Journal of Chromatography B, 908 (2012) 1–8

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

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

Characterization and identification of alanine to serine sequence variants in an IgG4 monoclonal antibody produced in mammalian cell lines

Jinmei Fu*, Jacob Bongers, Li Tao, Dan Huang, Richard Ludwig, Yunping Huang, Yueming Qian, Jonathan Basch, Joel Goldstein, Ramji Krishnan, Li You, Zheng Jian Li, Reb J. Russell

Biologics Manufacturing and Process Development, Global Manufacturing and Supply, Bristol-Myers Squibb, Princeton, NJ 08543, United States

ARTICLE INFO

Article history:
Received 3 July 2012
Accepted 13 September 2012
Available online 26 September 2012

Keywords:
Sequence variant
Monoclonal antibody
High performance liquid chromatography
Tandem mass spectrometry
Peptide mapping

ABSTRACT

Low levels of alanine to serine sequence variants were identified in an IgG4 monoclonal antibody by ultra/high performance liquid chromatography and tandem mass spectrometry. The levels of the identified sequence variants A183S and A152S, both in the light chain, have been determined to be 7.8–9.9% and 0.5–0.6%, by extracted ion currents of the tryptic peptides L16 and L14, respectively. The A183S variant was confirmed through tryptic map spiking experiments using synthetic peptide, SDYEK, which incorporated Ser at the position of native Ala in the tryptic peptide L16. Both mutations were also observed by endoproteinase Asp-N peptide mapping. The variant level of A183S was also quantified by LC–UV with detection at 280 nm and fluorescence detection of tyrosine residues on the tryptic peptides. The results from LC–MS, UV, and fluorescence detection are in close agreement with each other. The levels of the sequence variants are comparable among the antibody samples manufactured at different scales as well as locations, indicating that the variants' levels are not affected by manufacture scale or locations. DNA sequencing of the master cell bank revealed the presence of mixed bases at position 183 encoding both wild and mutated populations, whereas bases encoding the minor sequence variant at position 152 were not detected. The root cause for A152S mutation is not yet clearly understood at this moment.

1. Introduction

Sequence variants are protein molecules whose amino acid residues are replaced by other unintended ones due to various reasons. They are molecular variants of a desired product and have been identified and characterized in recombinant human monoclonal antibodies and other recombinant proteins in recent years [1–8]. Known mechanisms for the introduction of sequence variants include mutations at the DNA level [9] and mistranslation (through either tRNA misaminoacylation or misreading) at the protein expression level [10]. The impact of sequence variants on protein therapeutics' properties can be varied. Even though many of the variants do not affect target proteins' overall properties, some of them may alter proteins' catalytic constants, specificity, and stability [11–14]. For proteins produced for therapeutic use, sequence variants can potentially induce immune response or affect receptor–ligand interactions [15]. Therefore, sequence variants of the desired product are generally considered product-related impurities by regulatory authorities [16], unless proved otherwise. These variant proteins with only one or a few amino acid substitutions often retain overall chemical and physical properties similar to that of the native proteins and are difficult to remove from final product during downstream purification. Therefore, detection and identification of sequence variants during early phase clinical development is critical to ensure selection of an appropriate expression clone and cell line suitable for late stage development and commercial manufacture.

Detection and identification of amino acid sequence variants remains challenging. Although analytical methodologies at both DNA and protein levels can be used to assess sequence variants, DNA sequencing methods have sensitivity limits of 10–15% and do not necessarily reflect the changes in the final product [17]. Historically, protein N-terminal sequencing by Edman degradation has been used on either electrophoretic gel bands or isolated chromatographic fractions to determine the sites and types of the variants [18–21]. More recently, mass spectrometry (MS) has become the predominant method for variant analysis [22,23], due to its high

Abbreviations: CID, collision induced dissociation; DTT, dithiothreitol; EIC, extracted ion chromatogram; EPCB, end of production cell bank; FTMS, Fourier transform mass spectrometry; HPLC, high performance liquid chromatography; IAM, iodoactamide; LC/MS, liquid chromatography mass spectrometry; MCB, master cell bank; MS/MS, tandem mass spectrometry; MWCB, manufacturer's working cell bank; RP, reversed phase; RT-PCR, reverse transcription polymerase chain reaction; TFA, trifluoroacetic acid; UPLC, ultra performance liquid chromatography.

* Corresponding author at: Biologics Manufacturing & Process Development, Bristol-Myers Squibb Global Manufacturing and Supply, 6000 Thompson Road, East Syracuse, NY 13057, United States. Tel.: +1 315 431 7810; fax: +1 315 432 4827.
E-mail address: jinmei.fu@bms.com (J. Fu).

1570-0232/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jchromb.2012.09.023

sensitivity, resolving power, and efficiency. A typical enzyme digest of monoclonal antibody can be analyzed within 2–3 h by liquid chromatography mass spectrometry (LC–MS) and tandem mass spectrometry (MS/MS). Sequence variants with levels as low as 0.01% have been detected using ion trap mass spectrometers, such as LTQ-Orbitrap [5]. However, an inherent shortcoming of MS analysis is that it cannot detect variants with no change in mass, e.g., variation of Leu/Ile. In addition, only high resolution Fourier transform mass spectrometry (FTMS) can detect a variation of Lys/Gln (Δm = 0.0364 Da) [24]. To improve efficiency, analytical approaches combining LC/MS/MS peptide mapping with automatic Mascot or SEQUEST database search have been developed to detect and identify protein sequence variants [25,26]. These database searches generally focus on single amino acid variations resulting from single nucleotide mutations, for which there are potentially 75 pairs (a total of 150 types) of possible amino acid substitution derived from single nucleotide base mutation with absolute mass shifts ranging from 0 to 129 Da [27]. Although single base variations are most probable, sequence variants involving two-base substitutions have also been reported [2]. It is worth noting that multiple amino acid substitutions can cause same mass shift, e.g., a mass shift of 14 Da could be caused by mutations of Gly/Ala, Ser/Thr, Val/Leu, Val/Ile, Asp/Glu and Asn/Gln, etc. Furthermore, post-translational modifications (e.g., oxidation, deamidation, methylation, etc.) can also cause similar mass change as sequence variants, adding more complexity to the identification process. Therefore, manual assessment of the search results, which can be very time consuming, is usually required to confirm the automated assignment of sequence variants.

Here, we report the observation of Ala to Ser sequence variants at two locations in the light chain of an IgG4 monoclonal antibody by intact mass measurement of the reduced antibody molecule and peptide mapping with tandem mass spectrometry. Subsequent DNA sequencing of the master cell bank revealed one variant was caused by mutation at the DNA level. Regarding Ala to Ser variants, multiple mechanisms have been identified for this mutation which occur in nature through either tRNA ribosomal frameshift [28,29] or misaminoacylation [30,31], as well as mutation at the DNA levels [32,33]. However, no report has been published to date on Ala to Ser sequence variants in proteins produced by mammalian cell lines.

2. Materials and methods

2.1. Materials and reagents

The monoclonal antibody was expressed in recombinant CHO K1SV cells and produced using standard mammalian cell cultivation followed by chromatographic purification.

Guanidine hydrochloride (HCl), dithiothreitol (DTT) and iodoacetamide (IAM) were purchased from Sigma–Aldrich (St. Louis, MO). All solvents were HPLC grade and were purchased from Burdick and Jackson (Muskegon, MI). Formic acid and trifluoroacetic acid (TFA) were purchased from Pierce (Rockford, IL). Sequencing grade modified trypsin was purchased from Promega (Madison, WI). Sequencing grade endoproteinase Asp-N was purchased from Roche (Indianapolis, IN). Synthetic peptides were ordered from AnaSpec (Fremont, CA).

2.2. Reduced intact mass analysis

Antibody samples were incubated in the presence of 5 mM DTT at 37 °C for 20 min. The reduction was stopped by the addition of formic acid to a final concentration of 0.2% (v:v). The LC–MS system (Waters, Milford, MA) used was an Alliance HPLC with dual wavelength UV detection at 215 nm and 280 nm, coupled with a quadrupole time-of-flight (Q-ToF) mass spectrometer. Each reduced antibody sample (0.5 μg) was injected onto a 10 μm, 2.1 × 100 mm Poros® reversed phase (RP) column (Applied Biosystems, Foster City, CA) equilibrated with 20% acetonitrile containing 0.1% formic acid, followed by gradient elution from 20% to 50% acetonitrile in 25 min, at a flow rate of 0.25 mL/min. The column was washed with 80% acetonitrile and re-equilibrated to initial conditions prior to the next injection. The capillary voltage for the Q-ToF was set at 2300 V and the sample cone voltage at 45 V. The Q-ToF analyzer was set to scan from m/z 800 to 4000. The mass spectra were combined and deconvoluted using MaxEnt1 algorithm (Waters, Milford, MA).

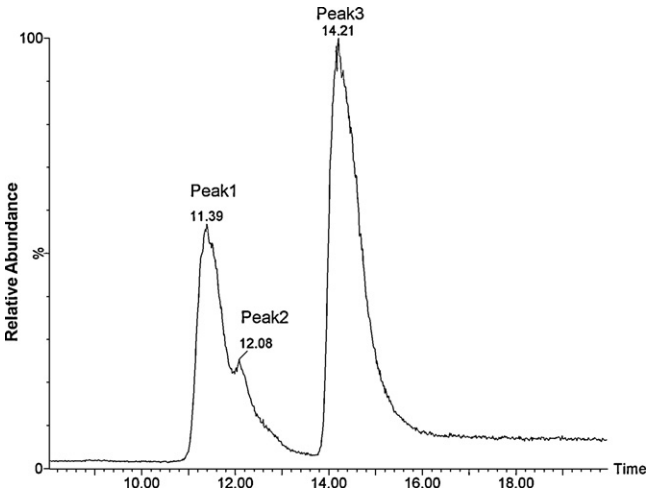

Fig. 1. Total ion chromatograms of reduced antibody by RPLC–MS.

2.3. Fractionation of antibody light chains and tryptic digestion

Two reversed phase separated fractions from peak 1 and peak 2, as shown in Fig. 1, were collected and pooled off a 10 μm, 4.6 × 100 mm Poros® reversed phase column from two 500 μg preparative runs of reduced/carboxyamidomethylated antibody. The fractions were evaporated to dryness under vacuum using a Savant SpeedVac centrifugal-evaporator with no heating. Sample residues were dissolved in buffer consisting of 0.1% RapiGest™ SF (Waters, Milford, MA), 50 mM Tris/HCl, 1.0 mM $CaCl_2$, pH 7.6 and digested with sequencing grade trypsin at 37 °C for 18 h.

2.4. LC–MS peptide mapping

Antibody samples were denatured with guanidine HCl followed by reduction with DTT and alkylation with IAM. The reduced and alkylated antibody was loaded onto a Zeba gel desalting 96-well spin-plate (Thermo Fisher Scientific, Rockford, IL) and eluted with a buffer containing 50 mM Tris/HCl, 1.0 M urea, 1.0 mM $CaCl_2$, pH 7.6. Sequencing grade trypsin (3%, w/w, enzyme to protein) or Asp-N (1%, w/w, enzyme to protein) was added to the sample prior to incubation at 37 °C for 1 h or 18 h, respectively. The digestion was stopped by adding 1.0 M HCl to ~pH 2–3. The tryptic digests were separated on an Acquity UPLC BEH300 C18 column (1.7 μm, 2.1 × 150 mm, Waters, Milford, MA) at a flow rate of 0.2 mL/min using a Waters Acquity UPLC instrument (Waters, Milford, MA). The column was equilibrated with 1% acetonitrile containing 0.02% TFA in water, and peptides were eluted with an increasing gradient of acetonitrile from 1 to 45% over 80 min.

The UPLC column was connected sequentially to a PDA UV detector, a fluorescence detector and a LTQ XL ion trap mass

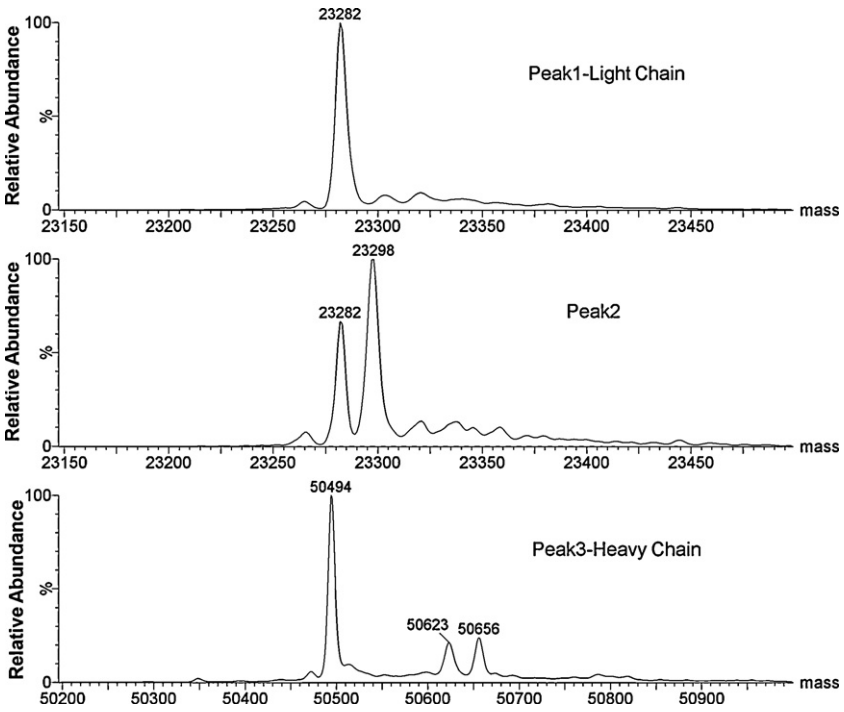

Fig. 2. Deconvoluted mass spectra of reduced antibody (peak numbers as assigned in Fig. 1).

spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with an electrospray ionization (ESI) interface for LC–MS and MS/MS analyses. The LC–MS data were acquired in positive ion mode in a mass to charge ratio (m/z) range of 200–2000 and a source voltage of 5000 V. Tandem mass spectra (MS/MS) data were acquired using 35% normalized collision energy, 0.25 activation Q and 30 ms activation time.

2.5. cDNA sequencing analysis of cell line

Frozen vials from the MCB (Master Cell Bank), MWCB (Manufacturer's Working Cell Bank), and EPCB (End of Production Cell Bank) were analyzed by cDNA sequencing in Beckman Coulter Genomics (Morrisville, NC). mRNA was isolated from cell pellets and a reverse transcription reaction and a reaction lacking reverse transcriptase ("No RT" negative control) were performed using the isolated mRNA. The resulting cDNA was used as template in PCR amplification reactions using primer pairs designed against the expected sequence. The purified PCR products were used as templates and sequenced using the BigDye® Terminator Cycle Sequencing Kit and analyzed by standard florescence detection automated DNA sequencing technology. Research grade quality (2-fold coverage) automated DNA sequencing was performed for each sample. The fluorescence data were analyzed using standard analysis software. The resulting consensus sequences were compared to the expected reference sequences for the heavy and light chains.

3. Results and discussion

3.1. Reduced intact mass analysis reveals light chain species with +16 Da mass increase

Under the reduction conditions used in this analysis, no denaturant was used and the IgG4 monoclonal antibody was partially reduced into its component light chains and heavy chains by reduction of the inter-chain disulfide linkages between the two heavy chains and the two heavy-light chain pairs while the intra-chain disulfide linkages were kept intact [34]. If successful, only two peaks (light chain and heavy chain) would be observed in the reversed phase profile of a monoclonal antibody. Fig. 1 shows the reversed phase liquid chromatography (RPLC) profile of reduced antibody, with three peaks observed. The deconvoluted mass spectra of each peak are shown in Fig. 2. The predominant observed mass of the peak eluted at 11.39 min is 23,282 Da, which within experimental error, is consistent with the theoretical mass of 23,283 Da corresponding to one light chain with two pairs of internal disulfide bonds. The predominant observed mass of the peak eluting at 14.21 min is 50,494 Da, which within experimental error, is consistent with the theoretical mass of 50,495 Da corresponding to the heavy chain with four internal disulfide bonds, N-terminal pyroglutamate modification, C-terminal lysine removal, and the G0F glycoform.

Two predominant observed masses, 23,282 Da and 23,298 Da, are observed in the peak eluting at 12.08 min. Because peak 1 and peak 2 are not completely resolved, the species with a mass of 23,282 Da is likely to be light chain co-eluted from peak1. The species with a mass of 23,298 Da is likely a modified light chain with +16 Da mass increase. Intact mass measurement of non-reduced antibody was also performed (data not shown). However, the 16 Da mass increase was not observed at the native protein level. With a mass resolution of ~8000–10,000 of the Q-ToF instrument used, an antibody of 150 KDa would have a theoretical peak width of ~25 Da [35] at half height, which precludes a variant species with mass change of 16 Da to be detected.

3.2. Analysis of two light chain fractions by tryptic peptide mapping

The fractions from peak 1 and peak 2 were collected and digested with trypsin in order to identify the nature and localization of this modification. Fig. 3 shows the tryptic peptide maps of the two RPLC light chain fractions. Visual comparison of the profiles shows a new peak at retention time 7.1 min in fraction 2 which is not seen in the fraction 1. The extracted mass spectrum shows that this peak

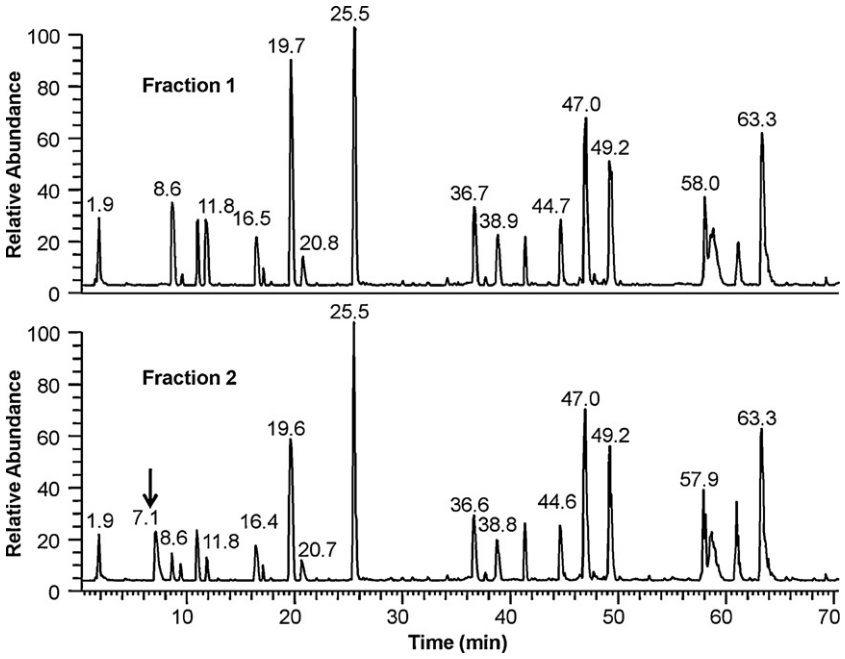

Fig. 3. Tryptic peptide mapping profiles of light chain fractions 1 and 2. A new peak (see arrow) at retention time 7.1 min was observed in fraction 2 but not seen in the fraction 1.

contains a peptide with *m*/*z* 641.2 (1+). This mass is +16 Da heavier than the 16th tryptic peptide of light chain (ADYEK, *m*/*z* 625.2, 1+). MS/MS analysis by collision induced dissociation (CID) of the native peptide and the variant peptide are shown in Fig. 4(a) and (b), respectively. The MS/MS spectrum of the variant shows no mass shift in the y2 to y4 ions and a mass shift of +16 Da in b3 and b4 ions compared to the fragment ions of the native peptide, indicating a modification at the N-terminal Ala residue of the native

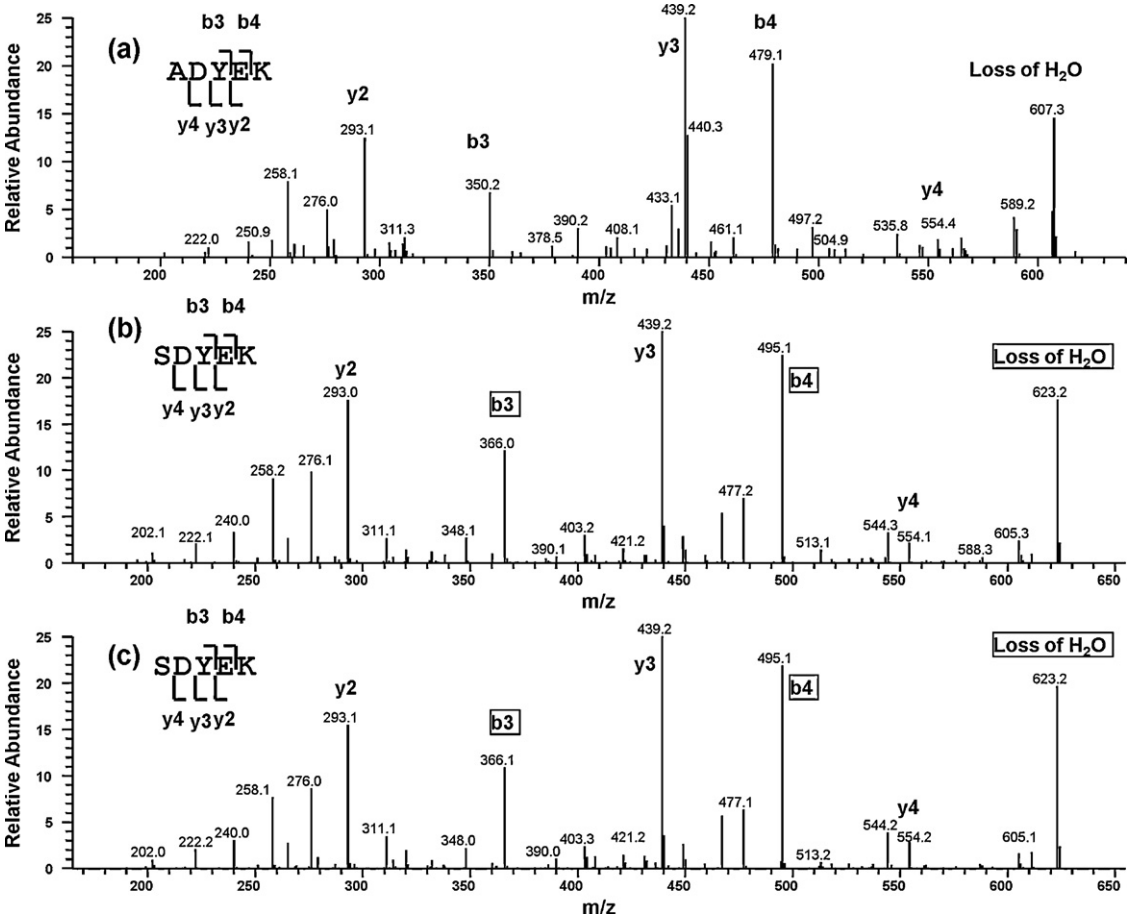

Fig. 4. CID MS/MS spectra of (a) *m*/*z* 625.2 (1+) for native peptide of ADYEK; (b) *m*/*z* 641.2 (1+) for variant peptide; and (c) *m*/*z* 641.2 (1+) for synthetic peptide of SDYEK.

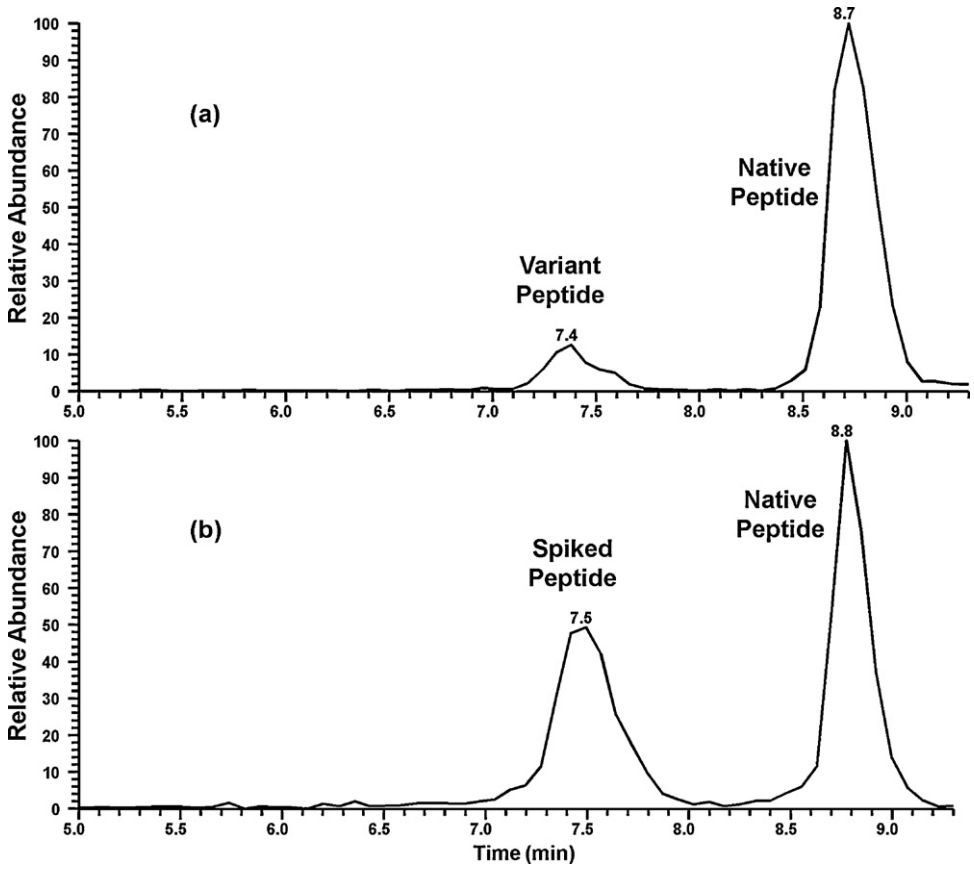

Fig. 5. Extracted ion chromatogram of native peptide, ADYEK, at *m*/*z* 625.2 (1+) and variant peptide, SDYEK, at *m*/*z* 641.2 (1+) in (a) non-spiked antibody tryptic digest and (b) spiked in SDYEK synthetic peptide.

peptide. The mass increase of 16 Da is consistent with addition of one oxygen atom and is often found as a result of oxidized protein products [36]. However, Ala residues are unlikely to be oxidized due to their relatively inert nature. As a result, replacement of the Ala by Ser was proposed. Consequently, a synthetic peptide SDYEK, was spiked into the antibody tryptic digest to confirm the identification of Ala to Ser variant. Fig. 5 shows that the synthetic peptide elutes at the same retention time as the variant peptide observed in the non-spiked sample digest. The variant peptide and synthetic peptide elute at ~7.4–7.5 min and the native peptide elutes at ~8.7–8.8 min. The variant peptide elutes earlier than the native peptide, which is consistent with decreased hydrophobicity when Ala is replaced by a Ser residue. The variant peptide showed the same MS/MS fragmentation pattern as those seen in the corresponding synthetic peptide as illustrated in Fig. 4(b) and (c).

3.3. Observation of Ala to Ser sequence variants at second site

Further analysis of LC–MS/MS data from the antibody tryptic digest revealed another possible Ala to Ser variant site in the light chain: Ala 152 in the 14th tryptic peptide VDNALQSGNSQESVTEQDSK. Fig. 6(a) and (b) shows the CID MS/MS spectra of the native peptide L14 and its corresponding variant peptide. The MS/MS spectrum of the variant peptide of *m*/*z* 1076.5 (2+) showed no mass shift in the y3 to y16 ions and a mass shift of +16 Da in y17 to y18 ions and b5 to b18 ions compared to the fragment ions of the native peptide of *m*/*z* 1068.5 (2+), also indicating a modification at the 17th amino acid residue, Ser, from the C-terminus. The light chain containing A152S was not enriched in the fraction 2 collected off RPLC column. Instead, A152S modification was observed at relatively similar levels in fraction 1 and fraction 2.

3.4. Relative quantitation of Ala to Ser variants by LC–MS, UV and fluorescence detection

The amounts of Ala to Ser variants in five lots of antibody samples were calculated by dividing the peak area of the variant peptide peak by the sum of the peak area for the native and variant peptide peaks based on their extracted ion chromatograms (EIC) and are listed in Table 1. The results showed that the Ala to Ser variant level is comparable among all the samples tested, with 7.8–9.9% for A183S and 0.5–0.6% for A152S. Within the precision of the applied methods, the levels of these variants are consistent among the five lots of samples which were produced at different scales and sites, indicating no change as a result of scale or site of manufacture.

As shown in Fig. 5, the variant peptide and native peptide for A183S are well separated. Although the variant peptide co-eluted with another peptide, the variant peptide contains a tyrosine residue which absorbs at 280 nm and the co-eluting peptide does not have absorption at 280 nm. Peak area percent values were also obtained for the A183S variants by 280 nm UV detection and tyrosine-specific native fluorescence traces (229 nm excitation and 305 nm emission) for the ADYEK and SDYEK tryptic peptides (Table 1). The A183S variant levels obtained from the three detection methods are in close agreement with each other. The peak resolution in the fluorescence trace (data not shown) was slightly less only because the UV and fluorescence detectors were connected in series, in that order, and there was a small amount of diffusional band-broadening that occurred in the tubing connecting the two detectors and in the fluorescence detector flow-cell. Unlike A183S, the A152S variant level could only be obtained from the LC–MS data of the tryptic peptide map due to co-elution of the A152S variant peptide with another peptide. This variant

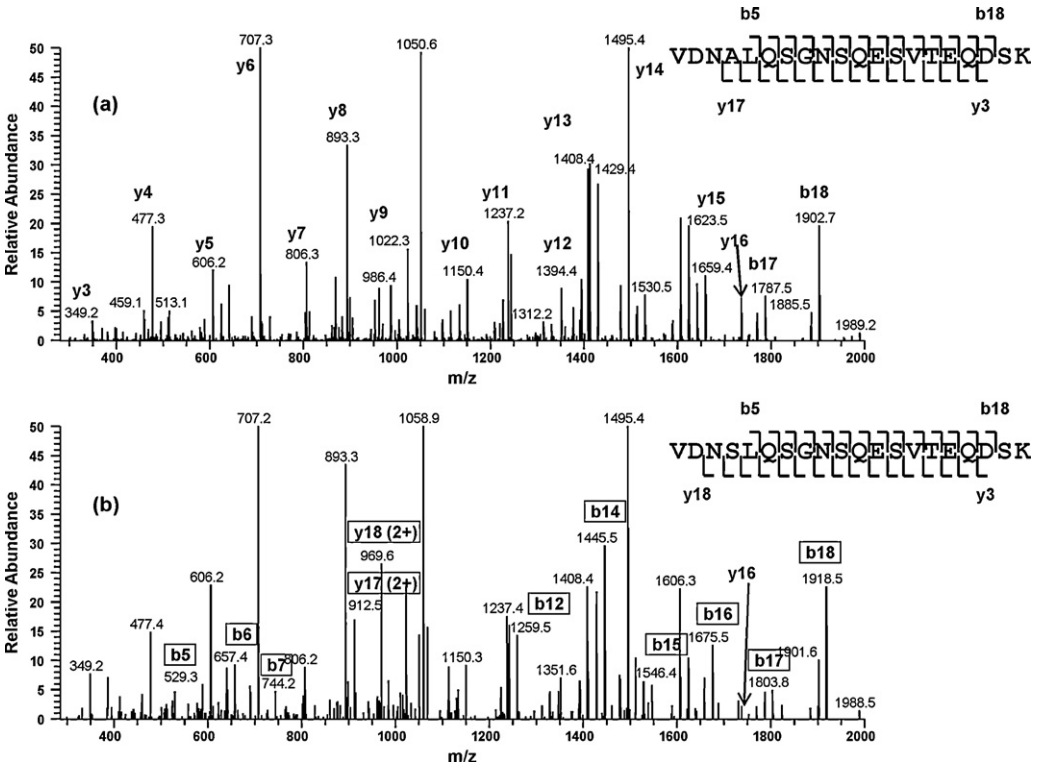

Fig. 6. CID MS/MS spectra of (a) *m*/*z* 1068.5 (2+) for native peptide of VDNALQSGNSQESVTEQDSK and (b) *m*/*z* 1076.5 (2+) for variant peptide of VDNSLQSGNSQESVTEQDSK.

peptide does not contain tyrosine or tryptophan for either 280 nm UV detection or fluorescence detection.

In this work, tryptic peptide mapping achieved 98% of overall sequence coverage. Several small peptides generated by tryptic digest are highly hydrophilic and did not retain on the C18 column, thus were not detected by MS. In order to detect potential sequence variants over the full amino acid sequence, endoproteinase Asp-N was used for the secondary map. As a result, full sequence coverage was obtained by combination of tryptic and Asp-N peptide mapping. Both A183S and A152S sequence variants observed in tryptic peptide mapping were confirmed by LC–MS endoproteinase Asp-N peptide mapping. The A183S variant was observed in the Asp-N peptide LD8 (native LD8 peptide: DSTYSLSSTLTLSKA, variant LD8 peptide: DSTYSLSSTLTLSKS) at a relative percentage of approximately 10% calculated by LC–MS peak area. This is consistent with the results obtained from the tryptic peptide mapping. The A152S variant was observed in the Asp-N peptide LD6 (native LD6 peptide: DNALQSGNSQESVTEQ, variant LD6 peptide: DNSLQSGNSQESVTEQ) at relative percentage of approximately 0.5% by LC–MS peak area. Both the variant peptide and native peptide LD6 were well separated from each other and from other peptides. The expanded UV trace at 215 nm displaying peaks from the variant peptide and native peptide LD6 is shown in Fig. 7. The A152S variant level was also obtained by 215 nm UV detection and the percentage was calculated to be approximately 0.5%.

3.5. cDNA sequencing

One-step reverse transcription polymerase chain reaction (RT-PCR) amplification of the sample RNAs resulted in the generation of PCR amplicons of the expected sizes. Sequencing of the PCR products generated 1404 base pairs (bp) of sequence data with 2× coverage for the heavy chain and 693 bp with 2× coverage data for the light chain. No differences were observed in the heavy chain sequences generated from the three samples MCB, MWCB, and EPCB when compared with the reference sequence. Two instances of mixed bases (T and G nucleotides together designated by the IUPAC code "K") were observed at positions 598 and 671 and one mixed base (A and G) at position 677 of the open reading frame in the light chain sequences generated from all 3 samples when compared with the reference sequence. Nucleotide position 598 corresponds to a wild-type alanine at position 183 (GCA) of the amino acid sequence. The presence of a mixed base results in a mutation to serine (TCA) and a potential mixed population at the peptide level. The presence of this amino acid sequence variant was confirmed by tryptic peptide map analysis described above. Two

Table 1
Estimated amounts of alanine to serine sequence variants detected in tryptic digests of different lots of antibody samples by LC–MS, LC–UV and tyrosine fluorescence detection.

Samples	Ala^{L183}-to-Ser L16 peptide			Ala^{L152}-to-Ser L14 peptide
	LC–MS area %	LC–UV 280 nm area %	Tyr fluorescence ex/em 229/305 nm area %	LC–MS area %
Lot A	8.8	8.7	7.8	0.6
Lot B	9.9	8.3	7.8	0.6
Lot C	7.8	7.9	7.3	0.5
Lot D	9.0	9.0	8.7	0.5
Lot E	9.5	10.0	9.6	0.5

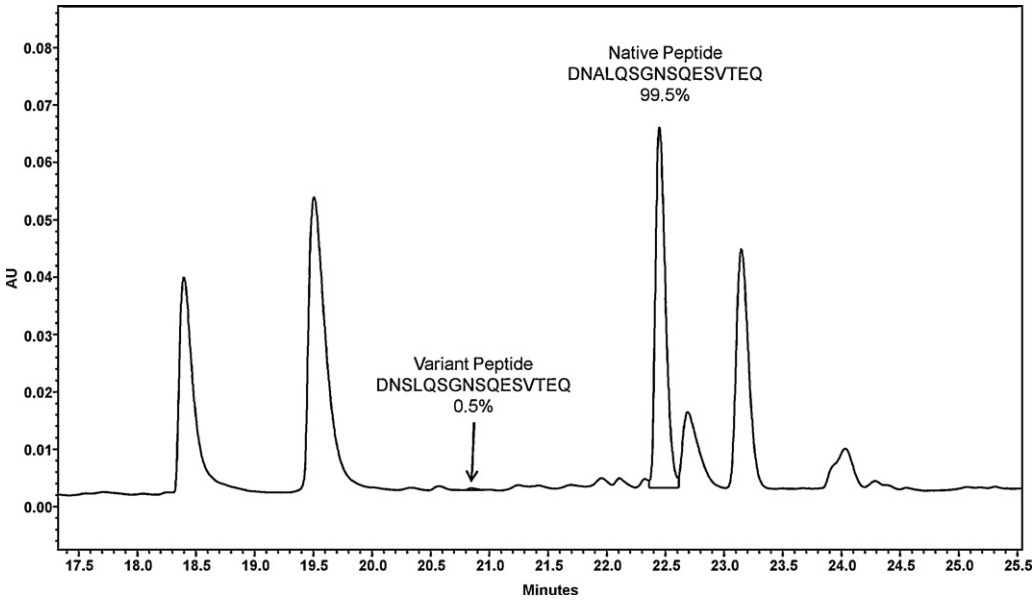

Fig. 7. Expanded UV trace of Asp-N peptide map of the antibody sample.

additional mixed base sequence were detected at position 671 and 677 of the light chain resulting in potential amino acid substitution at residue 207 (Ser to Ile) and substitution at residue 209 (Asn to Ser). However, these amino acid substitutions were not detected in the tryptic peptide map analysis. It is likely that the sequence containing Ser to Ile and Asn to Ser variants were not expressed. Even though theoretically possible, the possibility of the expressed population being cleared out during purification process should be very low.

Conversely, an Ala to Ser variant at position 152 of the light chain that was detected by tryptic peptide map analysis at a level of <1.0%, was not identified in the cDNA sequencing of the light chain. DNA sequencing only confirmed the expected codon (GCC) for Ala at position 152. Although a variation in DNA corresponding to position 152 was not detected, its existence at a level below 1%, as determined in expressed protein population, cannot be completely ruled out due to the sensitivity of cDNA sequencing. In addition, although the replication of DNA is much more accurate than mRNA transcription and tRNA translation in protein biosynthesis, with a spontaneous error rate of only 10^{-6}–10^{-10}, increased DNA mutation rate during transfection of DNA into mammalian cells has been reported [9]. Furthermore, translational errors are often found to randomly distribute in the proteins, for both codon-specific [5] (likely caused by misreading) and non codon-specific [6,37] (likely caused by mischarging of tRNA) mistranslations in recombinant proteins expressed in Chinese hamster ovary (CHO) cells. Among 10 Ala residues coded by the GCC codon in the light chain, only Ala152 was found to be replaced by Ser at a low level. Therefore, A152S variant is unlikely caused by mistranslation. To determine the root cause of this sequence variant, further investigation is required.

There were two possible sources of the mixed bases in the current cell line. One possibility was that the current MCB used to produce the antibody did not come from a pure single clone; the mixed populations had a higher chance to contain the mixed bases (sequence variants). Sub-cloning of the current MCB may not only confirm this possibility, but also provide an opportunity to generate a new cell line that does not have sequence variants. Another possibility was that the current cell line underwent spontaneous mutation in one or more copies, giving rise to microheterogeneity in the product; similar phenomena were reported in the other cell lines. The cell lines with higher copy numbers have been shown to have a higher probability of possessing mutated genes [3].

4. Conclusion

In summary, alanine to serine sequence variants were first observed in a reduced IgG4 monoclonal antibody as species with +16 Da mass difference from the expected mass of the light chain. Two variant sites were identified using LC–MS peptide mapping with tandem mass spectrometry sequencing. The A183S variation was determined to result from genetic mutation and confirmed by subsequence cDNA sequencing analysis. Minimization of sequence variants, particularly those resulting from mutation at the DNA level, is beneficial to lessen the possibility of negative impact to product development timelines from the standpoint of reduced concerns in process robustness as well as reduced risk in obtaining health authority approval. In the case reported here, identifying a significant sequence variant in the early stage of clinical development allowed for corrective actions to be taken in order to avoid costly corrective measures down the road. This work highlights the importance of comprehensive characterization of protein therapeutics during early stage of development. It also demonstrates the effectiveness of applying orthogonal analytical technologies (e.g., LC–MS intact mass measurement, LC–MS peptide mapping and tandem mass spectrometry were used in this study) in identification and characterization of protein variants.

References

[1] M. Wan, F.Y. Shiau, W. Gordon, G.Y. Wang, Biotechnol. Bioeng. 62 (1999) 485.
[2] 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.
[3] H. Dorai, T. Sauerwald, A. Campbell, Y.S. Kyung, J. Goldstein, A. Magill, M.J. Lewis, Q.M. Tang, D. Jan, S. Ganguly, G. Moore, BioProcess Int. 5 (2007) 66.
[4] D. Guo, A. Gao, D.A. Michels, L. Feeney, M. Eng, B. Chan, M.W. Laird, B. Zhang, X.C. Yu, J. Joly, B. Snedeco, A. Shen, Biotechnol. Bioeng. 107 (2010) 163.
[5] X.C. Yu, O.V. Borisov, M. Alvarez, D.A. Michels, Y.J. Wang, V. Ling, Anal. Chem. 81 (2009) 9282.
[6] D. Wen, M.M. Vecchi, S. Gu, L. Su, J. Dolnikova, Y. Huang, S.F. Foley, E. Garber, N. Pederson, W. Meier, J. Biol. Chem. 284 (2009) 32686.
[7] D. Ren, J. Zhang, R. Pritchett, H. Liu, J. Kyauk, J. Luo, A. Amanullah, J. Chromatogr. B 879 (2011) 2877.
[8] A.H. Que, B. Zhang, Y. Yang, J. Zhang, G. Derfus, A. Amanullah, BioProcess Int. 8 (2010) 52.

[9] J.S. Lebkowski, R.B. Dubridge, E.A. Antell, K.S. Greisen, M.P. Calos, Mol. Cell. Biol. 4 (1984) 1951.
[10] J. Parker, Microbiol. Rev. 53 (1989) 273.
[11] J. Langridge, Aust. J. Biol. Sci. 27 (1974) 309.
[12] V. Nene, R.E. Glass, Mol. Gen. Genet. 194 (1984) 166.
[13] J.R. Knowles, Science 236 (1987) 1252.
[14] C.G. Cupples, J.H. Miller, Genetics 120 (1988) 637.
[15] R.J. Harris, Dev. Biol. 122 (2005) 117.
[16] International Conference on Harmonization (ICH), Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products, 1999, August.
[17] Committee on Process Development and Manufacturing, Biotechnology Division, Biological Section, Pharmaceutical Manufacturers Association, BioPharm. Int. 4 (1991) 22.
[18] K. Ahrer, A. Jungbauer, J. Chromatogr. B 841 (2006) 110.
[19] Y. Wada, J. Chromatogr. B 781 (2002) 291.
[20] D. Allen, R. Baffi, J. Bausch, J. Bongers, M. Costello, J. Dougherty Jr., M. Federici, R. Garnick, S. Peterson, R. Riggins, K. Sewerin, J. Tuls, Biologicals 24 (1996) 255.
[21] P.M. Yuan, D.E. Ryan, W. Levin, J.E. Shively, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 1169.
[22] R. Kaneko, Y. Wada, M. Hisada, H. Naoki, T. Matsuo, J. Chromatogr. B 731 (1999) 125.
[23] A. Lim, T. Prokaeva, M.E. McComb, P.B. O'Connor, R. Théberge, L.H. Connors, M. Skinner, C.E. Costello, Anal. Chem. 74 (2002) 741.
[24] K. Tanaka, S. Takenaka, S. Tsuyama, Y. Wada, J. Am. Soc. Mass Spectrom. 17 (2006) 508.
[25] Y. Yang, A. Strahan, C. Li, A. Shen, H. Liu, J. Ouyang, V. Katta, K. Francissen, B. Zhang, mAbs 2 (2010) 285.
[26] C.L. Gatlin, J.K. Eng, S.T. Cross, J.C. Detter, J.R. Yates III, Anal. Chem. 72 (2000) 757.
[27] Y. Wada, Biol. Mass Spectrom. 21 (1992) 617.
[28] T.J. Dayhuff, J.F. Atkins, R.F. Gesteland, J. Biol. Chem. 261 (1986) 7491.
[29] A.G. Bruce, J.F. Atkins, R.F. Gesteland, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 5062.
[30] M. Guo, Y.E. Chong, R. Shapiro, K. Beebe, X.-L. Yang, P. Schimmel, Nature 462 (2009) 808.
[31] K. Beebe, M. Mock, E. Merriman, P. Schimmel, Nature 451 (2008) 90.
[32] R.H. Fillingame, M. Oldenburg, D. Fragas, J. Biol. Chem. 266 (1991) 20934.
[33] B. Saul, V. Schmiedenb, C. Kling, C. Mülhardt, P. Gassay, J. Kuhseb, C.-M. Beckera, FEBS Lett. 350 (1994) 71.
[34] L. Yu, R.L. Remmele, B. He, Rapid Commun. Mass Spectrom. 20 (2006) 3674.
[35] Z. Zhang, H. Pan, X. Chen, Mass Spectrom. Rev. 28 (2009) 147.
[36] H. Liu, G. Gaza-Bulsec, D. Faldu, C. Chumsae, J. Sun, J. Pharm. Sci. 97 (2008) 2426.
[37] A. Khetan, Y.-M. Huang, J. Dolnikova, N.E. Pederson, D. Wen, H. Yusuf-Makagiansar, P. Chen, T. Ryll, Biotechnol. Bioeng. 107 (2010) 116.