MS/MS, Rapid Commun. Mass Spectrom. 36 (2022) e9247.

[25] B.L. Duivelshof, A. Beck, D. Guillarme, V. D'Atri, Bispecific antibody characterization by a combination of intact and site-specific/chain-specific LC/MS techniques, Talanta 236 (2022) 122836.

[26] H. Stutz, Advances and applications of capillary electromigration methods in the analysis of therapeutic and diagnostic recombinant proteins – a Review, J. Pharmaceut. Biomed. Anal. 222 (2023) 115089–115133.

[27] L.D. Bastiaan, A. Murisier, J. Camperi, S. Fekete, A. Beck, D. Guillarme, V. D'Atri, Therapeutic Fc-fusion proteins: current analytical strategies, J. Separ. Sci. 1 (2020) 35–62.

[28] G. Rozing, Imaged capillary iso-electric focusing: background, status, and perspectives, Chromatography Today (2019) 8–14.

[29] L. Farmerie, R.R. Rustandia, J.W. Loughney, M. Dawoda, Recent advances in isoelectric focusing of proteins and peptides, J. Chromatogr. A 1651 (16) (2021) 462274.

[30] T. Chen, T. Kwok, A.R. Esmin, V. Li, G. Rozing, T.M. Huang, High-efficiency preparative imaged capillary isoelectric focusing (icIEF) and icIEF-MS protein charge variant characterisation, Chromatography Today (2021) 14–18.

[31] X.P. He, M. ElNaggar, M.A. Ostrowski, A. Guttman, E. Gentalen, J. Sperry, Evaluation of an icIEF-MS system for comparable charge variant analysis of biotherapeutics with rapid peak identification by mass spectrometry, Electrophoresis 43 (2022) 1215–1222.

[32] M. Scott, A. Don, B. Greg, B. Luc, D. Lieza, D. Vladislav, D. MaryAnn, G. Eric, H. Chris, J. Morten, J. Jennifer, L. Steve, R. Claudia, W. Ian, G. Erik, A novel microchip-based imaged CIEF-MS system for comprehensive characterization and identification of biopharmaceutical charge variants, Electrophoresis 40 (23–24) (2019) 3084–3091.

[33] L.Y. Wang, T. Bo, Z.Z. Zhang, G.B. Wang, W.J. Tong, D.D.Y. Chen, High resolution capillary isoelectric focusing mass spectrometry analysis of peptides, proteins, and monoclonal antibodies with a flow-through microvial interface, Anal. Chem. 90 (2018) 9495−9503.

[34] C. Montealegre, C. Neusüß C, Coupling imaged capillary isoelectric focusing with mass spectrometry using a nanoliter valve, Electrophoresis 39 (2018) 1151–1154.

[35] J.S. Mellors, V. Gorbounov, R.S. Ramsey, J.M. Ramsey, Fully integrated glass microfluidic device for performing high-efficiency capillary electrophoresis and electrospray ionization mass spectrometry, Anal. Chem. 80 (2008) 6881–6887.

[36] J. Schlecht, B. Moritz, S. Kiessig, C. Neusüß C, Characterization of therapeutic mAb charge heterogeneity by icIEF coupled to mass spectrometry (icIEF–MS), Electrophoresis 23 (2022) 1–9.

[37] T. Xu, L.J. Han, L.L. Sun, Automated capillary isoelectric focusing-mass spectrometry with ultrahigh resolution for characterizing microheterogeneity and isoelectric points of intact protein complexes, Anal. Chem. 94 (27) (2022) 9674–9682.

[38] T. Xu, L.J. Han, A.M. George Thompson, L.L. Sun, An improved capillary isoelectric focusing-mass spectrometry method for high-resolution characterization of monoclonal antibody charge variants, Anal. Methods 14 (4) (2022) 383–393.

[39] J. Dai, J. Lamp, Q.W. Xia, Y.R. Zhang, Capillary isoelectric focusing-mass spectrometry method for the separation and online characterization of intact monoclonal antibody charge variants, Anal. Chem. 90 (3) (2018) 2246–2254.

[40] G. Wu, C.F. Yu, W.B. Wang, J.L. Du, Z.H. Fu, G.L. Xu, M. Li, L. Wang, Mass spectrometry-based charge heterogeneity characterization of therapeutic mAbs with imaged capillary isoelectric focusing and ion-exchange chromatography as separation techniques, Anal. Chem. 95 (2023) 2548–2560.

[41] T. Kwok, M. Zhou, A. Schaefer, T. Bo, V. Li, T.M. Huang, T. Chen, Fractionation and online mass spectrometry based on imaged capillary isoelectric focusing (icIEF) for characterizing charge heterogeneity of therapeutic antibody, Anal. Methods 15 (4) (2023) 391–544.

[42] Y. Liang, J. Zhang, R.Y. Yuan, M.Y. Wang, P. He, J.G. Su, Z.B. Han, Y.Q. Jin, J. W. Hou, H. Zhang, X.F. Zhang, S. Shao, Y.N. Hou, Z.M. Liu, L.F. Du, F.J. Shen, W. M. Zhou, K. Xu, R.Q. Gao, F. Tang, Q.M. Li, Design of a mutation-integrated trimeric RBD with broad protection against SARS-CoV-2, Cell Discov 8 (1) (2022) 17.

[43] Q.M. Li, Y. Liang, J.G. Su, J. Zhang, Y.Q. Jin, J.W. Hou, et al., A Novel SARS-CoV-2 Trimeric S-RBD Protein Vaccine, its Preparation Method and Application, Chinese Patent, 2021. Publication number: CN113817029A. Publication date: 2021-12-21.

[44] H. Yao, Y.T. Song, Y. Chen, et al., Molecular architecture of the SARS-CoV-2 virus, Cell 183 (2020) 730–738.

[45] M. Sanda, L. Morrison, R. Goldman, N- and O-glycosylation of the SARS-CoV-2 spike protein, Anal. Chem. 93 (4) (2021) 2003–2009.

[46] Y. Watanabe, J.D. Allen, D. Wrapp, J.S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike, Science 369 (2020) 330–333.

[47] A. Beck, et al., Characterization of therapeutic antibodies and related products, Anal. Chem. 85 (2013) 715–736.

[48] R. Rogers, et al., A view on the importance of "multi-attribute method" for measuring purity of biopharmaceuticals and improving overall control strategy, AAPS J. 20 (2018) 7.

[49] K. Maráková, M. Opetová, R. Tomašovský, Capillary electrophoresis-mass spectrometry for intact protein analysis: pharmaceutical and biomedical applications (2018–March 2023), J. Separ. Sci. 46 (2023) 2300244–2300282.
[50] S. Hartung, R. Minkner, M. Olabi, H. Watzig, Performance of capillary electrophoresis instruments- State of the art and outlook, Trends Anal. Chem. 163 (2023) 117056–117069.
[51] L. Barry, Karger, Perspectives on capillary electrophoresis, Perspect. Capillary Electrophor. 167 (2023) 117215–117216.
[52] F. Krebs, H. Zagst, M. Stein, et al., Strategies for capillary electrophoresis: method development and validation for pharmaceutical and biological applications, Electrophoresis 44 (17–18) (2023) 1279–1342.