Journal of Chromatography B, 879 (2011) 2877–2884

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

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

Detection and identification of a serine to arginine sequence variant in a therapeutic monoclonal antibody

Diya Ren[a,*], Jian Zhang[a], Ross Pritchett[a], Hongbin Liu[b], Jennifer Kyauk[a], Jun Luo[a], Ashraf Amanullah[a]

[a] Oceanside Pharma Technical Development, Genentech, Oceanside, CA 92056, United States
[b] Protein Analytical Chemistry, Genentech, South San Francisco, CA 92080, United States

ARTICLE INFO

Article history:
Received 1 June 2011
Accepted 14 August 2011
Available online 22 August 2011

Keywords:
Monoclonal antibody
Sequence variant
HPLC–UV/MS/MS
Peptide mapping
Imaged capillary isoelectric focusing

ABSTRACT

Sequence variants, also known as unintended amino acid substitutions in the protein primary structure, are one of the critical quality attributes needed to be monitored during process development of monoclonal antibodies (mAbs). Here we report on analytical methods for detection and identification of a sequence variant in an IgG1 mAb expressed in Chinese hamster ovary (CHO) cells. The presence of the sequence variant was detected by an imaged capillary isoelectric focusing (ICIEF) assay, showing a new basic species in mAb charge variant profile. The new basic variant was fractionated and enriched by ion-exchange chromatography, analyzed by reduced light and heavy chain mass determination, and characterized by HPLC–UV/MS/MS of tryptic and endoproteinase Lys-C peptide maps. A Serine to Arginine sequence variant was identified at the heavy chain 441 position (S441R), and confirmed by using synthetic peptides. The relative level of the S441R variant was estimated to be in the range of 0.3–0.6% for several mAb batches analyzed via extracted ion chromatogram (EIC). This work demonstrates the effectiveness of using integrated analytical methods to detect and identify protein heterogeneity and the importance of monitoring product quality during mAb bioprocess development.

1. Introduction

Recombinant human monoclonal antibodies (rhumAbs) are rapidly being developed as therapeutics to treat cancers, infections, and autoimmune diseases [1]. To date, over 20 monoclonal antibodies have been approved by the Food and Drug Administration (FDA) for clinical use and hundreds are in the development or clinical trials for treatment of various diseases [1–3]. Through development stages, product quality attributes of therapeutic proteins are closely monitored from early phase pre-clinical development to late stage clinical development, and continuing in the commercial manufacturing process, to ensure product quality, safety, efficacy, and process consistency for human use. These attributes, such as identify, aggregates, fragments, charge variants, glycan content, primary structure, potency, etc., are routinely analyzed for batch release and characterization purpose.

One of the challenges in mAb characterization is to detect trace amount of sequence variants in the protein primary structure during clone selection and cell culture process development. Sequence variants may occur either at the DNA level by mutation [4,5] or at the protein level by misincorporation [6–8]. Although not as common as in recombinant proteins expressed in *Escherichia coli*, low levels of sequence variants have been observed in mAbs expressed in Chinese hamster ovary (CHO) cells [4–12]. Harris et al. [4] reported a Y376Q (tyrosine at position 376 replaced by glutamine) variant in the heavy chain of human epidermal growth factor receptor-2 (HER2). Study showed levels of the Y376Q variant were inversely proportional to cell age. A S167R (serine at position 167 replaced by arginine) variant [9] and a phenylalanine to leucine variant were also reported [12]. All these three antibodies were expressed in CHO cell lines and the variants were verified by polymerase chain reaction (PCR) analysis to be a genomic nucleotide mutation at the DNA level. Recently, two examples of misincorporation at the protein level were reported. One was the substitution of asparagine by serine due to the starvation for asparagine in the cell culture medium [6,7], and the other was codon-specific serine to asparagine mistranslation [8]. Therefore, we routinely perform sequence variant analysis during the clone selection phase by combining HPLC–UV/MS/MS characterization of tryptic and chymotryptic peptide maps with the Mascot Error Tolerant Search (ETS) data analysis package [10]. Sometimes sequence variants may be accompanied by changes in the electrophoretic

Abbreviations: mAb, monoclonal antibody; CHO, Chinese hamster ovary; ETS, error-tolerant search; EIC, extracted ion chromatogram; ICIEF, imaged capillary isoelectric focusing; IEC, ion exchange chromatography; HPLC–UV/MS/MS, high performance liquid chromatography with ultraviolet spectrophotometer and tandem mass spectrometry detection; CpB, carboxypeptidase B.

* Corresponding author at: Genentech, One Antibody Way, Oceanside, CA 92056, United States. Tel.: +1 760 231 3023; fax: +1 760 231 2464.
E-mail addresses: ren.diya@gene.com, diya.ren@gmail.com (D. Ren).

1570-0232/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2011.08.015

or chromatographic separation profile of the mAb. In this case, the identification of the root cause may require further investigation, typically through fraction isolation and enrichment of the new species followed by peptide mapping to establish the correlation between the changes in the protein primary structure and the observations in electrophoretic or chromatographic profiles of the mAb.

Here we report a study to identify an unintended amino acid substitution, namely Ser to Arg, in a therapeutic monoclonal antibody expressed in CHO cells during the cell culture development process. This study demonstrates the effectiveness of integrated analytical technologies to identify protein heterogeneity and highlights the necessity to monitor product quality throughout the mAb cell culture process development.

2. Materials and methods

2.1. Materials

The recombinant human monoclonal antibodies analyzed in this study were expressed in CHO and purified at Genentech (Oceanside, CA). Peptides RLSLSPG and LSLSPG were synthesized and purified at Genentech (South San Francisco, CA).

2.2. Reagents

Tris(hydroxymethyl)aminomethane (TRIS base), Tris–HCl, calcium chloride, dithiothreitol (DTT), idoacetic acid (IAA), 0.5 M ethylenediaminetetraacetic acid (EDTA) solution, sodium sulfite (anhydrous), sodium tetrathionate (dihydrate) and Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and urea were purchased from Sigma (St. Louis, MO). Sodium sulfate was obtained from Mallinckrodt (Phillipsburg, NJ). HPLC grade acetonitrile and water were acquired from J.T. Baker (Philipsburg, NJ). Sequence grade formic acid (FA), trifluoroacetic acid (TFA), and 8 M guanidine HCl were purchased from Thermo Scientific (Rockford, IL). Mass spectrometry grade trypsin (Gold) was from Promega (Madison, WI) and sequencing grade endoproteinase Lys-C were from Roche (Mannheim, Germany). DFP treated carboxypeptidase B (CpB) was purchased from Roche Diagnostics (Indianapolis, IN). Pharmalytes were obtained from GE Healthcare Biosciences (Uppsala, Sweden). 1% Methyl cellulose solution and pI markers were purchased from Convergent Bioscience (Toronto, Canada).

2.3. Imaged capillary isoelectric focusing (ICIEF) analysis

The ICIEF was performed on an iCE280 analyzer (Convergent Bioscience, Toronto, Canada) with a fluorocarbon-coated capillary cartridge (100 μm ID × 5 cm). The ampholyte solution contained a mixture of methyl cellulose (MC), urea, carrier pharmalytes (pH 5–8 and pH 8–10.5), and pI markers (5.5 and 9.77) in purified water. Samples were diluted and then treated with CpB enzyme at 1/100 w/w enzyme to protein ratio prior to mixing with the ampholyte solution. The final mAb concentration for ICIEF analysis was at 0.25 mg/mL. The ICIEF analysis was conducted by a potential of 3000 V for 10 min to allow mAb charge variants to migrate per their pIs. An image of the focused mAb charge variants was captured by passing 280 nm ultraviolet light through the capillary and into the lens of a charge-coupled device digital camera.

2.4. Ion exchange chromatography (IEC) analysis

IEC for mAb charge variant analysis was performed on a Dionex (Sunnyvale, CA) ProPac® WCX-10 (4.0 mm × 250 mm) column using an Agilent (Santa Clara, CA) 1200 HPLC equipped with a fraction collector. Mobile phase A was 20 mM Tris buffer, pH 8.5 (A) and mobile phase B was 100 mM Na_2SO_4 in 20 mM Tris, pH 8.5 (B). A linear gradient was used to separate mAb charge variant in a total of 55 min run time. The flow rate was 1 mL/min with the detection at 280 nm.

For IEC fractionation, the mAb was treated with CpB at a ratio of 1/100 w/w enzyme to protein to remove C-terminal lysine residues. The mAb main and basic charge variant fractions separated from IEC were collected and then concentrated using Amicon 10 kDa Ultracel®-10K membrane (Millipore, Billerica, MA) prior to ICIEF, reduced mass measurement and peptide map analysis.

2.5. Reduction, alkylation and tryptic digestion

The mAb samples were trypsin digested as described elsewhere [8,10,13]. Briefly, the mAb was diluted to about 1 mg/mL in a denaturing buffer (6 M guanidine hydrochloride, 360 mM Tris and 2 mM EDTA, pH 8.6). 20 μL of 1 M DTT was added to reduce disulfide bonds in the mAb. The solution was then incubated for 1 h at 37 °C, after which it was cooled to room temperature. The alkylation reaction was performed by adding 50 μL of freshly prepared 1 M IAA and incubating at room temperature for 15 min in the dark. 10 μL of 1 M DTT was added to stop the alkylation. The DTT-reduced and IAA-alkylated mAb was buffer exchanged into digestion buffer (25 mM Tris, 2 mM EDTA, pH 8.2) using Sephadex G-25M PD-10 columns (GE Healthcare, Uppsala, Sweden) prior to trypsin digestion, which was carried out with a 1/50 w/w ratio of trypsin to mAb at 37 °C for 5 h. The digestion was stopped by adding 10% TFA to a final concentration of 0.3%. Digests were stored at −70 °C prior to injection for HPLC–MS/MS analysis.

2.6. Sulfitolysis and endoproteinase Lys-C digestion

The sulfitolysis reaction was performed by incubating 50 μL of 20 mg/mL mAb with 950 μL sulfitolysis reagent (6 M guanidine–HCl, 360 mM Tris, 2 mM EDTA, 127 mM sodium sulfite, 26 mM sodium tetrathionate at pH 8.6) at 37 °C for 20 min. The reaction mixture was buffer exchanged into endoproteinase Lys-C digestion buffer (25 mM Tris, 1 mM EDTA, pH 8.3) using Sephadex G-25 M PD-10 columns. The digestion was performed by adding endoproteinase Lys-C at 1/80 w/w ratio of endoproteinase Lys-C to mAb and incubating at 37 °C for 5 h. 10%TFA was added at a final concentration of 0.2% to stop the digestion. Digests were stored at −70 °C prior to injection for HPLC–MS/MS analysis.

2.7. HPLC–MS analysis of reduced protein

Tris(2-carboxyethyl) phosphine hydrochloride was used to break the disulfide bond of the mAb to generate reduced light chain and heavy chain, which were separated by a Varian (Lake Forest, CA) PLRP-S (4.6 mm × 50 mm, 8 μm, 1000 Å) column using an Agilent 1200 HPLC. The flow rate was 0.5 mL/min with a column temperature of 75 °C. Mobile phase A was 0.025% TFA, 0.1% FA in H_2O and mobile phase B was 0.025% TFA, 0.1% FA in acetonitrile. A gradient elution over 15 min was used to separate the reduced light chain and heavy chain. The total assay run time was 25 min.

Mass spectrometric analysis was performed on Applied Biosystems/MDS Sciex (Framingham, MA) Qstar® Elite mass spectrometer equipped with an electrospray ionization (ESI) source. The flow from the HPLC column was spit in half prior to entering the ESI source. The instrument was operated in the positive mode with an ion spray voltage of 5500 V. Mass spectra of the reduced light chain and heavy chain of the mAb were deconvoluted using the Analyst QS/BioAnalyst 2.0 software package (Applied Biosystems/MDS Sciex) to obtain the reconstructed masses. The mass accuracy was typically within 30 ppm.

2.8. HPLC with ultraviolet spectrophotometer and tandem mass spectrometry (HPLC–UV/MS/MS) analysis of tryptic and endoproteinase Lys-C peptides

HPLC–UV/MS/MS analysis of tryptic and endoproteinase Lys-C peptides were performed with an Agilent 1200 HPLC coupled with a Thermo Fisher Scientific LTQ Orbitrap XL (San Jose, CA) mass spectrometer. Approximately 35 μg of digested proteins were injected for HPLC–UV/MS/MS analysis.

MAb tryptic digests were separated on a Phenomenex (Torrance, CA) Jupiter C18 column (2.0 mm × 250 mm, 5 μm, 300 Å) at a flow rate of 0.25 mL/min and a column temperature of 55 °C. Mobile phase A was 0.1% TFA in H_2O and mobile phase B was 90% acetonitrile in water with 0.09% TFA. The total HPLC run time was 190 min as previously reported [10,13]. The flow was subjected directly for HPLC–UV/MS/MS analysis.

MAb endoproteinase Lys-C digests were separated on a Agilent Zorbax 300SB-C8 column (4.6 mm × 150 mm, 3.5 μm, 300 Å) at a flow rate of 1 mL/min and a column temperature of 60 °C. Mobile phase A was 0.1% TFA in H_2O and mobile phase B was 0.08% TFA in acetonitrile. A 0–40%B gradient over 76 min was used to separate the digested peptides. A low-pressure post UV-detection microsplitter was used to direct a quarter of the flow from the column into the ESI source.

The LTQ Orbitrap XL mass spectrometer was operated in positive mode. The electrospray voltage was 3.5 kV and the capillary temperature was at 275 °C. The mass spectra were obtained using a full FTMS mode scan followed by MS/MS scan for the 5 most intensive peaks from the full FTMS scan. The dynamic exclusion feature was enabled. Collision energy of 35% was used to obtain fragmentation spectra. Peptides with unidentified MS/MS spectra were searched against a database containing all possible amino acid substitutions using a Mascot Error Tolerant Search (ETS) algorithm (Matrix Science, London, U.K.; http://www.matrixscience.com/help/error_tolerant_help.html).

3. Results and discussions

3.1. Observation of charge variant profile differences by ICIEF

Charge variant distribution is an important attribute of therapeutic mAbs. Several analytical methods, such as IEC, capillary isoeletric focusing (CIEF), and ICIEF, are often used to quantitatively monitor charge heterogeneity of therapeutic proteins [14–16]. Heterogeneity of C-terminal lysine residues is commonly observed in mAbs derived from mammalian cell culture and believed to be the result of proteolysis by endogenous CHO carboxypeptidase(s) during the cell culture process [17]. Since the presence of C-terminal lysine in mAbs will sometimes interfere with the detection of other small basic variant changes, a better assessment of mAb charge heterogeneity can be achieved by removing the C-terminal lysine residues. Thus, in this study, mAb materials were treated with carboxypeptidase B (CpB) to remove C-terminal lysine residues prior to ICIEF analysis.

In development of a new mAb, upon adjusting certain media components in a chemically defined media (CDM) and operation conditions, a slightly different charge variant profile was obtained between the reference material and the process development (PD) material, as shown in Fig. 1. ICIEF analysis yielded acidic, main, and basic variant percents of 37.9%, 57.9%, and 4.1% for the reference material and 33.6%, 61.8%, and 4.6% for the PD material, respectively. Although results from ICIEF analysis of CpB treated mAb demonstrated that the two materials had a consistent charge variant distribution, the presence of a basic shoulder peak in the PD material that was not observed in the reference material raised concerns regarding product quality and process consistency. Therefore, the following analytical efforts were made to identify the structural cause leading to the profile difference in the basic region between the reference and PD material.

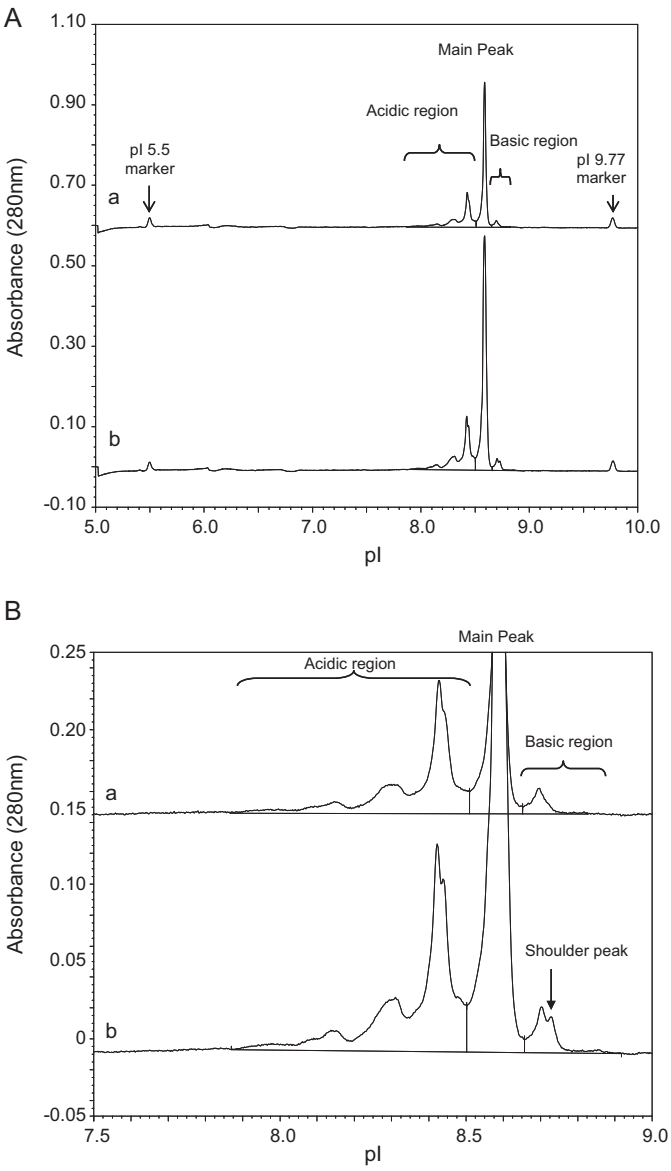

Fig. 1. ICIEF electropherograms of CpB-treated mAb (a) reference and (b) PD material in a (A) full-scale and (B) expanded view.

3.2. Collection of the charge variants by IEC

Although ICIEF delivers high-resolution separation of mAb charge species, it lacks the capability to be used as a preparative tool to collect enough variant species for further analysis. Several electrophoretic based fraction isolation instrument like BioRad Rotofor [18], Agilent OFFGEL 3100 [19,20], and continuous free-flow electrophoresis device [21] have been reported to successfully fractionate several proteins, peptides, and antibody charge variants. Due to experience and instrument availability, an IEC method was developed to isolate the mAb charge variant for characterization purpose. Typical IEC chromatograms of CpB-treated mAb reference and PD material were shown in Fig. 2, which exhibited similar profiles between the two samples, with two prominent

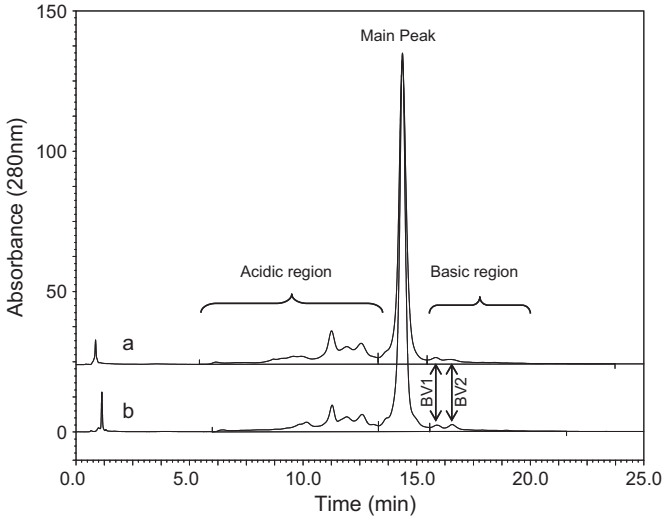

Fig. 2. IEC chromatograms of CpB-treated mAb (a) reference and (b) PD material. BV1 and BV2 are labeled as mAb basic variant peak 1 and 2, respectively.

peaks in the basic region denoted as Basic Variant Peak 1 (BV1) and Basic Variant Peak 2 (BV2), respectively. For the mAb basic variant fractionation, the reference and PD material were first treated with CpB, then subjected to IEC. The Main Peak, BV1, and BV2 fractions were collected, concentrated, and then analyzed by ICIEF to establish basic peak correlations between IEC and ICIEF methods.

Fig. 3 shows the expanded ICIEF electropherogram view of re-injected Basic Peak (BV1 & BV2) fractions collected from mAb reference and PD materials. Comparing the basic region, while there was no observed difference in BV1 fraction between two materials, the result clearly indicated the existence of a new species in the BV2 fraction from the PD material. It should be noted that a certain level of cross-fraction overlapping with the main peak was observed in the ICIEF electropherogram for collected basic peak fractions. While ICIEF separates mAb charge variants based on their pIs, IEC separates species based on the surface charge of solvent-exposed residues. The different separation principles and resolution between IEC and ICIEF make it difficult to obtain a single type of IEC species-containing fraction corresponding to a pure ICIEF peak at a given pI value. In addition, a wider collection time window was used in IEC fractionation to maximize the recovery of small amounts of basic variants. Taking all of the above into consideration, this cross-fraction overlapping was expected and also observed in other studies [22].

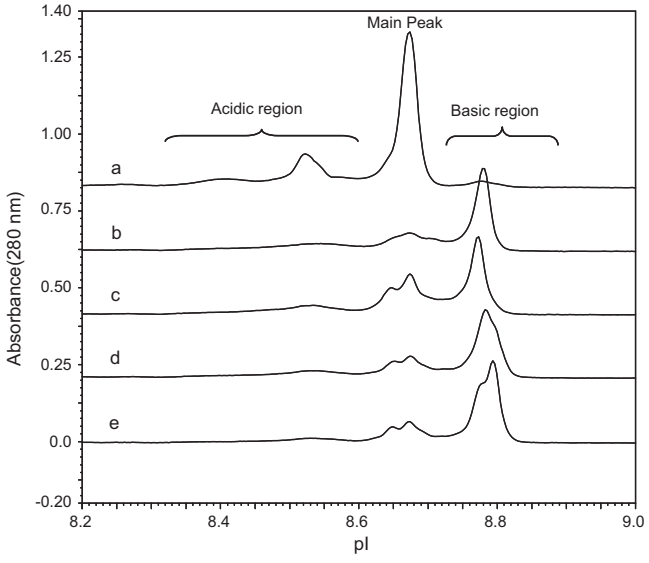

Fig. 3. ICIEF electropherograms of mAb basic charge variants collected from IEC (a) CpB-treated mAb reference material, (b) reference material BV1, (c) PD material BV1, (d) reference material BV2, and (e) PD material BV2.

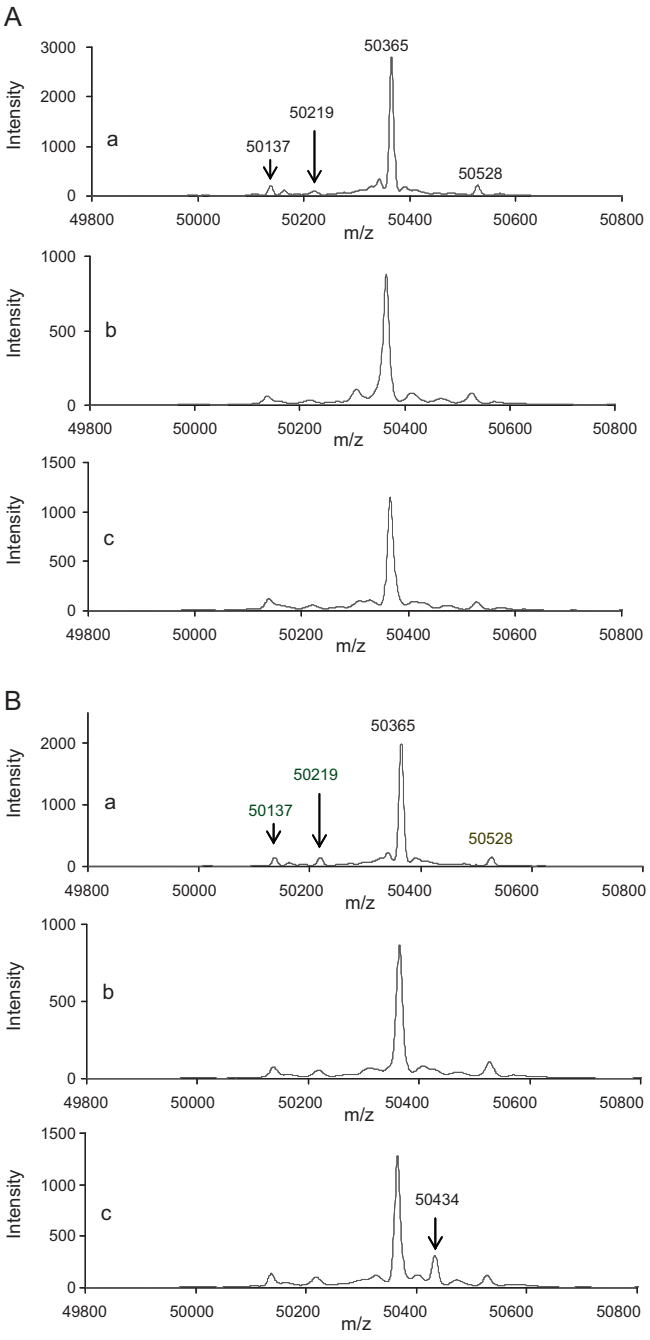

Fig. 4. Deconvoluted heavy chain mass spectra of the mAb (A) reference and (B) PD material from IEC collected (a) Main peak, (b) BV1, and (c) BV2 fractions. The calculated molecular mass is 50,365 Da for the glycosylated heavy chain without c-terminal lysine residue.

3.3. Observation of +69 Da signal by reduced heavy chain mass determination

Following the ICIEF findings on the BV2 fraction, the collected Main and Basic peak fractions were subjected to HPLC–MS for

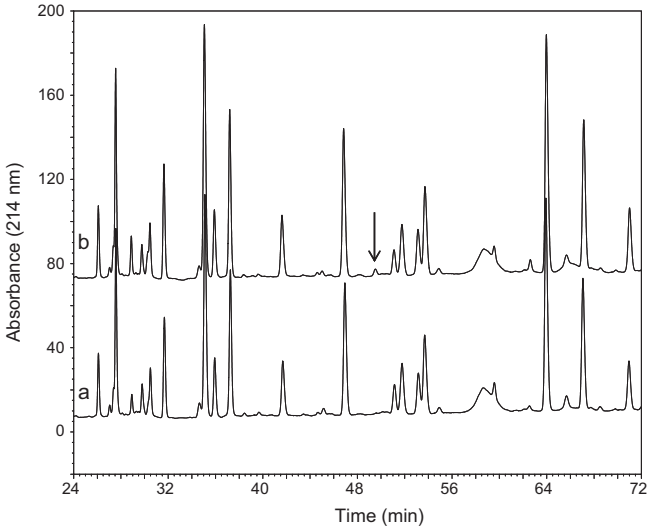

Fig. 5. HPLC-UV profiles (24–72 min) of mAb (a) reference material BV2 and (b) PD material BV2 tryptic peptide maps.

reduced light chain and heavy chain mass determination analysis. The average mass for the reduced light chain was 24,013 Da. There was no observed difference in the reduced light chain mass for all fractions collected from mAb reference and PD materials. Fig. 4 shows the deconvoluted heavy chain mass results for the IEC fractions (Main Peak, BV1, and BV2) from both materials. The typical level of carbohydrate heterogeneity was observed for the studied mAb where N-linked glycans accounted for most of the observed heterogeneity in the deconvoluted heavy chain mass data. The molecular weight of the main species in the heavy chain was 50,365 Da, which corresponded to the fucosylated biantennary glycan (G0) on the heavy chain. Species having molecular masses of 50,137, 50,219, and 50,528 Da were the heavy chain glycol-isoforms with G0-Man5, G0-Fucose, and G1 (G0 capped with 1 galactose residue), respectively. While there was no observed heavy chain mass difference for the Main Peak and BV1 fractions, the heavy chain mass of the BV2 fraction from the PD material revealed a new species with a mass shift of approximately +69 Da from the G0 isoform (Fig. 4B, trace c). The new species was not detectable in the reference material heavy chain (Fig. 4A, trace c). This result together with IEC and ICIEF data suggested that the new basic variant presented in mAb PD material was enriched in the BV2 fraction. Thus, analysis of a peptide map was carried out on BV2 fraction to identify the molecular origin of this new species.

3.4. Identification of the Ser to Arg variant by tryptic and endoproteinase Lys-C peptide mapping

HPLC–UV/MS/MS of tryptic digests of BV2 from mAb reference and PD materials was performed. The HPLC–UV profiles of these two samples were compared, which revealed the presence of an extra peak at 49.5 min (labeled by arrow) in the PD material (Fig. 5). The acquired MS data indicated this extra peak had a *m*/*z* of 573.33. MS/MS data analysis and Mascot search results (without enzyme restriction) suggested the *m*/*z* of 573.33 peptide's sequence was LSLSPG. This identification, together with the detection of the new species with a +69 Da mass shift on the heavy chain, and the observed basic shoulder peak by ICIEF, all lead to the possibility of a Ser to Arg (S441R) substitution in the original peptide SLSLSPG (underline indicates substitution site). The reasons are that a Ser to Arg substitution results in a mass increase of 69 Da, trypsin cleavage at R in the substituted peptide RLSLSPG generates LSLSPG peptide, and the basicity of the molecule is increased.

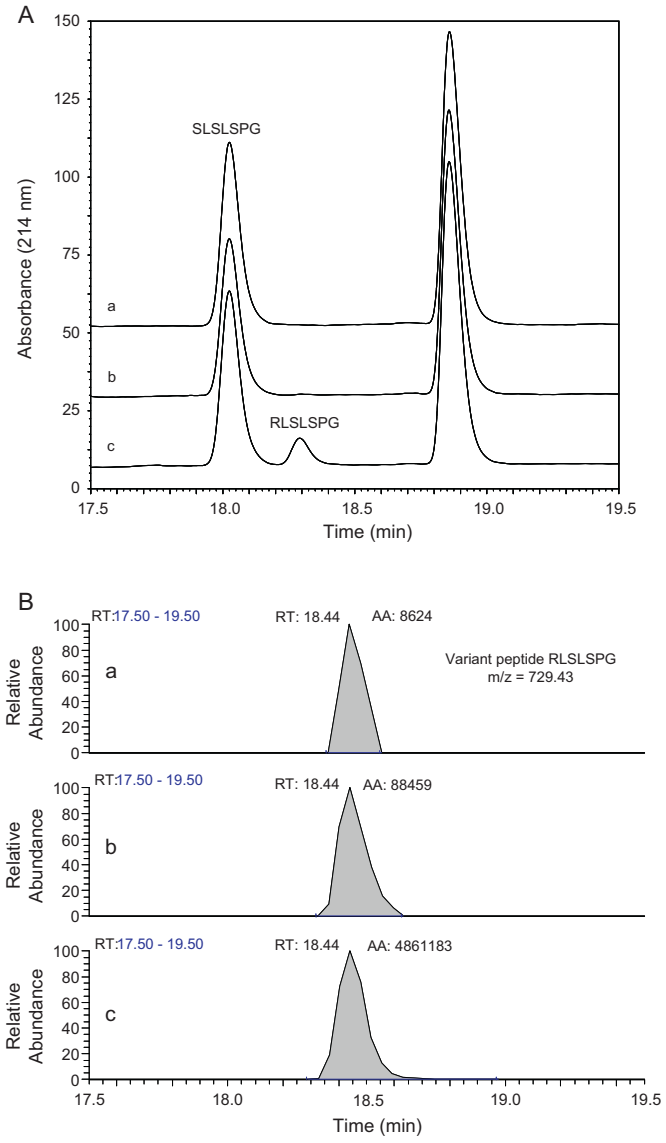

Fig. 6. (A) HPLC-UV profiles of endoproteinase Lys-C peptide maps (17.5–19.5 min) and (B) EIC of variant peptide RLSLSPG (*m*/*z* of 729.43) from mAb (a) reference material, (b) PD material, and (c) PD material BV2 fraction. Peaks in EIC are labeled with retention time (RT) and peak area (AA).

To confirm this sequence variant, i.e., the existence of variant peptide RLSLSPG in the PD material, an endoproteinase Lys-C peptide map was developed, which provided over 98% of amino acid sequence coverage. Endoproteinase Lys-C peptide maps of the reference material, PD material, and PD material BV2 fraction are overlaid in Fig. 6. The result clearly showed the presence of the variant peptide RLSLSPG that was separated from the wild-type peptide SLSLSPG with about 30 s retention time difference (Fig. 6A, trace c). It should be noted that without fraction collection and enrichment, the RLSLSPG peptide was not detected in the HPLC–UV profiles of either the reference or PD material (Fig. 6A, trace a and b). The extracted ion chromatogram (EIC) showed the presence of the variant peptide RLSLSPG (*m*/*z* of 729.43) in both the reference and PD materials (Fig. 6B, trace a and b). However, the MS signal of the RLSLSPG peptide in the reference material was about 10 times lower than that in the PD material (Fig. 6B), and there was no MS/MS spectra acquired for RLSLSPG peptide in the reference material. Without the MS/MS spectra, this S441R variant was not detected by the Mascot ETS during the initial clone selection.

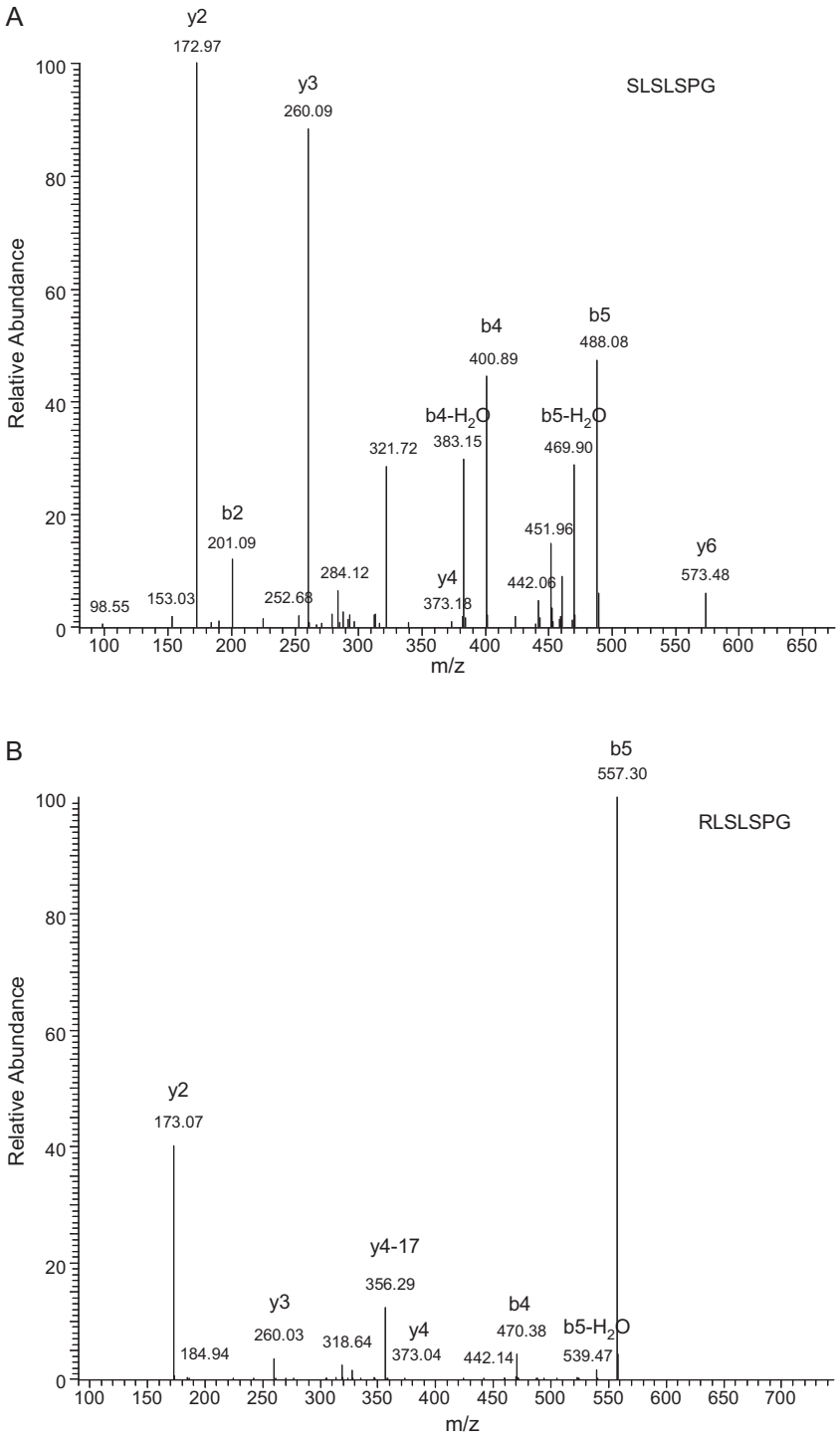

Fig. 7. MS/MS spectra of precursor ions at (A) *m*/*z* of 660.36 for expected peptide SLSLSPG and (B) *m*/*z* of 365.22 (2+) for variant peptide RLSLSPG.

With basic variant fractionation and enrichment, the identification of the variant peptide and its correlation to charge variant became achievable.

Tandem mass spectra of the two peptides SLSLSPG and RLSLSPG are compared and are shown in Fig. 7. Since there was only one amino acid difference at the peptide N-terminal, the y series of the fragment ions for the two peptides should be the same while the b series ions shifted by 69 Da. As expected, all observed *y* ions, i.e., *y*2, *y*3, and *y*4, were the same for both peptides, while the observed b4 and b5 ions from peptide RLSLSPG had a mass increase of 69 Da from those of the peptide SLSLSPG. This MS/MS result further supported the peptide identity.

3.5. Confirmation of Ser to Arg substitution with synthetic peptides

Identification of the variant peptides LSLSPG and RLSLSPG was further confirmed with the use of synthetic peptides, which were spiked into tryptic digest blank and analyzed with the same HPLC–UV/MS/MS conditions. The comparison of MS/MS spectra

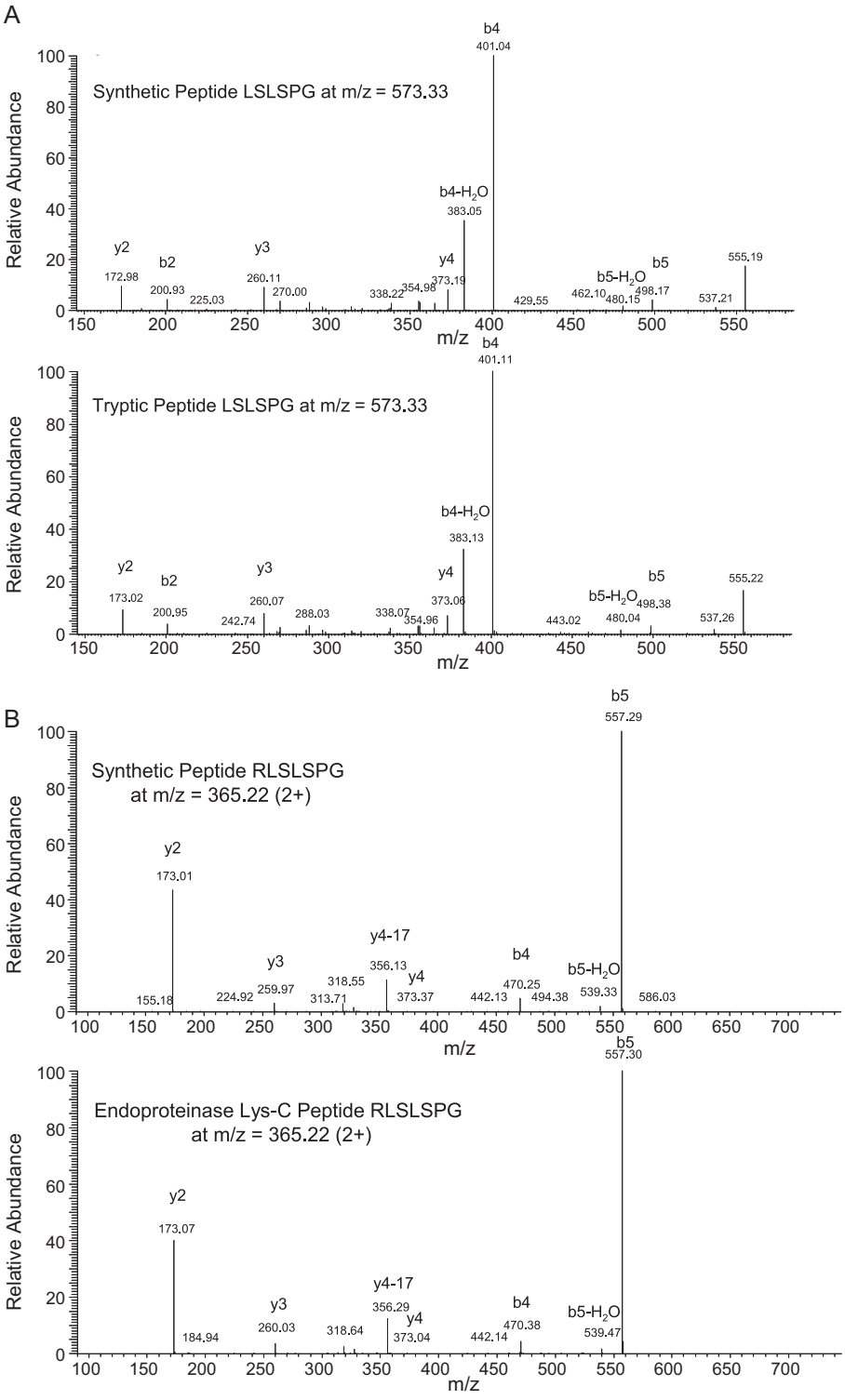

Fig. 8. MS/MS spectra of precursor ions at (A) *m*/*z* of 573.33 for tryptic peptide LSLSPG and (B) *m*/*z* of 365.22 (2+) for endoproteinase Lys-C peptide RLSLSPG.

between the determined variant peptides, i.e., LSLSPG and RLSLSPG, and their corresponding synthetic peptides are shown in Fig. 8. The data again supported the identity of the variant peptides presented in the mAb PD material, and the correct assignment of the Ser to Arg sequence variant.

Due to the very low levels of variant peptides LSLSPG and RLSLSPG, they were not detectable by HPLC–UV absorbance at 214 nm either in tryptic or endoproteinase Lys-C peptide maps. Thus, integration of EIC of variant (LSLSPG) and wild-type (SLSLSPG and SLSLSPGK) peptide peak areas from tryptic peptide map was used to estimate the percentage of sequence variant (%SV) in the studied mAb. The %SV was calculated to be 0.6% of LSLSPG in the PD material and less than 0.1% in the reference material. The analysis of an additional four mAb PD batches, which were from the same thaw but from different 2L bioreactor production runs and conditions, showed the presence of S441R sequence variant in the range of 0.3–0.5%, as summarized in Table 1.

Table 1
Percentage of sequence variant in a mAb determined by EIC.

mAb materials	%Sequence variant[a] for peptide (**R**)LSLSPG[b]
Reference material	0.03
PD material	0.57
PD batch 1	0.47
PD batch 2	0.44
PD batch 3	0.45
PD batch 4	0.30

[a] %Sequence variant is the peak area of variant ion divided by the sum of peak areas of wild-type and variant peptides multiplied by 100.
[b] Ser replaced by Arg is shown in bold. Parenthesis indicates the trypsin cleavage site.

3.6. Molecular origin of the sequence variant

The Ser to Arg sequence variant was reported previously in a mAb, where the substitution occurred at Ser167 in the light chain with the use of AGT codon [9]. The PCR analysis demonstrated that the S167R sequence variant is due to a genomic mutation where the nucleotide A in AGT is mutated to C. In our study, the identified Ser variant site correlated with AGC codon, where a single nucleotide mutation at A(CGC) or C (AGA, AGG) could lead to the coding of Arg. Misincorporation was reported to be another mechanism to cause sequence variant, where multiple AGC coded Ser sites were substituted by Asp in a mAb expressed in CHO cell line [8]. However, misincorporation is probably not the root cause for the study subject, since among the 17 Sers coded by the AGC codon in the heavy chain, only the heavy chain Ser441 was substituted by Arg. The root cause of this sequence variant is yet to be determined.

4. Conclusions

Due to the potential immunogenicity and safety risks posed by the presence of sequence variants in the primary structure of a therapeutic monoclonal antibody, it is critical to monitor unintended amino acid substitutions to ensure product safety, efficacy, and quality. During early development of a mAb, sequence variant analysis was performed based on in-house established practices in clone selection, and no detectable variants were found. However, in later cell culture development process, charge variant analysis by ICIEF revealed the appearance of a new basic variant in the materials generated by the new cell culture process. Extensive analytical investigation on this newly observed variant indicated the presence of a Ser to Arg sequence variant at the heavy chain Ser441 position. This substitution was further confirmed by using synthetic peptides. In summary, this work demonstrates the effectiveness of the integrated analytical technologies not only to identify protein heterogeneity, but also to correlate the changes between protein primary structure and charge variant.

Acknowledgements

We greatly thank Yajun Jennifer Wang and X. Christopher Yu for their helpful discussions and critical review of the manuscript.

References

[1] D.S. Dimitrov, J.D. Marks, in: A.S. Dimitrov (Ed.), Therapeutic Antibodies: Methods and Protocols, Humana Press, New York, 2009, p. 1.
[2] J.M. Reichert, mAbs 1 (2009) 86.
[3] J.M. Reichert, V.E. Valge-Archer, Nat. Rev. Drug Discov. 6 (2007) 349.
[4] R.J. Harris, A.A. Murnane, S.L. Utter, K.L. Wagner, E.T. Cox, G.D. Polastri, J.C. Helder, M.B. Sliwkowski, Nat. Biotechnol. 11 (1993) 1293.
[5] M. Wan, F.Y. Shiau, W. Gordon, G.Y. Wang, Biotechnol. Bioeng. 62 (1999) 485.
[6] D. Wen, M.M. Vecchi, S. Gu, L. Su, J. Dolnikova, Y.M. Huang, S.F. Foley, E. Garber, N. Pederson, W. Meier, J. Biol. Chem. 284 (2009) 32686.
[7] A. Khetan, Y.M. Huang, J. Dolnikova, N.E. Pederson, D. Wen, H. Yusuf-Makagiansar, P. Chen, T. Ryll, Biotechnol. Bioeng. 107 (2010) 116.
[8] X.C. Yu, O.V. Borisov, M. Alvarez, D.A. Michels, Y.J. Wang, V. Ling, Anal. Chem. 81 (2009) 9282.
[9] D. Guo, A. Gao, D.A. Michels, L. Feeney, M. Eng, B. Chan, M.W. Laird, B. Zhang, X.C. Yu, J. Joly, B. Snedecor, A. Shen, Biotechnol. Bioeng. 107 (2010) 163.
[10] Y. Yang, A. Strahan, C. Li, A. Shen, H. Liu, J. Ouang, V. Katta, K. Francissen, B. Zhang, mAbs 2 (2010) 285.
[11] A.H. Que, B. Zhang, Y. Yang, J. Zhang, G. Derfus, A. Amanullah, BioProcess Int. 8 (2010) 52.
[12] H. Dorai, T. Sauerwald, A. Campbell, Y.S. Kyung, J. Goldstein, A. Magill, M.J. Lewis, Q.M.J. Tang, D. Jan, S. Ganguly, BioProcess Int. 5 (2007) 66.
[13] B. Zhang, Y. Yang, I. Yuk, R. Pai, P. McKay, C. Eigenbrot, M. Dennis, V. Katta, K.C. Francissen, Anal. Chem. 80 (2008) 2379.
[14] V. Dolnik, Electrophoresis 29 (2008) 143.
[15] L.A. Khawli, S. Goswami, R. Hutchinson, Z.W. Kwong, J. Yang, X. Wang, Z. Yao, A. Sreedhara, T. Cano, D. Tesar, I. Nijem, D.E. Allison, P.Y. Wong, Y. Kao, C. Quan, A. Joshi, R.J. Harris, P. Motchnik, mAbs 2 (2010) 1.
[16] L.H. Silvertand, J.S. Torano, W.P. van Bennekom, G.J. de Jong, J. Chromatogr. A 1204 (2008) 157.
[17] R.J. Harris, J. Chromatogr. A 705 (1995) 129.
[18] R.K. Brobey, L. Soong, Proteomics 7 (2007) 116.
[19] C.D. Meert, L.J. Brady, A. Guo, A. Balland, Anal. Chem. 82 (2010) 3510.
[20] N.C. Hubner, S. Ren, M. Mann, Proteomics 8 (2008) 4862.
[21] H. Xie, S. Bandhakavi, T.J. Griffin, Anal. Chem. 77 (2005) 3198.
[22] Z. Sosic, D. Houde, A. Blum, T. Carlage, Y. Lyubarskaya, Electrophoresis 29 (2008) 4368.