

An antibody-based biomarker discovery method by mass spectrometry sequencing of complementarity determining regions

Lennard J. M. Dekker · Lona Zeneyedpour ·
Eric Brouwer · Martijn M. van Duijn ·
Peter A. E. Sillevs Smitt · Theo M. Luider

Received: 16 June 2010 / Revised: 8 October 2010 / Accepted: 19 October 2010 / Published online: 24 November 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Autoantibodies are increasingly used as biomarkers in the detection of autoimmune disorders and cancer. Disease specific antibodies are generally detected by their binding to specific antigens. As an alternative approach, we propose to identify specific complementarity determining regions (CDR) of IgG that relate to an autoimmune disorder or cancer instead of the specific antigen(s). In this manuscript, we tested the technical feasibility to detect and identify CDRs of specific antibodies by mass spectrometry. We used a commercial pooled IgG preparation as well as purified serum IgG fractions that were spiked with different amounts of a fully human monoclonal antibody (adalimumab). These samples were enzymatically digested and analyzed by nanoLC Orbitrap mass spectrometry. In these samples, we were able to identify peptides derived from the CDRs of adalimumab. These peptides could be detected at an amount of 110 attomole, 5 orders of magnitude lower than the total IgG concentration in these samples. Using higher energy collision induced dissociation (HCD) fragmentation and subsequent *de novo* sequencing, we could successfully identify 50% of the detectable CDR peptides of adalimumab. In addition, we demonstrated that an affinity purification with anti-dinitrophenol (DNP) monoclonal antibody

enhanced anti-DNP derived CDR detection in a serum IgG background. In conclusion, specific CDR peptides could be detected and sequenced at relatively low levels (attomole-femtomole range) which should allow the detection of clinically relevant CDR peptides in patient samples.

Keywords Mass spectrometry · Antibodies · CDRs · *de novo* sequencing

Introduction

Antibodies play an important role in many disorders, including autoimmune diseases and cancer [1, 2]. Proteins present in affected tissue of autoimmune diseases and tumors can differ from proteins in normal tissue in quantity, amino acid sequence, structure or post-translational modification. These changes can result in the production of autoantibodies against disease-related proteins [3]. The presence of specific autoantibodies in patients with autoimmune diseases and cancer can be of interest for diagnosis, prognosis [4, 5], drug targets [6] and for fundamental understanding of various disease processes.

In immune responses, activation of both T cells and B cells directed against the (auto)antigen can occur. The B cell activation (humoral response) will result in production of antibodies against the (auto)antigen. In autoimmune diseases, antibodies raised against (auto)antigens are present in the blood at relatively high concentrations of μg s per ml [7].

The genetic building blocks for antibodies are relatively constant while the enormous variation is created by recombination and somatic hypermutation that occur especially in the complementarity determining regions

Electronic supplementary material The online version of this article (doi:10.1007/s00216-010-4361-9) contains supplementary material, which is available to authorized users.

L. J. M. Dekker · L. Zeneyedpour · E. Brouwer ·
M. M. van Duijn · P. A. E. Sillevs Smitt · T. M. Luider (✉)
Department of Neurology, Lab. Neuro-Oncology,
Erasmus University Medical Center,
Dr. Molewaterplein 50,
3015 GE Rotterdam, The Netherlands
e-mail: t.luiders@erasmusmc.nl

(CDRs) of antibodies. The somatic hypermutations in antibody sequences occur during antibody maturation and results in the selection of B-cell clones that produce the antibodies with the highest affinity for the antigen. This selection, in theory, may lead to a convergent development of antigen binding regions, and antibodies with similar and identical CDRs in different individuals despite the enormous potential variation in antibodies [8–10]. In a recent publication, the complete antibody repertoires of 14 zebra fish were sequenced and compared at the mRNA level. The antibody repertoires of these zebra fish showed significantly more overlap than would be expected based on chance alone, suggesting that the production of antibodies is not a random process [11]. Our own research group investigated antibodies from rats immunized with purified antigens. Again, many similarities were found between samples from a treatment group immunized with a particular antigen. Based on these findings, we hypothesize that recurring antibody CDRs are also present in human diseases such as autoimmune diseases or cancer. The specific CDRs of these antibodies may thus be used as biomarkers.

Conventional antigen based biomarker discovery techniques have had only limited success in body fluids. The huge dynamic range of protein concentrations, rapid turnover and the fact that many proteins are not excreted into the circulation probably contribute to this. The technique described here approaches the problem from another angle, in which instead of antigens, peptides from the antibody are used as potential markers. The use of antibodies has a number of advantages; antibodies are excreted, present at high concentrations and purification or enrichment of specific antibodies can be performed with a number of well-established techniques [12, 13]. We have already shown that the analysis of enzymatic digests of purified IgG fractions with advanced mass spectrometry techniques can result in reproducible profiles of the immunome [12].

In the current study, we will focus on the analytical parameters related to the detection, quantification and identification of CDRs peptides. We tested two different approaches a non-targeted and an antigen targeted approach. For these experiments, the fully human monoclonal antibody adalimumab and a murine monoclonal antibody against dinitrophenol (DNP) were spiked into serum samples and a commercial pooled IgG sample. Adalimumab is an antibody which is used in the treatment of rheumatoid arthritis. Adalimumab binds to tumor necrosis factor alpha (TNF α), preventing it from activating TNF receptors. Adalimumab can be obtained in relatively large quantities and since the sequence of the antibody is known, it is well-suited to optimize CDR identification by mass spectrometry in a serum IgG background. In addition, we determined whether an affinity-purification (targeted ap-

proach) of a monoclonal antibody could be used to improve the sensitivity of detection of CDR peptides by mass spectrometry compared to CDRs present in a crude serum IgG background. Through such an approach, it may be possible to improve the detection of those CDRs that are responsible for affinity to particular disease related antigens. To test the latter approach, we used a mouse monoclonal antibody against the hapten dinitrophenol (DNP). We selected a mouse monoclonal antibody in a human antibody background to estimate non-specific binding by means of amino acid sequences of the constant regions of mouse IgGs, which differ from corresponding human sequences.

By conventional database searches, it is difficult to comprehensively identify CDR peptides because the number of IgG sequences in publicly available protein databases is relatively low. Identification of sequences that are not yet present in a database requires high quality MS/MS data for *de novo* sequencing. The quality of MS/MS data usually depends on the amount and purity of the samples. Since the amount of a serum sample is not a limiting factor (ml of serum are routinely taken for diagnostic purposes), samples can be extensively purified and enriched for IgG. Interpretation of the MS/MS spectra can be realized by specialized sequencing programs (both database-dependent sequencing and *de novo* sequencing as well as combinations thereof) [14–16]. We tested the applicability of *de novo* sequencing for a number of CDR peptides and used the higher energy collision induced dissociation (HCD) fragmentation of the Orbitrap mass spectrometer to obtain more accurate fragmentation spectra for the *de novo* sequencing.

Method

Materials and reagents

Water, acetonitrile, formic acid and trifluoroacetic acid were obtained from Biosolve, (Valkenswaard, The Netherlands). Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich, (Saint Louis, MO).

Samples

In a commercial pooled IgG preparation (Nanogam, Sanquin, Amsterdam, The Netherlands), we spiked five different amounts of adalimumab (Humira, Abbott, Chicago, IL): 0.0001, 0.001, 0.01, 0.1 and 1% (w/w adalimumab / total IgG), corresponding to a final concentration of 0.001, 0.01, 0.1, 1 and 10 μ g/ml of adalimumab, respectively. We also included a negative control in which no adalimumab was spiked. The sequence of adalimumab

has been described in a patent (United States patent US6090382 clone D2E7).

From five healthy volunteers, serum samples were obtained via the Sanquin Blood Bank Rotterdam, the Netherlands. The donors have given written consent for the use of serum for research purposes according to general guidelines of the Sanquin Blood Bank, Rotterdam. Immediately after collection the serum samples were centrifuged, aliquoted and stored at $-80\text{ }^{\circ}\text{C}$. Three aliquots of each sample were thawed, and two aliquots were spiked with 0.1 or 0.001% (w/w) of adalimumab / total IgG. For the calculation of spiked amounts of adalimumab, a total IgG concentration in serum of 13.5 g/l is taken as a standard value [7].

IgG purification

IgG was purified from serum samples by Melon Gel purification (Melon Gel IgG Purification Kit, Pierce, Rockford, IL) according to the manufacturer's protocol. Briefly, serum samples were diluted 1:10 and the diluted serum was added to a spin column containing the Melon Gel resin. After 30 minutes incubation, the purified IgG was collected in the flow through by centrifugation of the spin column. All IgG purified samples were aliquoted and stored at $-20\text{ }^{\circ}\text{C}$.

In-solution digestion

To 10 μL of each sample, 10 μL 0.2% Rapigest (Waters Corporation, Milford, MA), in 50 mM NH_4HCO_3 was added. The solution was reduced with 10 mM dithiothreitol (DTT) at $60\text{ }^{\circ}\text{C}$ for 30 min. After the mixture was cooled down to room temperature, it was alkylated in the dark with 50 mM iodoacetamide at ambient temperature for 30 min, and digested overnight with 1:100 (w/w) trypsin (Promega, Madison, WI). Five percent trifluoroacetic acid was added, to obtain a final concentration of 0.5% trifluoroacetic acid (pH<2). After 45 min of incubation at $37\text{ }^{\circ}\text{C}$ the samples were centrifuged at 13,000 g for 10 minutes.

Protein separation SDS-PAGE

Using standard methods supplied by the manufacturer, one-dimensional SDS-PAGE was performed. Ten percent polyacrylamide gels of 0.75 mm thickness were used. To each sample 3 μL loading buffer was added and samples were heated for 10 minutes at $90\text{ }^{\circ}\text{C}$. The gel was stained with Novex[®] Colloidal Blue Staining (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Subsequently, gels were scanned and densitometry was performed on the scan image using the ImageJ software package (<http://rsb.info.nih.gov/ij/>)

In-Gel digestion

The gel bands were excised and cut into plugs. Plugs were transferred into 1.5 mL Eppendorf tubes. The digestion was performed in Rapigest detergent solution according to the manufacturer's instructions. Briefly, the cut gel plugs were washed with water (10 min) and subsequently washed with a mixture of water/acetonitrile (1 : 1, v/v; for 10 min). A sufficient volume of water/acetonitrile was added to shrink the gel plugs. The solution was removed and the gel plugs were rehydrated with 100 mM NH_4HCO_3 for 5 min. An equal volume of acetonitrile was added to the 100 mM NH_4HCO_3 solution, and it was incubated for 15 min. The washing solution was then removed and the gel plugs were dried in a Speedvac (SPD 1010, Thermo Savant, Holbrook, NY). The proteins were reduced with 10 mM DTT in 100 mM NH_4HCO_3 (45 min at $56\text{ }^{\circ}\text{C}$). After removing the reducing solution, proteins were alkylated in the dark with 50 mM iodoacetamide in 100 mM NH_4HCO_3 (30 min). The alkylation was stopped by removing the reaction solution and by washing the gel plugs with a mixture of water/acetonitrile (1 : 1, v/v), acetonitrile and acetonitrile/100 mM NH_4HCO_3 . After the gel plug were dried in a Speedvac, 0.1% Rapigest in 50 mM NH_4HCO_3 was added and incubated at $37\text{ }^{\circ}\text{C}$ for 10 minutes. Subsequently, the gel plugs were dried in a vacuum centrifuge. The gel plugs were rehydrated at $4\text{ }^{\circ}\text{C}$ (45 min) with 10 ng trypsin/ μL 50 mM NH_4HCO_3 . A sufficient volume of 50 mM NH_4HCO_3 was added to completely cover the gel plugs. The digestion was carried out overnight at $37\text{ }^{\circ}\text{C}$. Before starting the extraction, the supernatant was transferred to an Eppendorf tube. After adding to each sample 50 μL 1% trifluoroacetic acid it was incubated at $65\text{ }^{\circ}\text{C}$ for 2 minutes. The supernatant was transferred into a new Eppendorf tube. To the gel plugs, 50 μL 50% acetonitrile and 50 μL 0.1% trifluoroacetic acid was added, the plugs were sonified for 2 minutes and subsequently incubated for 15 minutes at room temperature. The three resulting supernatants were pooled and dried in a vacuum centrifuge and redissolved in 50 μL 0.1% trifluoroacetic acid.

Affinity purification

Dot-blot

Amounts of 0.1 μg DNP BSA (US biological, Swampscott MA) were spotted onto a methanol-wetted PVDF membrane ($0.5\times 20\text{ mm}^2$). After blocking with BSA, the dot-blots were washed in Tris-buffered saline (TBS) (200 mM Tris, 1.5 M NaCl, 0.05% Tween-20 (v/v)). In the next step, IgG containing samples were applied on top of the DNP-BSA spots. The samples analyzed included 0.001, 0.01 and 0.1 μg of anti-DNP antibody (mouse monoclonal anti-DNP,

Acris antibodies GmbH, Herford, Germany), purified human IgG samples (1 μg) spiked with two different amounts of anti-DNP monoclonal 0.01, 0.1 μg and a negative control containing only the purified human IgG. The dot-blot was washed and incubated with alkaline phosphatase labeled anti-mouse IgG (1:500) (Polyclonal Rabbit anti-mouse IgG/AP, Dako, Glostrup, Denmark). Subsequently, the dot-blot was washed again and binding of antibodies was visualized by alkaline phosphatase conjugate substrate kit (Biorad, Hercules, Canada).

Magnetic bead purification of specific antibodies for MS analyses

DNP-BSA was coupled to magnetic beads (008Mag Plex™ C, Luminex Corporation, Austin, TX). After washing the beads with water, 160 μl 100 mM sodium phosphate buffer pH 6.2, 20 μl 50 mg/ml *N*-hydroxysulfosuccinimide, (Pierce, Rockford, IL), and 20 μl 50 mg/ml carbodiimide EDC (Pierce, Rockford, IL) (the latter two solutions both dissolved in water) were added. Subsequently, the beads were incubated at room temperature for 20 minutes. To the activated beads 930 μl 2-(*N*-morpholino) ethanesulfonic acid (MES) pH 5.0 and 70 μl DNP-BSA (2 mg/ml in water) was added, the beads were incubated for 2 hours at room temperature. Subsequently, the coupled beads were blocked with PBS-TBN (PBS, 0.1% BSA, 0.02% Tween-20, 0.05% NaN_3 , pH 7.4) for 30 minutes at room temperature. In the next step, beads were washed twice with PBS-TBN. Finally, the beads were resuspended in PBS-BN (PBS, 1% BSA, 0.05% NaN_3 , pH 7.4) and stored at 4 °C until use.

Amounts of 0.5 μg , 1 μg and 2 μg of anti-DNP were added to both 20 μl purified serum IgG and to 20 μl of PBS. Purified serum IgG without any spiked anti-DNP was used as a negative control. These mixtures were added to 6.10^6 beads coupled with DNP-BSA and beads that were not coupled to a protein. Subsequently, the beads were washed three times with PBS-T (PBS, 0.05% Tween-20). Before elution of the IgG from the beads, a fraction of each sample 6.10^5 beads was conjugated with PE labeled anti-mouse IgG (H+L) (Goat anti-mouse IgG (H+L), Caltag lab, Buckingham, UK), as secondary antibody. Subsequently, these fractions were measured and analyzed by the Luminex® 100™ (Luminex cooperation, Austin, TX). The remaining beads of each sample were incubated with 18 μl of 100 mM Glycine-HCl, pH 2.9. Subsequently, 2 μl of 1 M Tris-HCl, pH 9.0 was added to each sample to obtain a final pH of 8. All samples were digested according to the protocol described above (see section in-solution digest). In addition to the eluted samples, the material loaded on the beads was digested as well. Subsequently, all samples were measured by nanoLC LTQ-Orbitrap MS. In Fig. 1, a

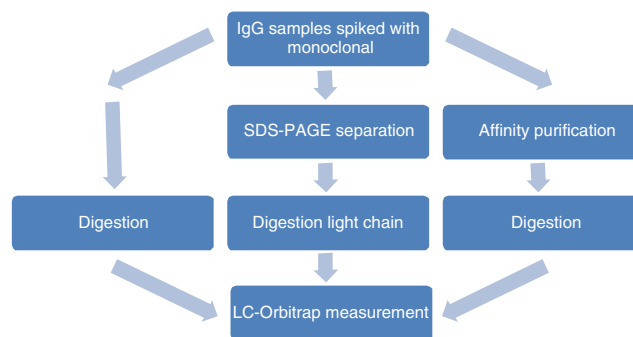


Fig. 1 Flow chart of the used sample preparation approaches

flowchart shows an overview of the different sample preparation methods that were applied.

Mass spectrometry

The completion of the digestion step was checked with reverse phase liquid chromatography and UV detection. For nanoLC LTQ-Orbitrap mass spectrometry measurements, five microliters (or higher volumes depending on the intensity of the peptide signals in UV) of peptide sample was injected on to a nanoliquid chromatography system (nanoLC Ultimate 3000; Dionex, Sunnyvale, CA). After preconcentration and washing of the sample on a C18 trap column (1 mm \times 300 μm i.d.), peptides were separated on a C18 PepMap column (250 mm \times 75 μm internal diameter) (Dionex, Amsterdam) using a linear 90 min gradient (4–40% acetonitrile/ H_2O ; 0.1% formic acid) at a flow rate of 300 nL/min. The separation of the peptides was monitored by a UV detector (absorption at 214 nm). The nanoLC was coupled to a nanospray source of a linear ion trap Orbitrap (LTQ-Orbitrap) mass spectrometer (LTQ Orbitrap XL, Thermo Electron, Bremen, Germany). All samples were measured in a data dependent acquisition mode. Each run was preceded by a blank MS run in order to monitor system background. The peptide masses are measured in a survey scan with a maximum resolution of 30,000 in the Orbitrap and an AGC (automatic gain control) setting of 1.00×10^6 . To obtain a maximum mass accuracy, a prescan is used to keep the ion population in the Orbitrap for each scan approximately the same. During the high-resolution scan in the Orbitrap, the 5 most intense monoisotopic peaks in the spectra were fragmented and measured in the linear ion trap with a normalized collision energy of 35% and an AGC setting of 1.00×10^4 . The fragment ion masses are measured in the linear ion trap to have a maximum sensitivity and the maximum amount of MS/MS data. For *de novo* sequencing, we also measured samples with HCD fragmentation. By this method the 3 most intense monoisotopic peaks in the spectra were fragmented with both CID and HCD. The HCD spectra were measured in the Orbitrap at a maximum

resolution of 7500 with a normalized fragmentation energy of 40% and an AGC setting of $2.00e^5$.

Data analyses

MS/MS spectra were extracted from raw data files and converted into mgf files using Extract-MSN (part of XCalibur version 2.0.7, Thermo Fisher Scientific Inc.). Mascot (version 2.2.06; Matrix Science Inc., London, UK) was used to perform database searches against the human subset NCBI database (version NCBIHuman_database_20090311; Homo sapiens species restriction; 222,071 sequences, to this database the sequence of adalimumab was added) of the extracted MS/MS data. For the database search the following settings were used: a maximum of two miss cleavages, oxidation as a variable modification of methionine, carbamidomethylation as a fixed modification of cysteine and trypsin was set as enzyme. A peptide mass tolerance of 10 ppm and a fragment mass tolerance of 0.5 Da were allowed. An ion score of 40 was used as a cut-off value.

In addition, the raw data files were also loaded into the software package Progenesis LCMS (Version 2.5; Non-linear Dynamics Ltd, New Castle, UK). The data files were aligned and feature selection was performed. Subsequently, the results of the database search were loaded into Progenesis. Finally, an export file was created in which for each feature the abundances in the different samples are displayed.

The IMGT website (<http://imgt.cines.fr/3Dstructure-DB/cgi/DomainGapAlign.cgi>) is used to perform an alignment of the adalimumab sequence to the closest germline sequence.

de novo sequencing was performed using Peaks studio 5.1 (Bioinformatics Solutions Inc., Waterloo, Canada). Orbitrap raw data files were loaded into Peaks and *de novo* sequencing was performed. The parent peptide mass accuracy was set at 10 ppm and for MS/MS fragments obtained by CID 0.5 Da tolerance was allowed and 0.01 Da for HCD data. For each spectrum maximal five *de novo* sequences were proposed, the resulting sequences were exported into a CSV file and the results were compared to database dependent results.

Results

Characterization of adalimumab

The amino acid sequence of adalimumab was largely confirmed by nanoLC Orbitrap mass spectrometry analyses using a shotgun approach. The sequence coverage was 91% for the light chain and 83% for the heavy chain. The

obtained amino acid sequence of adalimumab was added to the NCBI database and the mass spectrometry data was searched against this database using the Mascot search engine. For this specific antibody the number of arginine and lysine residues in and around the CDRs of the heavy chain was relatively low, resulting in relatively large tryptic fragments. Because these larger fragments could on average be detected at a lower sensitivity, the CDR peptides derived from the heavy chain had less intense MS signals than CDR peptides from the light chain. For this reason, we focused our analyses and measurements on the light chain of adalimumab. All peptides identified in the light chain CDR region are displayed in Table 1.

Spiking of adalimumab in IgG

We spiked a concentration range 0.0001%-1% w/w (11 attomole-110 femtomole absolute amounts analyzed) of adalimumab in a commercial IgG pool (Nanogam) to determine the lowest detectable concentration of the CDR peptides in a complex IgG background. These spiked samples were tryptically digested and analyzed by nanoLC Orbitrap mass spectrometry. A combination of IgG peptides derived from the constant regions was used for normalization, although the influence of this normalization was minimal. Four of the 11 CDR light chain peptides that were identified during sequencing of adalimumab were also detected in the spiked samples (see Table 2). The lowest spiked concentrations at which specific CDR peptides could be observed differed between the CDR peptides (Table 2). The peptide LLIYAASLTQSGVPSR of the CDR 2 in the light chain was not specific and could also be observed in the IgG pool without any spiked adalimumab. This could be expected because this CDR did not contain any mutation compared to IgG germline sequences (<http://imgt.cines.fr/3Dstructure-DB/cgi/DomainGapAlign.cgi>). The results of the alignment of the CDR peptides of the light chain of adalimumab are displayed in Figure S1 of the Electronic Supplementary Material. The linear dynamic range of detection of spiked adalimumab is at least three orders of magnitude for (R)NYLAWYQQKPGK(A) (mass 748.38822^{2+}) (Fig. 2). This experiment is performed in duplicate which resulted in a similar linear dynamic range as shown in the figure. The technical reproducibility of the measurements is determined by running triplicate measurements of the concentrations above the detection limit. This resulted in an average CV of 18 % ranging from 10 – 22 % for the normalized intensities of the peptide displayed in Fig. 1. If the CV is calculated for all the measured peptides an average CV is obtained of 27 %. More than 80 % of the masses have a CV below 20 % which is comparable to the results we obtained in a previous study in which we investigated the reproducibility of Orbitrap measurements

Table 1 Identified CDR peptide sequences of the light chain of adalimumab

	MH ⁺	Start	End	Missed cleavage	Sequence	CDR
	1304.7103	19	30	1	(R)VTITCRASQGI R (N)	1
	2108.1035	25	42	1	(R)ASQGI R NYLAWYQQKPGK(A)	1
	1791.954	31	45	1	(R) NYLAWYQQKPGKAPK(L)	1
	1495.7692	31	42	0	(R) NYLAWYQQKPGK(A)	1
	1972.1225	43	61	1	(K)APKLLIYAAS T LQSGVPSR(F)	2
	1675.9377	46	61	0	(K)LLIYAAS T LQSGVPSR(F)	2
	3563.6383	62	93	1	(R)FSGSGSGTDFLT I SSLPED VATYYCQ R Y N R(A)	3
	3130.4309	62	90	0	(R)FSGSGSGTDFLT I SSLPEDVA TYYCQ R (Y)	3
	1502.7386	91	103	1	(R) Y N R A PYTFGQGTK(V)	3
Mutations from the germline sequence are bold and underlined in the sequence	1538.8213	94	107	1	(R) APYTFGQGTK V EIK(R)	3
	1069.5313	94	103	0	(R) APYTFGQGTK(V)	3

of IgG samples in more detail [12]. In Fig. 3, chromatographic traces of the sample spiked with 0.01% Adalimumab are shown. In these chromatograms the complexity of the sample and the intensity of the background in relation to the peptides of interest can be observed. At this ratio of 0.01% (1.1 fmol adalimumab), MS/MS CID spectra can be assigned to the correct CDR sequence by means of a database dependent search.

Based on the above measurements in an artificial pool of antibodies (Nanogam), two spiked concentrations (0.1% or 11 fmole yielded four detectable peptide peaks and 0.001% or 110 attomole yielded one peptide peak, respectively) were selected for studying the feasibility of sequencing CDRs in IgG fractions from serum. Spiked percentages were calculated based on an average IgG concentration in serum of 13.5 mg/ml. The spiked samples were IgG purified, tryptically digested and analyzed by nanoLC Orbitrap mass spectrometry. During the data analyses we focused on the three specific light chain's CDR peptides of adalimumab. In all samples the peptides could be detected at a spiked ratio of 0.1%. At the ratio of 0.001% none of the peptides could be detected. In addition, we separated the spiked samples by SDS-PAGE (Fig. 4). Tryptic peptides were extracted from the excised gel bands of the light chain and measured by nanoLC-Orbitrap MS. In these samples,

all the CDR peptides displayed in Table 2 could be detected at a 0.1% concentration level. The peptide (R)NYLAWYQQKPGK(A) at mass 748.38822²⁺ could be detected in three out of the five samples at the 0.001% spiked concentration. In addition, one sample was IgG purified by Melon gel, analyzed on SDS-PAGE, tryptic digested and subsequently measured by nanoLC-Orbitrap MS in triplicate to determine the variation of the complete sample preparation procedure. For these three measurements an average CV of 29 % was obtained on the normalized intensities of all MS signals.

In addition, we analyzed gels by densitometry to determine the yield of IgG and its purity obtained by the Melon Gel fractionation. On average the Melon Gel purified IgG fractions contained 10.8 µg/µl of IgG as well as 30-40% (w/w) of impurities. The commercial IgG pool (Nanogam) contained 5-10% of impurities.

de novo sequencing of CDR peptides

An in-solution tryptic digest of 10 and 100 fmol/µl adalimumab in buffer (equivalent to 1% and 0.1% spiked concentration) was measured by nanoLC Orbitrap mass spectrometry. In addition to the standard CID fragmentation and detection in the LTQ, the peptides were also frag-

Table 2 Detected CDR peptide sequences of an in-solution digest of adalimumab in a background of a commercial IgG pool (Nanogam)

MH ⁺	Start	End	Missed cleavage	Sequence	CDR	ratio total IgG/adalimumab	absolute spiked amount (fmol)
2108.1035	25	42	1	(R)ASQGI R NYLAWYQQKPGK(A)	1	0.001	11
1495.7692	31	42	0	(R)NYLAWYQQKPGK(A)	1	0.00001	0.11
1675.9377	46	61	0	(K)LLIYAAS T LQSGVPSR(F)	2	N.A.*	N.A.*
1069.5313	94	103	0	(R)APYTFGQGTK(V)	3	0.0001	1.1

*Sequence is not specific for the adalimumab antibody and is also present in commercial IgG pool. N.A. = not applicable

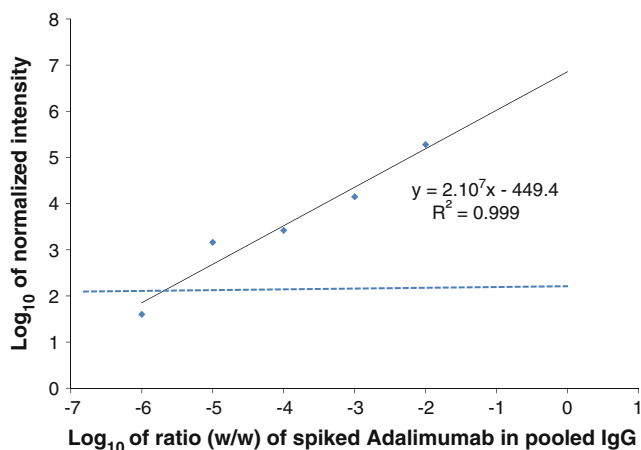


Fig. 2 Dilution series of adalimumab in a pooled IgG sample. Log₁₀ of the intensities for the CDR1 peptide (R)NYLAWYQQKPGK(A) of the light chain of adalimumab are plotted against the Log₁₀ of the spiked ratios of adalimumab in pooled IgG. The dashed horizontal line indicates the background intensity. This intensity corresponds to the signal that is observed at the position of the peptide in non-spiked samples. This plot shows that the correlation between the peptide intensity and concentration of spiked adalimumab is linear over at least three orders of magnitude

mented in the HCD cell and subsequently measured at high resolution by Orbitrap MS. For adalimumab, we focused our analyses on three peptides of the CDRs of the light chain. All three peptides at both concentrations could be identified using Mascot database search. The mass spectra of the three specific peptides of the CDRs of the light chain were *de novo* sequenced using the computer program Peaks. For the three peptides, two could be correctly identified based on both fragmentation spectra. As an example two fragmentation spectra of one peptide are displayed (CID and HCD) in Fig. 5. The confidence score for the HCD results were higher, and the difference in scoring between the correct proposal and other incorrect proposals was larger compared to the CID approach. The quality of the HCD fragmentation spectra performed on a 10 fmol/μl concentration was not sufficient to obtain a correct *de novo* sequence; at a 100 fmol/μl concentration a complete *de novo* sequence could be obtained. In contrast, the CID spectra for both concentrations were comparable in quality and did not result in different *de novo* results.

Affinity-purification

We used a dot-blot to test the binding between mouse anti-DNP antibodies and DNP-BSA, visualized by a secondary alkaline phosphatase-labeled antibody. Dot-blot were incubated with 0.1 and 0.01 μg of anti-DNP antibody spiked into human IgG as well as human IgG alone as a negative control. In addition, we used three amounts of anti-DNP in buffer (0.1, 0.01 and 0.001 μg). Densitometry

of the scanned image of the dot-blot showed that the binding of anti-DNP is related to concentration (Fig. 6). For the spiked samples, the observed color intensity is comparable to the spots incubated with anti-DNP alone. Negative controls using anti-DNP show a negligible background, comparable to that of a negative control containing BSA.

For preparative extraction of anti-DNP antibodies from serum IgG and subsequent analyses by mass spectrometry we used magnetic beads coupled with BSA-DNP. Three amounts of anti-DNP (0.5, 1 and 2 μg) were used, both in buffer and spiked in IgG purified from human serum (20 μg). We used the most abundant specific peptides of the mouse monoclonal anti-DNP to monitor the purification procedure. As an example, the intensities of one peptide of the constant region (VNSAAFPAPIEK) is plotted against the amount of anti-DNP used (see Fig. 7). The intensity of this peptide increases linearly with the amount of used anti-DNP antibodies. This linear correlation is also observed when the peptide is spiked into human IgG (see Fig. 7). Unspecific binding is negligible on uncoated beads and beads coated with BSA-DNP incubated in human IgG. The observed intensities of peptides of human IgG and BSA are at least 100 times lower than specific signals observed for affinity purified anti-DNP antibody at the lowest concentrations measured. The intensities of IgG peptides measured on the beads by the fluorescence-based Luminex approach correspond to the mass spectrometry measurements i.e. similar linear curves are observed. In addition, we measured anti-DNP spiked into IgG without subsequent affinity purification. The amount of sample that could be loaded is limited by the loading capacity of the column. In samples with anti-DNP antibody in a background of 20 μg of human IgG, the signal intensity of the mouse IgG decreased an order of magnitude, in part because of the amount of material that could be loaded on the system (see Fig. 7). Therefore, we also analyzed the same amounts of anti-DNP antibody in only 2 μg of human IgG. Using these samples we could inject the same amounts of the spiked anti-DNP samples. For the highest concentration (2 μg of anti-DNP) we observed similar intensities as for the affinity-purified samples. The lower concentrations (0.5 and 1 μg) showed a decrease in intensity (20-30%) in comparison with the affinity-purified samples. The decrease in intensity is probably related to ion suppression of the signal in the mass spectrometer caused by the presence of relative high concentrations of human IgG peptides.

Discussion

We determined the sensitivity for the detection and identification of specific CDR peptides in complex IgG

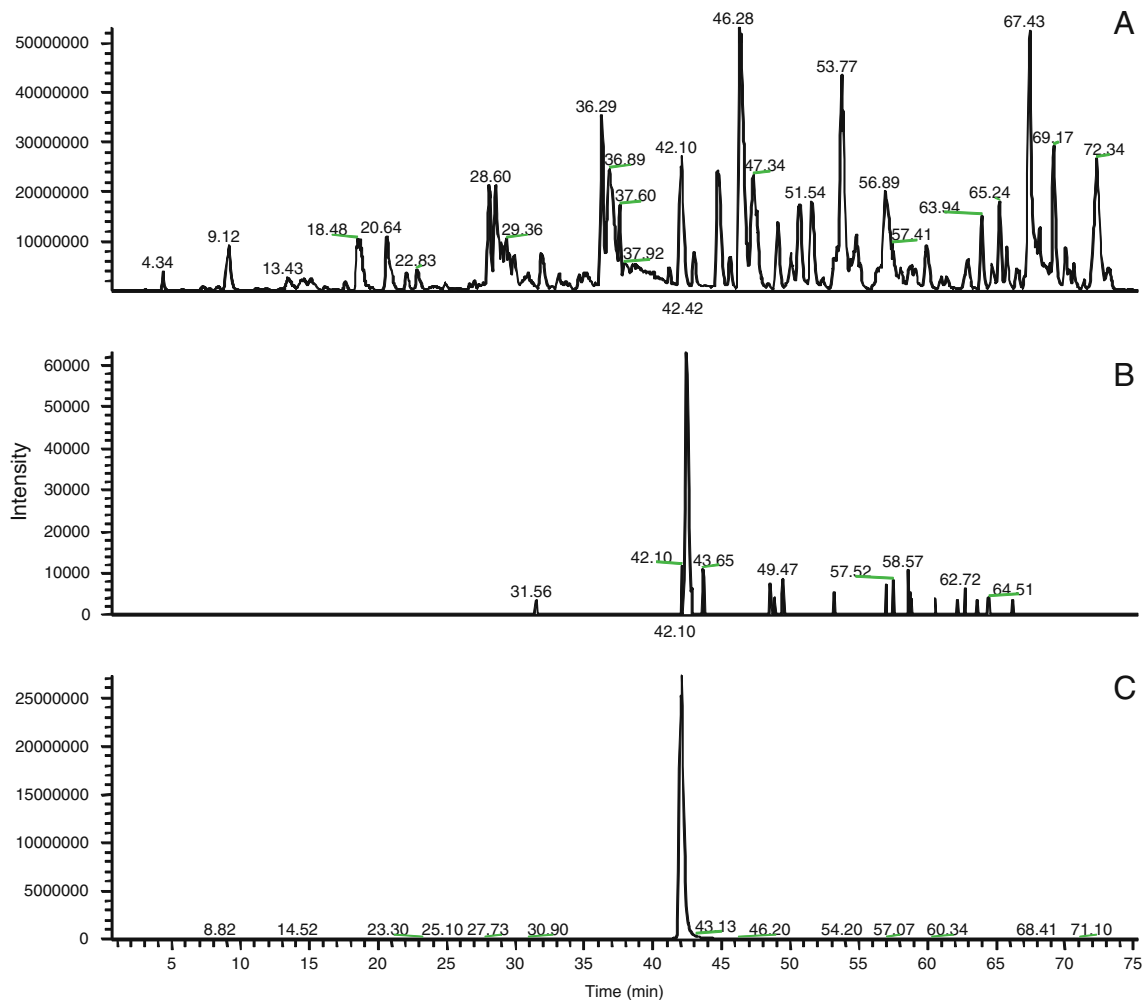


Fig. 3 Chromatographic traces of a nanogam sample spiked with 0.1% Adalimumab. Panel A shows the base peak intensity chromatograph, panel B shows the extracted ion chromatogram of CDR1

peptide (R)NYLAWYQQKPGK(A) and panel C the extracted ion chromatogram of a peptide of the constant region that is eluting at the same retention time as the CDR peptide of interest

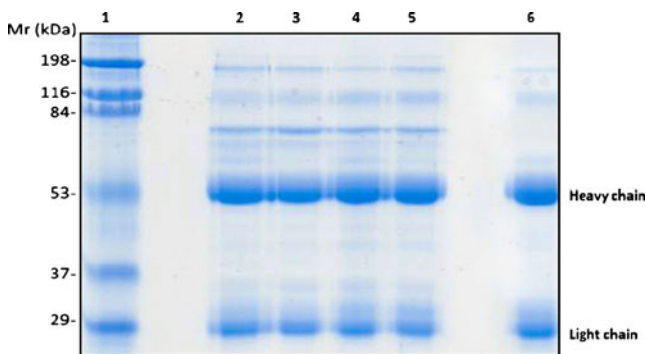


Fig. 4 SDS-PAGE loaded with Melon gel purified IgG from four serum samples (loaded amount is equivalent to 0.5 µl of serum) and 5 µg of the commercial IgG pool (Nanogam). Lane 1 contains a protein marker, lanes 2-5 Melon gel purified IgG and lane 6 contains 5 µg of the commercial IgG pool. Based on gel densitometry, the IgG purified with Melon Gel is more contaminated with serum proteins (50-60%) than the commercial IgG pool (10-15%). The recovery of the IgG was about 100% for the Melon Gel purification

pools as well as in purified serum IgG fractions. We showed that specific CDR peptides of a spiked antibody could be detected at sub-femtomole amounts and in concentrations that are 5 orders of magnitude lower than the total IgG concentration in serum.

The detectable concentrations of CDR peptides of adalimumab when spiked into Nanogam IgG pool were at least a factor 10 lower than after spiking into IgG purified from donor sera by Melon Gel. A possible explanation for this is that the Melon Gel purified samples are more contaminated than the commercial IgG pool. These contaminations could be observed on gel, where in addition to the light and heavy chain other bands are visible (see Fig. 4). One of the major contaminants revealed by nanoLC Orbitrap analyses was serotransferrin. By means of the UV signal of the liquid chromatography test runs we prevented overloading of the column and ensured the best possible separation of the peptides during the Orbitrap measurement.

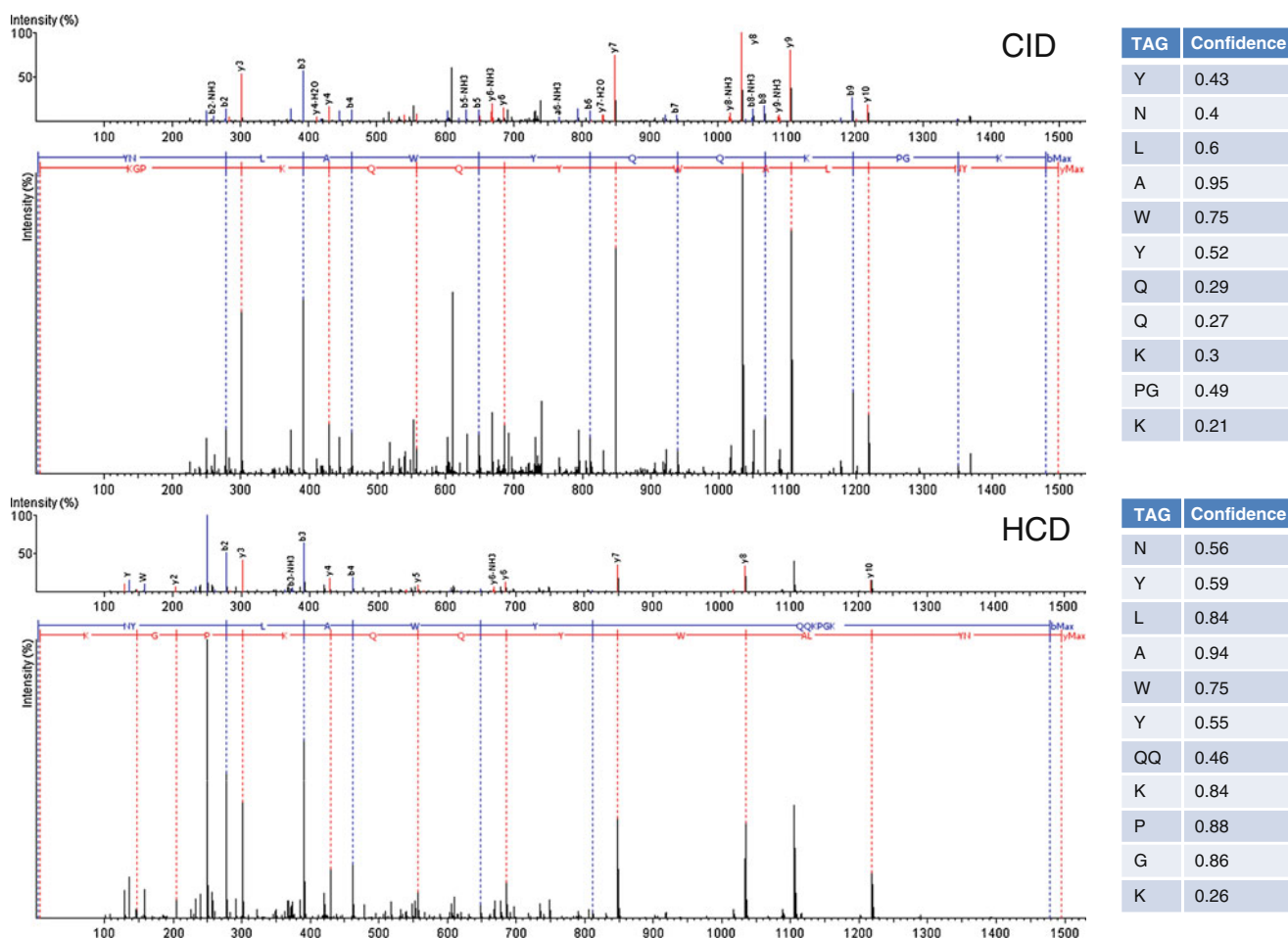


Fig. 5 Annotated MS/MS spectra of mass 748.38822²⁺ (R)NYLA-WYQQKPGK(A) obtained by CID and HCD fragmentation. The signal to noise in the CID spectrum is much higher compared to the HCD spectrum. Because of the higher mass accuracy in the HCD spectra and the possibility of detecting the lower mass fragments the

confidence score for most of the amino acid sequence has increased in HCD. Combined interpretation of both spectra results in *de novo* sequences with a higher confidence compared to HCD spectra or CID spectra alone

Contaminations in the samples result in less IgG being loaded onto the column for the Orbitrap analyses, which will decrease the sensitivity. In addition, the contaminations will also affect the detection and could result in ion-suppression of the MS signals if peptides of contaminants

co-elute with peptides of IgG. These possible factors might explain the improved sensitivity that is observed when the SDS-PAGE step is included for the spiked Melon Gel IgG samples. Using this step, the impurities in the IgG sample were considerably reduced. In addition, the light and heavy

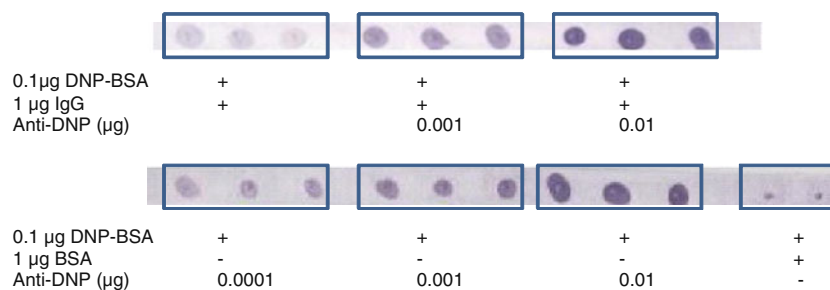


Fig. 6 Dot-blot of DNP-BSA incubated with purified IgG spiked with monoclonal anti-DNP. These dot-blot shows that the intensity of the signal increases with the amount of spotted anti-DNP. The intensities of anti-DNP spiked into IgG are comparable to those of

pure anti-DNP. The background signal of IgG alone is less than the signal observed for 0.001 µg of anti-DNP antibody. For the three replicates per measurement an average CV of 4% for the various concentrations was obtained by densitometry

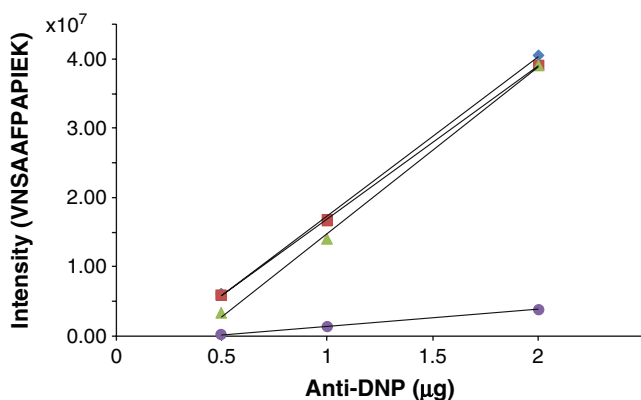


Fig. 7 Intensity of a mouse constant region peptide as a function of the amount of mouse anti-DNP used. The blue diamonds show the intensities of the affinity purified anti-DNP in buffer, the red squares show the intensities of affinity purified anti-DNP spiked in 20 µg of IgG, the green triangles show the intensities of anti-DNP spiked in two µg of IgG without affinity purification, the purple circles show the intensities of anti-DNP spiked in twenty µg of IgG without affinity purification. The intensities of anti-DNP constant region peptide VNSAAFPAPIEK in buffer and spiked in human IgG purified on magnetic beads show similar intensities and both are linearly correlated ($R^2=0.99$). When no magnetic bead purification is applied, the addition of 2 µg human IgG results in a suppression of the peptide signal at lower concentrations of anti-DNP. Spiking anti-DNP into 20 µg human IgG results in a much more pronounced decrease in intensity

chain were separated using this process, which further decreases the complexity of the sample and therefore increases the expected number of identifications.

The success of identifying a peptide based on *de novo* sequencing depends on spectra quality and mass accuracy. A combination of a CID spectrum in the LTQ of the Orbitrap and HCD spectrum measured in the Orbitrap gives complementary information which benefited peptide identification as can be seen in Fig. 4. We have shown that even at low femtomol levels, database interpretable CID spectra can still be obtained. HCD is a more energetic fragmentation which results in additional fragments. Using this method, the low mass region can be measured as well because it does not suffer from the one third mass cut-off as is the case in iontrap CID fragmentation. In addition, the mass accuracy in HCD is better for the fragments generated than in CID which helps to confirm the predicted sequence considerably. For the peptides of interest, we typically observed that, on average, a concentration of 100 femtomol of peptide is required to obtain *de novo* interpretable HCD fragmentation spectra. When acquiring HCD spectra, standard CID spectra can be obtained without additional analysis time because it is performed in parallel to HCD fragmentation and spectra acquisition. This means that for low intensity signals with a poor HCD spectrum, the more sensitive CID is still available. For two out of three CDR peptides of the light chain of adalimumab the top scoring

proposed sequence by the *de novo* sequencing software was completely correct. This is in agreement with data from a larger test set in which a comparison of a database search and *de novo* sequencing results is performed. In this test, the same sequences were obtained for the database search and the *de novo* sequencing in 50-70% of the cases depending on the settings used (data not shown).

This percentage can be further increased when in addition to this method ETD fragmentation is used. Another alternative to improve the coverage of CDRs is to use different enzymes [17–20] and to obtain peptides that are cut at different positions which can potentially fragment or ionize better or complement the sequencing information obtained of tryptic peptides.

We have shown that it is possible to detect specific CDR peptides at attomole levels at spiked concentrations which were 5 orders of magnitude lower than the concentration of the IgG serum background. This method can be further improved by including additional and more efficient separation steps. We demonstrated in this study that an affinity-purification is a useful step to enrich antibodies specific for an antigen. This purification results in an increase of sensitivity of at least one order of magnitude for the mass spectrometry analyses comparing enriched to not enriched samples. An advantage of this approach is that the total column capacity can be used for the peptides of specific IgGs resulting in an increased sensitivity. In addition, the amount of suppression is reduced as can be observed in Fig. 7. Using this purification method, relatively large quantities of specific IgG can be obtained: from only 100 µl of serum several µg's of specific IgG can probably be harvested. We have demonstrated that a µg of affinity-purified IgG is sufficient for the identification of specific CDRs by our approach. This approach shows that specific antigens with unknown identity gives possibilities to identify specific CDRs. It may also be possible to extend this method to the affinity enrichment of antibodies against a tissue or protein mixture containing antigens that have not been fully purified, or antigens in such a mixture that are still unidentified.

Conclusions

We have shown that relatively high sensitivities for the detection of specific CDR peptides can be obtained in serum IgG fractions. If this approach is applicable to detect CDR peptides of disease related antibodies in patients samples has to be further investigated. However, based on our own data [13], we expect that 1-3 % of total IgG can have affinity to an antigen. In addition, a polyclonal antibody response against a protein consists of approximately 100 B-cell clones [21] (these clones can undergo

further mutations during maturation). Therefore a particular antibody would be present at a fraction of at most 0.01 % of total IgG depending on the intensity of the immunological reaction against the antigen and diversification in the B cell clones. The expectation that CDRs of specific antibodies at these concentrations can be detected and identified seems to be realistic based on the data presented and previously published data [12, 13]. This would make it possible to use the described technique for an antibody-based biomarker approach. In addition, we have shown that if the antigen is known and is available in a purified form, an immune affinity capturing can be used to further increase the sensitivity of the method. For many autoimmune diseases the antigen is known and can be used for enrichment of specific autoantibodies to study the antibody production in patients in more detail.

Acknowledgment The authors acknowledge financial support from a NWO (Nederlandse organisatie voor Wetenschappelijke Onderzoek) Horizon Breakthrough grant (9351808).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Tan EM, Zhang J (2008) Autoantibodies to tumor-associated antigens: reporters from the immune system. *Immunol Rev* 222:328–340
2. Tan EM (2001) Autoantibodies as reporters identifying aberrant cellular mechanisms in tumorigenesis. *J Clin Invest* 108 (10):1411–1415
3. Canelle L, Bousquet J, Pionneau C, Deneux L, Imam-Sghiouar N, Caron M, Joubert-Caron R (2005) An efficient proteomics-based approach for the screening of autoantibodies. *J Immunol Meth* 299(1–2):77–89
4. Imafuku Y, Omenn GS, Hanash S (2004) Proteomics approaches to identify tumor antigen directed autoantibodies as cancer biomarkers. *Dis Markers* 20(3):149–153
5. Caron M, Choquet-Kastylevsky G, Joubert-Caron R (2007) Cancer immunomics using autoantibody signatures for biomarker discovery. *Mol Cell Proteomics* 6(7):1115–1122
6. Zhang Z, Pan H, Chen X (2009) Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom Rev* 28 (1):147–176
7. Janeway CA, Travers P, Walport M, Shlomchik MJ (2005) *Immuno biology: The immune system in health and disease*. Garland Science Publishing, New York
8. Tonegawa S (1976) Reiteration frequency of immunoglobulin light chain genes: further evidence for somatic generation of antibody diversity. *Proc Natl Acad Sci USA* 73(1):203–207
9. de Wildt RM, van Venrooij WJ, Winter G, Hoet RM, Tomlinson IM (1999) Somatic insertions and deletions shape the human antibody repertoire. *J Mol Biol* 294(3):701–710
10. Matthyssens G, Hozumi N, Tonegawa S (1976) Somatic generation of antibody diversity. *Ann Immunol* 127(3–4): 439–448
11. Weinstein JA, Jiang N, White RA, 3rd, Fisher DS, Quake SR (2009) High-throughput sequencing of the zebrafish antibody repertoire. *Science (New York, NY)* 324 (5928):807–810
12. de Costa DD, Broodman I, Vanduijn MM, Stingl C, Dekker LJ, Burgers PC, Hoogsteden HC, Sillevs Smitt PA, van Klaveren RJ, Luidert TM (2010) Sequencing and quantifying igg fragments and antigen-binding regions by mass spectrometry. *J Proteome Res* 9 (6):2937–2945
13. Vanduijn MM, Dekker LJ, Zeneyedpour L, Sillevs Smitt PA, Luidert TM (2010) Immune responses are characterized by specific shared immunoglobulin peptides which can be detected by proteomic techniques. *J Biol Chem* 285(38): 29247–29253
14. Spengler B (2004) *de novo* sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* 15(5):703–714
15. Bandeira N, Pham V, Pevzner P, Arnott D, Lill JR (2008) Automated *de novo* protein sequencing of monoclonal antibodies. *Nat Biotechnol* 26(12):1336–1338
16. Pham V, Henzel WJ, Arnott D, Hymowitz S, Sandoval WN, Truong BT, Lowman H, Lill JR (2006) *de novo* proteomic sequencing of a monoclonal antibody raised against ox40 ligand. *Anal Biochem* 352(1):77–86
17. Kang SU, Fuchs K, Sieghart W, Pollak A, Csaszar E, Lubec G (2009) Gel-based mass spectrometric analysis of a strongly hydrophobic gabaa-receptor subunit containing four transmembrane domains. *Nat Protoc* 4(7):1093–1102
18. Swaney DL, Wenger CD, Coon JJ (2010) Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J Proteome Res* 9(3):1323–1329
19. Taouatas N, Drugan MM, Heck AJ, Mohammed S (2008) Straightforward ladder sequencing of peptides using a lys-n metalloendopeptidase. *Nat Meth* 5(5):405–407
20. van Breukelen B, Georgiou A, Drugan MM, Taouatas N, Mohammed S, Heck AJ (2010) Lysndenovo: An algorithm enabling *de novo* sequencing of lys-n generated peptides fragmented by electron transfer dissociation. *Proteomics* 10(6):1196–1201
21. Poulsen TR, Meijer PJ, Jensen A, Nielsen LS, Andersen PS (2007) Kinetic, affinity, and diversity limits of human polyclonal antibody responses against tetanus toxoid. *J Immunol* 179 (6):3841–3850