International Journal of Mass Spectrometry 312 (2012) 61–69

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

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

Primary sequence determination of a monoclonal antibody against α-synuclein using a novel mass spectrometry-based approach

Eric Sousa[a,1], Stephane Olland[a,1], Heather H. Shih[a,1], Kim Marquette[a], Robert Martone[b], Zhijian Lu[a], Janet Paulsen[a], Davinder Gill[a], Tao He[a,*]

[a] Global Biotherapeutics Technologies, Pfizer, 87 Cambridge Park Dr., Cambridge, MA 02140, USA
[b] Neuroscience Research Unit, Pfizer, Groton, CT, USA

ARTICLE INFO

Article history:
Received 22 February 2011
Received in revised form 2 May 2011
Accepted 9 May 2011
Available online 14 May 2011

Keywords:
de novo sequencing
Antibody
Stable isotope labeled amino acids in cell culture (SILAC)
α-Synuclein
Collision induced dissociation (CID)
Higher energy collision dissociation (HCD)

ABSTRACT

A novel mass spectrometry-based workflow for *de novo* sequencing of a full-length monoclonal antibody (LB509) against α-synuclein is presented. This approach combines chemical modification of cysteine residues, multiple enzymatic digestion, stable isotope labeled amino acids in cell culture (SILAC), liquid chromatography coupled tandem mass spectrometry (LC/MS/MS) analysis including collision-induced dissociation (CID), higher energy collision-induced dissociation (HCD), and *de novo* sequencing software for data interpretation. The introduction of aminoethyl-8 (A-8) reagent which renders cysteine residues in the antibody to become substrates for trypsin digestion significantly increases the sequence coverage of the antibody, especially in the complementarity determining regions (CDRs). Incorporation of SILAC helps to distinguish minor sequence variations between original and deduced sequences. Methods in obtaining detailed sequence characterization are described, highlighting the advantage of using multiple fragmentation schemes, CID in ion trap and HCD in C-trap. Complete primary sequence for LB509 has been construed and the antibody generated exhibits specific binding to α-synuclein. The development and implementation of this technology enables an alternative approach for generating therapeutic candidate and reagent antibodies.

1. Introduction

Parkinson's disease (PD) and dementia with Lewy bodies (DLB) are neurodegenerative disorders where accumulation of pathologic α-synuclein-containing Lewy bodies in brain tissue has been shown to be one of the major morphologic and molecular events in disease progression [1]. LB509 is a mouse monoclonal antibody raised against Lewy bodies purified from the brains of deceased patients with DLB [2]. Since its generation, LB509, an antibody that specifically recognizes α-synuclein, has been widely used as a research tool to study the function of α-synuclein and its pathological association with diseases such as PD and DLB [3].

Monoclonal antibodies have been accepted as an important therapeutic modality for a variety of human disease treatment, including cancer, immunological disorders and transplant rejection [4]. Antibody-based therapeutics offer high specificity to selected targets, owing to complementarity determining regions (CDRs) in variable region. Recombination of variable (V), diversity (D), and junctional (J) gene segment of the heavy chain (HC) and VJ gene segments of the light chain (LC) as well as somatic hypermutation define the specificity of the variable region in antibodies [5]. Therefore, without cDNA information (such as loss of hybridoma or externally acquired antibodies), it is extremely challenging to obtain the complete primary sequence of monoclonal antibodies. Edman degradation [6] is a widely accepted approach to delineate unknown protein sequence. However, it is a low throughput process where large quantities of highly purified samples are needed. Mass spectrometry-based peptide and protein analysis has become an indispensable tool with the development of soft ionization techniques, namely, electrospray [7] and matrix-assisted laser desorption/ionization (MALDI) [8,9]. One of the important factors that facilitate unknown protein identification for proteomics studies is the availability of protein database, where tandem mass spectra are searched against known theoretical sequences [10–12]. Development of stable isotope labeling reagents, such as stable-isotope labeling by amino acids in cell culture (SILAC) [13] and isotope-

Abbreviations: SILAC, stable isotope labeled amino acids in cell culture; CID, collision induced dissociation; HCD, higher-energy collisiondissociation; CDR, complementary determining region; HC, heavy chain; LC, light chain; MALDI, matrix-assisted laser desorption/ionization; LC/MS/MS, liquid chromatography coupled tandem mass spectrometry; RE, reverse-engineered; A-8, aminoethyl-8; IAA, iodoacetamide; ELISA, enzyme-linked immunosorbent assay.

* Corresponding author. Tel.: +1 617 665 8021.
E-mail address: tao.he@pfizer.com (T. He).

[1] Contributed equally to this work.

1387-3806/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2011.05.005

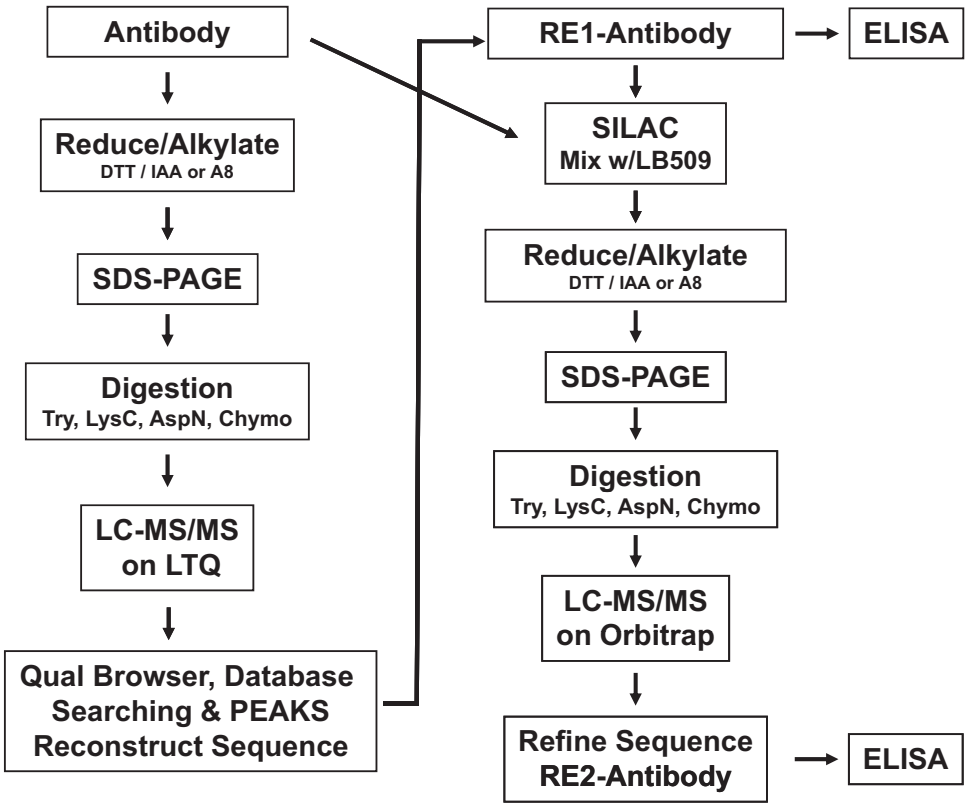

Fig. 1. Overall workflow for antibody reverse engineering. The procedure will be performed for both heavy and light chain of LB509.

encoded affinity tags (ICAT) [14], has enabled the emergence of quantitative proteomics for large-scale kinetic studies of a proteome.

Although it is routine that hundreds if not thousands of proteins could be identified by mass spectrometry-based proteomics approach [15], it is still an area of active research to develop workflows for *de novo* protein sequencing when a relevant protein database is not available [16]. For delineating the primary sequence of antibodies, CDR regions are extremely challenging due to the incomplete sequence coverage for CDRs in protein databases. Pham et al. presented the first *de novo* sequence study on a monoclonal antibody against OX40 ligand, applying both Edman sequencing and mass spectrometry in combination with proteolytic and chemical digestion [17]. Recently, Perdivara et al. deduced primary sequence of an antibody against β-amyloid (clone 6E10) mainly based on liquid chromatography coupled with tandem mass spectrometry [18]. In this study, *de novo* tandem mass spectrometry analysis provided majority of the variable domain sequence information. Additionally, advances in software development were reported in support of *de novo* sequencing effort for monoclonal antibodies, which was exemplified by comparative shotgun protein sequencing (CSPS) approach [19]. However, it is still a challenge to obtain complete protein sequence of a monoclonal antibody and to demonstrate the retention of activity of a reverse-engineered (RE) antibody, especially in distinguishing minor sequence variability between original and RE antibodies.

In this study, we describe a novel workflow for *de novo* protein primary sequence analysis of monoclonal antibodies. This approach combines chemical modification of cysteine residues to become tryptic/LysC proteolysis substrates, multiple enzymatic digestion, SILAC for detailed sequence comparison, high resolution LC/MS/MS analysis and *de novo* sequencing software for data interpretation. We applied this *de novo* sequencing method to deduce the primary sequence of LB509. Importantly, we showed that the reverse engineered LB509 antibody retained its specific α-synuclein binding activity.

2. Experimental

2.1. Materials

Mouse anti-α-synuclein monoclonal antibody (LB509, SIG-39725) was purchased from Covance (Princeton, NY). Aminoethyl-8 was obtained from Pierce (Rockford, IL). Heavy Lysine and Arginine amino acids were purchased from Cambridge Isotope Lab Inc (Andover, MA). Enzymes (Trypsin, LysC, Chymotrypsin, AspN) were acquired from Roche Diagnostics (Indianapolis, IN). All other reagents were acquired from Sigma–Aldrich (St. Louis, MO) of the highest grade.

2.2. SDS-PAGE and in-gel digestion

The monoclonal antibody LB509 was reduced with DTT and alkylated with either A-8 or iodoacetamide (IAA) and separated by SDS-PAGE (10% Tricine). The light and heavy chain bands were excised and digested with different enzymes (trypsin, LysC, chymotrypsin and AspN) on a digestion workstation DigestPro (ABimed GmbH, Langenfeld, Germany). Briefly, bands were washed with 100 mM ammonium bicarbonate after shrinking with acetonitrile. The enzyme was added (0.5 μg/gel band) and incubated at 37 °C for 8 h. The resulting peptides were extracted with acetonitrile washes followed by 5% formic acid washes. In the case of double digestion, AspN followed by trypsin, the AspN digest was carried out first and 5% formic acid was replaced with 100 mM ammonium bicarbonate. The resulting AspN peptides were further digested with trypsin.

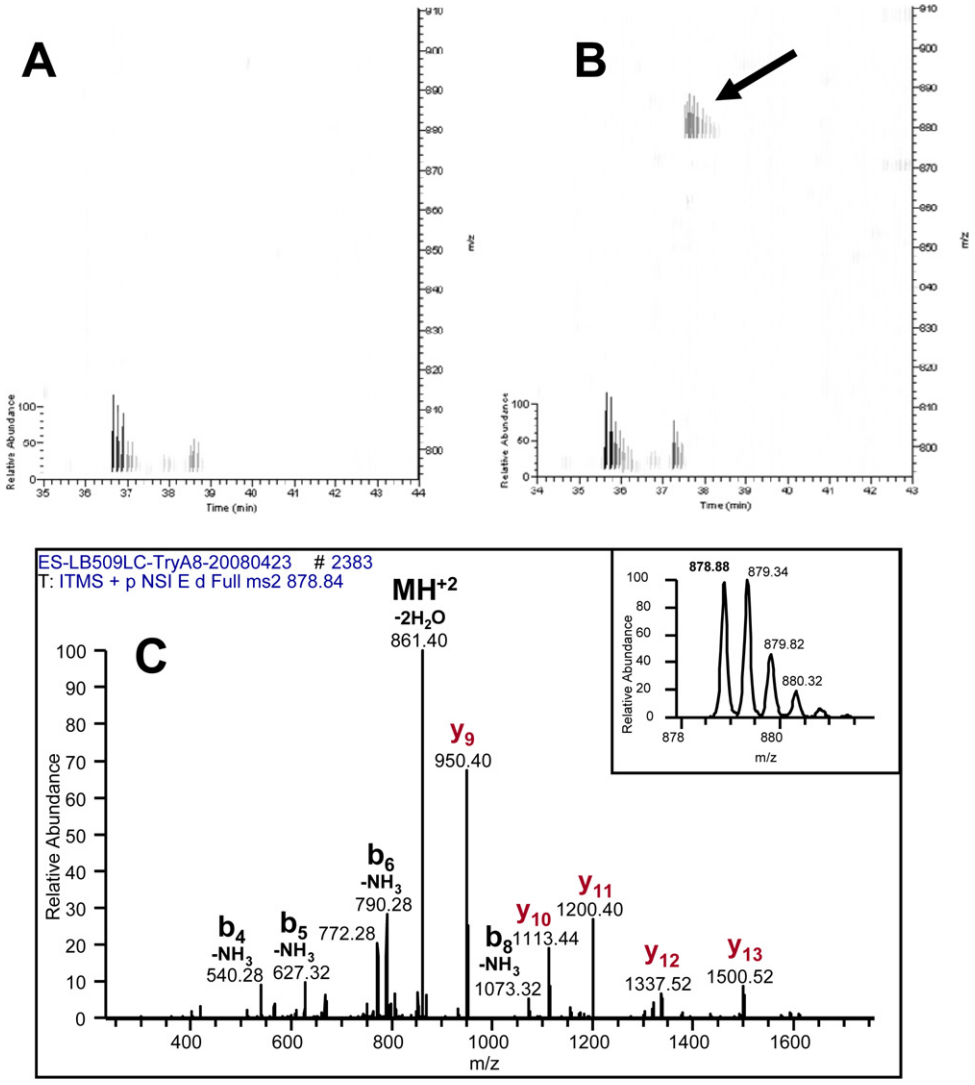

Fig. 2. Use of visualization tool to aid identification of CDR-containing peptides after A-8 modification and trypsin digestion. (A) 2D map of LB509 sample alkylated with IAA and digested with trypsin; (B) 2D map of LB509 sample alkylated with A-8 and digested with trypsin, the arrow highlights the additional peptide ion; (C) MS/MS spectrum of highlighted peptide ion in 2B, the peptide was determined to be light chain CDR3-containing: QQYHSYPWTFGGGTK. Insert shows MS spectrum of the peptide.

2.3. Liquid chromatography mass spectrometry

Recovered peptide samples from in-gel digest were loaded onto a Pico-frit column (New Objective, Woburn, MA) packed with reversed-phase Magic (Michrom, Auburn, CA) C18 material (5 μm, 200 Å, 75 μm × 10 cm) and coupled to an LTQ or LTQ-Orbitrap XL mass spectrometer (ThermoElectron, Waltham, MA). Chromatographic methods were identical for all the samples analyzed. Peptides were separated at a flow rate of 0.2 μl/min using a 90 min linear gradient ranging from 2% to 40% B (mobile phase A: 0.1% formic acid/2% ACN; mobile phase B: 90% ACN/0.1% formic acid). Electrospray voltage was 1.8 kV. The instrumental method consisted of a full MS scan (scan range 375–1550 *m/z*, with 30 K fwhm resolution @ *m/z* 400, target value 2×10^6, maximum ion injection time of 500 ms) followed by data-dependent CID scans of the four most intense precursor ions. Peptide precursor ions were selected with an isolation window of 2.5 Da and a target value of 1×10^5. Dynamic exclusion was implemented with a repeat count of 2 and exclusion duration of 75 s. The normalized collision energy (NCE) was set to 35% for CID and 40% for HCD, respectively.

2.4. Database search, visualization software and de novo sequencing

The mass spectra were searched against a public antibody database using Bioworks 3.3.1 SP1 with SEQUEST seach algorithm (ThermoElectron, Waltham, MA). The mass accuracy was set to 5 ppm for precursor ions and to 0.5 Da tolerances for fragment ions. The search parameters took into account two missed cleavages for trypsin, static modification of carboxamidomethylation at cysteine (+57.0215 Da) or A-8 (+43.0420), and dynamic modification for methionine oxidation (Met +15.9949 Da). Data from A-8 and IAA modifications were displayed in Qual Browser (ThermoElectron, Waltham, MA). The mass spectra were also searched with *de novo* software Peaks (Bioinformatics Solutions Inc, Waterloo, ON) for potential matches and all hits were manually inspected, interpreted and the light chain and heavy chain amino acid sequences were constructed.

2.5. Gene construction for deduced LB509

The protein sequences of LB509 heavy chain and light chain were provided to GeneArt (Burlingame, CA) for conversion to

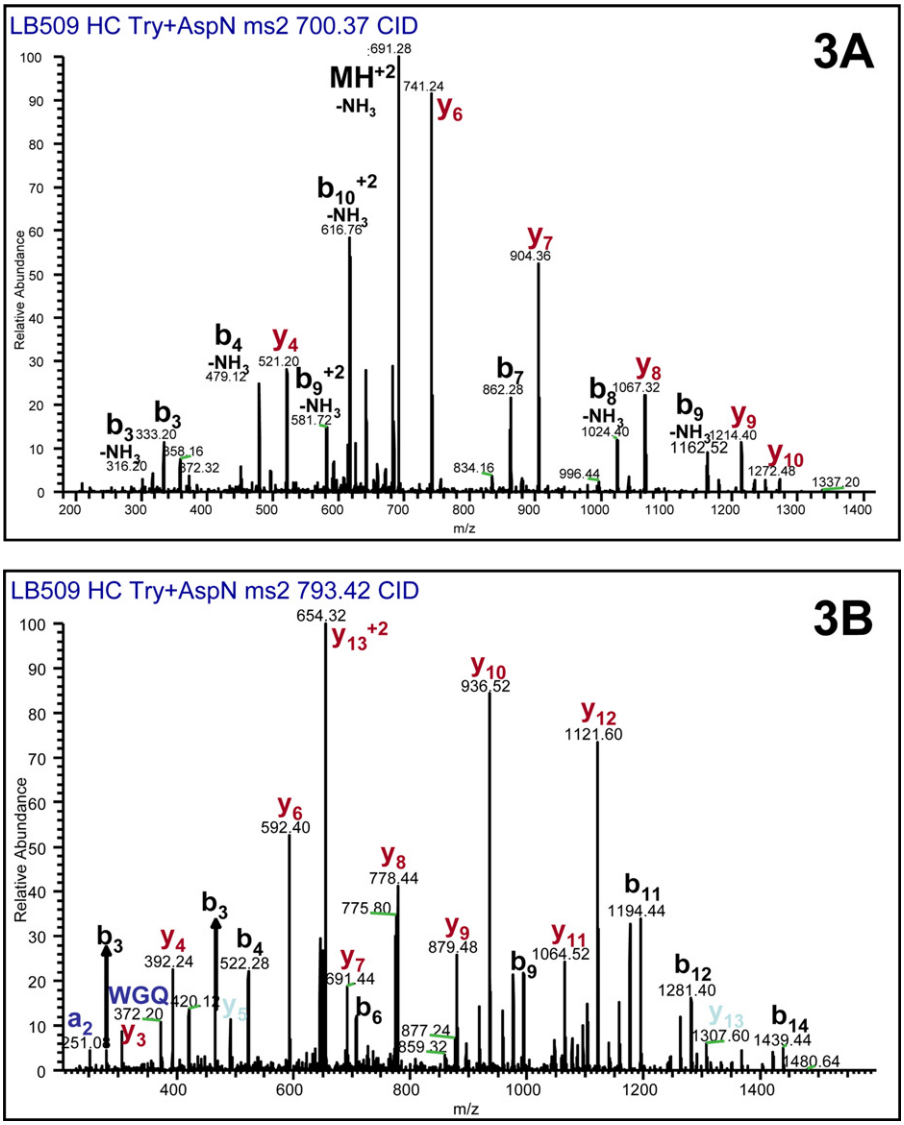

Fig. 3. MS/MS spectra of double digestion with AspN and trypsin for deducing heavy chain CDR3-containing peptides, (A) peptide QGFYYGYYHAM; and (B) peptide DYWGQGTSVTVSSAK.

DNA coding sequences with codons optimized for mammalian cell expression. The DNA fragments encoding the heavy and light chains of the deduced LB509 (RE-LB509) were cloned into the mammalian expression vectors pSMED and pSMEN to express the heavy and light chains, respectively.

2.6. Expression and purification of RE-LB509

The DNA constructs encoding RE-LB509 heavy and light chains were co-transfected into HEK293 cells using Lipofectamin2000 (InVitrogen, Carlsbad, CA). In 24 h post transfection, the media was changed to either 293 serum-free media (InVitrogen) or 293 serum-free media containing stable isotope labeled amino acids Lysine and Arginine (Cambridge Isotope Lab Inc), and were harvested in 48 h. The conditioned media expressing RE-LB509 was purified on Protein G sepharose column to capture the RE-LB509 antibody. The antibody was subsequently eluted with 20 mM citric acid and 150 mM NaCl at pH 2.5 and immediately neutralized with 1 M Tris pH 8. The antibody was buffer exchanged into PBS. The purified antibody was analyzed by size exclusion chromatography and SDS-PAGE followed by staining with SimpleBlue according to the manufacturer's instructions (InVitrogen).

2.7. Western blot analysis

Proteins in the conditioned media were separated on SDS/PAGE gels and subsequently transferred onto nitrocellulose membranes. The membranes were probed with peroxidase-conjugated anti-mFc antibody. The heavy chain of the antibody was detected with Western Lightning Chemiluminescence reagents (PerkinElmer, Waltham, MA).

2.8. α-Synuclein binding ELISA

Full-length a-synuclein was coated on high-binding ELISA plate (CoStar, Lowell, MA) at 1 μg/ml overnight at 4 °C. The plate was washed with washing buffer (PBS/0.1% Tween-20) and subsequently blocked with assay buffer (PBS/0.1% Tween-20/1% BSA) for 1 h at 25 °C. Conditioned media or purified antibodies diluted in assay buffer were added to the blocked plate and incubated for 1 h at 25 °C. After washing, the plate was incubated with peroxidase-conjugated anti-mFc antibody (Jackson's Laboratories, Bar Harbor, ME) for 1hr at 25 °C. The plate was then washed and developed by incubating with tetramethylben-

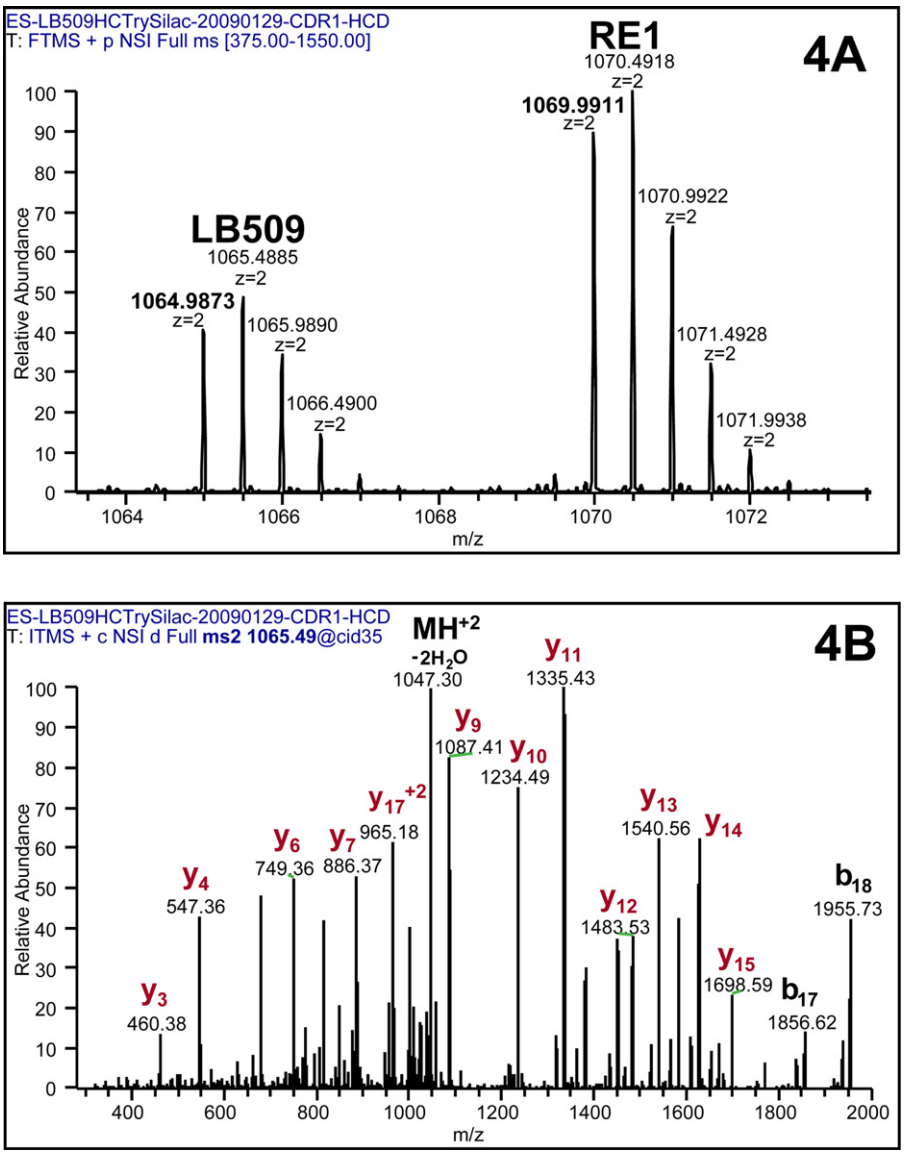

Fig. 4. SILAC peptide pair of heavy chain CDR1-containing peptide. (A) Mass spectrum showing unlabeled light (LB509) and labeled heavy (RE1) peptide LSCAASGFTFSNHAMSWVR; and (B) MS/MS spectrum of light (LB509) peptide LSCAASGFTFSNHAMSWVR.

zidine substrate, and the optical density was measured on a UVmax microplate reader (Molecular Devices, Sunnyvale, CA) at 450 nm.

3. Results and discussion

Majority of the constant region sequences of a mouse IgG1 can be found in current databases. However, the variable domain containing CDRs are not complete in sequence databases. In order to overcome the challenges posted by these CDRs, we developed a strategy combining (a) A-8 reagent to convert conserved cysteines present adjacent to CDR1 and CDR3 in both light and heavy chains to tryptic and LysC sites; (b) employing multiple enzymes to generate overlapping sequences; (c) incorporating SILAC to determine minor differences between original and deduced sequences; (d) applying both CID and HCD for detailed sequence characterization. The overall workflow for reverse engineering the anti α-synuclein antibody LB509 is shown in Fig. 1.

3.1. Advantage of introducing A-8 in decoding CDRs

It has been shown that cysteine is a conserved residue before CDR1 and CDR3 in light chain and several amino acids before CDR1 and CDR3 in heavy chain [20]. To identify peptides containing CDRs, A-8 reagent was used to modify these conserved cysteine residues proximal to the various CDRs. A-8 is an alkylating reagent specific to free thiols that introduces a primary amine on cysteine's side chain, rendering it into a trypsin and LysC substrate [21]. LB509 antibody was reduced by DTT and alkylated with either IAA or A-8, and heavy and light chains were separated by SDS-PAGE. In-gel digestions were performed on both samples and recovered peptides were subjected to LC/MS/MS analysis. Additional peaks would be expected in A-8 alkylated sample, compared to an IAA alkylated sample, since the modification should introduce extra tryptic/LysC sites. Qual Browser software is used as a visualization tool for locating the cysteine-containing peptide. Fig. 2A and B shows the peptide ion containing LC CDR3 displayed in Qual Browser software (arrow). Database searching identified the peptide sequence as C.QQYHSYPWTFGGGTK.L (Fig. 2C). The use of A-8 reagent evidently facilitated the identification of peptides containing CDRs

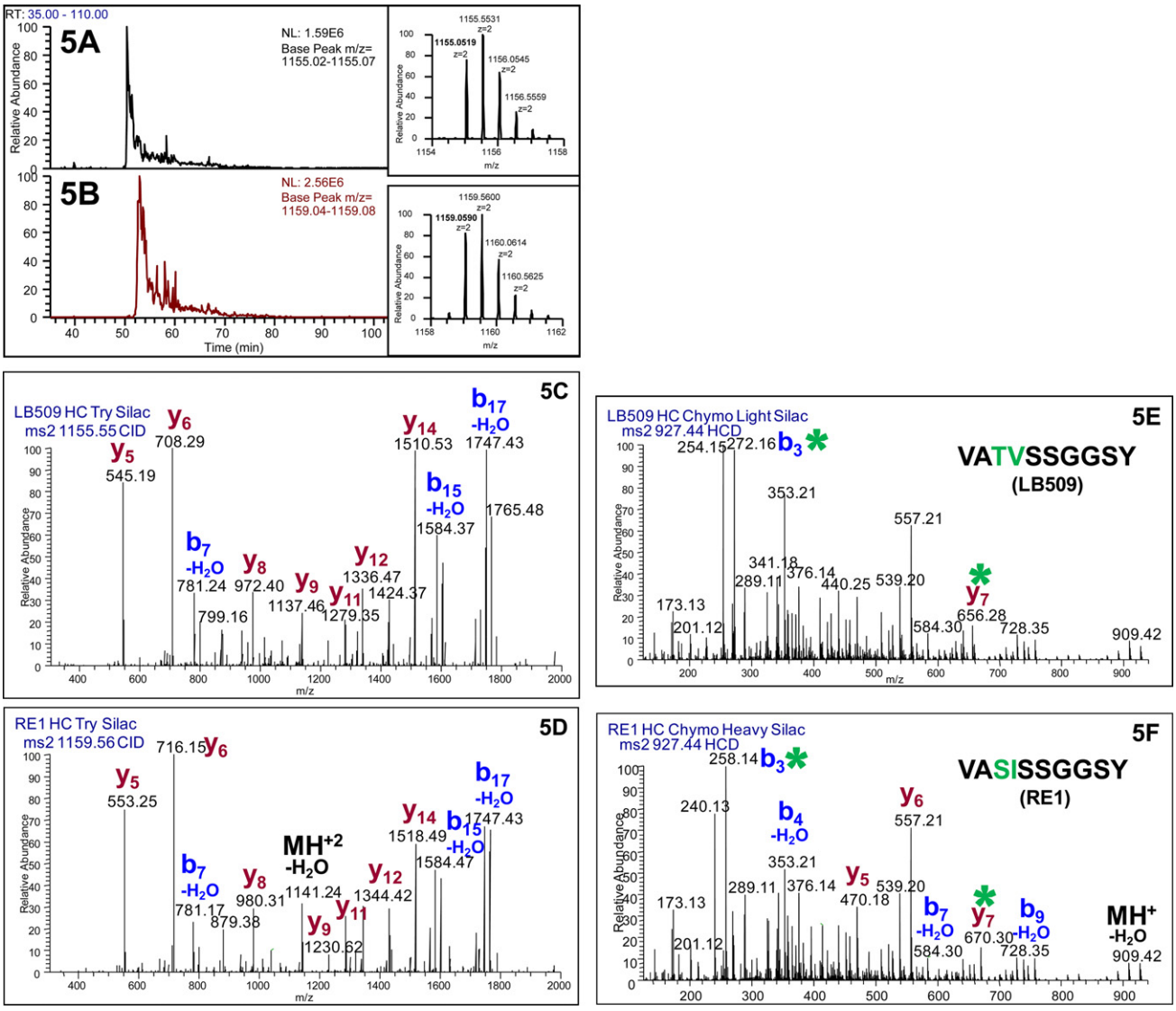

Fig. 5. Determination of heavy chain CDR2-containing peptide sequence by SILAC and combination of CID and HCD. Extracted ion chromatograms of heavy chain CDR2-containing peptide of (A) light (LB509) and (B) heavy (RE1), inserts show the mass spectra of light and heavy peptide (assigned as LEWVASISSGGSYTYYPDSVK in RE1) with a mass deviation of 0.35 ppm and 0.26 ppm, respectively. CID MS/MS spectra are shown for (C) light and (D) heavy peptide with matching ions labeled. HCD MS/MS spectra are displayed for chymotrypsin digested peptide, (E) light (LB509) and (F) heavy (RE1) samples (no heavy amino acid Arg or Lys due to chymotrypsin digest). *: denotes the different fragment ions (b_3 and y_7) which could be explained as Thr-Val in LB509 (light) and Ser-Ile in RE1 (heavy) samples.

by generating additional tryptic peptides in lengths suitable for adequate MS spectra interpretation.

3.2. *De novo* sequencing of heavy chain CDR3

It is known that heavy chain CDR3 typically starts with CXX (normally CAR) and ends with WGXG which provides us with sequence tags to locate the potential hits from database search results. Several hits showed up in database search results with WGXG sequence tag, but with low confidence score. Initial attempts to use *de novo* software Peaks to decode the sequences containing WGXG were not completely successful due to the size of the peptide being too large (2965 Da) and only the C-terminal region of the peptide could be interpreted as AMDYWGQGTSVTVSSAK (Supplementary Fig. 1A). Next overlapping enzymatic coverage was attempted to complete the sequence for heavy chain CDR3. Peptides identified from chymotrypsin digest revealed that part of heavy chain CDR3 sequence could contain CARQGFY (Supplementary Fig. 1B). To delineate the missing amino acid sequence within heavy chain CDR3, a combination of AspN and trypsin digestion was performed and sequence tag QGFY was used to highlight potential heavy chain CDR3 containing peptide. *De novo* software Peaks identified a peptide sequence of QGFYYGYYHAM (m/z = 700.29, z = 2). Upon manual inspection, nearly complete sequence coverage was obtained for this peptide (Fig. 3A). Furthermore, the C-terminal portion of the heavy chain CDR3 was determined as DYWGQGTSVTVSSAK (m/z = 793.38, z = 2) (Fig. 3B). The successful deduction of heavy chain CDR3 sequence (CAR.QGFYYGYYHAMDY.WGQG) clearly demonstrated the power of multiple enzymatic digestion combined with LC/MS/MS to produce overlapping peptides for *de novo* sequencing of unknown CDRs.

3.3. Generation of version 1 of reverse-engineered LB509

Combining data from multiple enzymatic digestions, the first version of the deduced sequence for LB509 (dubbed RE1) by low resolution mass spectrometer LTQ is assembled (Supplementary Fig. 2). The DNA for both the heavy and light chain was synthesized,

RE2 Light chain:
QIVLTQSPAIMSASPGEKVTISCSASSSVSYMYWYQQKPGSSPKLL
PWRTSNLASGVPARFSGSGSGTSYSLTISSMEAEDAATYYCQQYH
SYPWTFGGGTKLELKRADAAPTVSIFPPSSEQLTSGGASVVCFLN
NFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLT
KDEYERHNSYTCEATHKTSTSPIVKSFNRNEC

RE2 Heavy chain:
EVMLVESGGGLVKPGGSLKLSCAASGFTFSNHAMSWVRQTPEKR
LEWVATVSSGGSYTYYPDSVKGRFTISRDNAKNTLYLQMSSLRSE
DTAMYYCARQGFYYGYYHAMDYWGQGTSVTVSSAKTTPPSVYPL
APGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAV
LQSDLYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRDC
GCKPCICTVPEVSSVFIFPPKPKDVLTITLTPKVTCVVVDISKDDPEV
QFSWFVDDVEVHTAQTQPREEQFNSTFRSVSELPIMHQDWLNGK
EFKCRVNSAAFPAPIEKTISKTKGRPKAPQVYTIPPPKEQMAKDKV
SLTCMITDFFPEDITVEWQWNGQPAENYKNTQPIMDTDGSYFVYSK
LNVQKSNWEAGNTFTCSVLHEGLHNHHTEKSLSHSPGK

Fig. 6. Deduced sequences for the anti-α-synuclein antibody LB509 light and heavy chains. Underlined sequences are amino acids corrected in RE2 compared with RE1.

cloned into expression vectors. The RE1 antibody was expressed from transiently transfected 293 cells. Good expression level was achieved for the antibody as assayed by Western (Supplementary Fig. 3A). However, no binding to α-synuclein was observed by ELISA (Supplementary Fig. 1B).

3.4. Incorporation of SILAC in sequence determination of LB509

One apparent explanation for the lack of α-synuclein binding activity by RE1 would be sequence variations between RE1 antibody and the original LB509 antibody. To determine whether this was the case, stable isotope labeled amino acids (heavy Lys and heavy Arg) were incorporated into RE1 which was subsequently analyzed by a high resolution mass spectrometer, LTQ-Orbitrap XL. This SILAC labeled "heavy" RE1 antibody was mixed with the unlabeled "light" original LB509 antibody for LC/MS/MS analysis. Correctly deduced sequences would generate peptide ion pairs in mass spectra with exactly the same retention time. In contrast, unpaired spectra are expected for those incorrectly assigned sequences. Fig. 4A displays an example of a correctly assigned peptide sequence with SILAC ion pairs showing both light (original LB509) and heavy (RE1 antibody) CDR1 (heavy chain) corresponding to peptide LSCAASGFTFSNHAMSWVR (cysteine was modified by iodoacetamide). Tandem mass spectrum of the SILAC light (original LB509) is shown in Fig. 4B.

Heavy chain CDR2-containing tryptic peptide pair was found in the SILAC chromatogram where the correct *m*/*z* was detected for both SILAC heavy and light species by LC/MS/MS on Orbitrap (mass error less that 1 ppm); however, slightly a different retention time was observed for the SILAC peptides (Fig. 5A and B). Through database search, initial identification of the peptide based on CID assigned the peptide sequence as LEWVASISSGGSYTYYPDSVK. Examination of the CID spectra for both SILAC heavy and light peptides confirmed majority of the sequence assignment, with multiple ions missing in the low mass region, as reported previously due to the low mass cutoff in ion trap mass spectrometers (Fig. 5C and D). Further analysis was carried out by chymotrypsin digestion and heavy and light samples were analyzed separately

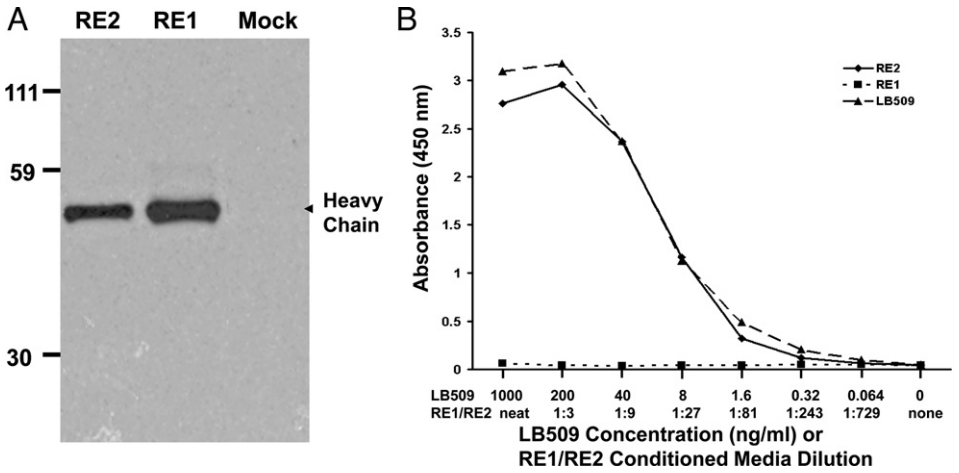

Fig. 7. Expression and activity confirmation of reverse-engineered LB509. (A) Western blot shows good expression for both RE1 and RE2 versions of reverse-engineered LB509; (B) RE2-LB509 was found to be active in specific binding to α-synuclein in ELISA assay, where RE1-LB509 was inactive.

by LC/MS/MS. Again presumably correct *m*/*z* and CID fragmentation patterns were identified for the peptide VASISSGGSY, but slightly different retention times (about 2 min) from heavy and light samples were observed, suggesting a sequence variation (data not shown). To determine the sequence variation, HCD fragmentation scheme [22] was applied to generate MS/MS spectra for both RE1 and original LB509 peptides. Upon examination of the HCD fragmentation spectra for RE1 and original LB509 peptide, most of the MS/MS peptide ions matched between the two peptides with the exception of *m*/*z* 240 (b_3-H_2O), 258 (b_3) and 670 (y_7) from heavy SILAC peptide sample (RE1). In the original (light SILAC) LB509 sample, *m*/*z* of those ions were 254, 272, and 656. The discrepancy in *m*/*z* of 14 between these ions could be attributed to a methylene group which led us to assign the corrected sequence as VA**TV**SSGGSY instead of VA**SI**SSGGSY in RE1 (Fig. 5E and F). Using the BLASTp program no matches to the VATVSSGGSY peptide are found in the NCBI public databases.

Other modifications to RE1 sequence found through the SILAC scheme include a change before light chain CDR2 from amino acids IY to PW, a D to N alteration within light chain CDR2, and an N to K amendment in the framework of the heavy chain (data not shown). A revised sequence for the second version of reverse engineered LB509 (dubbed RE2) is shown in Fig. 6.

The incorporation of SILAC in confirming the amino acid assignment by *de novo* sequencing of antibodies, especially CDR regions, significantly improved our ability to correctly deduce the incorrectly assigned sequences between the reverse-engineered and the original antibodies. It is close to impossible to distinguish the subtle differences in heavy chain CDR2 without the retention time discrepancy between the heavy labeled RE1 and unlabeled original antibodies in the SILAC experiment. The complete sequence of LB509 was derived solely from a mass spectrometry-based approach, thus circumventing the sample-consuming Edman sequencing analysis. Furthermore, this workflow is also applicable to terminally blocked antibodies, another advantage over Edman. The use of high resolution/mass accuracy mass spectrometry in combination with SILAC sample analysis further enhances the assignment confidence of identified peptide pairs. With the introduction of additional fragmentation scheme, such as ETD, into *de novo* peptide sequencing workflow, better sequence coverage is expected [23]. Though various software programs are reported to aid *de novo* peptide sequencing efforts [24,25], further software development, especially data processing capabilities optimized for multiple fragmentation techniques, will aid efficient data interpretation.

3.5. Activity confirmation of the second version of reverse-engineered LB509

After the amino acid sequence for RE2 LB509 was deduced, the DNA encoding the heavy and light chains was synthesized and cloned into expression plasmids. The antibody was subsequently expressed in the conditioned media of 293 cells transiently transfected with RE2 LB509 expressing plasmids. The expression level of RE2 was slightly lower than RE1 as measured by Western blot analysis (Fig. 7A). When tested in α-synuclein binding ELISA, the antibody showed specific binding activity (Fig. 7B), indicating that the second round of reverse engineering was successful to create an active antibody molecule.

The reverse engineered LB509 exhibited weaker binding activity than the parental antibody (data not shown), which could be due to imperfect sequence deduction. As a result, this technology should be conceptualized as a method mainly to recover an antibody similar enough to the parental antibody to exhibit specific antigen binding activity, rather than to re-create the perfect clone of the parental antibody. A partially active reverse-engineered antibody can be subsequently subjected to affinity maturation for the generation of a high-affinity antibody.

4. Conclusion

A novel workflow for *de novo* sequencing of antibodies where a genomic sequence is not available is presented. It combines site specific modification of conserved cysteine residues to aid interpretation of CDR sequences, enzymatic digestion with multiple enzymes, introduction of SILAC to label reverse-engineered antibody for detailed structural analysis and integration of data produced from CID and HCD fragmentation schemes. The resulted reverse-engineered antibody was shown to bind α-synuclein specifically, providing an alternative route for therapeutic candidate and reagent antibody generation.

Acknowledgements

We thank Wei Cao and Ying Sun for technical support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijms.2011.05.005

References

[1] K.A. Jellinger, Formation and development of Lewy pathology: a critical update, J. Neurol. 256 (Suppl. 3) (2009) 270–279.
[2] M. Baba, S. Nakajo, P.H. Tu, T. Tomita, K. Nakaya, V.M. Lee, J.Q. Trojanowski, T. Iwatsubo, Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies, Am. J. Pathol. 152 (1998) 879–884.
[3] J.L. St Martin, J. Klucken, T.F. Outeiro, P. Nguyen, C. Keller-McGandy, I. Cantuti-Castelvetri, T.N. Grammatopoulos, D.G. Standaert, B.T. Hyman, P.J. McLean, Dopaminergic neuron loss and up-regulation of chaperone protein mRNA induced by targeted over-expression of alpha-synuclein in mouse substantia nigra, J. Neurochem. 100 (2007) 1449–1457.
[4] J.M. Reichert, C.J. Rosensweig, L.B. Faden, M.C. Dewitz, Monoclonal antibody successes in the clinic, Nat. Biotechnol. 23 (2005) 1073–1078.
[5] A. Li, M. Rue, J. Zhou, H. Wang, M.A. Goldwasser, D. Neuberg, V. Dalton, D. Zuckerman, C. Lyons, L.B. Silverman, S.E. Sallan, J.G. Gribben, Utilization of Ig heavy chain variable, diversity, and joining gene segments in children with B-lineage acute lymphoblastic leukemia: implications for the mechanisms of VDJ recombination and for pathogenesis, Blood 103 (2004) 4602–4609.
[6] P Edman, Method for determination of the amino acid sequence in peptides, Acta Chem. Scand. 4 (1950) 283–293.
[7] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, Science 246 (1989) 64–71.
[8] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers, Anal. Chem. 63 (1991) 1193A–1203A.
[9] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of flight mass spectrometry, Rapid Commun. Mass Spectrom. 2 (1988) 151–153.
[10] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, C. Watanabe, Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 5011–5015.
[11] J.K. Eng, A.L. McCormack, J.R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989.
[12] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, Electrophoresis 20 (1999) 3551–3567.
[13] S.E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, M. Mann, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, Mol. Cell. Proteomics 1 (2002) 376–386.
[14] S.P Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, Nat. Biotechnol. 17 (1999) 994–999.
[15] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates 3rd, Direct analysis of protein complexes using mass spectrometry, Nat. Biotechnol. 17 (1999) 676–682.
[16] J. Seidler, N. Zinn, M.E. Boehm, W.D. Lehmann, De novo sequencing of peptides by MS/MS, Proteomics 10 (2010) 634–649.

[17] V. Pham, W.J. Henzel, D. Arnott, S. Hymowitz, W.N. Sandoval, B.-T. Truong, H. Lowman, J.R. Lill, De novo proteomic sequencing of a monoclonal antibody raised against OX40 ligand, Anal. Biochem. 352 (2006) 77–86.
[18] I. Perdivara, L. Deterding, A. Moise, K.B. Tomer, M. Przybylski, Determination of primary structure and microheterogeneity of a beta-amyloid plaque-specific antibody using high-performance LC-tandem mass spectrometry, Anal. Bioanal. Chem. 391 (2008) 325–336.
[19] N. Bandeira, V. Pham, P. Pevzner, D. Arnott, J.R. Lill, Automated de novo protein sequencing of monoclonal antibodies, Nat. Biotechnol. 26 (2008) 1336–1338.
[20] D.R. Livesay, S. Subramaniam, Conserved sequence and structure association motifs in antibody-protein and antibody-hapten complexes, Protein Eng. Des. Sel. 17 (2004) 463–472.
[21] M. Thevis, R.R. Ogorzalek Loo, J.A. Loo, In-gel derivatization of proteins for cysteine-specific cleavages and their analysis by mass spectrometry, J. Proteome Res. 2 (2003) 163–172.
[22] J.V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, Higher-energy C-trap dissociation for peptide modification analysis, Nat. Methods 4 (2007) 709–712.
[23] A. Bertsch, A. Leinenbach, A. Pervukhin, M. Lubeck, R. Hartmer, C. Baessmann, Y.A. Elnakady, R. Muller, S. Bocker, C.G. Huber, O. Kohlbacher, De novo peptide sequencing by tandem MS using complementary CID and electron transfer dissociation, Electrophoresis 30 (2009) 3736–3747.
[24] H. Chi, R.X. Sun, B. Yang, C.Q. Song, L.H. Wang, C. Liu, Y. Fu, Z.F. Yuan, H.P. Wang, S.M. He, M.Q. Dong, pNovo: de novo peptide sequencing and identification using HCD spectra, J. Proteome Res. 9 (2010) 2713–2724.
[25] L. Mo, D. Dutta, Y. Wan, T. Chen, MSNovo: a dynamic programming algorithm for de novo peptide sequencing via tandem mass spectrometry, Anal. Chem. 79 (2007) 4870–4878.