scientific reports

Check for updates

OPEN

Rapid identification of antibody impurities in size-based electrophoresis via CZE-MS generated spectral library

Quan Liu[2,4], Jiaying Hong[2,4], Yukun Zhang[2], Qiuyue Wang[2], Qiangwei Xia[2], Michael D. Knierman[3], Jim Lau[3], Caleen Dayaratna[1], Benjamin Negron[1], Hirsh Nanda[1] & Harsha P. Gunawardena[1]✉

Methods for the reliable and effective detection and identification of impurities are crucial to ensure the quality and safety of biopharmaceutical products. Technical limitations constrain the accurate identification of individual impurity peaks by size-based electrophoresis separations followed by mass spectrometry. This study presents a size-based electrophoretic method for detecting and identifying impurity peaks in antibody production. A hydrogen sulfide-accelerated degradation method was employed to generate known degradation products observed in bioreactors that forms the basis for size calibration. LabChip GXII channel electrophoresis enabled the rapid (<1 min) detection of impurity peaks based on size, while capillary zone electrophoresis-mass spectrometry (CZE-MS) facilitated their accurate identification. We combine these techniques to examine impurities resulting from cell culture harvest conditions and forced degradation to assess antibody stability. To mimic cell culture harvest conditions and the impact of forced degradation, we subjected samples to cathepsin at different pH buffers or exposed them to high pH and temperature. Our method demonstrated the feasibility and broad applicability of using a CZE-MS generated spectral library to unambiguously assign peaks in high throughput size-based electrophoresis (i.e., LabChip GXII) with identifications or likely mass of the antibody impurity. Overall, this strategy combines the utility of CZE-MS as a high-resolution separation and detection method for impurities with size-based electrophoresis methods that are typically used to detect (not identify) impurities during the discovery and development of antibody therapeutics.

Keywords Antibody, Impurity, Degradation, LabChip, CE-SDS, CZE-MS

The detection and identification of impurities have always been crucial in the development of biopharmaceutical products. Impurities linked to size variants (clips, fragments, aggregates) are important quality attributes that can directly impact the efficacy or safety of protein therapeutics[1–3]. Resulting from peptide bond cleavage via chemical or enzymatic reactions, fragments or clips are generated during cell culture, bioprocessing and storage[4,5]. Therefore, size heterogeneity is an important indicator of the entire antibody production process, necessitating an effective monitoring strategy to assess the purity and integrity of the protein. Having been widely applied for decades to study impurities, capillary electrophoresis (CE) has been an important technique for analyzing biological therapeutic molecules with unique features yet complementary to liquid chromatography[6,7].

For the analysis of biotherapeutic antibodies, the determination of protein purity has transitioned from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) slab gel technology to capillary electrophoresis-sodium dodecyl sulfate (CE-SDS) methods[5,8–10]. The robustness, reliability, and accuracy of the CE-SDS method underwent extensive research and investigation in a 2006 cross-organization collaboration lead by Nunnally et al.[11]. The results demonstrated that CE-SDS is more suitable than traditional SDS-PAGE for the characterization and quality control of monoclonal antibody drugs, exhibiting significantly higher robustness and

[1]Johnson & Johnson Innovative Medicine Research & Development, 1400 McKean Road, Spring House, PA 19477, USA. [2]CMP Scientific Corp, 760 Parkside Ave, STE 211, Brooklyn, NY 11226, USA. [3]Agilent Technologies, 5301 Stevens Creek Blvd, Santa Clara, CA 95051, USA. [4]These authors contributed equally: Quan Liu and Jiaying Hong. ✉email: hgunawar@ITS.JNJ.com

reliability. CE-SDS has undergone significant development over the years and has been incorporated into several pharmacopoeias worldwide[12] and established itself as a conventional method for molecular size estimation and purity analysis in biopharmaceuticals[5,9,12,13]. The continuous development within the CE market has led to the progress of the microchip electrophoresis method due to its automation, high throughput, increased sensitivity, and resolution[14,15]. Herein, we have chosen commercial microchip electrophoresis instrument (LabChip GXII) for examining degradation products of antibodies. However, strategies based on capillary based SDS electrophoresis or microchip techniques can only address the detection of impurity peaks but not the identification of these peaks. The assignment of impurity peaks in antibody products and understanding the root causes of degradation remain challenging based on current techniques since direct identification of impurity peaks observed in CE-SDS is oftentimes an ambiguous task[16].

Mass spectrometry (MS) has been widely used as a protein characterization technique during research and development of biopharmaceutical products[17–24]. Due to strong ion suppression effects of SDS and other separation buffer components, direct analysis by coupling CE-SDS separated proteins with electrospray ionization (ESI) mass spectrometers has not been feasible[25–28]. However, Sarkozy et al. recently described a method to detect small proteins via MS by coupling CE-SDS using a coaxial-sheath flow reactor[29]. Orthogonal analytical methods have been established for indirect analysis of CE-SDS peak identification and characterization. However, workflow based on liquid chromatography coupled with MS tend to rely on the combination of intact and peptide level analysis, and multi-dimensional separation to remove MS-incompatible components before detection, which can be tedious and sometimes makes data analysis more complex[28,30–32].

With the advancement of electrospray ionization (ESI) ion sources, capillary electrophoresis-mass spectrometry (CE-MS) now offers higher sensitivity[33,34] capillary zone electrophoresis-mass spectrometry (CZE-MS) and capillary isoelectric focusing-mass spectrometry (cIEF-MS) coupling[27,35–37]. CZE-MS holds promise for facilitating efficient protein separation and potentially enhancing the detection sensitivity of low-abundance impurity proteins[38–40].

In this study, we developed a method for detecting and identifying impurities that can form during antibody production. Previous studies have revealed that cysteine in the feed and the bioreactor media produces hydrogen sulfide which leads to the insertion of a sulfur atom to a disulfide bond or formation of a trisulfide bond in the biological product[41,42]. Thus, we employed a hydrogen sulfide-accelerated degradation experiment to simulate bioreactor-induced antibody degradation to generate specific antibody fragments that can be used as a calibration standard. Microchip electrophoresis LabChip GXII was used to detect impurity peaks in the antibody product, and capillary zone electrophoresis-mass spectrometry (CZE-MS) was used for generating spectral libraries of antibody impurities that are used for the identification and assignment of LabChip GXII peaks. After successful method development and validation of the approach for assigning GXII peaks to a known antibody clip site, we subjected antibodies to Cathepsin and pH-induced forced degradation experiments to emulate enzymatic or chemical reactions in cell culture or during down-stream processing[43,44]. With the established CZE-MS-based methodology, we obtained spectra that consists of the accurate molecular mass of each impurity peak in GXII electropherograms and proposed possible explanations for their occurrence. Further validation of the GXII impurity peaks was performed by using CZE-MS/MS. Overall, this method demonstrates the feasibility and broad applicability of detecting and mapping impurity peaks during antibody production.

Experimental section

Reagents

Iodoacetamide (IAA), N-ethylmaleimide (NEM) and NIST IgG K1 (NISTmAb), Rituximab, Inflixmab, Pembrolizumab, Vedolizumab, Sigmamab ADC mimic, Cathepsins D, Cathepsin L, formic acid, ammonium solution, water, ammonium acetate, methanol, acetic acid, formic acid, isopropanol, and hydrogen sulfide were all purchased from Sigma Aldrich (St. Louis, MO). HT Pico Protein Express kits and HT Protein Express LabChips were obtained from Revvity (Waltham, MA), and Zeba spin desalting columns and plates 7 kDa MWCO, and dithiothreitol (DTT) were bought from Thermo Fisher Scientific (Waltham, MA).

Degradation of antibodies

Three stock solutions of 1 mg/mL of NIST IgGK1 (NISTmAb) antibody are prepared by buffer exchanging to 200 mM sodium phosphate solution buffered at a pH of 7, 500 mM citrate solution buffered at a pH of 3 and 200 mM acetate solution buffered at a pH of 5. The NISTmAb at pH 3 and 7 was digested with Cathepsin L at 100:1 ratio (w/w) while NISTmAb at pH 5 was digested Cathepsin D at 100:1 ratio (w/w). Three 100 µL aliquots from each mixture was taken after 1 day or after 2 days of incubation at room temperature for LabChip GXII and CZE-MS experiments. H_2S gas was bubbled into a 5 mL vial containing 1 mg/mL NISTmAb buffered with 500 mM citrate at pH 3 for 20 min. 100 µL of the reaction mixture was used for LabChip GXII and CZE-MS experiments. 1 mg/mL Rituximab stock solution was prepared in 1xPBS. Two stock solutions of 1 mg/mL Rituximab were prepared in 25% ammonia solution (pH ~ 10) that was stored at 4 °C or 37 °C for two days. Three 100 µL aliquots from each condition was taken for LabChip GXII and CZE-MS experiments.

LabChip® GXII

Purity analysis on the LabChip GXII Touch (PerkinElmer Inc.) was performed according to the manufacturer's instructions with a few modifications. In brief, samples were diluted to 0.5 mg/mL and desalted (Zeba 7 kD desalting plates) prior to analysis. For non-reduced analysis, 5 µL of sample was incubated with 5 µL of a working fluorescent dye solution (40 µM). For the ladder preparation, 4 µL ladder, 2 µL of a 5 × non-reduced labeling buffer and 4 µL water were mixed by pipetting, and 5 µL of the diluted ladder was added to 5 µL working dye solution. Both samples and ladder were incubated in the dark at room temperature for 1 h, followed by addition

of 1 µL stop solution; 18 µL non-reduced sample buffer (0.2 M Bis–Tris Citrate SDS buffer containing 25 mM NEM) was added to 5 µL of the sample, and 11 µL ladder, and denatured at 70 °C for 10 min before the addition of 77 µL water to the sample, and 137 µL water to the ladder. Samples were placed in the 96-well BioRad fully skirted plates for analysis and testing was performed on the HT Protein Express LabChip (P/N 760499) with the Pico Protein Express Reagent Kit (P/N 760498). The LabChip was prepared according to the manufacturer's instructions using Protein Gel Matrix. During the 40 s run time, the chip was primed with reagents and warmed to 30 °C. The samples were injected at − 100 to − 3000 V and separated on the approximately 12 mm separation channel with excitation by the red laser diode (635 nm) and emission detected by a bandpass filter (670–725 nm). An internal lower molecular weight marker was added to each sample to align the time. The ladder was used to quantitate protein size (kDa) and brackets every 12 injections. The data was analyzed by LabChip GX Reviewer Software ver. 5.8.84.0.

CZE-MS

The CE-MS system consists of an Agilent 7100 CE system (Agilent Technologies, Santa Clara, CA), an EMASS-II CE-MS ion source (CMP Scientific Corp, Brooklyn, NY), an Agilent 6230 TOF mass spectrometer (Agilent Technologies, Santa Clara, CA), and an Agilent 6545XT AdvanceBio Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA). Electrospray emitters and separation capillaries (100 cm, 50 µm i.d. neutral coating PS2 capillary) were from CMP Scientific. The background electrolyte (BGE) for the CZE separation was 5% acetic acid and the sheath liquid was 10% methanol and 0.2% formic acid. In the CZE-MS experiment, the sample was injected using 50 mbar injection pressure for 20 s, followed by a 50-mbar injection pressure for 3 s of the BGE. The CZE separation was performed with a voltage of + 30 kV for 20 min. The CZE capillary was flushed with BGE for 2 min between runs.

MS parameters

The regular LC–MS spray shield was used on the Agilent mass spectrometers. The drying gas temperature was 365 °C. Flow rate of drying gas was 4 L/min. The nanospray was established with a voltage on the TOF capillary at 2500 V. Fragmentor voltage was set to 380 V and the skimmer was set to 45 V. The acquisition rate was collected at two spectra per second from 400 to 3200 m/z. MS/MS experiments are performed with an isolation window of 4 m/*z* and collision energy of 45 V.

MS data analysis

The Agilent Mass Hunter software (ver. 10.0) was used for instrument control, data acquisition, and data analysis. Data analysis was performed with Agilent BioConfirm software (ver. 10.0). Mass spectra were deconvoluted using MaxEnt. MS/MS spectra were interpreted manually.

Results and discussion

Clipping variants of recombinant protein therapeutics are closely monitored during cell culture and process development as the existence of clipping impurities are a high-risk critical quality attribute. Although effective approaches have been routinely applied to detect and quantify such impurities, there are unmet needs in the biopharmaceuticals industry for efficient means of identifying impurities detected by size-based separation methods. By obtaining a more accurate understanding of the identity of the clipping impurities, optimization procedures in process development may then be developed to improve safety, efficacy, and quality of biologics.

Overview of the study

In the production of biologics, impurities can be generated due to degradation caused by various factors. These processes can occur in both upstream cell culture all the way through down-stream purification and formulation development (Fig. 1A). To emulate a few of these process development scenarios where degradation of biologics is observed, we subjected several commercial IgG1 antibodies to both chemical and biochemical stresses i.e., to H_2S gas or to Cathepsin D and L or to high-pH buffer leading to IgG degradation. For the identification of electrophoretic peaks resulting from the degradation of mAb, we initially utilized a degradation standard that was generated via the reduction of a NISTmAb with H_2S. This standard, comprising several known degradants, is the basis of a mass ladder that was subjected to LabChip GXII for electrophoretic separation detection and to CZE-MS for identification (Fig. 1B). A GXII peak identification strategy, illustrated in the workflow, shows a linear fit of the log values of molecular weight of each degradant and the inverse of their migration times. The regression analysis enables accurate estimation of molecular weights of each impurity peak in the GXII electropherogram, via the impurity masses derived by CZE-MS (Fig. 1C). A given GXII electrophoretic peak that was initially assigned a molecular weight based CZE-MS analysis, is further validated based on the accuracy of migration time of the peak which is the difference in migration times between the predicted migration and the observed migration time. This allows previously questionable GXII peaks to be reassigned with the most accurate impurity species (Fig. 1D). The CZE-MS spectra forms the basis of this investigation for accurate identifications of the impurity peaks, confirming and validating the identity LabChip GXII impurity peak. In some instances, top-down sequencing was performed via CZE-MS/MS of impurity fragment ions. This approach provided detailed structural and sequence information, confirming the reliability of our CZE-MS results.

H_2S-induced degradation products of IgG as a GXII mobility ruler

We used PBS buffer that was sparged with hydrogen sulfide to emulate the environment of a bioreactor that has high levels of cysteine and whereby the specific fragments can be used as a calibration standard. Cysteine in the

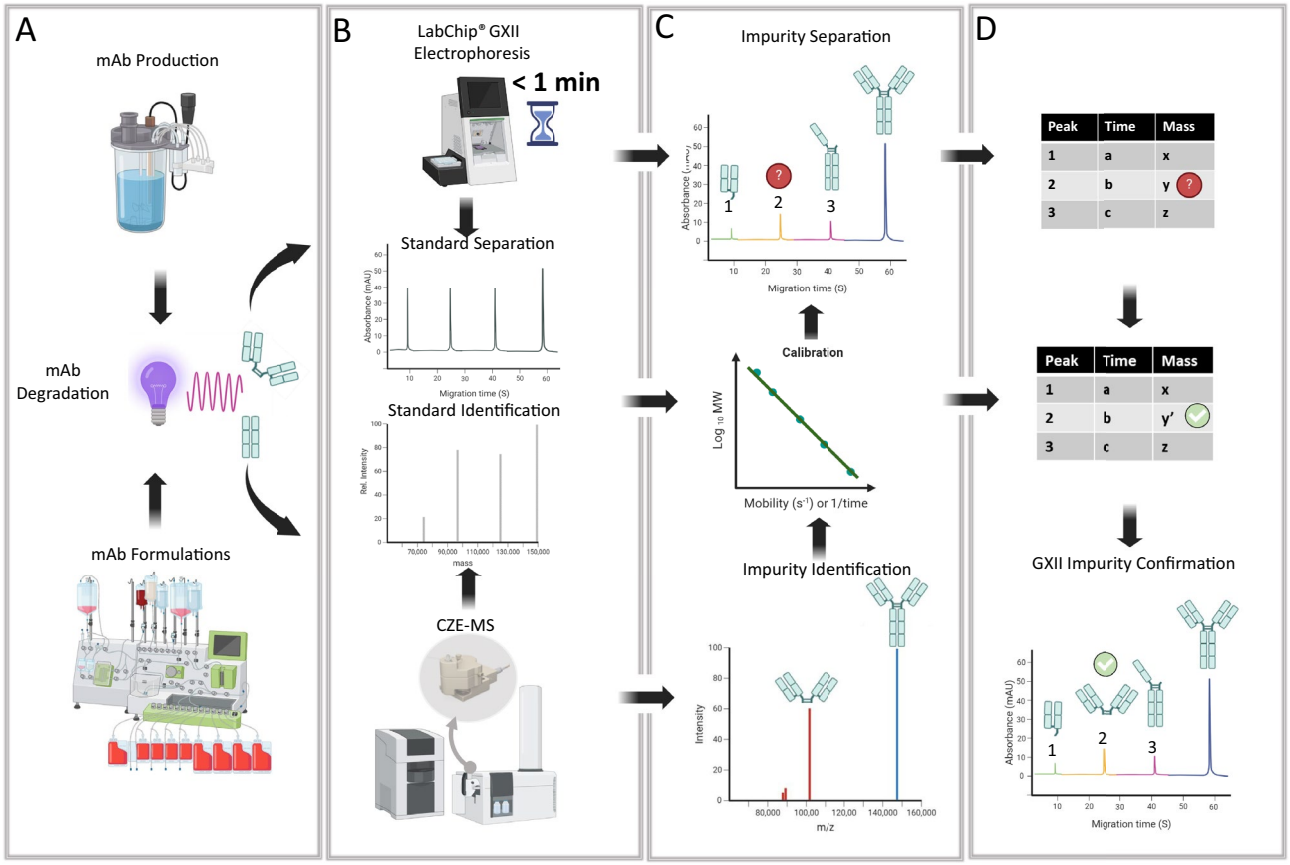

Fig. 1. A workflow for mAb impurity analysis combining electrophoretic separation and accurate assignment of masses (**A**) mAb impurities derived by forced degradation conditions emulating cell culture and formulation screening. (**B**) Analysis of impurity mass ladder derived from mAb using LabChip GXII electrophoresis and CZE-MS (**C**) CZE-MS spectral mass identifications of mAb impurities are mapped to the GXII peaks in the electropherogram via the calibration curve. (**D**) Reassignment of GXII peak identifications based on the accuracy of migration times using a lookup table of accurate CZE-MS masses. Note: graphics were created with BioRender.com.

feed and bioreactor media may lead to the generation of hydrogen sulfide. It is commonly observed that hydrogen sulfide reaction with mAbs may result in the insertion of a sulfur atom at disulfide linkage and leading to a trisulfide linkage. The disulfide linkage between Light chain and Heavy chain is most susceptible to trisulfide formation[45]. Trisulfides are unstable and leads to monoclonal antibody (mAb) degradation to several products. i.e., Heavy–Heavy–Light chain (HHL), Heavy–Heavy (HH), Half IgG1 or Heavy–Light (HL), Heavy (H), and Light (L) chains. The GXII electropherogram of native NISTmAb shows low intensity peaks that correspond to these degradation products.

We next incubated NISTmAb with PBS buffer saturated with hydrogen sulfide, which induced several H_2S-induced degradation products. Specifically, peaks 1–5 showed a significant increase in peak area relative to the unstressed NISTmAb (Fig. 2A). When attempting to calculate and infer the molecular weights of the impurity peaks, estimations were made based on the calibration procedure on LabChip GXII, that resulted in high degree of mass inaccuracy (Table S1) due to a consideration of low molecular weight markers to infer higher masses (Fig. S1A). CZE-MS analysis revealed the exact identity of most degradation product with high mass accuracy. The native NISTmAb is composed of two heavy chains and two light chains (H2L2) that is the most abundant peak in the deconvoluted mass spectrum. Upon close examination, H_2S stressed NISTmAb shows several trisulfide adducts. In addition, several other impurity peaks (*i.e.,* HHL, HL, H, and L) were confidently assigned based on the CZE-MS results (Fig. 2B). By obtaining accurate molecular weights of all impurities from the mass deconvolution of CZE-MS raw data resulting in significantly high mass accuracy (< 40 Da) and calibrating the observed GXII migration times, a highly reliable linear curve is then established. This enabled the annotation of each peak in the LabChip GXII analysis based on the linear calibration obtained by plotting the logarithm molecular weight of the standards versus their electrophoretic mobility (Fig. 2C). Consequently, even in cases where the molecular weight of Heavy chain dimer (HH) was not directly determined by CZE-MS, the mass of the HH impurity observed by GXII was accurately estimated, using the calibration curve and the migration time provided by LabChip GXII (Fig. 2D) analysis. The main electrophoretic peak was the intact H2L2 while peak with an asterisk was confirmed as aglycosylated H2L2 (Fig. S1B). The H_2S-induced NISTmAb degradation products allow the effective, accurate calibration of the GXII masses where masses of impurities are validated using the

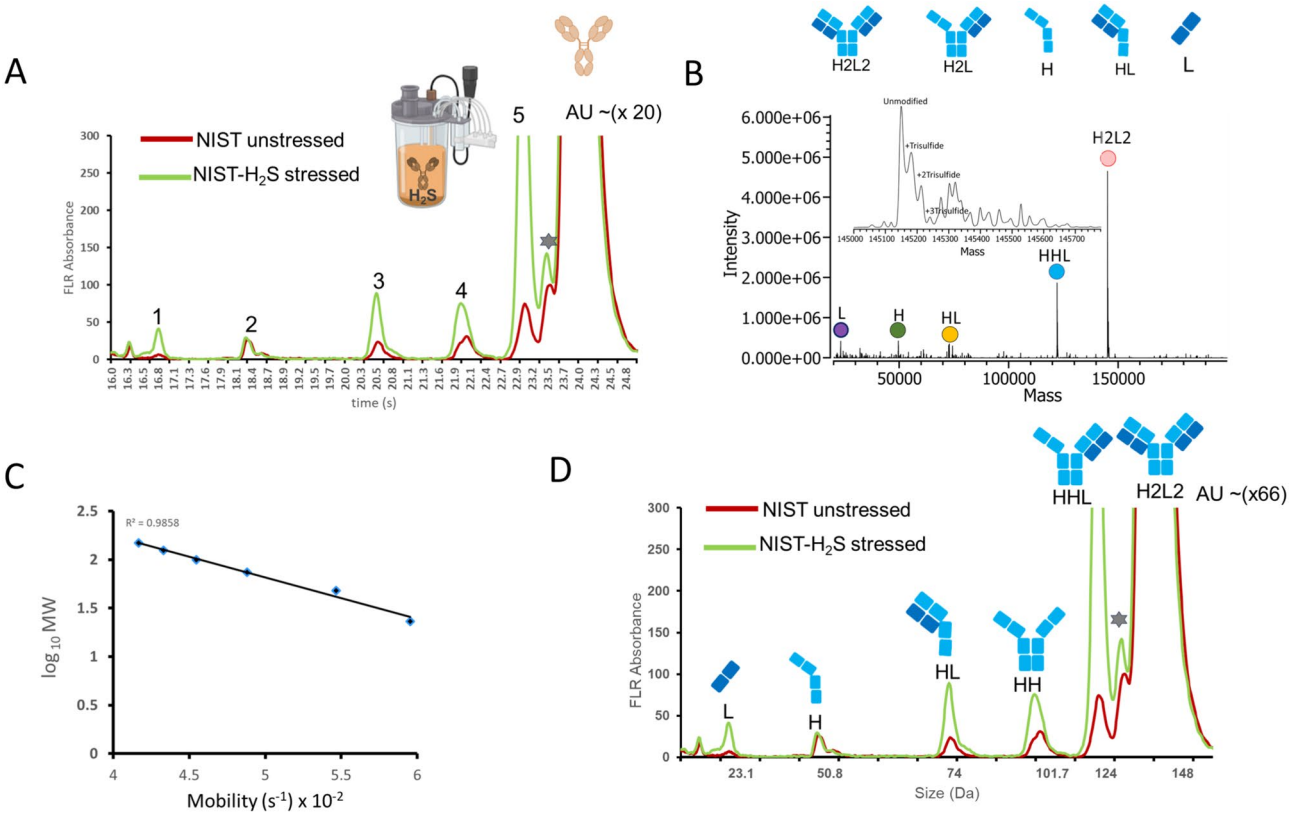

Fig. 2 . (**A**) Analysis of degradation products of NISTmAb A) LabChip GXII electropherogram of unknown mAb fragments peak 1–5 under unstressed (native condition), red trace and H2S stressed conditions green trace of NISTmAb. Inset is the intact H2L2 with with up to 3 trisulfide adducts. (**B**) Deconvoluted mass spectrum resulting from H_2S-induced degradation products of NISTmAb. (**C**) Linear regression analysis via NISTmAb fragments, plotted as log10 of the NISTmAb fragment molecular weights (MW) versus the mobility or inverse of migration time. (**D**) Assignment of masses to GXII peaks based on calibration in C and assignment of peak 4 corresponding to HH. Note: NISTmAb intact molecule is denoted as H2L2, * is aglycosylated H2L2 and degradation products are denoted as L, H, HL, HH, HHL. *Note*: graphics were created with BioRender.com.

CZE-MS. Having low-level degradation products even under native conditions, the use of NISTmAb as a standard provides a basis for accurate molecular weight determination of observed GXII masses. NISTmAb calibration curve when applied to several mAbs having different molecular weights was robust enough to unambiguously annotate fragments with a larger mass error (Table S2 and Figure S4).

Identification and assignment of pH-induced degradation products of IgG

We first applied our approach to examine Rituximab degradation. Rituximab is a therapeutic IgG that is prone to degradation due to Asparagine-Proline (NP) clipping under basic conditions[46]. We identify the putative NP clip site in a GXII electropherogram using the size-calibration obtained from the H_2S-induced degradation products. Commercially available rituximab when subjected to pH-induced degradation, under high-pH buffers, is susceptible to cleavage at the NP site of the light chain (Fig. 3A). The high-pH stress samples as well as the unstressed samples were subjected to CZE-MS to detect clip sites of Rituximab. Under high-pH stress we identified the putative NP clip site (Fig. 3B). Untreated rituximab sample stored at 4 °C did not exhibit significant impurity peaks. However, after one day of high-pH stress at both 4 °C and 37 °C, there was a noticeable increase in Peaks 2 and 7 (Fig. 3B) relative to other peaks (Fig. 3B Inset).

Based on the data obtained from CZE-MS, and via the size calibration of the GXII electropherogram using the H2S-induced NISTmAb calibration curve, we were able to assign most peaks (Table S3), including the impurities resulting from the NP Clip site (Fig. 3C and D). It is noteworthy that increasing temperature in pH 10 buffer promotes significant NP cleavage and removal of light chain resulting in Peak 7 as assignment as H2L-NP Clip with a concomitant increase of Peak 2 assigned to light chain (L), Peak 8, which was assigned to H2L2-2xNP Clip, also shows a temperature dependent increase in intensity.

Identification of cathepsin-induced degradation products of IgG

We employed cathepsins D and L to induce clipping of mAbs. As commonly seen proteases in cell culture harvest conditions, Cathepsins D and L clipping variants have been identified as tenacious impurities, which are hard to be removed during downstream bioprocessing of mAbs. These proteases cleave antibodies under specific

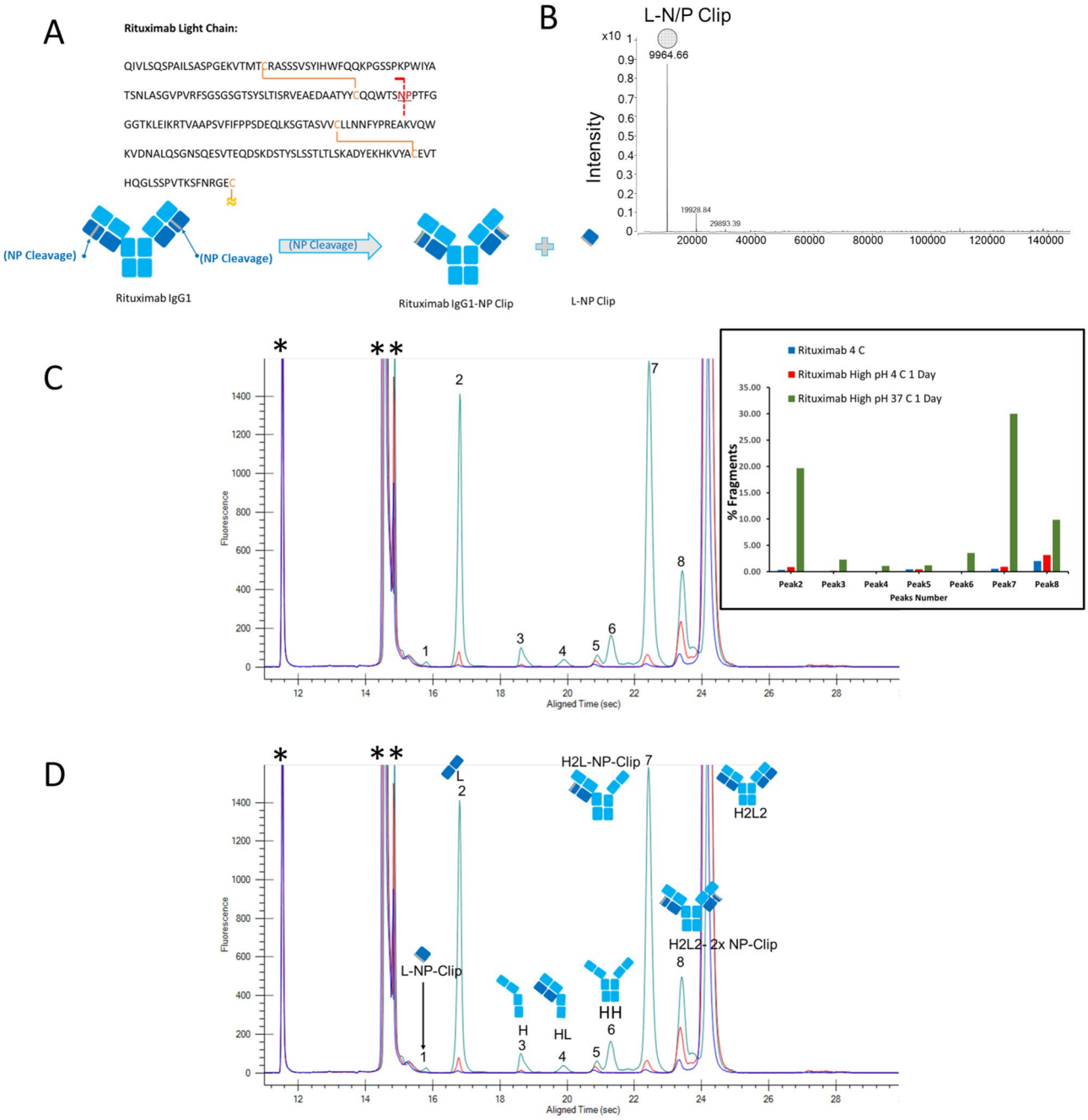

Fig. 3. Analysis of pH-induced degradation products of Rituximab (**A**) Rituximab light chain sequence with site-specific NP cleavage. (**B**) CZE-MS mass spectrum of L-N/P Clip. (**C**) Overlayed LabChip GXII electropherogram corresponding to pH- and temperature-induced degradation products of Rituximab (Inset shows the intensity of peaks 1–8 under native Rituximab stored at 4 °C (blue), incubated at high-pH of 10 at 4 °C for 1 day (red) and incubated at high-pH of 10 at 4 °C for 2 days (green). (**D**) Assignment of NP cleavage products of Rituximab to peaks in the electropherogram with identities or masses obtained from the regression analysis. *Note*: Peak 6~90 kDa, * ~ 1 kDa standard and ** is a system peak from the fluorescent dye.

conditions, resulting in structural changes and the formation of different-sized fragments, which contribute to the impurity profile of IgG1.

IgG1 light chain undergoes site-specific cleavage between AA and FI residues to form L-AA Clip and L-FI Clip respectively. L-AA Clip is formed by Cathepsin L at pH 3 while L-FI Clip is formed due to either digestion conditions of Cathepsin L at pH 7 and Cathepsin D at pH 5[47]. Prediction of the fragment masses in silico due to enzymatic cleavage facilitated the accurate prediction peak identities in the GXII electropherograms. The FI Clip (Mass = 12,306) and AA Clip (Mass = 11,806) of the NIST mAb (H2L2) can take place on a single light chain giving rise to two complementary fragments, H2L2-FI Clip and H2L2-AA Clip. In addition, when FI and AA clipping occurs in both light chains simultaneously, results in two additional complementary fragments:

H2L2-2FI Clip and H2L2-2AA Clip of the NIST mAb sequence (Fig. 4A). To mimic production and processing effects we explored various digestion conditions, including treatment by cathepsin D at pH 5, and cathepsin L at pH 3 and 7. As expected, this strategy generated a wide spectrum of clipping variant peaks, 1–10, which are detected in the overlaid GXII electropherograms (Fig. 4B). We observe a time dependent increase in some peaks when comparing the peak area for each species after incubation with respective enzyme and pH condition for an additional day (after day 2 vs. day 1) (Fig. 4B Inset). The assignment of peaks associated with likely fragments is challenging as the corresponding masses are unknown. GXII electropherograms are, therefore, typically used as a stability indicating assay and not to identify fragment species. Herein, we can assign peaks with most probable fragments using a look up table that was created using the calibration curve generated with CZE-MS masses and GXII migration time.

Using the strategy described above, we can now annotate GXII peaks 1–10 (Table S4) based on our predicted fragment masses due to the activity of Cathepsin L at pH 3 and 7 and endogenous fragments found in NIST mAb. We assigned of Peak 3 and Peak 4 as light (L) and heavy (H) chain. Peak 10 and Peak 5 were assigned as HHL and HH based on the notion that they were also present in the unstressed NIST mAb. Cathepsin induced clips were next categorized based on the most likely fragment masses derived from the linear regression. We tentatively assigned Peak 1 as an L-AA Clip by mass although it was exclusively detected when IgG1 was incubated with Cathepsin L at pH 3. Peak 2 corresponds to an L-FI Clip found exclusively when IgG1 was incubated with Cathepsin D at pH 5. Finally, we assigned complementary fragments of the AA and FI clips. Incubation of IgG1 with Cathepsin L at pH 3 produced Peak 9, which was assigned as the H2L2-2 AA Clip. Exposing IgG1 to Cathepsin D at pH 5 exclusively generated Peak 8, which was categorized as the H2L2- 2 FI Clip. Analogously, Peaks 7 and 6 were assigned as the HHL-FI Clip and HHL-AA Clip, respectively. The initial assignment of heavy chain fragments at Peak 4 (H) and Peak 5 (HH) were tentative annotations as there was no CZE-MS evidence, these assignments still require further verification (Fig. 4C).

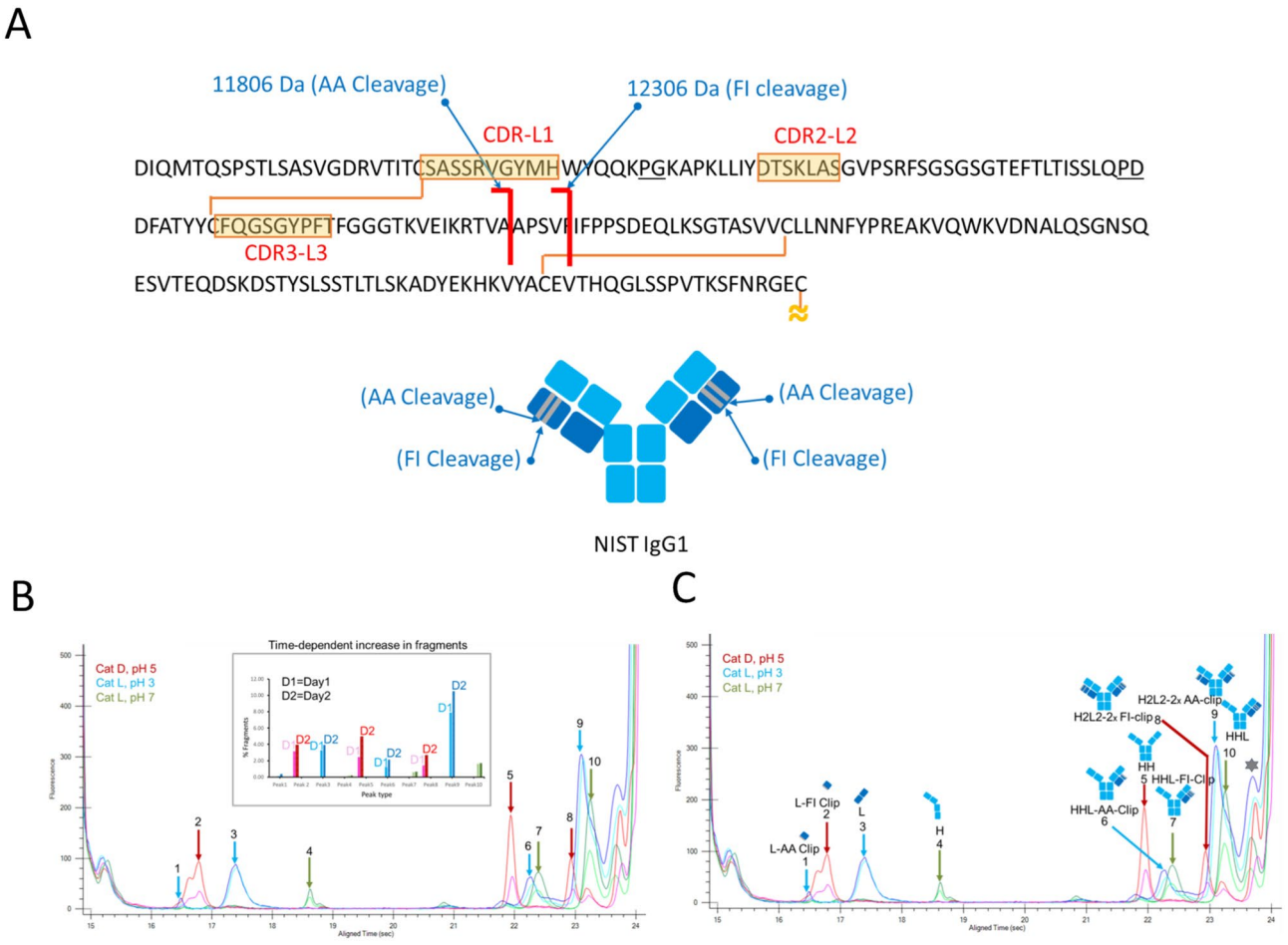

Fig. 4. Analysis of Cathepsin D- and Cathepsin L-induced degradation products of NISTmAb (**A**) NISTmAb light chain sequence with site-specific AA cleavage and FI cleavage. Cathepsin-D and -L cleavages of light chain. (**B**) Overlayed LabChip GXII electropherogram corresponding to incubation of NISTmab with Cathepsin D at pH 5 (red), Cathepsin L at pH 3 (blue) and Cathepsin L at pH 7 (green) for 1 day (D1) and 2 days (D2) (Inset shows the intensity of peaks 1–10 for cathepsin degradation products at D1 and D2) (**C**) Assignment of Cathepsin cleavage products of NISTmAb to peaks in the electropherogram with masses obtained from the regression analysis. *Note*: In addition to peaks 1–10, additional peaks are observed in the electropherogram.

CZE-MS identification of cathepsin-induced clips

CZE-MS provides the unambiguous identification of degradation products. The high-resolution of CE separations coupled with accurate mass spectrometry detection makes CZE-MS a powerful technique in detecting low-abundant clips. To validate the composition of the impurity peaks observed in GXII, we conducted CZE-MS experiments using samples treated with cathepsin L (at pH 3 and 7) or cathepsin D (at pH 5) to generate a library of MS1 and some instances MS2 spectra. We confirmed the identities of the AA and FI clips of the light chain (L) by mass deconvolution of the extracted ion electropherograms of fragments corresponding to deconvoluted masses at 11.8 kDa, 12.3 kDa respectively. The L-AA Clip was identified from cathepsin L digestion (pH 3) while L-FI Clip was identified in Cathepsin L (pH7) and Cathepsin D (pH5) digestion. In addition, a larger 97.6 kDa fragment was identified, which is the result of FL Clip at each heavy chain below the hinge disulfides (Fig. 5A–D). It is important to note that CZE separations partially resolved L-AA, L-FI and H-FL clips but were fully resolved by their extracted ion electropherograms.

By performing tandem mass spectrometry using CZE-MS/MS analysis, we unambiguously identified both as an AA Clip and a FI Clip. The sample treated with Cathepsin L at pH 3 shows multiple charge sates of the AA Clip Fig. 5E). The + 8-charge state corresponding to the AA Clip was first mass selected and subjected to collision-induced dissociation (CID) to obtain fragment ion spectra for further characterization. The observed fragment ions in the MS/MS spectra matched with theoretical fragmentation pattern of the respective AA Clip sequence.

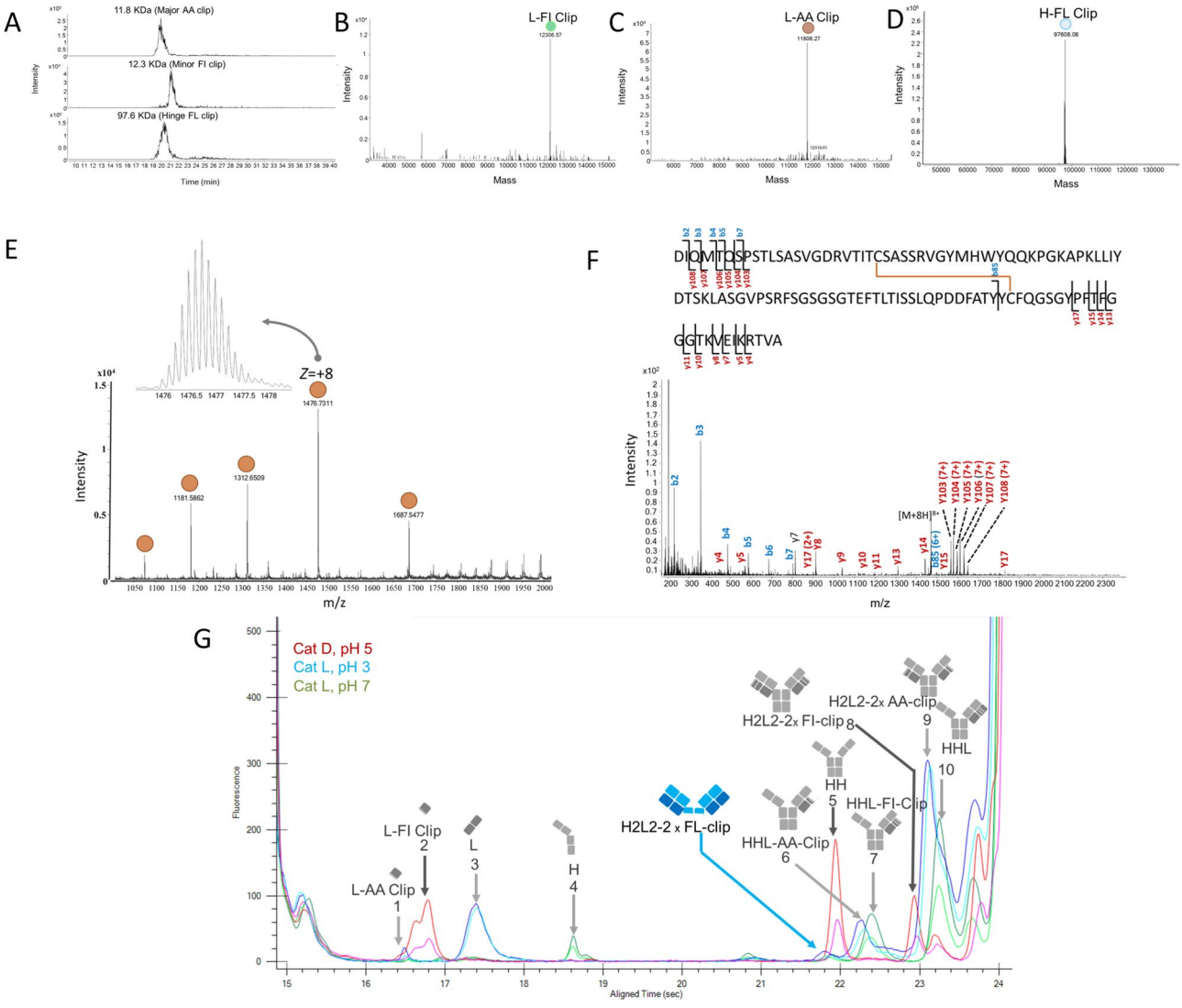

Fig. 5. GXII peak assignment via CZE-MS identification of cathepsin degradation products of NISTmAb (**A**) Extracted ion chromatograms of AA, FI and FL clips. (**B**) Deconvoluted mass spectrum of FI Clip (**C**), Deconvoluted mass spectrum of AA Clip (**D**) Deconvoluted mass spectrum of and FL Clip (**E**). Mass spectrum of Cathepsin L-induced AA Clip shown in brown circles. Inset shows the isotope distribution of Z = + 8 ion). (**F**) CZE-MS/MS product ion spectrum of AA Clip resulting from collision induced dissociation of Z = + 8. Inset shows the fragment ion map of the AA Clip. (**G**). Assignment of a new peak associated with FL Clip along with electropherogram peak assignments from Fig. 4C.

The *b* and *y* ions derived from the cleavage sites were identified and agreed with the expected sequence (Fig. 5F). Similarly, samples treated with Cathepsin L at pH 7 show a suspected FI Clip mass, which we subjected to CID. Fragment ion spectra confirmed the sequence corresponding to the FI Clip (Fig. S2). This observation validates the mass of the in silico fragment generated via Cathepsin L digestion at pH 3 and 7, respectively. The masses of these clips were obtained using *in silco* predictions and were validated by mass and sequence confirmation using CZE-MS. This data further reinforces the value of CZE-MS for the identification of impurity peaks in LabChip GXII. In addition, during the analysis of the CZE-MS results, a species with a molecular weight of 97,608.08 Da was observed (Fig. S3). The deconvolution mass of this ion corresponds to H2L2 of FL Clip (2x) that we were able to assign a low intense GXII peak (Fig. 5G).

Reassignment of GXII impurity peaks based on CZE-MS identifications

We used CZE-MS data to build a comprehensive list of identifications to further improve LabChip GXII peak assignments. Interestingly, a species with a molecular weight of 47,628.20 Da that was identified by CZE-MS, was confidently assigned as a Fab' fragment generated via two HT Clips (H2L2-2 × HT-clip) in IgG treated with Cathepsin L at pH 7 (Fig. 6A). The mass of the HT Clip, when applied to the regression analysis, closely matched Peak 4. However, in the original GXII results, Peak 4 was assigned as heavy chain. Unlike the heavy chain that is glycosylated, Peak 4 does not contain any glycans which further invalidated our initial assignment. This discrepancy highlights potential errors that could arise from GXII peak assignments based on the GXII calibration that relies on the ladder of low molecular weight markers. Consequently, we performed reassignments for peaks that were possibly misidentified, aiming to ensure accurate identification of the impurity peaks in GXII via assigning masses that were identified by using the spectral library obtained via CZE-MS. CZE-MS data was used to revalidate previously identified peaks as well as correctly assign misidentified peaks. The CZE-MS data of the sample treated with cathepsin L at pH 7 and Cathepsin D at pH 5, showed a species with a charge state envelope centered at $Z = +44$ corresponding to a deconvoluted molecular mass of 100,593.14 Da. We assigned that mass as a H2L2-2 × L-SS Clip (Fig. 6B) Close examination shows Fc glycans on this fragment (Fig. 6B inset). This discovery revealed the possibility of a misassignment of the GXII peak 5 as HH, which could potentially be reassigned as H2L2-2 × L-SS Clip. Additionally, in the data from the sample treated with Cathepsin D at pH 5, the species we identified as L-FI Clip with molecular weight of 12,306.57 Da (Fig. 4B), can now be further evaluated by including identifications or masses derived from the CZE-MS spectral library. For example, we observed two clusters of adjacent mass spectral signals with deconvoluted molecular weights of 18,371.08 Da and 18,503.06 Da (Fig. 6C) exclusively in the sample treated with cathepsin D at pH 5. These adjacent species may explain for the unresolved doublet of peak 2. With accurate molecular weights being obtained by the CZE-MS analyses, we identified the most likely assignment for each mass for the observed peaks in the GXII electropherogram. The most likely assignment is derived by minimizing the time errors (theoretical migration time—observed migration time for an observed mass) obtained by the linear calibration of masses and 1/time (mobility) in (Fig. 2C). For example, peak 4 was assigned with fragment the mass of the H2L2-2 × HT-Clip derived from the spectral library having the lowest migration time error. Notably, our initial assignment of Peak 4 as heavy chain (H) had the second lowest migration time error. Likewise, peak 5 was assigned to H2L2-2 × L-SS Clip despite peak 5 also have a similar migration time error for HH. However, HH was not observed as an impurity in the CZE-MS derived spectral library. The doublets of Peak 2 most accurately matched 18,371.08 Da and 18,503.06 Da masses detected in CZE-MS as a doublet. It is conceivable that Peak 2 is composed of L-FI Clip and additional clipped species at masses 18.3 kDa and 18.5 kDa that are poorly resolved. In addition, peaks 1, 8 and 10 were validated for their assigned identifications of Light chain-AA Clip, H2L2-2SS Clip and HHL respectively as also having the minimum migration time errors for the assigned mass (Fig. 6D). Utilizing the CZE-MS data, we achieved a highly accurate assignment of all impurity peaks in the GXII electropherogram (Fig. 6E).

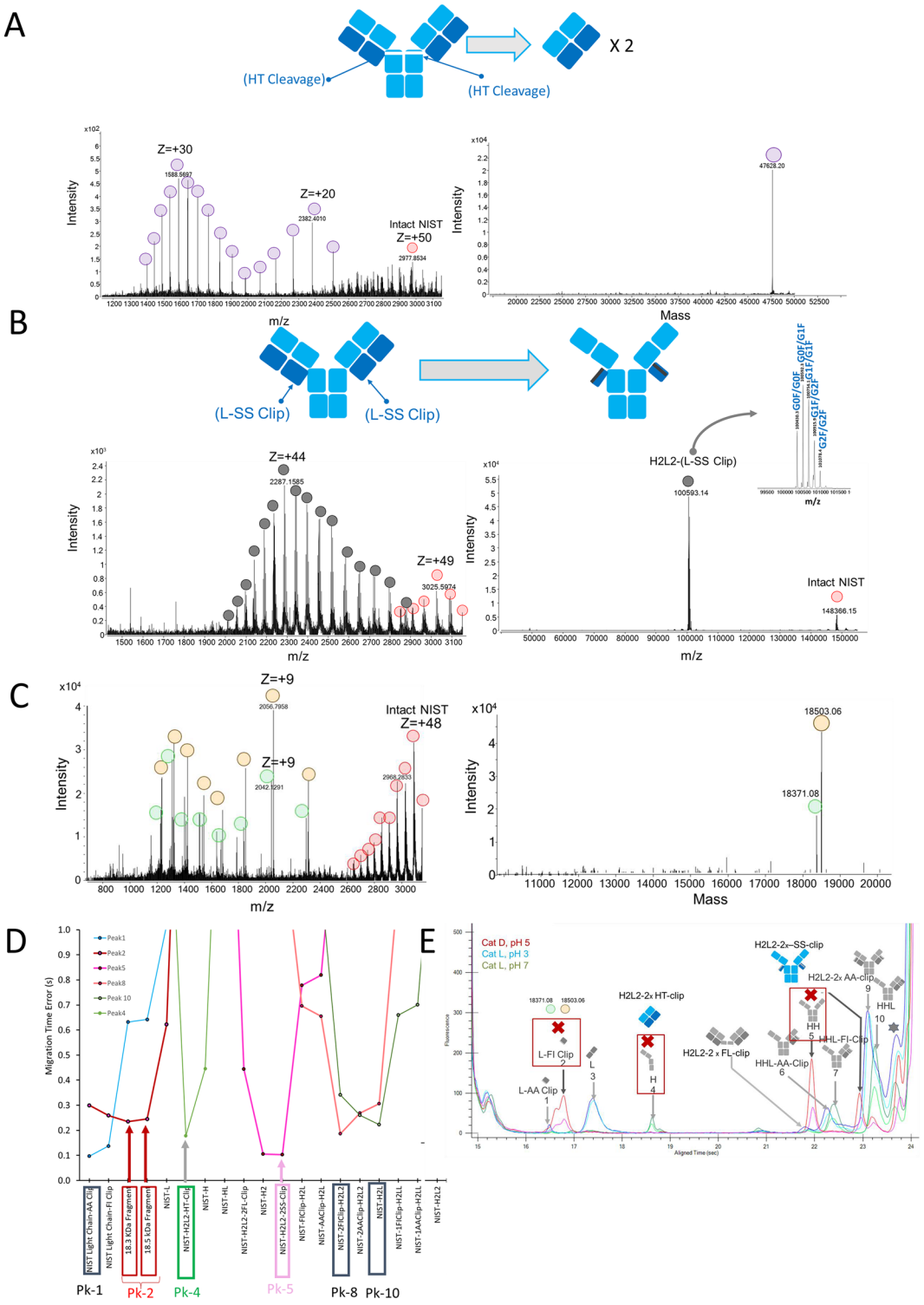

Fig. 6. Reannotation of GXII NISTmAb cathepsin-induced degradation products by accurate assignment of MS identifications (**A**) CZE-MS identification of 47.6 kDa F_{ab}' fragment due to heavy chain HT cleavage with Cathepsin L incubated at pH 7. (**B**) CZE-MS identification of 100.5 kDa (H2L2-SS Clip) degradation product due to light chain SS cleavage formed by Cathepsin D incubated at pH 5. Inset shows Fc glycans (**C**) CZE-MS mass spectrum and corresponding deconvolution spectrum of 18.3 kDa and 18.5 kDa fragment masses formed by cathepsin D incubation at pH 5. (**D**) Migration time error versus degradation products obtained from the spectral library. Note: migration time errors are shown for GXII peaks 1, 2, 4, 5, 8 and 10. The identity of the library species is assigned to the peak in electropherogram when migration time error is minimized. (**E**) Reassignment of peaks 2, 4 and 5 in GXII electropherogram with accurate identities. Peaks 1, 8 and 10 are validated to the initial assignments in Fig. 4C.

Conclusion

This study presents a method for detecting and identifying product and process impurity peaks in size-based electrophoretic separation of antibodies using CZE-MS. This approach provides a means to identity of certain size-based impurities, complementing the information obtained from the electrophoretic separation alone. Antibody degradation products were generated under forced stress to mimic conditions that are prevalent during bioprocessing and storage of biotherapeutic antibodies. LabChip GXII electrophoresis enabled the rapid (< 1 min) size separation and detection of impurity peaks, while capillary zone electrophoresis-mass spectrometry (CZE-MS) facilitated the generation of a spectral library of antibody impurities via high-resolution separation and accurate identification. CZE-MS identification approaches are complementary to chromatographic methods hyphenated with MS. Wang et al., demonstrated a HILIC-MS method for the direct detection of antibody impurities. However, this approach required deglycosylation of antibodies while peak resolution was an impediment to accurately quantify impurities compared to electrophoretic separation[48]. In our method, a spectral library was generated using accurate mass information of antibody impurities that increased the accuracy of the calibration curve allowing improved peak assignments in GXII electropherograms. The method was applied to several antibodies in systems that emulate cell culture harvest conditions and forced degradation. Our results demonstrate the feasibility of this approach for identifying and assigning impurity peaks in GXII size-based electropherograms. This method holds potential for ensuring the quality and safety of antibody therapeutics and improving the production processes. Future explorations are needed to evaluate its broad applicability towards comprehensive characterization of a variety of biotherapeutic modalities. Herein, partial validation of impurity peaks was performed using CZE-MS/MS via collision induced dissociation (CID). However, new advancements in mass spectrometry dissociation methods, such electron capture dissociation (ECD), will enhance the confidence in the identification of degradation products. High-resolution CZE separations will enable the generation of comprehensive MS/MS spectral libraries that will further improve in the precise annotation of peaks in size-based electrophoretic methods. We envision that this method will have broad applicability in unambiguous and rapid assignment of a clipped sites, and importantly, for monitoring mispaired chains in the growing number of multi-specific antibodies that are emerging as anticancer therapeutics.

Data availability

The CZE-MS datasets and LabChip GXII data sets generated during the current study is available in the figshare repository. https://doi.org/https://doi.org/10.6084/m9.figshare.25777530 https://doi.org/https://doi.org/10.6084/m9.figshare.25777749 https://doi.org/https://doi.org/10.6084/m9.figshare.25777785 https://doi.org/https://doi.org/10.6084/m9.figshare.25777818.

Received: 9 May 2024; Accepted: 22 August 2024
Published online: 30 August 2024

References

1. Ha, T. K., Kim, D., Kim, C. L., Grav, L. M. & Lee, G. M. Factors affecting the quality of therapeutic proteins in recombinant Chinese hamster ovary cell culture. *Biotechnol. Adv.* **54**, 107831 (2022).
2. Atsumi, Y. *et al.* Clip formation in the complementarity determining region of bevacizumab lowers monomer stability and affinity for both FcRn and FcγR: A comprehensive characterization of the clipped variant including its higher order structure. *J. Pharm. Sci.* **111**(12), 3243–3250 (2022).
3. Atsumi, Y., Sakurai, N., Nishimura, K., Yamazaki, K. & Wakamatsu, K. Identification and characterization of a monoclonal antibody variant species with a clipping in the complementarity determining region isolated by size exclusion chromatography under native conditions. *J. Pharm. Sci.* **110**(10), 3367–3374 (2021).
4. Vlasak, J. & Ionescu, R. Fragmentation of monoclonal antibodies. *mAbs* **3**(3), 253–263 (2011).
5. Rustandi, R. R. & Wang, Y. Use of CE-SDS gel for characterization of monoclonal antibody hinge region clipping due to copper and high pH stress. *Electrophoresis* **32**(21), 3078–3084 (2011).
6. Alhazmi, H. A. & Albratty, M., Analytical techniques for the characterization and quantification of monoclonal antibodies. *Pharmaceuticals (Basel)* **16**(2) (2023).
7. Lechner, A. *et al.* Insights from capillary electrophoresis approaches for characterization of monoclonal antibodies and antibody drug conjugates in the period 2016–2018. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1122–1123**, 1–17 (2019).
8. Breadmore, M. C. Capillary and microchip electrophoresis: Challenging the common conceptions. *J. Chromatogr. A* **1221**, 42–55 (2012).
9. Rustandi, R. R., Washabaugh, M. W. & Wang, Y. Applications of CE SDS gel in development of biopharmaceutical antibody-based products. *Electrophoresis* **29**(17), 3612–3620 (2008).
10. Wiesner, R., Scheller, C., Krebs, F., Wätzig, H. & Oltmann-Norden, I. A comparative study of CE-SDS, SDS-PAGE, and Simple Western: Influences of sample preparation on molecular weight determination of proteins. *Electrophoresis* **42**(3), 206–218 (2021).
11. Nunnally, B. *et al.* A series of collaborations between various pharmaceutical companies and regulatory authorities concerning the analysis of biomolecules using capillary electrophoresis. *Chromatographia* **64**(5–6), 359–368 (2006).
12. Kumar, R., Guttman, A. & Rathore, A. S. Applications of capillary electrophoresis for biopharmaceutical product characterization. *Electrophoresis* **43**(1–2), 143–166 (2022).
13. Filep, C. & Guttman, A. Capillary sodium dodecyl sulfate gel electrophoresis of proteins: Introducing the three dimensional Ferguson method. *Anal. Chim. Acta* **1183**, 338958 (2021).
14. Yagi, Y., Kakehi, K., Hayakawa, T. & Suzuki, S. Application of microchip electrophoresis sodium dodecyl sulfate for the evaluation of change of degradation species of therapeutic antibodies in stability testing. *Anal. Sci.* **30**(4), 483–488 (2014).
15. Kahle, J., Maul, K. J. & Watzig, H. The next generation of capillary electrophoresis instruments: Performance of CE-SDS protein analysis. *Electrophoresis* **39**(2), 311–325 (2018).
16. Duhamel, L. *et al.* Therapeutic protein purity and fragmented species characterization by capillary electrophoresis sodium dodecyl sulfate using systematic hybrid cleavage and forced degradation. *Anal. Bioanal. Chem.* **411**(21), 5617–5629 (2019).
17. Rathore, D. *et al.* The role of mass spectrometry in the characterization of biologic protein products. *Expert Rev. Proteom.* **15**(5), 431–449 (2018).
18. Liu, S. & Schulz, B. L. Biopharmaceutical quality control with mass spectrometry. *Bioanalysis* **13**(16), 1275–1291 (2021).

19. Gunawardena, H. P., Jayatilake, M. M., Brelsford, J. D. & Nanda, H. Diagnostic utility of N-terminal TMPP labels for unambiguous identification of clipped sites in therapeutic proteins. *Sci. Rep.* **13**(1), 18602 (2023).
20. Sandra, K., Vandenheede, I. & Sandra, P. Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization. *J. Chromatogr. A* **1335**, 81–103 (2014).
21. Duivelshof, B. L. *et al.* Therapeutic Fc-fusion proteins: Current analytical strategies. *J. Sep. Sci.* **44**(1), 35–62 (2021).
22. Tassi, M. *et al.* Advances in native high-performance liquid chromatography and intact mass spectrometry for the characterization of biopharmaceutical products. *J. Sep. Sci.* **41**(1), 125–144 (2018).
23. Qu, M. *et al.* Qualitative and quantitative characterization of protein biotherapeutics with liquid chromatography mass spectrometry. *Mass Spectrom. Rev.* **36**(6), 734–754 (2017).
24. Rashid, M. H. Full-length recombinant antibodies from *Escherichia coli*: Production, characterization, effector function (Fc) engineering, and clinical evaluation. *MAbs* **14**(1), 2111748 (2022).
25. Gahoual, R., Beck, A., Leize-Wagner, E. & Francois, Y. N. Cutting-edge capillary electrophoresis characterization of monoclonal antibodies and related products. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1032**, 61–78 (2016).
26. Lu, J. J., Zhu, Z., Wang, W. & Liu, S. Coupling sodium dodecyl sulfate-capillary polyacrylamide gel electrophoresis with matrix-assisted laser desorption ionization time-of-flight mass spectrometry via a poly(tetrafluoroethylene) membrane. *Anal. Chem.* **83**(5), 1784–1790 (2011).
27. Li, M. *et al.* Identification of a monoclonal antibody clipping variant by cross-validation using capillary electrophoresis—Sodium dodecyl sulfate, capillary zone electrophoresis—Mass spectrometry and capillary isoelectric focusing - mass spectrometry. *J. Chromatogr. A* **1684**, 463560 (2022).
28. Romer, J., Stolz, A., Kiessig, S., Moritz, B. & Neususs, C. Online top-down mass spectrometric identification of CE(SDS)-separated antibody fragments by two-dimensional capillary electrophoresis. *J. Pharm. Biomed. Anal.* **201**, 114089 (2021).
29. Sarkozy, D. & Guttman, A. Analysis of peptides and proteins by native and SDS capillary gel electrophoresis coupled to electrospray ionization mass spectrometry via a closed-circuit coaxial sheath flow reactor interface. *Anal. Chem.* **95**(18), 7082–7086 (2023).
30. Li, W. *et al.* Identification and characterization of monoclonal antibody fragments cleaved at the complementarity determining region using orthogonal analytical methods. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1048**, 121–129 (2017).
31. Wang, W. H., Cheung-Lau, J., Chen, Y., Lewis, M. & Tang, Q. M. Specific and high-resolution identification of monoclonal antibody fragments detected by capillary electrophoresis-sodium dodecyl sulfate using reversed-phase HPLC with top-down mass spectrometry analysis. *MAbs* **11**(7), 1233–1244 (2019).
32. Cao, M. *et al.* Identification of a CE-SDS shoulder peak as disulfide-linked fragments from common C(H)2 cleavages in IgGs and IgG-like bispecific antibodies. *MAbs* **13**(1), 1981806 (2021).
33. Wojcik, R., Dada, O. O., Sadilek, M. & Dovichi, N. J. Simplified capillary electrophoresis nanospray sheath-flow interface for high efficiency and sensitive peptide analysis. *Rapid Commun. Mass Spectrom.* **24**(17), 2554–2560 (2010).
34. Sun, L., Zhu, G., Zhang, Z., Mou, S. & Dovichi, N. J. Third-generation electrokinetically pumped sheath-flow nanospray interface with improved stability and sensitivity for automated capillary zone electrophoresis-mass spectrometry analysis of complex proteome digests. *J. Proteome Res.* **14**(5), 2312–2321 (2015).
35. Dai, J., Lamp, J., Xia, Q. & Zhang, Y. Capillary isoelectric focusing-mass spectrometry method for the separation and online characterization of intact monoclonal antibody charge variants. *Anal. Chem.* **90**(3), 2246–2254 (2018).
36. Xu, T., Han, L., George Thompson, A. M. & Sun, L. An improved capillary isoelectric focusing-mass spectrometry method for high-resolution characterization of monoclonal antibody charge variants. *Anal. Methods* **14**(4), 383–393 (2022).
37. Lamp, J., Ikonomova, S. P., Karlsson, A. J., Xia, Q. & Wang, Y. Online capillary electrophoresis—Mass spectrometry analysis of histatin-5 and its degradation products. *Analyst* **145**(14), 4787–4794 (2020).
38. Shen, X., Liang, Z., Xu, T., Yang, Z., Wang, Q., Chen, D., Pham, L., Du, W. & Sun, L. Investigating native capillary zone electrophoresis-mass spectrometry on a high-end quadrupole-time-of-flight mass spectrometer for the characterization of monoclonal antibodies. *Int. J. Mass Spectrom.* 462 (2021)
39. Sun, L., Zhu, G., Yan, X. & Dovichi, N. J. High sensitivity capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry for the rapid analysis of complex proteomes. *Curr. Opin. Chem. Biol.* **17**(5), 795–800 (2013).
40. Han, M., Rock, B. M., Pearson, J. T. & Rock, D. A. Intact mass analysis of monoclonal antibodies by capillary electrophoresis-Mass spectrometry. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1011**, 24–32 (2016).
41. Kshirsagar, R., McElearney, K., Gilbert, A., Sinacore, M. & Ryll, T. Controlling trisulfide modification in recombinant monoclonal antibody produced in fed-batch cell culture. *Biotechnol. Bioeng.* **109**(10), 2523–2532 (2012).
42. Seibel, R. *et al.* Impact of S-sulfocysteine on fragments and trisulfide bond linkages in monoclonal antibodies. *MAbs* **9**(6), 889–897 (2017).
43. Luo, H. *et al.* Cathepsin L causes proteolytic cleavage of Chinese-Hamster-Ovary cell expressed proteins during processing and storage: Identification, characterization, and mitigation. *Biotechnol. Prog.* **35**(1), e2732 (2019).
44. Cui, T. *et al.* Removal strategy on protein A chromatography, near real time monitoring and characterisation during monoclonal antibody production. *J. Biotechnol.* **305**, 51–60 (2019).
45. Gu, S. *et al.* Characterization of trisulfide modification in antibodies. *Anal. Biochem.* **400**(1), 89–98 (2010).
46. Tarelli, E. & Corran, P. H. Ammonia cleaves polypeptides at asparagine proline bonds. *J. Pept. Res.* **62**(6), 245–251 (2003).
47. Caval, T. *et al.* The lysosomal endopeptidases Cathepsin D and L are selective and effective proteases for the middle-down characterization of antibodies. *FEBS J.* **288**(18), 5389–5405 (2021).
48. Wang, S., Liu, A. P., Yan, Y., Daly, T. J. & Li, N. Characterization of product-related low molecular weight impurities in therapeutic monoclonal antibodies using hydrophilic interaction chromatography coupled with mass spectrometry. *J. Pharm. Biomed. Anal.* **154**, 468–475 (2018).

Acknowledgements

The authors would like to thank the support of Johnson & Johnson Innovative Medicine. We thank Joshua Justice from Johnson & Johnson for critically reviewing the manuscript. Authors also thank John Sausen from Agilent Technologies and Bob Hepler from Johnson & Johnson for assistance in the collaboration.

Author contributions

Conceived the idea, designed research and supervised the project: H.P.G. Performed the CZE-MS experiments: J.L., M.D.K., Y.Z., Q.L., J.H., Q.X Prepared samples: J.L., Y.Z., Analyzed the MS data: H.P.G., J.L., M.K. Performed LabChip experiments: C.D., B.N. Contributed reagents/materials/analysis tools: H.P.G., H.N., Q.X, J.L., Wrote the paper: H.P.G., Q.L., J.H., with input from all authors. All authors reviewed the manuscript and approved. Agilent approval DE0141219.

Funding
This research received no specific grant from any funding agency in the public, commercial, or non-for-profit sectors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-70914-5.

Correspondence and requests for materials should be addressed to H.P.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.