



A food safety control low mass-range proteomics platform for the detection of illicit treatments in veal calves by MALDI-TOF-MS serum profiling

Lorenza Della Donna, Maurizio Ronci, Paolo Sacchetta, Carmine Di Ilio, Bartolomeo Biolatti, Giorgio Federici, Carlo Nebbia, Andrea Urbani

► To cite this version:

Lorenza Della Donna, Maurizio Ronci, Paolo Sacchetta, Carmine Di Ilio, Bartolomeo Biolatti, et al.. A food safety control low mass-range proteomics platform for the detection of illicit treatments in veal calves by MALDI-TOF-MS serum profiling. *Biotechnology Journal*, 2009, 4 (11), pp.1596. 10.1002/biot.200900085 . hal-00527652

HAL Id: hal-00527652

<https://hal.science/hal-00527652>

Submitted on 20 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**A food safety control low mass-range proteomics platform
for the detection of illicit treatments in veal calves by
MALDI-TOF-MS serum profiling**

Journal:	<i>Biotechnology Journal</i>
Manuscript ID:	biot.200900085.R2
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	26-Jun-2009
Complete List of Authors:	Della Donna, Lorenza; Faculty of Veterinary Medicine University of Turin, Department of Animal Pathology Ronci, Maurizio; University of Rome Tor Vergata, Internal Medicine Sacchetta, Paolo; University G. d'Annunzio Chieti-Pescara, Biomedical science Di Ilio, Carmine; University G. d'Annunzio Chieti-Pescara, Biomedical science Biolatti, Bartolomeo; Faculty of Veterinary Medicine University of Turin, Department of Animal Pathology Federici, Giorgio; University of Rome Tor Vergata, Internal Medicine Nebbia, Carlo; Faculty of Veterinary Medicine University of Turin, Department of Animal Pathology Urbani, Andrea; University of Rome Tor Vergata, Internal Medicine
Keywords:	FOOD SAFETY, MASS SPECTROMETRY, GROWTH PROMOTING AGENTS, PROTEOMICS, BIOMARKER



Research Article

A food safety control low mass-range proteomics platform for the detection of illicit treatments in veal calves by MALDI-TOF-MS serum profiling

Lorenza Della Donna^{1*}, Maurizio Ronci^{2,3*}, Paolo Sacchetta^{4,5}, Carmine Di Ilio^{2,4,5}, Bartolomeo Biolatti¹, Giorgio Federici^{2,3}, Carlo Nebbia¹ and Andrea Urbani^{2,3}

¹Department of Animal Pathology, Division of Pharmacology and Toxicology, Faculty of Veterinary Medicine, Università di Torino, Grugliasco (Turin), Italy

²IRCCS-Fondazione Santa Lucia Centro di Ricerca sul Cervello, Rome, Italy

³Dipartimento di Medicina Interna, Università di Roma “Tor Vergata”, Rome, Italy

⁴Centro Studi sull’Invecchiamento (Ce.S.I.), Chieti, Italy

⁵Dipartimento di Scienze Biomediche, Università “G. D’Annunzio” di Chieti-Pescara, Italy

Correspondence: Professor Andrea Urbani, Department of Internal Medicine, University of Rome “Tor Vergata”, Via del Fosso di Fiorano 65, 00149 Roma, Italy

E-mail: andrea.urban@uniroma2.it

Fax: +39-06-50170332

Keywords: Biomarker / Food safety / Growth promoting agents / Mass spectrometry / Proteomics

* Both authors have contributed equally to this work.

Abbreviations: **17 β E**, 17 β -estradiol; **17 β T**, 17 β -testosterone; **IS**, internal standard; **PCA**, principal components analysis; **PCs**, principal components; **PEAs**, performance enhancing agents; **PLS-DA**, partial least square-discriminant analysis; **SPE**, solid phase extraction; **SS**, sum of squares; **TOF**, time-of-flight

Performance enhancing agents (PEAs) are illegally used in cattle and other meat producing species to increase food conversion and lean meat production. Due to the very short breeding cycle, veal calves represent the meat producing bovine category mostly subjected to illicit treatments. These chemical agents are difficult to detect by conventional analytical approaches due to the employment of synergistic formulations at very low dosage and given the use of uncharacterized novel compounds. Such a scenario has fostered a strong interest in the discovery of functional molecular biomarkers for the detection of growth promoting agents in meat producing species. A multivariate MALDI-TOF-MS proteomics platform has been developed using bovine serum samples. Analytical performances have been thoroughly evaluated in order to enable reproducible profiles from 10 μ L sera samples. We propose univariate and multivariate discrimination models capable to identify calves undergoing illicit treatments. In particular, we found a strong discrimination power associated with a polypeptide fragment from β 2-glycoprotein-I. We provide a fundamental proof of concept in the potential application of MALDI-TOF-MS proteomics profiling in the food safety control.

1 Introduction

Despite the ban by the EU, a number of performance enhancing agents (PEAs) are still illegally used in cattle and other meat producing species to increase food conversion, lean meat production, and to improve carcass quality traits. The effect of the residues of these

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

chemicals in the resulting edible products is extremely worrying for health risk of consumers. According to the 2005 residue monitoring plans, in the European Union Member States the number of noncompliant PEAs, such as, sexual steroids, β -agonists, and corticosteroids, is reported to be less than 0.3% of the examined samples. However, the results from the histological investigations performed in target organs and analyses of seized black market preparations suggest that the chemical manipulation of cattle growth is still a matter of concern [1]. The remarkable discrepancy between the reported official data and the real situation may be due to the ever changing strategies for animal illegal treatment. For example, these would involve: (i) the “week-end schedule,” consisting of administration of rapidly metabolized active principles on Fridays, (ii) the use of hormone precursors, (iii) the use of a combination of different active principles at very low dosages, (iv) the use of principles of unknown chemical structure or not included in the national residue monitoring plans [2]. Any of these procedures can result in the generation of (false) negative data. This is especially true when samples are screened for the presence of PEA residues with immunochemical assays, often characterized by a quite low sensitivity. There is therefore the urgent need to develop biological assays able to ascertain exposure of food producing animals to a wide number of illegally administrated PEAs [3]. Proteomics represents a powerful tool to detect the modifications of protein expression resulting from the exposure of living animals to such agents, which are known to act at molecular level by altering the rate of gene transcription. A previous study from our laboratory already proved the usefulness of the proteomic approach in detecting biomarkers of illicit treatments [4]. However, we are still facing the lack of proteomics investigations in biological samples collected from living animals. Among body fluids, serum may constitute a potential sample for the detection of illegal treatments. Recently, serum has gained an increasing interest in proteomic research for

molecular biomarker discovery, especially for human clinical issues, where it is considered potentially the most easily accessible sample specimen for biomarker elucidation [5–7].

To date, a limited number of proteomic studies on bovine tissues and biological fluids have been published [8–15].

The aim of this study was to develop and characterize a MALDI time-of-flight MS (MALDI-TOF-MS) proteome profile of the bovine serum in order to provide a suitable biotechnological framework for on field investigations of food safety. In particular, focusing on the low molecular proteome profile ranging between 3 and 20 kDa, we developed a biomarker detection strategy to highlight exposure of food producing species to illegal drugs and PEAs. The achieved classification models have been benchmarked on the discrimination of calves undergoing treatment with different PEA protocols currently employed on bovine illicit treatment.

2 Methods

2.1 Animal treatments and source of serum samples

Serum samples were obtained from 4- to 6-month-old crossbred male veal calves subjected to different experimental treatments or from animals ($n = 17$) reared in a breeding farm and officially declared untreated (C_{UT}). In all cases, blood samples were collected by venipuncture in Falcon vessels, allowed to clot and centrifuged at room temperature for 10 min at 2000 rpm ($814 \times g$). The resulting sera were stored at -80°C in 1.5 mL aliquots until further use.

Twenty-four calves were randomly allotted to four groups of six individuals each. One group consisted in untreated animals (C_2); the other groups were subjected to one of the following

treatments: (i) 17 β -estradiol (17 β E), six i.m. administrations at increasing dose rates (20–40 mg i.m. *per capita*) at weekly intervals; (ii) 17 β -testosterone (17 β T), six i.m. administrations at increasing dose rates (150–200 mg i.m. *per capita*) at weekly intervals; (iii) glyburide 2.5 mg + androsterone 1 mg (GLY–AND) *per capita per day* for 37 days. Blood samples were collected 24 h after the last treatment.

2.2 Serum preparation for MALDI-TOF-MS analysis

An amount of 10 μ L of crude equine myoglobin (5 pmol/ μ L) was added to 10 μ L crude sera samples as internal standard (IS). Samples were then desalted, concentrated, and spotted on MALDI-TOF ground steel plate using different solid phase extraction (SPE) procedures. ZipTips C₄ and C₁₈ (tip size P10; Millipore, USA) were activated and equilibrated using 50% ACN/H₂O and 0.1% TFA, respectively. The proteins were let bound to the resin, desalted by washing with 0.1% TFA and eluted directly on a plate together with 2 μ L of sinapinic acid saturated solution in 30% ACN/0.1% TFA. Strong cationic exchanger ZipTips (SCX) were activated and washed with 0.1% TFA, charged with the sample and eluted on a plate with 2 μ L of 5% ammonium hydroxide in H₂O/MeOH 70:30. After air drying, 1 μ L of sinapinic acid saturated solution in 30% ACN/0.1% TFA was added to the samples. Metal chelating surface ZipTips (MC) were equilibrated with 0.1% acetic acid in H₂O/ACN 50:50, charged with 200 mM FeCl₃ in 10 mM HCl, washed with Milli-Q grade water and washed again with 1% acetic acid in H₂O/ACN 80:20. Next, they were equilibrated with 0.1% acetic acid in H₂O/ACN 90:10, charged with the sample and eluted in a clean vial with 2 μ L of 0.3 M ammonium hydroxide. An amount of 1 μ L of the eluted solution was spotted on the plate, allowed to partially dry and overlaid with 1 μ L of sinapinic acid saturated solution in 30% ACN/0.1% TFA.

2.3 MALDI-TOF-MS analysis

All sera samples were analyzed using a MALDI-TOF mass spectrometer. Spectra acquisitions were performed on a Reflex IV and on an UltraFlex III TOF/TOF (Bruker-Daltonics, Germany) MALDI-TOF mass spectrometer instruments equipped with nitrogen laser ($\lambda = 337$ nm) and SmartBeam laser, respectively, both working in positive linear mode. The acquisition range was set to 3–20 kDa; laser power and total number of shots were adjusted for each acquisition to obtain no more than 10^3 ion counts *per* shot train and between 2×10^4 and 3×10^4 total ion counts for the base peak. External calibrations were performed using singly and doubly charged β -hemoglobin chain (at 15 868 and 7934.5 Th, respectively), and singly and doubly charged α -hemoglobin chain (at 15 127 and 7564 Th, respectively) from human blood. Heights of MS peaks were normalized to the peak intensity of the equine myoglobin (IS). Spectra were smoothed before analysis. All spectra processing for the validation step was performed using FlexAnalysis 2.0 software (Bruker-Daltonics) with an S/N threshold of 10 and a centroid evaluation of the peaks.

2.4 Serum fractionation

An amount of 50 μ L of a pool of sera from untreated animals was diluted 1:1 with 0.1% TFA and processed in HPLC using an Agilent 1100 series Liquid Chromatograph (Agilent Technologies, USA) equipped with a reverse phase semiprep-column (C18 HYPERSIL ODS 5 μ m, 25 mm i.d. \times 200 mm 0.4) and a multidiode array detector monitoring at 280 nm. The flow rate was maintained at 3.8 mL/min and a stepwise gradient from 95% solvent A (0.1% TFA in water) to 95% solvent B (0.1% TFA in acetonitrile) was used. Fractions were

1
2
3 automatically collected in 30 s intervals (1.9 mL) and concentrated by vacuum centrifugation
4
5
6 to approximately 150 μ L. Concentrated fractions were analyzed by linear MALDI-TOF-MS
7
8 and subsequently used for protein identification by nLC-Q-TOF-MS/MS after reduction,
9
10 alkylation, and digestion with DTT 100 mM, iodoacetamide 200 mM, and trypsin 0.5 μ g/ μ L,
11
12 respectively.
13
14
15
16
17

18 **2.5 LC-ESI-MS/MS analysis**

19
20
21
22 LC-MS/MS analysis was carried out on a Q-TOF Ultima (Waters, USA) equipped with a
23
24 nano-Lock-Spray ion source and coupled with a CapLC chromatography system (Waters). A
25
26 volume of 6 μ L of digested fractions was injected and first concentrated on a Trap Column
27
28 (Waters Symmetry300 C₁₈ 5 μ m OPTI-PAK) at high flow rate, then, eluted on a C₁₈ column
29
30 (LCPackings DIONEX PepMap 5 μ m, 100 Å 75 mm i.d. 250 mm) with a H₂O/ACN gradient
31
32 in the presence of 0.2% of formic acid using a precolumn splitter at 200 nL/min. On-line
33
34 recalibration was performed by constantly infusing a GFP solution at 250 fmol/ μ L in
35
36 H₂O/ACN 0.1% formic acid. The voltage applied to the capillary and the desolvation gas
37
38 flow rate were tuned up before each analysis to obtain a stable spray. Argon was used as
39
40 collision gas. The collision energy was automatically chosen from control software as a
41
42 function of the precursor ion mass. MS/MS spectra were acquired by automatic switching
43
44 between MS and MS/MS mode. Acquired MS/MS data were converted in a centroid format
45
46 by ProteinLinx 2.0 software (Waters).
47
48
49
50
51
52
53
54
55

56 **2.6 Statistical data analysis**

57 **2.6.1 Univariate data analysis**

58
59
60

ClinProTools (Bruker-Daltonics) was used for differential peaks selection between the following groups: C_2 , $17\beta E$, $17\beta T$, and GLY + AND; Box and Whisker graphs and Mann–Whitney tests were carried out by means of MedCalc 9. Kruskal–Wallis one way analysis of variance on ranks was performed with SigmaStat version 3.5 with the Dunn’s method.

2.6.2 Multivariate data analysis

Collected spectra from experiment 2 and from the independent control set were divided in two groups: belonging to the first group (untreated animals) were C_2 and C_{UT} ; belonging to the second group were $17\beta E$, $17\beta T$, or GLY + AND. Spectra were processed in order to extract a consistent data matrix of masses and intensities by means of the MatLab routine LIMPIC [16]. To execute the multivariate data analysis, this matrix was exported to Simca-P+ (Umetrics, Sweden). All data were scaled to unit variance and centered in the preprocessing step. Principal components analysis (PCA) was run on the complete dataset. For projection to latent structure partial least square-discriminant analysis (PLS-DA), two independent datasets were generated by selecting even and odd observations in the original dataset: all even observations were included in the training set used for the generation of model and all odd observations were integrated in the prediction set used to validate the model itself.

3 Results

3.1 Extraction strategy

We initially addressed the development of a robust proteomics platform in order to explore biomarker discovery investigations in bovine sera. We chose to implement linear MALDI-TOF-MS molecular profiling coupled with protein SPE to obtain reproducible and appropriate sized datasets, in order to possibly classify calves subjected to illicit treatments. We selected Sinapinic Acid as MALDI matrix given the high ionization properties on intact proteins [17].

We tested different type of SPE strategies on packed pipette tips (ZipTip) encompassing reverse phase materials such as C₄ and C₁₈ [18], cationic exchange (SCX) and metal affinity (MC) resins to find the best performing strategy in terms of unique MS signals and repeatability. Figure 1 shows the average spectrum ($n = 6$) for each type of extraction applied to officially declared untreated control samples (C_{UT} Section 2**█Please check the change made█**). Table 1 reports the number of corresponding signals and SD for the same set of samples. The average number of peaks with S/N>10 for C₄ extraction is 79 with a SD 5.8 and with uniform distribution of signals over the investigated mass range. C₁₈ showed an unbalanced extraction capacity resulting in a few MS signals in the high mass range and high abundance in the low mass range. SCX and MC were clearly not suitable for our purpose since they gave rise to an extremely low number of proteins, mainly identified in a restricted molecular weight interval ranging between 6500 and 9000 Th.

((Fig.1))

((Table 1))

Considering the number of distinct signals peaks, their distribution over a wide m/z range (3–20 kDa) and the reproducibility in extracting the IS for ZipTip C₄ SPE, this method was selected to develop our MALDI-TOF-MS sera profiling platform.

3.2 Protein signals identification

The C₄ SPE method allowed us to reproducibly profile 79 signals, which resulted in a typical MALDI-TOF mass spectrum of bovine serum proteins. In order to identify the main molecular features of this profile, we used an LC-MALDI-TOF-MS fractionation approach coupled with an nLC-Q-TOF-MS/MS protein identification strategy. Figure 2 shows the separation profile of the semipreparative chromatographic run and the resulting MALDI-TOF-MS spectra for “nonempty” fractions. The analysis of the collected fractions by Q-TOF-MS/MS gave rise to 28 identified proteins (Table 2). Some of them were attributed to fragments of higher molecular weight proteins while others match, in the investigated mass range, with tentative multiple charge species of large proteins. Nevertheless, few identified proteins from the Q-TOF-MS/MS experiments did not correlate with any of the reported MALDI-TOF-MS signals. Such behavior might have different explanations: it can be due to the limited mass range explored in LC-MALDI-MS fractionation experiments (which, however, is super-imposable with the expected C₁₈ HPLC column retention [19]), to nonpredictable post-translational modifications, or to fragmentation of proteins.

((Fig.2))

((Table 2))

3.3 Analytical performances

Reproducibility is a key aspect of using MALDI-TOF-MS as a semiquantitative analytical instrument. Variability could arise from sample handling, sample preparation, and mass spectrometric platform. We followed the guideline for full analytical validation imposed by several international organizations, including FDA.

In order to consolidate the MALDI-TOF-MS bovine serum investigation, we assessed the analytical performances of such a profiling strategy and MS analysis. Mass accuracy of MALDI-TOF mass spectra, after external calibration, resulted in approximately 100 ppm, with resolution (FWHM) higher than 1000. The mass accuracy of the measurements was calculated on six signals chosen between the identified and annotated protein signals. The mass deviation from the calculated average molecular mass for the IS was equal to 0.02 Da, in excellent agreement for MALDI-TOF-MS data in this mass range (Table 3).

((Table 3))

The relative ratio of the representative endogenous peaks *versus* the ISs has been investigated in order to evaluate the repeatability of the method. Table 4 is reported the MALDI intra- and interday repeatability, showing a variability lower than 15% ($n = 18$) for intraday series and lower than 16% ($n = 18$) for interday series.

((Table 4))

All the acquisitions for the figure of merit referring to the analytical validation study have been performed on the Reflex IV instrument. Nevertheless we found similar or better performances on the Ultraflex series; therefore, this latest instrument has been used for the study of groups.

3.4 Application to PEA's illicit treatment

To evaluate a direct application of bovine sera proteins profiling, we focused our interest on biomarker identification for PEA's illicit treatment. To this end, serum samples were used, including samples from controls (C_2) and treated groups (17 β E, 17 β T, or GLY + AND). Samples were independently analyzed by MALDI-TOF-MS after C_4 SPE. Each MALDI-TOF mass spectrum showed approximately 80 distinct peaks at $S/N > 10$. Descriptive

nonparametric statistic (Fig. 3, Box and Whisker plots are reported at the side of each signal) was used to evaluate the significance of the differences between the classes after signal intensities normalization, based on the added IS (heart horse myoglobin 16 952.3 Th). This type of normalization flattened the differences observed in the total ion count weighted profile, as shown in Fig. 3e depicting the IS average signal in the different groups. Due to both the low number of observations and their non-normal distribution inside the groups, a Mann–Whitney test for independent samples was chosen in order to assess the statistical significance of differences. A more appropriate test, when more than two groups must be compared, is the Kruskal–Wallis one way analysis of variance on ranks. The results of both statistical tests are very similar. Resulting p values for the differences between groups at 95% of confidence are shown in Table 5. This direct comparison highlighted four signals that specifically discriminate the control calves from those belonging to the different treatments (6950, 7451, 8733, and 14 900 Th; Fig. 3). Signal at 6950 Th was significantly higher for each treatment group and it was assigned to the sushi domain 3–4 of the β_2 -glycoprotein-I (β_2 GPI, apolipoprotein H, NCBI#: P17690), by combining the described LC-MALDI-TOF-MS fractioning with nLC-MS/MS protein identification (Table 2). Signals at 7451 and 14 900 Th resulted significantly different from the ones highlighted in the untreated animals only for two out of three groups (*i.e.*, 17 β T and GLY + AND); signal at 8733 Th was found significantly different only for the 17 β T treated group. To calculate an ROC curve based on the normalized intensity of the signal of sushi 3–4 of β_2 GP-I, the number of control samples was increased by adding to C_2 ($n = 6$) a set of 17 sera samples collected from a homogeneous group of veal calves reared in the same farm, formally considered as untreated animals (C_{UT}). All treated animals were assigned to the same class. The ROC curve for the signal at 6950 Th demonstrated high sensitivity and specificity (Fig. 4). In particular, adopting the optimal selection criterion

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

$$\frac{I_{6950}^i}{I_{IS}^i} > 0.551_5$$

where I_{6950}^i is the intensity of signal at 6950 for i th sample and I_{IS}^i is the internal standard intensity for the same sample, we can achieve 97.7% specificity with 95% confidence limits at 85.4–99.5%, and 97.2% sensitivity with 95% confidence limits at 87.7–99.6% for the classification of groups. The area under the curve is 0.992 and the significance level p is 0.0001.

((Figs.3,4))
((Table 5))

Despite the statistical significance of single signals differences, a better interpretation of such a large dataset could be provided by a multivariate approach. In order to combine the overall MS molecular sera profile, we have applied a multivariate analysis using both PCA and PLS-DA. All data were scaled to unit variance and the PCA performed on all samples resulted in a nine principal components (PCs) model with an explained variance of 76.1% (Table 6). Two strong outliers were present in the overall cohort of 65 data points; nevertheless the distribution of residues was indicative of a homogeneous dataset. The first component explains the 22.4% of the variance; on the x -axis three distinct areas are visible (Fig. 5a), suggesting a partial unsupervised clustering of the samples. Even if a separation between control and treated calves is visible in the PCA model, no significant clustering was obtained between the different growth promoting agents treatments.

((Fig.5))
((Table 6))

A more powerful statistical analysis for class discrimination is the PLS-DA, in which group information is specified *a priori* and helps drive the separation. PLS-DA was carried out on all samples. Observations were divided in a training dataset, employed to construct the model, and a prediction dataset was used to perform a crossvalidation. In order to reach a significant number of observations we kept the treated group as a single cohort: this assumption originated by considering the similar biological effects of the treatments and the results of the unsupervised PCA analysis. The generated PLS-DA model showed five significant components and gave rise to a complete clustering of the two groups for all replicates of investigated samples (Fig. 5b). The built mathematical model had a predictive power of 69.4% (Table 7). The loading plot (Fig. 5c) highlights the strong statistical weight of sushi 3–4 fragment of the β_2 GP-I, which is the first variable in the “importance of variables in the projection” (VIP) plot (Fig. 5e); in effect the signal of the β_2 GP-I at 6950 Th resulted absent in all control from the validation group (C_{UT}). The other signals, resulted significantly different in the univariate analysis (7451, 8733, and 14 900 Th), but apparently did not show a very high influence upon the combined class clustering for all the growth promoting agents treatments.

((Table 7))

The crossvalidation of the PLS-DA model was obtained by classifying the prediction dataset (Fig. 5d). The samples were distributed in two different clusters, both in the region of the control samples and in the area of PEA’s treated calves. The region of the plane in which sera from untreated calves are clustered is highlighted. By considering the prediction made by this model, it can be noted that 3 out of 17 treated samples resulted misclassified (false negatives), while only 1 out of 13 control samples was misclassified (false positive).

4 Discussion

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Due to the very short breeding cycle, veal calves represent the meat producing bovine category mostly subjected to illicit treatments with PEAs. Their residues can accumulate in meat and offals and potentially affect consumer's health. Body fluids, like serum, represent one of the most convenient samples to collect and investigate; serum characterization is a fundamental milestone to detect functional molecular biomarkers of illicit treatment.

Our study focuses on the low range proteome repertoire of sera in the peptide and proteins range detectable between 3 and 20 kDa. This specific range represents a serum fraction of great biological relevance since it encodes information on protein and peptides, proteolytic maturation, and post-translational modifications [20, 21]. The high analytical performances achieved with a MALDI-TOF-MS platform in this mass window allow a fine microcharacterization of molecular species. The study of the low molecular range was also suggested by the intent to investigate proteins in a range not detectable from other common proteomic techniques, such as 2-DE and SDS-PAGE. Moreover, the electrophoretic approach followed by MS analysis is laborious and time-consuming, and in SDS-PAGE the signal often results spurious, since a given zone could be an envelope of bands with closely spaced mass molecular values.

The choice of employing serum rather than plasma (EDTA or citrate), as suggested by HUPO, was driven by the absolute requirement to have an easy collectable sample in any farms. In particular, plasma collection was not pursued as it would require special care in avoiding erythrocyte hemolysis, which might return massive hemoglobin signal contamination in the MALDI-TOF-MS profile.

Despite the criticism and setbacks [22–25], MALDI-TOF technique has the potential for enormous impact on the field of biomarker discovery as it can offer a high-throughput analysis method applicable to complex biological mixtures. The analytical performances

should be validated as required by several international organizations regardless of the instrumental technique applied. In this light, it is necessary to perform a validation study both of the extraction strategies and of the MS technology employed before using any quantitative or semiquantitative platform to pursue an analytical molecular investigation. Nonetheless any open platform MS profiling approach is different from the traditionally accepted analytical measurement. In fact, critical parameters like LOD, LOQ, precision, or linearity have to be carefully evaluated when considering pure quantitative methods on specific and well-defined molecules but they become senseless in a semiquantitative analysis especially when dealing with biological macromolecules in complex mixtures and no information on target molecular species is available. For this concern, if parameters like resolution, mass accuracy, and repeatability are inside an acceptable range we can consider the method adequate to conduct such a proteomics investigation.

Our results concerning inter- and intraday repeatability confirm CV% lower than the limits for full analytical validation imposed by several international organizations, including FDA [25]. The high mass resolution (FWHM>1000) and accuracy (~100 ppm) allows to clearly identify specific protein products characteristic of calf sera. The overall evaluation of these analytical features plainly demonstrates how a fine tuned proteomic platform, based on MALDI-TOF-MS, can be employed for semiquantitative molecular biomarker investigations. The number and the quality of proteomic approaches aimed to investigate biological fluids and tissues is constantly increasing but to keep pace with the rapid development of new unknown substances used for growth promotion still urges the development of screening and confirmatory methods. These methods must be suitable to deal with the relatively large numbers of samples and to guarantee high reliability of results.

In our work, we undertook an initial annotation effort in order to set a reference molecular framework for further biomarker investigations. A number of reference MALDI-TOF-MS

peak signals have been identified by employing HPLC reverse phase fractionation [26] on semipreparative set-up. Such a fractionation scale of the intact protein level was conveniently performed with a relatively large quantity of serum. Several linear MALDI-TOF-MS peaks were observed in each fraction and, as expected, more than one peak might correspond to the same gene product. Moreover, considering the lack of complete genome sequence databases, signals annotations might suffer from the poor prediction of all the presumed maturation process and post-translational modifications.

Predicting the possible application of serum proteomic signature in the detection of illicit treatments, we compared the low molecular weight serum protein profile, of untreated calved to the one of calves exposed to different compounds. In particular, the protocols used for the treated groups mocked the most commonly used illegal treatment schedules. These include: the most effective natural estrogen (17 β E), one of the most active androgens (17 β T), and an association of moderately active anabolic agent (AND) and insulin-like sulfanylurea-derived synthetic agent (GLY). The collected data have been processed both by classical univariate statistical approach and by the more powerful multivariate approach. The latter is considered a valuable tool to discriminate among different classes and predict the behavior of new observations. We highlighted one putative protein fragment significantly different from the comparison of sera between treated and untreated animals. The normalized intensity of the signal at 6950 Th resulted significantly higher for every treatment group, thus, validating the relationship among altered metabolic pathways and serum MS profile. The signal corresponding to the β_2 GP-I, at 6950 Th, was not found in any of the samples from 4-month-old calves declared “untreated” by the farmer (C_{UT}). In fact, the sampled animals did not show overt clinical signs of treatment, like abnormal muscle hypertrophy, nervousness, polypnea, or increased heart frequency.

Multivariate PLS-DA analysis also confirmed the high discriminatory load of this molecular species. Nevertheless, the separation in two large classes of treated and untreated animals accomplished with the PLS-DA analysis result in a predictive power of only 69.4%, as explained by Q^2Y value, which is the measure of how well the model predicts new data. Such a lower discrimination power might be due to the characteristic of multivariate approaches since they might be influenced by every slight change in the data profile. A biological sample-to-sample variation was expected considering how intrinsic differences due to genotypic and phenotypic characteristic, health condition, *etc.* could influence the overall protein profile and, consequently, the serum low range proteome.

The three protein signals associated to a specific type of PEAs treatment in the univariate analysis, did not find confirmation by global PLS-DA. Given the sample number required to consolidate a multivariate data analysis we could not apply this procedure to single class treatments. The importance of the fragment of β_2 GP-I in the construction of the model can be noted in the loading plot (Fig. 5c).

In order to detect illegal treatments, the most convenient criterion is the one that gives 100% sensitivity, thus excluding false negative cases, while doubtful cases undergo further analysis by classical toxicological confirmation methods. Our results highlight the role of sushi 3–4 fragment of the β_2 GP-I (apolipoprotein H) as a potential reproducible marker of PEAs treatment. The ROC curve for this signal demonstrates high sensitivity and specificity: choosing 100% sensitivity results in 74.4% specificity.

The bovine β_2 GP-I is a plasma glycoprotein, identical to the human counterpart for the 81% of the sequence and circulating as a free protein or associated to lipoproteins. Although quite a lot is known about the structure of β_2 GP-I, its biological function is still unclear. It is known that β_2 GP-I can bind to negatively charged substances, such as DNA, heparin and negatively charged phospholipids *in vivo* [27], but the meaning of such interaction remains to

be elucidated. It is also known that β_2 GP-I may serve as a major factor in clearing the plasma liposome [28] as well as an anticoagulant in blood [29]. The lipid-binding and transportation functions are considered as a basic mechanism related to its physiological and pathogenic functions. Its first isolation from the bovine serum, dating back to the 1990s, was due to its strong inhibiting effect in the incorporation of thymidine into fetal calf erythroid cells [30]. More recently, its presence as a low abundant protein was highlighted in colostrum, suggesting an eventual special physiologic relevance to the health and development of calves in lactation [31]. However, apart from his anticoagulant action, its role, as well as the one of the human form, is still uncertain. Due to the recent increase in the use of anabolic steroids among athletes, the modulation of the apolipoproteins upon administration of anabolic steroids has already been established [32].

The interpretation of the observed increase in serum β_2 GP-I expression in calves from all experimentally treated groups is difficult and goes beyond the scopes of our study.

Interestingly, another truncated form of a bovine apolipoprotein Apo-A1, was detected in plasma samples from calves experimentally administered a combination of boldenone and boldione, two androgenic sexual steroids [33]. Taken together, all these data suggest a linkage between administration of anabolic steroids and/or hormones and apolipoprotein expression. However, more investigations should still be conducted to understand the underlying mechanism and the actual existence of a direct link.

The possibility to export such platform to a territory monitoring plan and to extend this approach to other food safety issues, such as dioxin food contamination, can be considered one main goal of this state-of-the-art study in the development of Proteomics functional biomarker. The high throughput features of the proposed MALDI-TOF-MS analysis combined to a well defined and simplified classification model provide a powerful tool for the detection of PEAs usage.

This paper describes a step forward to the detection of biomarkers of illicit treatment in veal calves serum in the low molecular weight range (3–20 kDa) through a postgenomic approach. Although further research is warranted to corroborate the data obtained in the present study, serum proteomics seems to be a promising technology to enhance the food safety controls in living animals.

We are grateful to Dr. Luisa Pieroni for her critical discussion and corrections. This work was supported by the Piemonte Regional Grant CIPE “Identification of molecular biomarkers of illicit treatment on calves by proteomics and genomic investigations.”

The authors have declared no conflict of interest.

((Funded by:

- Piemonte Regional Grant CIP))

5 References

- [1] De Brabander, H. F., Poelmans, S., Schilt, R., Stephany, R. W. *et al.*, Presence and metabolism of the anabolic steroid boldenone in various animal species: A review. *Food Addit. Contam.* 2004, 21, 515–525.
- [2] Courtheyn, D., Le Bizec, B., Brambilla, G., De Brabander, H. F. *et al.*, Recent developments in the use and abuse of growth promoters. *Anal. Chim. Acta* 2002, 473, 71–82.
- [3] Kuiper, H. A., Noordam, M. Y., van Dooren-Flipsen, M. M., Schilt, R. *et al.*, Illegal use of beta-adrenergic agonists: European Community. *J. Anim. Sci.* 1998, 76, 195–207.
- [4] Gardini, G., Del Boccio, P., Colombatto, S., Testore, G. *et al.*, Proteomic investigation in the detection of the illicit treatment of calves with growth-promoting agents. *Proteomics* 2006, 6, 2813–2822.

- [5] Ebert, M. P., Korc, M., Malfertheiner, P., Röcken, C., Advances, challenges, and limitations in serum-proteome-based cancer diagnosis. *J. Proteome Res.* 2006, 5, 19–25.
- [6] Villanueva, J., Philip, J., Chaparro, C. A., Li, Y. *et al.*, Correcting common errors in identifying cancer-specific serum peptide signatures. *J. Proteome Res.* 2005, 4, 1060–1072.
- [7] Matharoo-Ball, B., Miles, A. K., Creaser, C. S., Ball, G. *et al.*, Serum biomarker profiling in cancer studies: A question of standardisation? *Vet. Comp. Oncol.* 2008, 6, 224–247.
- [8] Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D. *et al.*, Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 1996, 13, 19–50.
- [9] D'Ambrosio, C., Arena, S., Talamo, F., Ledda, L. *et al.*, Comparative proteomic analysis of mammalian animal tissues and body fluids: Bovine proteome database. *J. Chromatogr. B* 2005, 815, 157–168.
- [10] Ardekani, A. M., Liotta, L. A., Petricoin, E. F., III, Clinical potential of proteomics in the diagnosis of ovarian cancer. *Expert. Rev. Mol. Diagn.* 2002, 2, 312–320.
- [11] Sanchez, J. C., Appel, R. D., Golaz, O., Pasquali, C. *et al.*, Inside SWISS-2DPAGE database. *Electrophoresis* 1995, 16, 1131–1151.
- [12] Sanchez, J. C., Chiappe, D., Converset, V., Hoogland, C. *et al.*, The mouse SWISS-2D PAGE database: A tool for proteomics study of diabetes and obesity. *Proteomics* 2001, 1, 136–163.
- [13] Kaji, H., Isobe, T., Protein database of *Caenorhabditis elegans*. *J. Chromatogr. B* 2003, 787, 91–99.
- [14] Wait, R., Miller, I., Eberini, I., Cairoli, F. *et al.*, Strategies for proteomics with incompletely characterized genomes: The proteome of *Bos taurus* serum. *Electrophoresis* 2002, 23, 3418–3427.

- [15] Talamo, F., D'Ambrosio, C., Arena, S., Del Vecchio, P. *et al.*, Proteins from bovine tissues and biological fluids: Defining a reference electrophoresis map for liver, kidney, muscle, plasma and red blood cells. *Proteomics* 2003, 3, 440–460.
- [16] Mantini, D., Petrucci, F., Pieragostino, D., Del Boccio, P. *et al.*, LIMPIC: A computational method for the separation of protein MALDI-TOF-MS signals from noise. *BMC Bioinformatics* 2007, 26, 101.
- [17] Urbani, A., Gemeinhardt, S., Warne, A., Saraste, M., Properties of the detergent solubilised cytochrome c oxidase (cytochrome cbb(3)) purified from *Pseudomonas stutzeri*. *FEBS Lett.* 2001, 508, 29–35.
- [18] Biroccio, A., Del Boccio, P., Panella, M., Bernardini, S. *et al.*, Differential post-translational modifications of transthyretin in Alzheimer's disease: A study of the cerebral spinal fluid. *Proteomics* 2006, 6, 2305–2313.
- [19] Wang, Y., Balgley, B. M., Rudnick, P. A., Lee, C. S., Effects of chromatography conditions on intact protein separations for top-down proteomics. *J. Chromatogr. A* 2005, 1073, 35–41.
- [20] Yi, J., Kim, C., Gelfand, C. A., Inhibition of intrinsic proteolytic activities moderates preanalytical variability and instability of human plasma. *J. Proteome Res.* 2007, 6, 1768–1781.
- [21] Tammen, H., Schulte, I., Hess, R., Menzel, C. *et al.*, Peptidomic analysis of human blood specimens: Comparison between plasma specimens and serum by differential peptide display. *Proteomics* 2005, 5, 3414–3422.
- [22] Baggerly, K. A., Morris, J. S., Edmonson, S. R., Coombes, K. R., Signal in noise: Evaluating reported reproducibility of serum proteomic tests for ovarian cancer. *J. Natl. Cancer Inst.* 2005, 97, 307–309.

- [23] Diamandis, E. P., Analysis of serum proteomic patterns for early cancer diagnosis: Drawing attention to potential problems. *J. Natl. Cancer Inst.* 2004, 96, 353–356.
- [24] Ransohoff, D. F., Lessons from controversy: Ovarian cancer screening and serum proteomics. *J. Natl. Cancer Inst.* 2005, 97, 315–319.
- [25] Guidance for Industry, Bioanalytical Method Validation, U.S. Food and Drug Administration, Center for Drug Evaluation and Research 2001 (CDER).
- [26] Lopez, M. F., Mikulskis, A., Kuzdzal, S., Bennett, D. A. *et al.*, High-resolution serum proteomic profiling of Alzheimer disease samples reveals disease-specific, carrier-protein-bound mass signatures. *Clin. Chem.* 2005, 51, 1946–1954.
- [27] Kochl, S., Fresser, F., Lobentanz, E., Baier, G. *et al.*, Novel interaction of Apolipoprotein (a) with Beta 2 glycoprotein I mediated by the Kringle IV domain. *Blood* 1997, 90, 1482–1489.
- [28] Chonn, A., Semple, S. C., Cullis, P. R., Beta 2 glycoprotein I is a major protein associated with very rapidly cleared liposomes *in vivo*, suggesting a significant role in the immune clearance of “nonself” particles. *J. Bio. Chem.* 1995, 270, 25845–25849.
- [29] Brighton, T. A., Hogg, P. J., Dai, Y. P., Murray B. H. *et al.*, Beta 2-glycoprotein I in thrombosis: Evidence for a role as a natural anticoagulant. *Br. J. Haematol.* 1996, 93, 185–194.
- [30] Li, Q. G., Blacher, R., Esch, F., Congote, L. F., Isolation from fetal bovine serum of an apolipoprotein-H-like protein which inhibits thymidine incorporation in fetal calf erythroid cells. *Biochem. J.* 1990, 267, 261–264.
- [31] Yamada, M., Murakami, K., Wallingford, J. C., Yuki, Y., Identification of low-abundance proteins of bovine colostrum and mature milk using two-dimensional electrophoresis followed by microsequencing and mass spectrometry. *Electrophoresis* 2002, 23, 1153–1160.

[32] Hartgens, F., Kuipers, H., Effects of androgenic-anabolic steroids in athletes. *Sports Med.* 2004, 34, 513–554.

[33] Draisci, R., Montesissa, C., Santamaria, B., D'Ambrosio, C. *et al.*, Integrated analytical approach in veal calves administered the anabolic androgenic steroids boldenone and boldione: Urine and plasma kinetic profile and changes in plasma protein expression. *Proteomics* 2007, 7, 3184–3193.

Figure 1. Average MALDI-TOF spectra of bovine sera samples after different types of ZipTip extraction. (a) C₄; (b) C₁₈; (c) SCX; (d) MC. As expected, ZipTip C₁₈ was best suited to extract a high number of species in the low mass range whereas ZipTip C₄ resulted in a more homogeneous extraction profile. SCX and MC were clearly inappropriate for our purpose. IS signal (heart horse myoglobin) was clearly detectable only in C₄ ZipTip spectrum.

Figure 2. Average spectrum with identified signals and corresponding gel/stack view of not-empty collected fractions. Numbered signals in the spectrum correspond to the proteins identified by LC-MS/MS experiments in Table 2.

Figure 3. Superimposed average spectra from control (red line) and treated groups (green, blue, yellow lines), processed by smoothing and baseline subtraction. Beside each signal is shown the Box and Whisker representation of normalized single spectrum signal intensity. Panel (e) shows the IS signal distribution.

Figure 4. (a) ROC curve for the normalized signal intensity of the peak at 6950 Th; dotted lines represent the confidence limits at 95%. (b) Box and Whisker graph for the normalized signal intensity of the peak at 6950 Th in the different groups.

Figure 5. (a) PCA on all samples return a nine PCs model with an explained variance of 76.1%; in the distribution of scores in the PCs plane (PC1 vs. PC2) three regions of clustering are visible and highlighted: the blue region refers to the control group (K and K_2006 corresponding to C_2 and C_{UT} , respectively), the yellow one to the overlapping area and the red region to the treated groups (D, E, and F corresponding to GLY + AND, 17 β E and 17 β T treatments' groups, respectively). (b,c) Score scatter plot and loadings scatter plot of the calculated PLS-DA model. Black and red items represent control and treated animals, respectively. The loading related to the signal at 6950 Th is red circled. (d) Distribution of score residues for the prediction set (gray triangles) in the PLS-DA components plane. The region of samples from untreated animals is gray circled. Three out of 14 misclassified treated samples (false negatives) are green circled; one out of 13 misclassified control sample (false positive) is red circled. (e) Plot of the importance of variables on the projection (VIP value) for the PLS-DA analysis.

Table 1. Average number of signals extracted by each ZipTip type (S/N>10)

ZipTip	Sample	Sample	Sample	Sample	Sample	Sample	Average	SD
type	1	2	3	4	5	6		
C ₄	81	90	76	76	79	74	79.3	5.8
C ₁₈	59	77	80	58	79	106	76.5	17.5
MC	11	4	8	7	6	8	7.3	2.3
SCX	6	5	7	8	6	9	6.8	1.5

Table 2. List of identified proteins by nano-LC-nano-ESI-Q-TOF-MS/MS fragmentation

experiments

N	Proteins	Mass value	Expected mass of	Mass of un
o.		(Th)	fragments (Da)	precursor
1	Apolipoprotein C-II short form	7945.0	7944.9	86
2	Apolipoprotein C-III	8218.7	8217.0	10 6
3	Apolipoprotein A-II short form 1–76	8566.4	8566.6	11 2
4	β-2-Microglobulin	11 635.6	11 635.2	13 6
5	Hemoglobin subunit alpha	15 055.0	15 053.2	15 0
6	Transthyretin	13 559.0	13 557.3	15 7
7	Hemoglobin subunit beta	15 957.0	15 954.4	15 9
8	Complement component 3 res38–113	8776.8	8777.1	17 0
9	α-1-Acid glycoprotein	n.a.	21 253.0	21 7
1	Coagulation factor XIII A1 subunit [M + 2H ⁺]	11 369.6	22 744.5	22 7
1	Hypothetical protein VIIa protein	n.a.	–	24 6
1	Apolipoprotein A-I [M + 2H ⁺]	13 775.5	27 549.1	30 2
1	β-2-Glycoprotein 1 (Apolipoprotein H) sushi 3–4	6949.0	6949.0	38 2
1	α-2-HS-glycoprotein	n.a.	13 253.7–12 531.5	38 4
1	Hypothetical protein Apolipoprotein A-IV	n.a.	41 032.2	43 0
1	Pigment epithelium-derived factor precursor	n.a.	44 056.3	46 3
1	α-1-Antiproteinase	n.a.	43 693.9	46 4
1	Endopin 2C	n.a.	–	46 8
1	Glial fibrillary acidic protein, Coil 1B	11 835.0	11 832.1	49 4
2	LOC507464 protein (Fragment)	n.a.	–	51 1
2	Antithrombin-III	n.a.	–	52 3
2	α-1-B glycoprotein	n.a.	–	53 5
2	Fibrinogen alpha chain	n.a.	–	67 0
2	Kininogen-1, cystatin 2, cystatin 3	n.a.	13 876.5–13 285.2	68 8
2	Coagulation factor II	n.a.	–	70 4
2	Serotransferrin (transferrin) (siderophilin)	n.a.	–	77 7
2	Complement factor H sushi 6 sushi 11sushi 15	6995.0	6997.8	140
		6737.0	6737.8	
		7651.0	7647.5	
2	Complement C3 C3a anaphylatoxin	8931.0	8933.2	187

n.a., not assigned.

Table 3. Comparison of SD and mass error in ppm ($n = 20$) for homologous peaks in different samples for the acquisitions on the Reflex IV and Ultraflex III mass spectrometers

	Average mass	RSD (Th)	Mass error (ppm)
	(Th)		
Reflex IV	7945.2	1.25	158

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Ultraflex III	7945.7	0.55	70
Reflex IV	8217.7	0.54	66
Ultraflex III	8217.9	0.53	65
Reflex IV	8933.0	1.11	125
Ultraflex III	8933.5	0.90	101
Reflex IV	11 633.4	1.02	87
Ultraflex III	11 633.8	0.35	30
Reflex IV	13 558.00	0.83	61
Ultraflex III	13 557.84	0.54	40
Reflex IV	15 054.5	0.27	18
Ultraflex III	15 053.7	0.49	32
Reflex IV	<i>*16 952.27</i>	<i>0.03</i>	<i>1.8</i>
Ultraflex III	<i>*16 952.25</i>	<i>0.05</i>	<i>2.9</i>

Average mass of IS (heart horse myoglobin theoretical mass = 16 952.3).

α and β hemoglobin singly and double charged were used to perform quadratic calibration of TOF. Additional internal calibration was performed on the mass of IS. **Please explain the significances of italic and asterisk used in the table**;

Table 4. Inter- and intraday repeatability

<i>m/z</i>	Intensity	SD	RSD
Repeatability intraday (<i>n</i> = 18)			
6 820	180.13	18.85	10.46
6 839	249.08	33.51	13.45
6 883	116.81	18.06	15.46
12 634	323.67	51.00	15.76
12 792	215.31	31.68	14.71
13 413	241.54	36.12	14.96
13 558	7117.17	1094.73	15.38
13 639	1142.42	158.29	13.86

13 677	1380.93	147.33	10.67	
13 735	444.84	67.08	15.08	
13 765	567.07	84.59	14.92	
17 159	61.16	8.46	13.82	
Repeatability interday (<i>n</i> = 18)				
Twelve signals present in at least 75% of samples were chosen to evaluate repeatability.	6 820	182.08	22.20	12.19
Average, SD and RSD were calculated on the normalized intensity for each signal.	6 839	267.51	37.93	14.18
	6 883	127.23	18.99	14.93
	12 634	321.50	41.16	12.80
Table 5. <i>p</i> -Values for Kruskal–Wallis one way analysis of variance on ranks	12 792	227.09	26.73	11.77
	13 413	226.71	31.65	13.96
	13 558	7109.82	854.39	12.02
	13 639	1171.92	153.55	13.10
	13 677	1439.08	172.98	12.02
	13 735	424.90	63.13	14.86
	13 765	548.81	74.06	13.49
	17 159	56.14	8.32	14.81
<i>m/z</i>	<i>C</i> ₂ versus GLY + AND	<i>C</i> ₂ versus 17βE	<i>C</i> ₂ versus 17βT	
6950	<0.05	<0.05	<0.05	

7451	<0.05	N.S.	<0.05
8733	N.S.	N.S.	<0.05
12 590	N.S.	N.S.	N.S.
14 900	<0.05	N.S.	N.S.

Multiple comparisons *versus* control group (Dunn’s method).

N.S., not significant.

Table 6. PCA-X; observations (N) = 58; Variables (K) = 85 ($X = 85$, $Y = 0$)

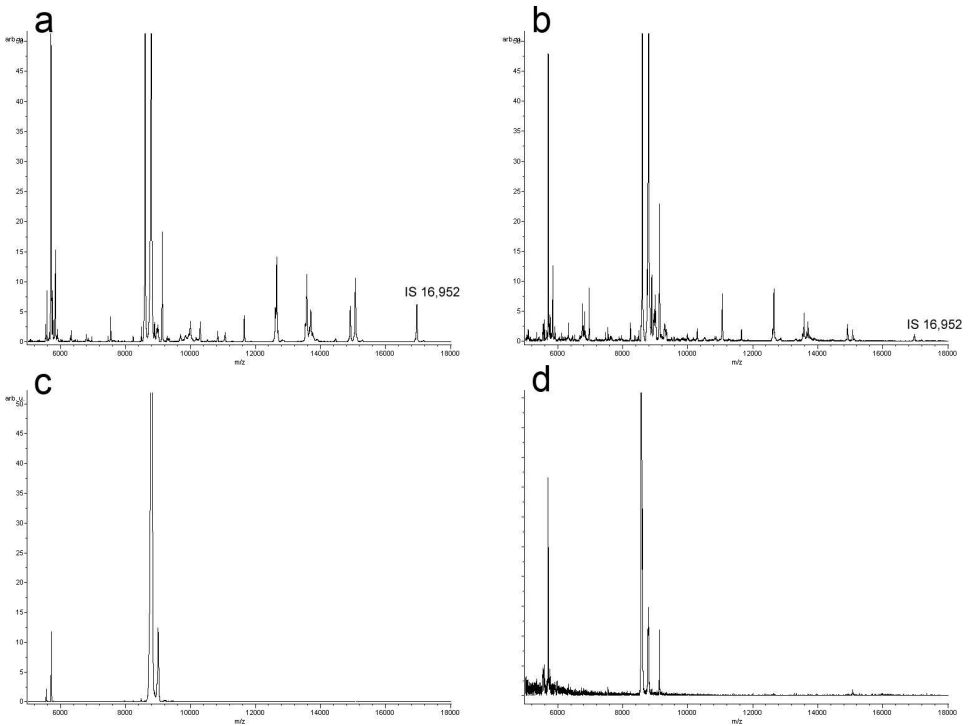
PC	Fraction of sum of squares (SS) of the entire X ’s explained by the current component		Fraction of the total variation of Y ’s that can be predicted by the current component	
	R^2X	$R^2X(\text{cum})$	Q^2	$Q^2(\text{cum})$
1	0.224	0.224	0.117	0.117
2	0.151	0.374	0.0119	0.127
3	0.103	0.478	−0.0314	0.0998
4	0.0709	0.548	−0.0246	0.0776
5	0.0612	0.61	0.0565	0.13
6	0.0505	0.66	−0.0303	0.103
7	0.038	0.698	−0.0306	0.0759
8	0.0341	0.732	−0.0149	0.0622
9	0.0283	0.761	−0.0487	0.0165

The PCA model returned nine significative PCs. The rate of the total variance explained by each PC is reported in the R^2X column; the cumulative value is reported in the $R^2X(\text{cum})$ column. The total explained variance of the model is 76.1%.

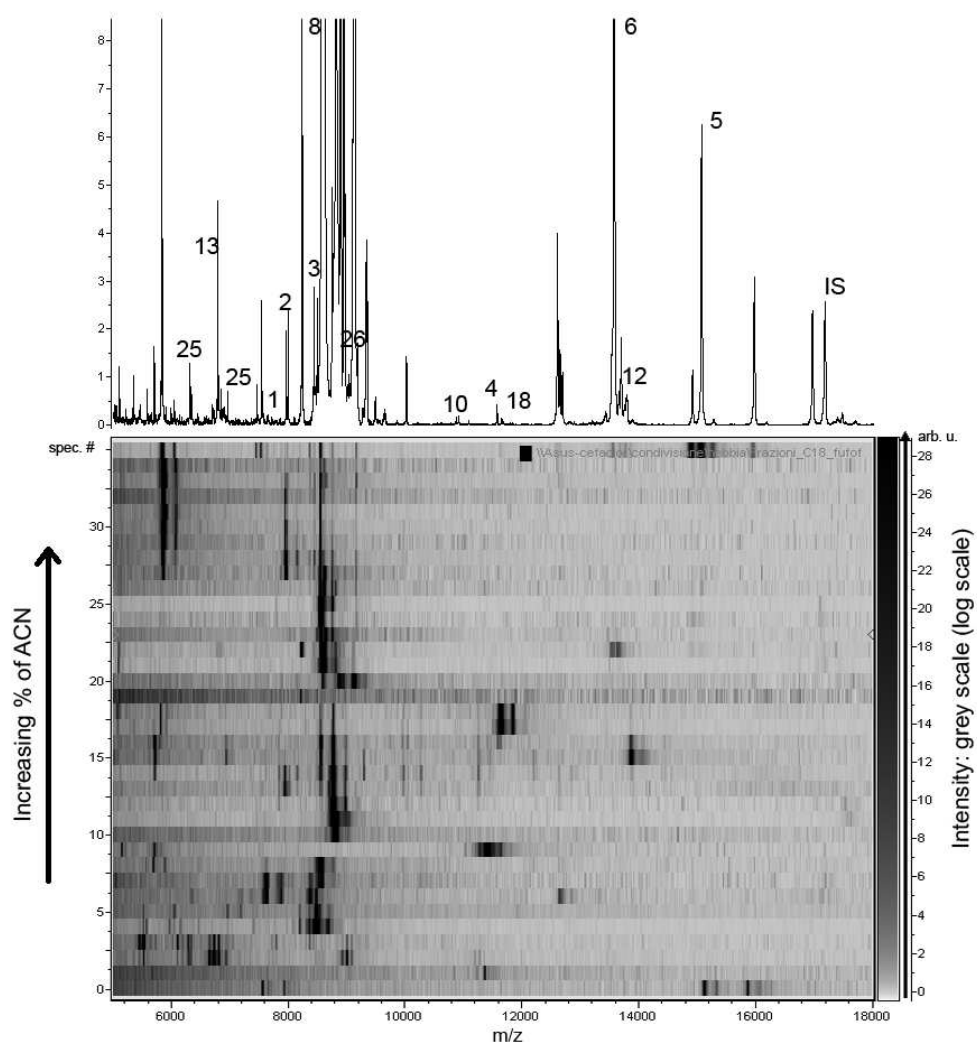
Table 7. PLS-DA; observations (N) = 30; variables (K) = 87 (X = 85, Y = 2)

PC	Fraction of SS of the entire X 's explained by the current component			Fraction of SS of all the Y 's explained by the current component		Fraction of the total variance of Y 's that can be predicted by the current component	
	R^2X	$R^2X(\text{cum})$	Eigenvalues	R^2Y	$R^2Y(\text{cum})$	Q^2	$Q^2(\text{cum})$
1	0.229	0.229	6.88	0.529	0.529	0.445	0.445
2	0.081	0.31	2.43	0.253	0.781	0.135	0.576
3	0.0888	0.399	2.66	0.112	0.894	0.43	0.606
4	0.11	0.509	3.31	0.0603	0.954	0.359	0.665
5	0.0663	0.575	1.99	0.0184	0.973	0.0955	0.694

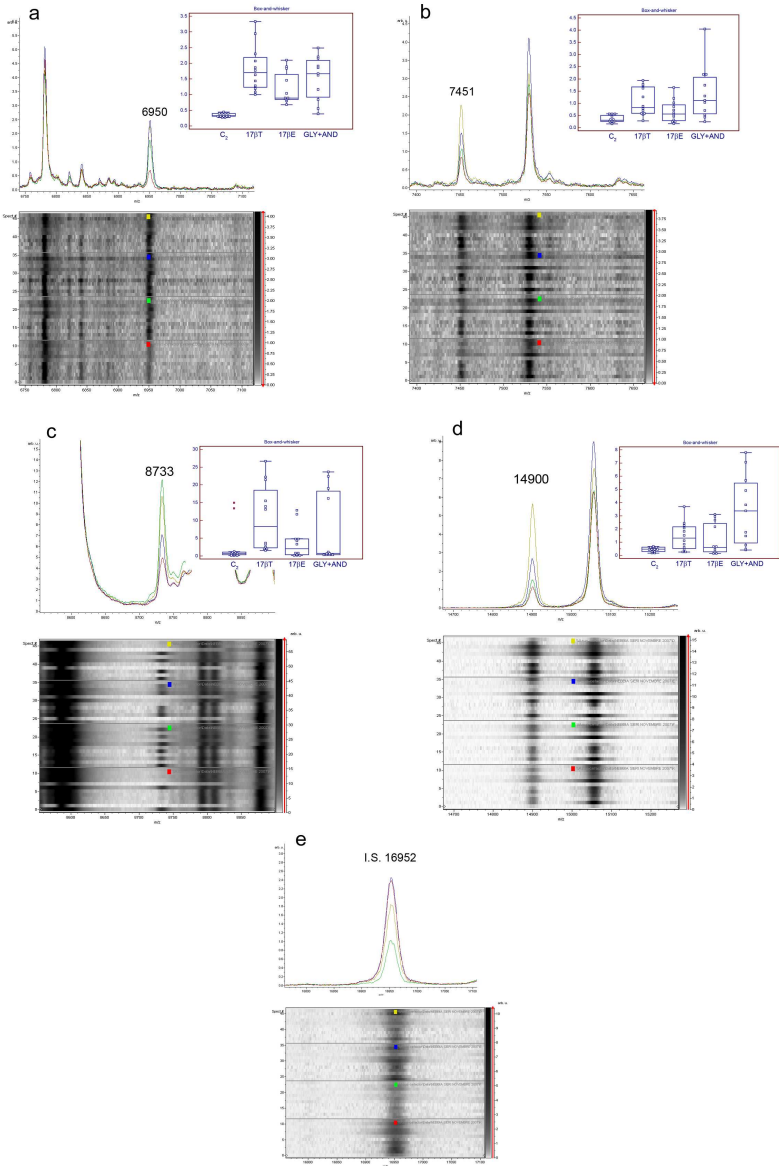
The PLS-DA model returned five significative PCs. The cumulative predictive power is reported for each PC in the $Q^2(\text{cum})$ column. The mathematical model has a total predictive power of 69.4%.



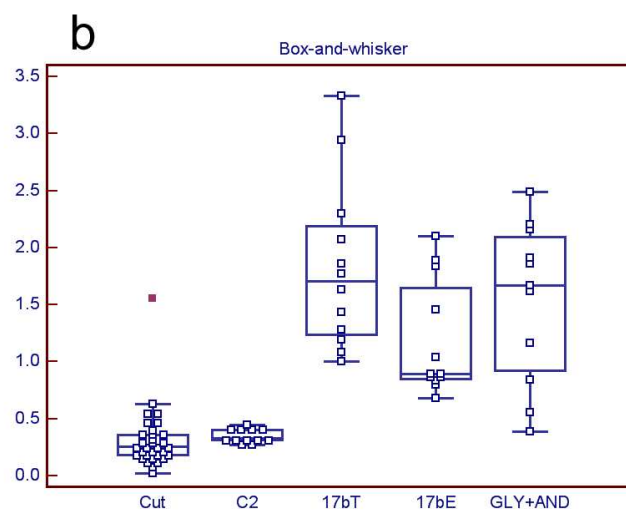
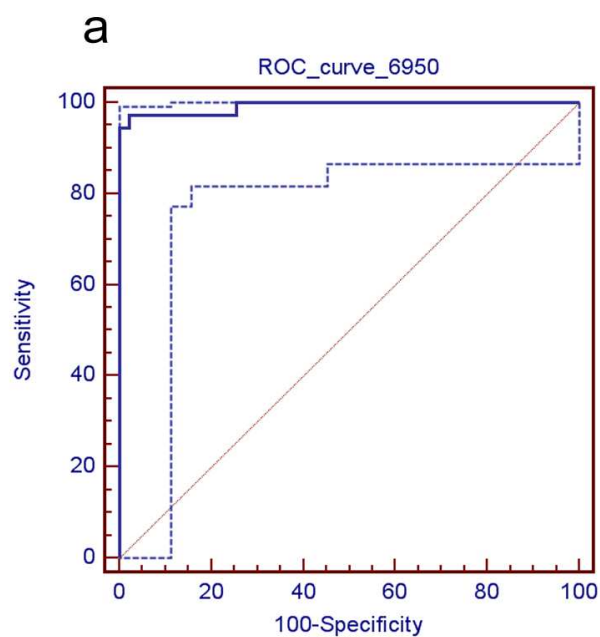
136x102mm (300 x 300 DPI)



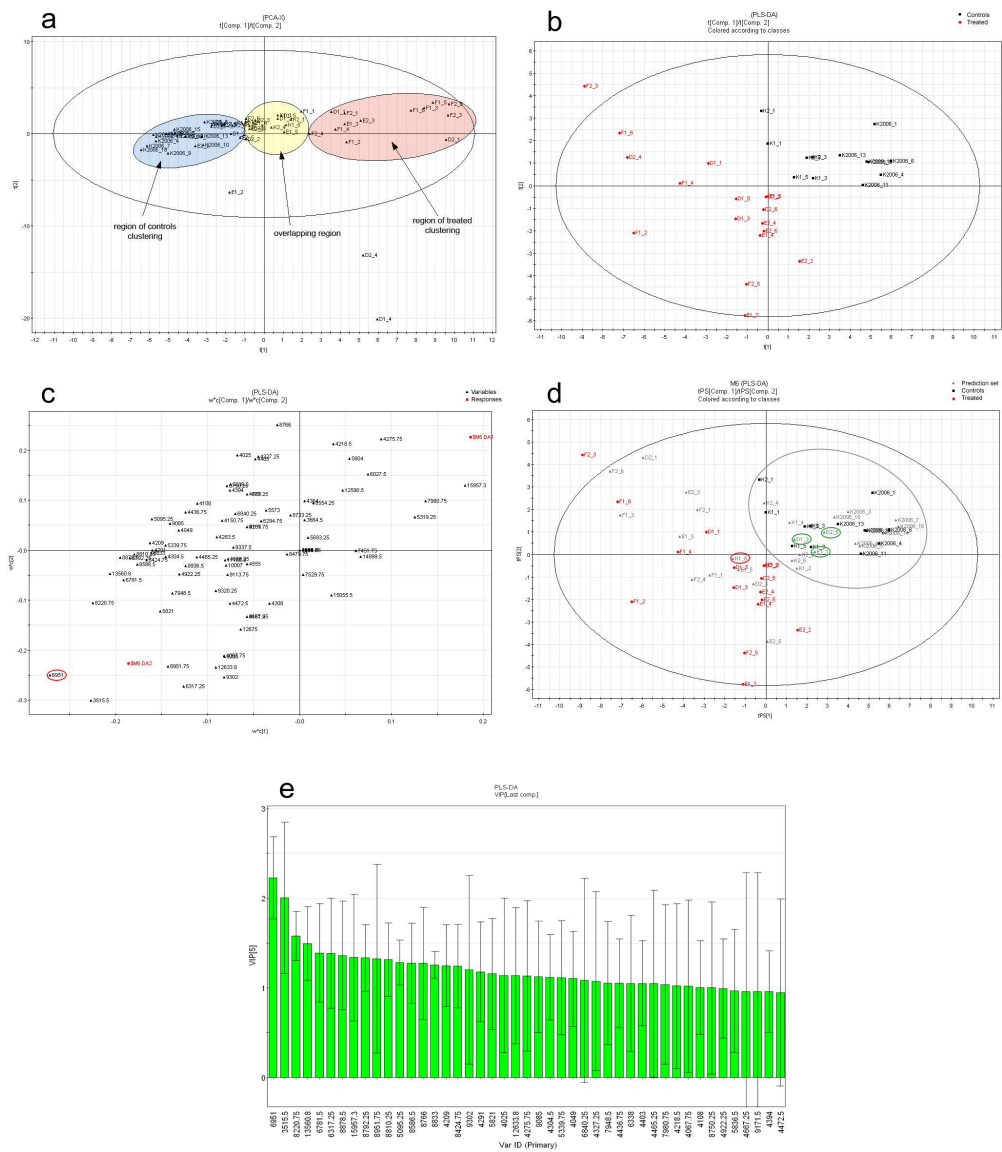
75x79mm (300 x 300 DPI)



196x294mm (300 x 300 DPI)



73x121mm (300 x 300 DPI)



173x200mm (300 x 300 DPI)