Development of multiplex mass spectrometric immunoassay for detection and quantification of apolipoproteins C-I, C-II, C-III and their proteoforms

Olgica Trenchevska, Matthew R. Schaab, Randall W. Nelson, Dobrin Nedelkov*

The Biodesign Institute at Arizona State University, Tempe, AZ, 85287

*Corresponding author:

Molecular Biomarkers Laboratory, Biodesign Institute, Arizona State University

P.O. Box 876601

Tempe, AZ 85287-6601

Tel: (480) 727-2280

e-mail:dobrin.nedelkov@asu.edu

Abstract

The impetus for discovery and evaluation of protein biomarkers has been accelerated by recent development of advanced technologies for rapid and broad proteome analyses. Mass spectrometry (MS)-based protein assays hold great potential for *in vitro* biomarker studies. Described here is the development of a multiplex Mass Spectrometric Immunoassay (MSIA) for quantification of apolipoprotein C-I (apoC-I), apolipoprotein C-II (apoC-II), apolipoprotein C-III (apoC-III) and their proteoforms. The multiplex MSIA assay was fast (~40 min) and high-throughput (96 samples at a time). The assay was applied to a small cohort of human plasma samples, revealing the existence of multiple proteoforms for each apolipoprotein C. The quantitative aspect of the assay enabled determination of the concentration for each proteoform individually. Low-abundance proteoforms, such as fucosylated apoC-III, were detected in less than 20% of the samples. The distribution of apoC-III proteoforms varied among samples with similar total apoC-III concentrations. The multiplex analysis of the three apolipoproteins C and their proteoforms using quantitative MSIA represents a significant step forward toward better understanding of their physiological roles in health and disease.

Keywords: apolipoprotein C-I, C-II, C-III, proteoforms, biomarkers, mass spectrometric immunoassay, quantification.

1. Introduction

As a result of alternative splicing, single-nucleotide polymorphisms (SNPs), and posttranslational modifications (PTMs), proteins can exist as multiple variants (or "proteoforms" - [1]) *in vivo*. These proteoforms may play important physiological roles, with possible implications in disease development. Thus, development of methods for detection and quantification of proteoforms is considered an important, albeit challenging task.

Mass spectrometry (MS)-based protein assays hold great potential for *in vitro* detection of protein biomarkers. Intrinsic to the MS-based methods is the measurement of molecular mass – a unique property of each protein. Several MS-based methodologies employ a combination of immuoaffinity enrichment and MS detection. Among them are MSIA [2, 3], SELDI [4], SISCAPA [5, 6], iMALDI [7] and SILAC [8]. Some of these methods are geared towards detection of proteolytic peptides as surrogate measures for protein quantification (e.g., SISCAPA, SILAC), leaving a large part of the protein sequence un-assessed. Hence, those approaches cannot detect proteoforms without *a priori* knowledge of their existence. Top-down MS-based approaches are better suited for proteoforms detection because they detect the mass of the intact proteins, thus covering putative modifications in the entire protein sequence [9, 10].

The mass spectrometric immunoassay (MSIA) is one such top-down proteomics method [2, 3, 11-14]. MSIA combines micro-scale immunoaffinity separation with mass spectrometric detection. Antibodies towards targeted proteins are attached to porous microcolumns fitted at the entrance of a pipette tip, and used for proteins affinity extraction directly from the biological sample. After the affinity capture, proteins are eluted either directly onto a target plate for MALDI-TOF MS analysis [15], or eluted with a small volume of elution solution for subsequent LC-ESI MS analysis [16]. When optimized, MSIA can provide detailed insights into the intrinsic

protein properties, in both qualitative and quantitative manner [17-21]. The capability for multiplexing is another advantage of MSIA, because it enables simultaneous determination of several proteins in a single analysis [22, 23]. In relative quantification mode, the multiplex MSIA approach yields a protein profile consisting of all proteoforms, which can be used to assess the protein function. In fully quantitative mode (using internal reference standards), multiplex MSIA provides information about the concentration of the native (intact) protein and the various proteoforms that may exist alongside.

Apolipoprotein C-I (apoC-I), apolipoprotein C-II (apoC-II) and apolipoprotein C-III (apoC-III) are members of the same apolipoprotein family, and are synthesized in the liver. They are carried primarily in the very-low density lipoprotein (VLDL) fraction [24] although they become associated with other lipoproteins during normal lipid metabolism. They play important roles as co-factors and inhibitors in triglyceride metabolism. For example, apoC-II is required for efficient lipolysis of triglyceride (TG)-rich lipoproteins through lipoprotein lipase activation [25]. In contrast, apoC-III inhibits the lipolysis of TG-rich lipoproteins and is associated with elevated TG-rich lipoproteins and possibly CVD [26]. Interest in apoCs, particularly apoC-III, has grown considerably since recent studies indicate loss-of-function mutations in the apoC-III gene are associated with reduced triglyceride levels and cardiovascular risk [27-29]. In fact, there is intensive new drug development targeting apoC-III levels [30, 31]. Importantly, several proteoforms of apoCs exist [32] which reflect their function *in vivo*. MS-based methods provide a unique opportunity to study apoCs proteoforms [33-37], and correlate their structural modifications with function in physiological and pathological states.

Presented here is the development and application of a multiplex MSIA for detection and quantification of apoC-I, apoC-II, and apoC-III and their proteoforms. The main scope was to

investigate the quantitative distribution and the heterogeneity of the apoCs in human plasma, and to provide insight into the potential role of these proteoforms as biomarkers.

2. Material and methods

2.1. Reagents

Polyclonal goat anti-human antibodies to apoC-I (Cat. No. 31A-G1b), apoC-II (32A-G2b), apoC-III (33A-G2b), and ultra-pure human apolipoprotein C-I (31P-UP201) were obtained from Academy Bio-medical Co. (Houston, TX, USA). Native human apoC-II (Cat. No. MD-26-0012P), native human apoC-III (MD-26-0013P), and mouse anti-hen egg lysozyme (128-10094) were obtained from RayBiotech (Norcross, GA, USA). Protein calibration standard I (Cat. No. 206355) was purchased from Bruker (Billerica, MA). Phosphate buffered saline (PBS) buffer (Cat. No. 28372), MES buffered saline (28390), acetonitrile solution (ACN; A955-4), hydrochloric acid (HCl; A144-212), *N*-methylpyrrolidinone (NMP; BP1172-4, 1,1' Carbonyldiimidazole (97%) (CDI, 530-62-1) and affinity pipettes fitted with porous microcolumns (991CUS01) were obtained from Thermo Scientific (Waltham, MA, USA). Tween20 (Cat. No. P7949), trifluoracetic acid (TFA, 299537), sinapic acid (85429-5G), sodium chloride (S7653), HEPES (H3375), ethanolamine (ETA; 398136), albumin from bovine serum (BSA; A7906) and lysozyme from chicken egg white (L7651) were obtained from Sigma Aldrich (St. Louis, MO, USA). Acetone (Cat. No. 0000017150) was obtained from Avantor Performance Materials (Center Valley, PA, USA).

2.2. Human samples

A small set of 82 human Na_2EDTA plasma samples was used for the method development and validation. These samples were received from the National University of Singapore, with signed informed consent, without any identifiers, and labeled only with a

barcode. The samples were centrifuged for 5 min at 3,000 rpm, and then aliquoted into 96-well microplates (V = 90 µL) and stored at -80°C. Prior to analysis, the samples were thawed and diluted by adding 3 µL of each sample aliquot into 117 µL PBS, 0.1%Tween (40x dilution).

2.3. Antibody immobilization

Activation of the microcolumns in the affinity pipettes was executed using a Multimek 96 channel pipettor (Beckman Coulter, Brea, CA). The affinity pipettes were first rinsed with 200 mM HCl (2 x 20 aspiration/dispense cycles, 100 µL volume), followed by acetone rinse (20 aspiration/dispense cycles, 100 µL). The pipettes were then immersed into a microplate containing 100 mg/mL 1,1' Carbonyldiimidazole (in methyl-1pyrrolidone-2 (NMP)), and 450 cycles of 100 µL aspirations and dispenses through each affinity pipette were performed. Two rinses with NMP (10 cycles each, 150 µL) followed. The affinity pipettes were then immediately immersed into the wells of a 96-microwell microplate which contained the antibody mixture: 0.32 µg anti-apoC-I; 2.25 µg anti-apoC-II; 2.5 µg anti-apoC-II; and 0.40 µg anti-Lys; dissolved into 50 µL of 10 mM MES buffer (these amounts of antibodies were determined empirically to yield comparable signals of the targeted proteins in the mass spectra; they differed because of the unique affinities of each antibody toward the targeted protein) . A total of 750 cycles of aspirations and dispenses of the antibody mixture were performed, followed by rinses with ethanolamine (50 cycles, 100 µL) and twice with HBS-N buffer (100 cycles, 100 µL). The derivatization process took 2 hours, for 96 pipettes. The antibody-derivatized pipettes were stored at +4°C until used (storage was good for > 6 months).

2.4. Sample preparation

For the standard curves, a mixture containing equal mass of the three purified human apolipoproteins was prepared to a final concentration of 10 mg/L (total apolipoproteins). The

protein standards mixture was then serially diluted to 5 mg/L, 2.5 mg/L, 1.25 mg/L, 0.625 mg/L, 0.313 mg/L and 0.156 mg/L, in PBS buffer containing 3 g/L BSA. Analytical samples were prepared by combining 30 µL of the 40x diluted plasma (or the protein standards mixture) with 30 µL of a 0.1 mg/mL solution of lysozyme, and 60 µL PBS with 0.1% Tween, yielding a total of 120 µL analytical sample volume. Standards and analytical samples were run simultaneously on the same multi-channel robotic pipettor.

2.5. Assay execution

The antibody-derivatized affinity pipettes were mounted onto the head of the Multimek 96-channel pipettor. The MSIA started with an assay buffer rinse (PBS, 0.1% Tween, 10 aspirations and dispense cycles, 100 µL each). Next, the affinity pipettes were immersed into a microplate containing the analytical samples and 250 aspirations and dispense cycles were performed (100 µL each) allowing for affinity capture of all the proteins from the samples. The affinity pipettes were then rinsed with assay buffer (100 cycles, 100 µL volumes), and twice with water (10 cycles and 20 cycles, 100 µL) to remove non-specifically bound proteins from the microcolumns. For elution of the captured proteins, 6 µL aliquots of MALDI matrix (15 g/L sinapic acid in aqueous solution containing 33 % (v/v) ACN, and 0.4 % (v/v) TFA) were aspirated into each affinity pipette, and after a 10 second delay (to allow for the dissociation of the proteins from the capturing antibodies), the eluates were dispensed directly onto a 96-well formatted MALDI target. Following drying, linear mass spectra were acquired from each sample spot using Bruker's Ultraflex III MALDI-TOF instrument (Bruker, Billerica, MA). The instrument was operated in positive ion mode, in the mass range from 5 to 30 kDa, with 200 ns delay, 20.00 kV and 18.45 kV ion source voltages, and signal suppression up to 500 Da. Approximately 5,000 laser-shot mass spectra were saved for each standard and sample.

2.6. Data analysis

Mass spectra were internally calibrated using the protein calibration standard I, and further processed with Flex Analysis 3.0 software (Bruker Daltonics). All peaks representing apolipoproteins and their proteoforms, along with the lysozyme IRS peak, were integrated baseline-to-baseline using Zebra 1.0 software (Intrinsic Bioprobes Inc), and the obtained peak area values were tabulated in a spreadsheet. To distinguish between noise and low intensity signals, the peak areas were corrected individually with baseline noise-bin signals (10 Da wide), picked from regions in the mass spectra in close proximity to the integrated protein signals. The corrected apolipoprotein peak areas were then normalized (divided) with the IRS peak area. For the plasma samples, the normalized peak areas for each apolipoprotein and its proteoforms were summed up, and the total apolipoprotein C-I, C-II, and C-III concentrations was determined using the corresponding apo C-I, C-II and C-III standard curves. The concentrations of the individual proteoforms were determined based on their percentage of the total concentrations for each individual apoC protein.

3. Results

3.1. Multiplex MSIA of human apolipoproteins C-I, C-II and C-III and their proteoforms

Multiplex assays have great allure in proteomics [38]. Currently, only a limited number of MS-based methodologies are amenable of multiplex protein analyses, including MSIA [39] and selected peptide-based proteomics approaches [40-42]. In this work we demonstrate a multiplex MSIA for the apoC-I, apoC-II, apoC-III and their proteoforms.

Presented in **Fig. 1** is an example of mass spectra resulting from the multiplex apoCs MSIA. Signals from native apoC-I (MW 6,630 Da), apoC-II (MW 8,915 Da), apoC-III (MW 8,765 Da) and the IRS, lysozyme (MW 14,305 Da), were clearly identified in all samples, along

with several proteoforms. To identify the proteoforms, two internal calibrations were performed – first one using protein calibration standard I (Bruker, Billerica, MA), and second, using the IRS in the spectrum. Masses were calibrated up to 0.001 Da in m/z, in order to achieve accurate mass assignment for the peaks in the mass spectra. Using this approach, and accounting for the selectivity of the apoCs antibodies, we were able to identify several proteoforms, as listed in **Table 1.** The table shows the theoretical and observed mz/values of the proteoforms, their distribution (relative percent abundance, as calculated from the total for each apoC), and the frequency of the proteoforms in the cohort. Most of the proteoforms were detected in all samples, with the exception of the low abundance fucosylated apoC-III proteoforms.

For apoC-I, beside the native full-length form, one additional proteoform was detected and identified as a sequence lacking two *N*-terminal amino acids (des-TP apoC-I; MW 6,432 Da). This proteoform was identified in all samples and is a well-known product of dipeptidyl-peptidase IV (DPP-4) enzymatic cleavage [43, 44].

For apoC-II, a proteoform lacking six *N*-terminal amino acids (des-TQQPQQ apoC-II; MW 8,204 Da) was identified in all samples. This proteoform is termed mature apoC-II [45], and it is of low abundance (accounting for <10 % of the total apoC-II). The full-length apoC-II, also called pro-apoC-II, represents ~ 90% of the total apoC-II.

For apoC-III, we detected a total of 12 proteoforms, including the full-length native apoC-III and numerous glycosylated proteoforms. ApoC-III is highly heterogeneous protein because of the extensive protein glycosylation [34]. The most abundant proteoform is apoC-III containing one *N*-acetyl neuraminic acid (NeuAc; sialic acid) attached to a glycan core of one galactose (Gal) and one *N*-acetylgalactosamine (GalNAc): apoC-III-$(Gal)_1(GalNAc)_1(NeuAc)_1$. This proteoform is commonly referred to as apoC-III_1 (one sialic acid) in the literature, with the

other two most abundant forms being apoC-III-$(Gal)_1(GalNAc)_1$ and apoC-III-$(Gal)_1(GalNAc)_1(NeuAc)_2$, commonly referred to as apoC-III_0 (no sialic acid) and apo$CIII_2$ (two sialic acids), respectively [32].

3.2 Quantitative MSIA analysis of apoC-I, apoC-II, apoC-III and their proteoforms

Quantification of the apoCs in this multiplex MSIA was achieved through the generation of three standard curves (one each for apoC-I, apoC-II and apoC-III), utilizing the corresponding protein standards. Because all three apolipoproteins were analyzed in the same assay, a single, exogenous protein (chicken-egg lysozyme) was added in the analytical samples as an internal reference standard (IRS) for quantification. The choice of IRS was critical step in the method development, because it had to satisfy all the necessary requirements in these "fit-for-purpose" approaches [46]. Chicken-egg lysozyme, a non-human protein, was a good choice for an IRS because it is not present in human plasma (therefore, avoiding cross-reactivity), and its signal in the mass spectrum (at m/z = 14,305 Da) is in close proximity to the signals from the apoCs (thus ensuring similar MS instrument settings use). The IRS was added in constant concentration in all analytical samples (standards and plasma), and produced a constant signal in the mass spectra.

The physiological concentrations of the three apoCs is in the mg/L range, with apoC-III having the highest range (80 – 150 mg/L), and apoC-I and apoC-II exhibiting similar concentration ranges (40 – 60 mg/L and 30 – 50 mg/L, respectively). The three standard curves, therefore were generated with an identical concentration range (0.156 – 5 mg/L) that was sufficient for quantitation of all the proteoforms (the plasma samples were diluted 40-fold). Examples of the standard curves, along with the corresponding mass spectra are presented in **Fig. 2**. The response for all standard curves is linear across the entire range (r^2 = 0.9988, SEE = 0.0928 for apoC-I; r^2 = 0.9940, SEE = 0.0262 for apoC-II; and r^2 = 0.9996, SEE = 0.1605 for

apoC-III). A control sample with known apoC-I, apoC-II and apoC-III concentrations (1 mg/L each) was analyzed in triplicate each day in order to assess the inter-day variability, and presented CVs of 8.82 % (apoC-I), 7.88 % (apoC-II) and 10.1 % (apoC-III).Linearity of the assay was evaluated by analyzing several admixtures of plasma samples with various concentrations of the apoCs, yielding a recovery of 89-124%.

Using these standard curves we determined the concentrations of all apoCs and their proteoforms in the 82 human plasma samples (**Fig. 3**). We first determined the protein/lysozyme peak areas ratios for each proteoform, summed up the ratios of all proteoforms for an individual protein, determined the total individual protein concentration using the corresponding standard curve, and then determined the concentration of each proteoform based on its percentage of the total specific apoC concentration. This approach worked especially well for those proteoforms whose proteoform/lysozyme peak ratios were below those of the standard curve. The total concentrations were in the range of 13.8 to 148 mg/L for apoC-I (mean = 59.2 mg/L), 30.2 to 161 mg/L for apoC-II (77.8 mg/L) and 20.4 to 188 mg/L for apoC-III (76.7 mg/L), which is within the range of normal concentrations as determined by conventional immunoassays. The concentration ranges for the individual proteoforms are shown in **Fig. 3**.

4. Discussion

Delineating protein heterogeneity is an important undertaking because of its implications in physiological and pathological processes. The MSIA methodology can unambiguously distinguish between different proteoforms in a single assay. The approach is time efficient (~40 min) and high-throughput, allowing us to study the range of apoCs proteoforms from a small cohort of 82 human plasma samples in a single run on the automated 96-channel pipettor.

The apoC-I des-TP proteoform, a product of a DPP-4 enzymatic cleavage, was identified in all samples. Analyzing this proteoform is significant for several reasons. First, the des-TP apoC-I proteoform accounts for approximately a quarter of the total apoC-I in normal plasma (**Table 1**), making a significant contribution to the overall apoC-I concentration. Second, this proteoforms could potentially be used as a marker of DPP-4 inhibitor drugs activity. DPP-4 inhibitor drugs, such as vildagliptin, sitagliptin and sexagliptin, are prescribed for the treatment of type 2 diabetes (T2D) for controlling the HbA1C levels and preservation of the progressive loss of insulin-secreting capacity [47, 48]. DPP-4 cleaved apoC-I can therefore be a useful marker for the effect of DPP-4 inhibitor drugs and monitoring of therapy.

ApoC-II is a relatively homogenous protein exhibiting just two proteoforms: pro-apoC-II (the dominant proteoform), and mature apoC-II - a proteoform resulting from the removal of an *N*-terminal hexapeptide (TQQPQQ). ApoC-II has been primarily associated with activation of the lipoprotein lipase (LPL), and promoting triglyceride hydrolysis in lipid particles [24]. Deficiency of apoC-II was shown to cause hypertriglyceridemia, which can be reversed by infusion with normal plasma or synthetic apoC-II in affected individuals [49]. The MSIA approach measures both proteoforms - the pro- and mature apoC-II, and can be used for monitoring and predicting the kinetics of maturation of this apolipoprotein *in vivo*.

ApoC-III heterogeneity is a result of its glycosylation and presence of *O*-linked glycan motifs w/o and w/ one and/or two sialic acid molecules. These glycosylated forms (termed apoC-III_0, apoC-III_1 and apoC-III_2) derive from glycosylation on threonine 74 (T74) in the amino acid sequence, and account for ~ 90 % of the total apoC-III concentration (**Table 1**). Our results indicate that apoC-III exhibits additional proteoforms: the low abundance *C*-terminal alanine cleaved proteoforms (present in majority of the samples, at 1 to 5 %), and also glycosylated

proteoforms containing fucose in the glycan motif. The presence of fucosylated apoC-III has initially been detected by Balog *et al*., in relation to *Schistosoma mansoni* infection [50]. Nicolardi *et al*. confirmed the presence of fucosylated apoC-III proteoforms in their work [33], and went a step further in analyzing these glycan structures using MALDI-FTICR MS [51]. Their approach, however, is time consuming, and requires sample pretreatment using magnetic beads-based solid-phase extraction prior to MALDI and ESI-FTICR MS analysis. In our work we were able to apply the MSIA approach and detect these low abundance apoC-III proteoforms in a single run alongside the other apoCs, therefore significantly reducing the time and cost of the analysis. The concentrations of the fucosylated apoC-III proteoforms were determined to be in the very low mg/L range (**Fig. 3**). With the exception of apoC-III-$(Gal)_2(GalNAc)_2(Fuc)_3$, they were detected in less than 20 % of samples. Finally, the distribution of the apoC-III proteoforms was noticeably different among the samples, even in samples with similar total apoC-III concentrations. This was especially evident for the fucosylated proteoforms of apoC-III, suggesting that fucosylation follows different pathways. Additional studies are needed in order to better understand these differences in the context of the physiological significance. Gaining an insight into the structural heterogeneity of apoC-III will certainly contribute to new findings and correlations, especially when assessed together with apoC-I and apoC-II, as done here with the multiplex MSIA.

5. Conclusion

The multiplex apoCs assay is based on the mass spectrometric immunoassay (MSIA) methodology. The assay's two step approach is similar to that of well-established enzymatic immunoassays, with the secondary step of MALDI-MS detection an enabling factor in differentiating between multiple proteoforms. The quantification aspect further strengthens the

methodology because the concentration of each proteoforms can be determined individually. The multiplex quantitative MSIA is a unique and high-throughput method for measuring the three apoproteins C and their proteoforms. The ability to detect and quantify all proteoforms in a single assay represents a significant step forward toward better understanding of their physiological roles in health and disease.

Acknowledgements

The authors would like to acknowledge Dr. Tai E Shyong and Dr. Kao Shih Ling from National University Hospital, Singapore for providing us with the human plasma samples used in this analysis. This work was supported in part by Award Numbers R01DK082542 and R24DK090958 from the National Institute of Diabetes and Digestive and Kidney Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Table 1. ApoC-I, apoC-II and apoC-III proteoforms, with their relative percent abundance (RPA) and frequency in a cohort of 82 human EDTA-plasma samples

Proteoforms	Theoretical m/z value	Observed m/z value	RPA (%) Mean (range)	Frequency (%)
des-TP apoC-I	6432.344	6432.104	21.3 (13.6 – 27.0)	100
apoC-I native	6630.565	6630.572	78.7 (72.9 – 86.4)	100
des-TQQPQQ apoC-II	8204.162	8205.290	8.51 (4.09 – 13.7)	100
apoC-II native	8914.906	8915.760	91.5 (86.3 – 95.9)	100
apoC-III native	8764.652	8764.587	6.85 (1.06 – 19.0)	100
apoC-III-$(Gal)_1(GalNAc)_1$	9135.800	9137.925	29.9 (25.4 – 36.8)	100
des-A apoC-III-$(Gal)_1(GalNAc)_1(NeuAc)_1$	9350.171	9351.781	1.17 (0.336 – 3.65)	93.9
apoC-III-$(Gal)_1(GalNAc)_1(NeuAc)_1$	9422.249	9422.562	40.6 (29.1 – 48.3)	100
des-A apoC-III-$(Gal)_1(GalNAc)_1(NeuAc)_2$	9641.429	9642.267	5.96 (4.32 – 9.26)	100
apoC-III-$(Gal)_1(GalNAc)_1(NeuAc)_2$	9712.507	9713.492	12.1 (6.63 – 19.4)	100
apoC-III-$(Gal)_2(GalNAc)_2(Fuc)_3$	9933.768	9933.941	2.61 (0.477 – 6.83)	100
apoC-III-$(Gal)_3(GalNAc)_3(Fuc)_2$	10152.961	10153.559	0.949 (0.241 – 1.42)	12.2
apoC-III-$(Gal)_4(GalNAc)_2(Fuc)_3$	10258.056	10259.700	1.74 (0.139 – 3.78)	20.7
apoC-III-$(Gal)_2(GalNAc)_4(Fuc)_3$	10340.158	10341.098	1.40 (0.837 – 2.00)	17.1
apoC-III-$(Gal)_3(GalNAc)_3(Fuc)_4$	10445.869	10446.206	0.507 (0.209 – 0.771)	15.9
apoC-III-$(Gal)_5(GalNAc)_3(Fuc)_4$	10769.541	10770.404	0.924 (0.593 – 1.53)	12.2

Gal – Galactose; GalNAc - *N*-acetylgalactosamine; NeuAc - *N*-acetyl neuraminic acid (sialic acid); Fuc – Fucose

References

[1] L.M. Smith, N.L. Kelleher, C.f.T.D. Proteomics, Nat Methods, 10 (2013) 186-187.
[2] R.W. Nelson, J.R. Krone, A.L. Bieber, P. Williams, Anal Chem, 67 (1995) 1153-1158.
[3] R.W. Nelson, C.R. Borges, J Am Soc Mass Spectrom, 22 (2011) 960-968.
[4] H.J. Issaq, T.D. Veenstra, T.P. Conrads, D. Felschow, Biochem Biophys Res Commun, 292 (2002) 587-592.
[5] N.L. Anderson, N.G. Anderson, L.R. Haines, D.B. Hardie, R.W. Olafson, T.W. Pearson, J Proteome Res, 3 (2004) 235-244.
[6] M. Razavi, L.E. Frick, W.A. LaMarr, M.E. Pope, C.A. Miller, N.L. Anderson, T.W. Pearson, J Proteome Res, 11 (2012) 5642-5649.
[7] J.D. Reid, D.T. Holmes, D.R. Mason, B. Shah, C.H. Borchers, J Am Soc Mass Spectrom, 21 (2010) 1680-1686.
[8] M. Mann, Nat Rev Mol Cell Biol, 7 (2006) 952-958.
[9] I. Messana, T. Cabras, F. Iavarone, F. Vincenzoni, A. Urbani, M. Castagnola, J Sep Sci, 36 (2013) 128-139.
[10] F. Weiß, B.H. van den Berg, H. Planatscher, C.J. Pynn, T.O. Joos, O. Poetz, Biochim Biophys Acta, 1844 (2014) 927-932.
[11] D. Nedelkov, K.A. Tubbs, E.E. Niederkofler, U.A. Kiernan, R.W. Nelson, Anal. Chem., 76 (2004) 1733-1737.
[12] D. Nedelkov, U.A. Kiernan, E.E. Niederkofler, K.A. Tubbs, R.W. Nelson, Proc Natl Acad Sci U S A, 102 (2005) 10852-10857.
[13] D. Nedelkov, D.A. Phillips, K.A. Tubbs, R.W. Nelson, Mol Cell Proteomics, 6 (2007) 1183-1187.
[14] O. Trenchevska, D.A. Phillips, R.W. Nelson, D. Nedelkov, PLoS One, 9 (2014) e100713.
[15] P.E. Oran, O. Trenchevska, D. Nedelkov, C.R. Borges, M.R. Schaab, D.S. Rehder, J.W. Jarvis, N.D. Sherma, L. Shen, B. Krastins, M.F. Lopez, D.C. Schwenke, P.D. Reaven, R.W. Nelson, PLoS One, 9 (2014) e92801.
[16] E.E. Niederkofler, D.A. Phillips, B. Krastins, V. Kulasingam, U.A. Kiernan, K.A. Tubbs, S.M. Peterman, A. Prakash, E.P. Diamandis, M.F. Lopez, D. Nedelkov, PLoS One, 8 (2013) e81125.
[17] O. Trenchevska, E. Kamcheva, D. Nedelkov, J Proteome Res, 9 (2010) 5969-5973.
[18] O. Trenchevska, D. Nedelkov, Proteome Sci, 9 (2011) 19.
[19] O. Trenchevska, E. Kamcheva, D. Nedelkov, Proteomics, 11 (2011) 3633-3641.
[20] U.A. Kiernan, D.A. Phillips, O. Trenchevska, D. Nedelkov, PLoS One, 6 (2011) e17282.
[21] N.D. Sherma, C.R. Borges, O. Trenchevska, J.W. Jarvis, D.S. Rehder, P.E. Oran, R.W. Nelson, D. Nedelkov, Proteome Sci, 12 (2014) 52.
[22] U.A. Kiernan, D. Nedelkov, R.W. Nelson, J Proteome Res, 5 (2006) 2928-2934.
[23] U.A. Kiernan, R. Addobbati, D. Nedelkov, R.W. Nelson, J Proteome Res, 5 (2006) 1682-1687.
[24] M.C. Jong, M.H. Hofker, L.M. Havekes, Arterioscler Thromb Vasc Biol, 19 (1999) 472-484.
[25] A.A. Kei, T.D. Filippatos, V. Tsimihodimos, M.S. Elisaf, Metabolism, 61 (2012) 906-921.
[26] C. Zheng, C. Khoo, J. Furtado, F.M. Sacks, Circulation, 121 (2010) 1722-1734.
[27] R.A. Hegele, P.W. Connelly, A.J.G. Hanley, F. Sun, S.B. Harris, B. Zinman, Arteriosclerosis, Thrombosis, and Vascular Biology, 17 (1997) 2753-2758.

[28] A.B. Jørgensen, R. Frikke-Schmidt, B.G. Nordestgaard, A. Tybjærg-Hansen, New England Journal of Medicine, 371 (2014) 32-41.
[29] J. Crosby, G.M. Peloso, P.L. Auer, D.R. Crosslin, N.O. Stitziel, L.A. Lange, Y. Lu, Z.Z. Tang, H. Zhang, G. Hindy, N. Masca, K. Stirrups, S. Kanoni, R. Do, G. Jun, Y. Hu, H.M. Kang, C. Xue, A. Goel, M. Farrall, S. Duga, P.A. Merlini, R. Asselta, D. Girelli, O. Olivieri, N. Martinelli, W. Yin, D. Reilly, E. Speliotes, C.S. Fox, K. Hveem, O.L. Holmen, M. Nikpay, D.N. Farlow, T.L. Assimes, N. Franceschini, J. Robinson, K.E. North, L.W. Martin, M. DePristo, N. Gupta, S.A. Escher, J.H. Jansson, N. Van Zuydam, C.N. Palmer, N. Wareham, W. Koch, T. Meitinger, A. Peters, W. Lieb, R. Erbel, I.R. Konig, J. Kruppa, F. Degenhardt, O. Gottesman, E.P. Bottinger, C.J. O'Donnell, B.M. Psaty, C.M. Ballantyne, G. Abecasis, J.M. Ordovas, O. Melander, H. Watkins, M. Orho-Melander, D. Ardissino, R.J. Loos, R. McPherson, C.J. Willer, J. Erdmann, A.S. Hall, N.J. Samani, P. Deloukas, H. Schunkert, J.G. Wilson, C. Kooperberg, S.S. Rich, R.P. Tracy, D.Y. Lin, D. Altshuler, S. Gabriel, D.A. Nickerson, G.P. Jarvik, L.A. Cupples, A.P. Reiner, E. Boerwinkle, S. Kathiresan, N.H. TG and HDL Working Group of the Exome Sequencing Project, L.ng, and Blood Institute, N Engl J Med, 371 (2014) 22-31.
[30] M.W. Huff, R.A. Hegele, Circ Res, 112 (2013) 1405-1408.
[31] M.J. Graham, R.G. Lee, T.A. Bell, W. Fu, A.E. Mullick, V.J. Alexander, W. Singleton, N. Viney, R. Geary, J. Su, B.F. Baker, J. Burkey, S.T. Crooke, R.M. Crooke, Circ Res, 112 (2013) 1479-1490.
[32] P.V. Bondarenko, S.L. Cockrill, L.K. Watkins, I.D. Cruzado, R.D. Macfarlane, J Lipid Res, 40 (1999) 543-555.
[33] S. Nicolardi, Y.E. van der Burgt, M. Wuhrer, A.M. Deelder, Proteomics, 13 (2013) 992-1001.
[34] Y. Wada, M. Kadoya, N. Okamoto, Glycobiology, 22 (2012) 1140-1144.
[35] S.K. Flood-Nichols, D. Tinnemore, M.A. Wingerd, A.I. Abu-Alya, P.G. Napolitano, J.D. Stallings, D.L. Ippolito, Mol Cell Proteomics, 12 (2013) 55-64.
[36] M. Rezeli, Á. Végvári, F. Donnarumma, O. Gidlöf, J.G. Smith, D. Erlinge, G. Marko-Varga, J Proteomics, 87 (2013) 16-25.
[37] W. Jian, R.W. Edom, D. Wang, N. Weng, S.W. Zhang, Anal Chem, 85 (2013) 2867-2874.
[38] S.F. Kingsmore, Nat Rev Drug Discov, 5 (2006) 310-320.
[39] S. Peterman, E.E. Niederkofler, D.A. Phillips, B. Krastins, U.A. Kiernan, K.A. Tubbs, D. Nedelkov, A. Prakash, M.S. Vogelsang, T. Schoeder, L. Couchman, D.R. Taylor, C.F. Moniz, G. Vadali, G. Byram, M.F. Lopez, Proteomics, 14 (2014) 1445-1456.
[40] E. Kuhn, T. Addona, H. Keshishian, M. Burgess, D.R. Mani, R.T. Lee, M.S. Sabatine, R.E. Gerszten, S.A. Carr, Clin Chem, 55 (2009) 1108-1117.
[41] H. Keshishian, T. Addona, M. Burgess, E. Kuhn, S.A. Carr, Mol Cell Proteomics, 6 (2007) 2212-2229.
[42] J.D. Chapman, D.R. Goodlett, C.D. Masselon, Mass Spectrom Rev, 33 (2014) 452-470.
[43] K. Augustyns, G. Bal, G. Thonus, A. Belyaev, X.M. Zhang, W. Bollaert, A.M. Lambeir, C. Durinx, F. Goossens, A. Haemers, Curr Med Chem, 6 (1999) 311-327.
[44] A.M. Lambeir, C. Durinx, S. Scharpé, I. De Meester, Crit Rev Clin Lab Sci, 40 (2003) 209-294.
[45] S.S. Fojo, L. Taam, T. Fairwell, R. Ronan, C. Bishop, M.S. Meng, J.M. Hoeg, D.L. Sprecher, H.B. Brewer, J Biol Chem, 261 (1986) 9591-9594.
[46] S.A. Carr, S.E. Abbatiello, B.L. Ackermann, C. Borchers, B. Domon, E.W. Deutsch, R.P. Grant, A.N. Hoofnagle, R. Hüttenhain, J.M. Koomen, D.C. Liebler, T. Liu, B. MacLean, D.R.

Mani, E. Mansfield, H. Neubert, A.G. Paulovich, L. Reiter, O. Vitek, R. Aebersold, L. Anderson, R. Bethem, J. Blonder, E. Boja, J. Botelho, M. Boyne, R.A. Bradshaw, A.L. Burlingame, D. Chan, H. Keshishian, E. Kuhn, C. Kinsinger, J.S. Lee, S.W. Lee, R. Moritz, J. Oses-Prieto, N. Rifai, J. Ritchie, H. Rodriguez, P.R. Srinivas, R.R. Townsend, J. Van Eyk, G. Whiteley, A. Wiita, S. Weintraub, Mol Cell Proteomics, 13 (2014) 907-917.
[47] R.E. Pratley, A. Salsali, Curr Med Res Opin, 23 (2007) 919-931.
[48] A. Barnett, Int J Clin Pract, 60 (2006) 1454-1470.
[49] I.J. Goldberg, C.A. Scheraldi, L.K. Yacoub, U. Saxena, C.L. Bisgaier, J Biol Chem, 265 (1990) 4266-4272.
[50] C.I. Balog, O.A. Mayboroda, M. Wuhrer, C.H. Hokke, A.M. Deelder, P.J. Hensbergen, Mol Cell Proteomics, 9 (2010) 667-681.
[51] S. Nicolardi, Y.E. van der Burgt, I. Dragan, P.J. Hensbergen, A.M. Deelder, J Proteome Res, 12 (2013) 2260-2268.

Figures legends:

Figure 1: Mass spectra resulting from multiplex MSIA of apoC-I, apoC-II and apoC-III utilizing lysozyme as an IRS. Labeled on the mass spectra are the major C apolipoprotein proteoforms. The inlets represent mass spectra obtained from samples where the specific proteoforms were more prominent.

Figure 2: Quantitative apoC-I, apoC-II and apoC-III mass spectrometric immunoassay (MSIA). **a)** Representative MSIA mass spectra for the six standard apoCs samples containing lysozyme as an IRS; **b)** Representative standard curves generated with the multiplex apoCs MSIA, with an r^2= 0.9988 and SEE = 0.0928 for apoC-I; r^2 = 0.9940 and SEE = 0.0262 for apoC-II; and r^2 = 0.9996 and SEE = 0.1605 for apoC-III.

Figure 3: Individual concentrations of apolipoprotein C-I, C-II and C-III and their proteoforms in a cohort of 82 human EDTA-plasma samples. Box - 25-75th percentile; Solid line – median concentration; Short dash line – mean concentration; Error bars – 10th and 90th percentile; Symbols – outlying points.

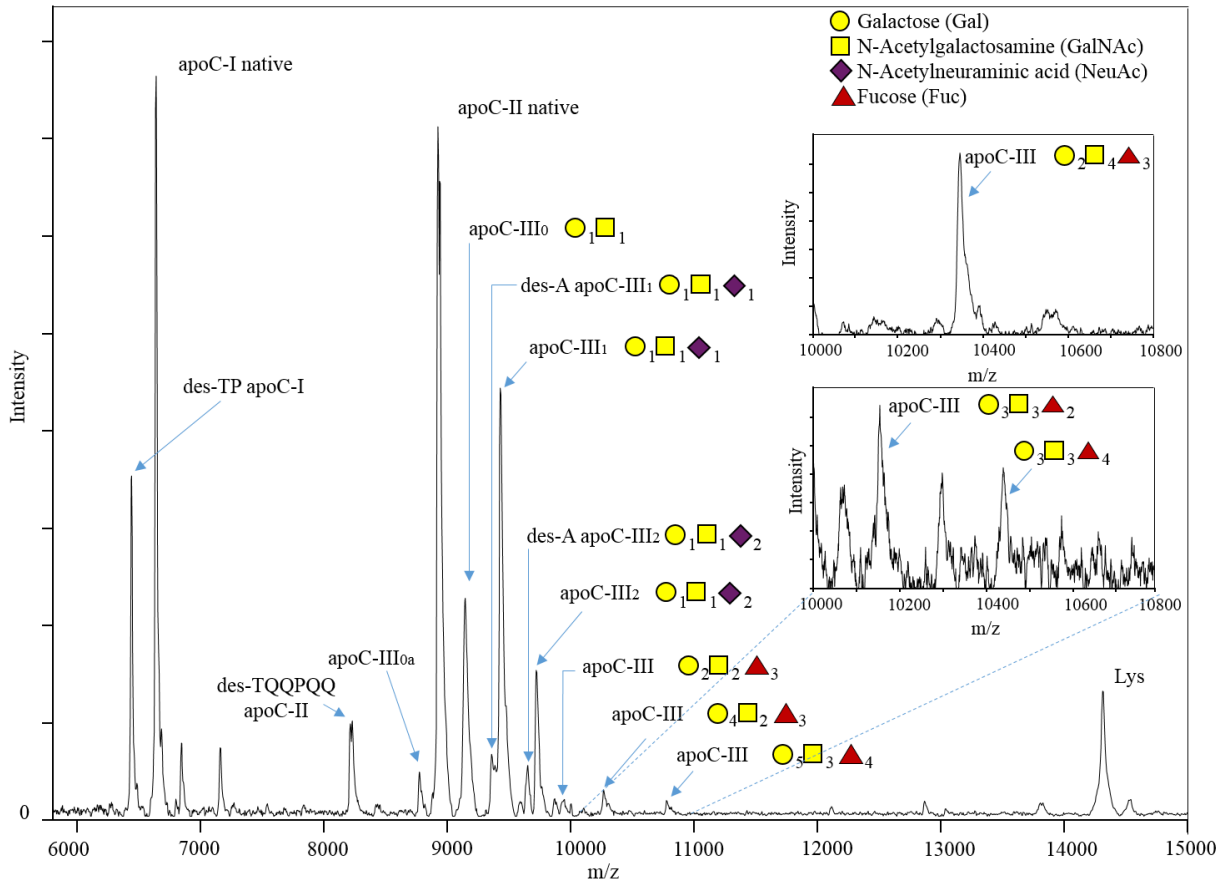
Galactose (Gal)
N-Acetylgalactosamine (GalNAc)
N-Acetylneuraminic acid (NeuAc)
Fucose (Fuc)
apoC-I native
apoC-II native
apoC-III0
des-A apoC-III1
apoC-III1
des-TP apoC-I
des-A apoC-III2
apoC-III2
apoC-III0a
des-TQQPQQ apoC-II
apoC-III
apoC-III
apoC-III
Lys
Intensity
m/z
0
6000
7000
8000
9000
10000
11000
12000
13000
14000
15000
10000
10200
10400
10600
10800

Figure 1

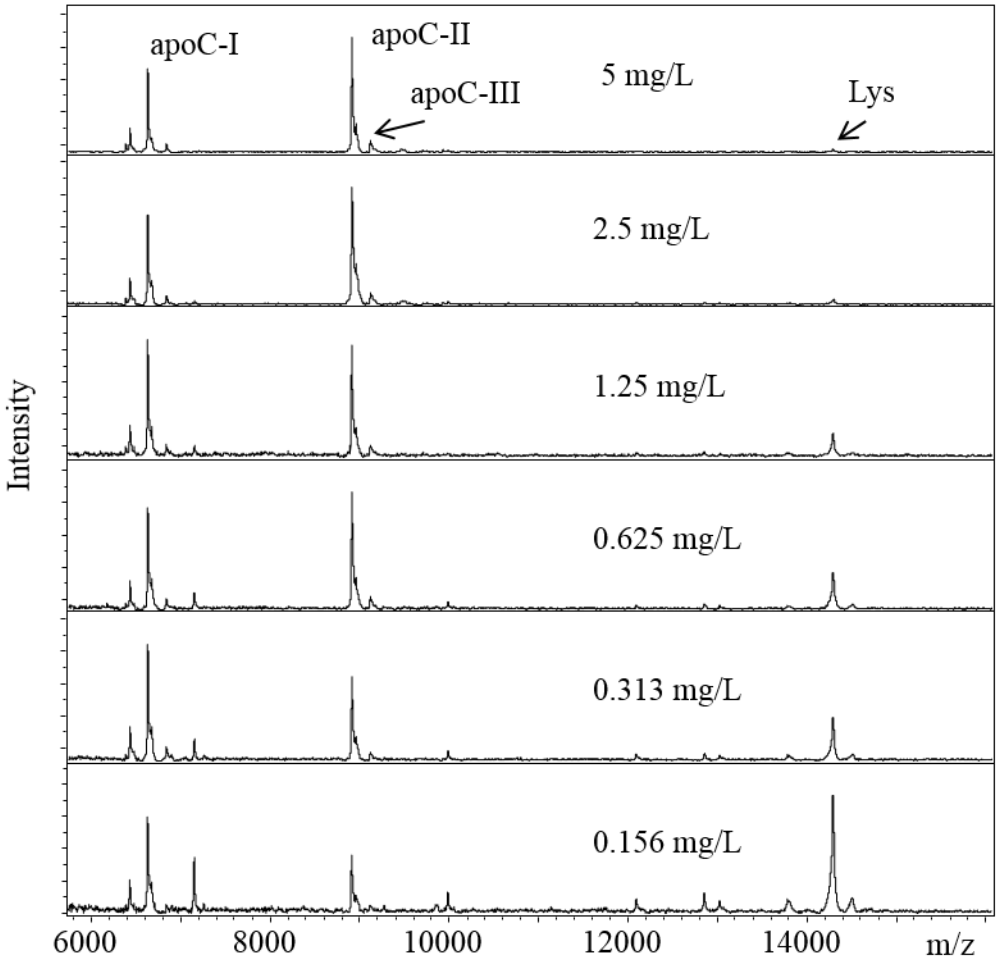
apoC-I
apoC-II
apoC-III
5 mg/L
Lys
2.5 mg/L
1.25 mg/L
0.625 mg/L
0.313 mg/L
0.156 mg/L
Intensity
6000
8000
10000
12000
14000
m/z

Figure 2.a)

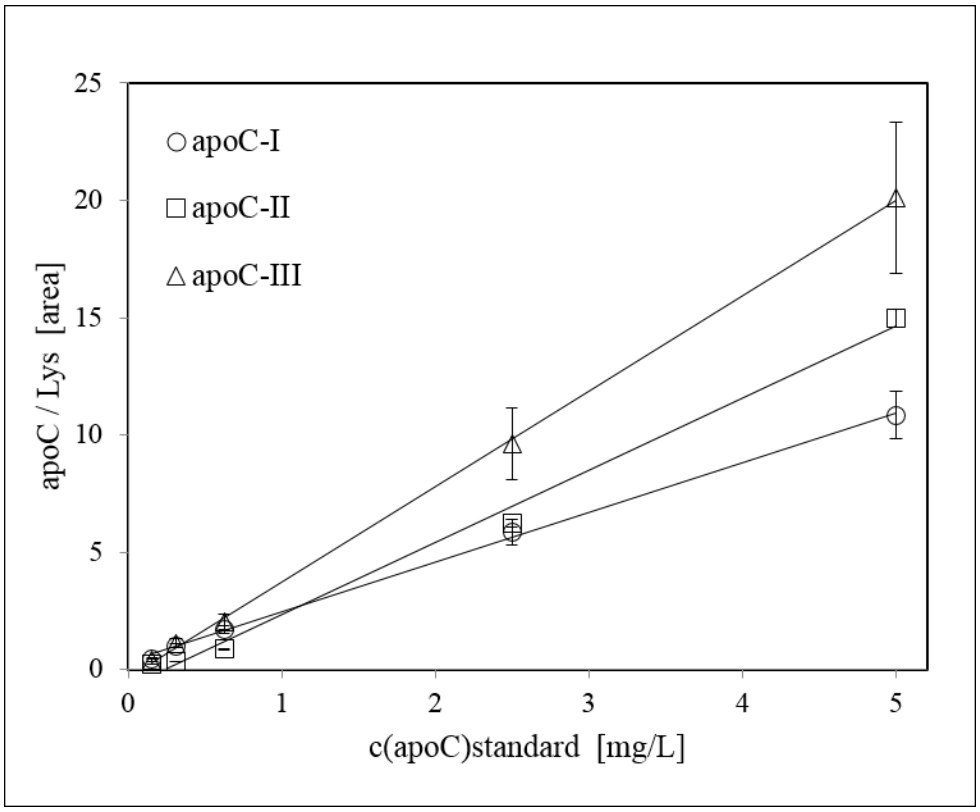
25
20
15
10
5
0
0
1
2
3
4
5
apoC / Lys [area]
c(apoC)standard [mg/L]
apoC-I
apoC-II
apoC-III

Figure 2.b)

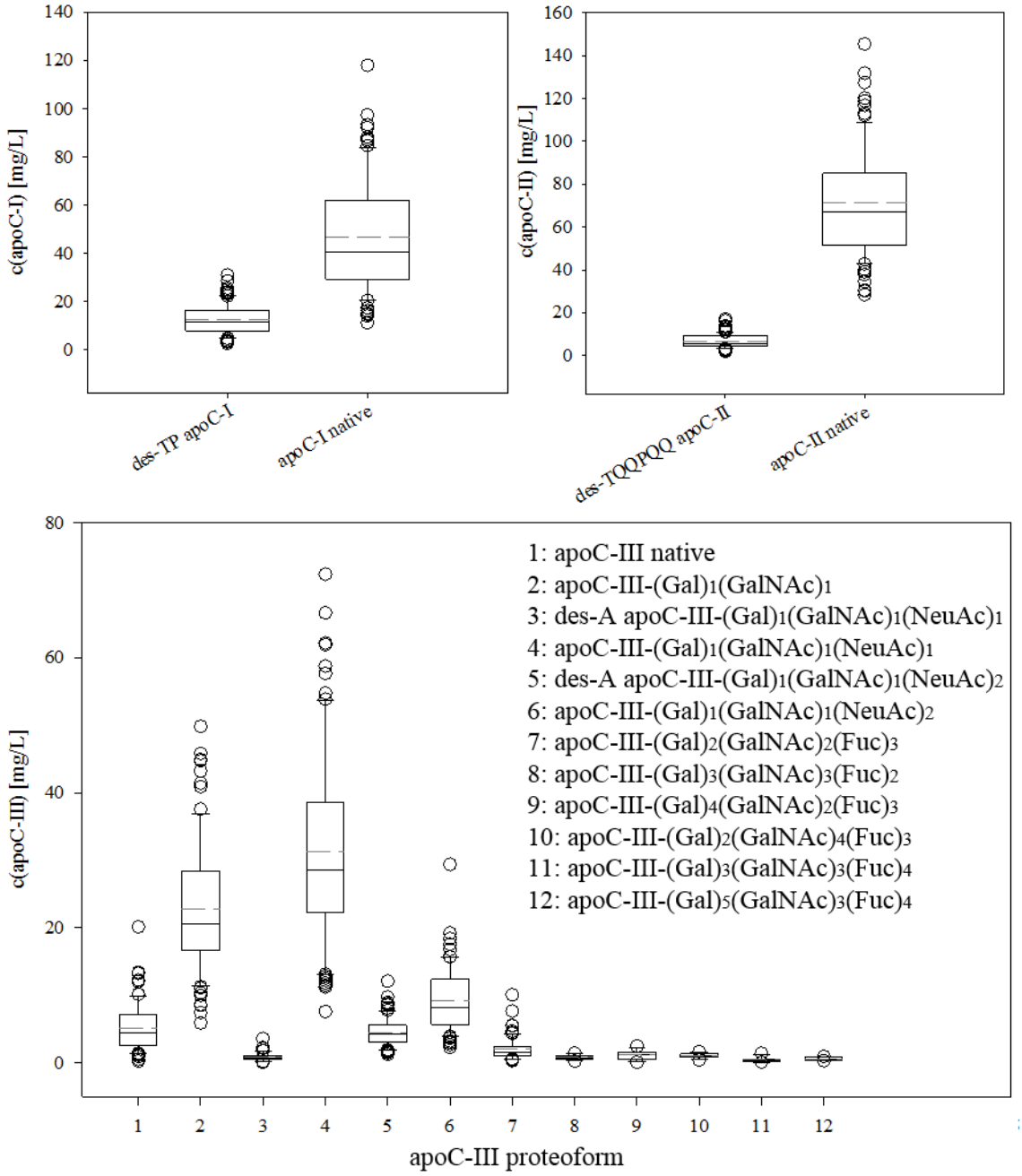
c(apoC-I) [mg/L]
des-TP apoC-I
apoC-I native
c(apoC-II) [mg/L]
des-TQQPQQ apoC-II
apoC-II native
c(apoC-III) [mg/L]
1: apoC-III native
2: apoC-III-(Gal)1(GalNAc)1
3: des-A apoC-III-(Gal)1(GalNAc)1(NeuAc)1
4: apoC-III-(Gal)1(GalNAc)1(NeuAc)1
5: des-A apoC-III-(Gal)1(GalNAc)1(NeuAc)2
6: apoC-III-(Gal)1(GalNAc)1(NeuAc)2
7: apoC-III-(Gal)2(GalNAc)2(Fuc)3
8: apoC-III-(Gal)3(GalNAc)3(Fuc)2
9: apoC-III-(Gal)4(GalNAc)2(Fuc)3
10: apoC-III-(Gal)2(GalNAc)4(Fuc)3
11: apoC-III-(Gal)3(GalNAc)3(Fuc)4
12: apoC-III-(Gal)5(GalNAc)3(Fuc)4
apoC-III proteoform

Figure 3