Human tear peptide/protein profiling study of ocular surface diseases by SPE-MALDI-TOF mass spectrometry analyses

Nerea Gonzáleza,1, Ibon Iloroa,1, Javier Soriaa, Juan A. Duranc,d, Alaitz Santamaríae, Felix Elortza b, Tatiana Suáreza,∗
a Biofimlik Applied Research, Science and Technology Park of Bizkaia, Ed. 800, 2nd Floor, E-48160 Derio, Bizkaia, Spain
b Proteomics Platform, CIC bioGUNE, CIBERehd, ProteoRed-ISICIII, Science and Technology Park of Bizkaia, 48160 Derio, Bizkaia, Spain
c Instituto Clínico Quirúrgico de Oftalmología (ICQO), Virgen de Begoña N° 34, E-48006 Bilbao, Bizkaia, Spain
d Department of Ophthalmology, School of Medicine, University of Basque Country, E-48940 Leioa, Bizkaia, Spain
e Department of Ophthalmology, Cruces Hospital, Plaza Cruces 12, E-48903 Baracaldo, Bizkaia, Spain

A R T I C L E   I N F O
Article history:
Received 1 August 2013
Received in revised form 31 January 2014
Accepted 28 February 2014
Available online 21 March 2014

Keywords:
Dry eye
Meibomian gland dysfunction
Tears
MALDI-TOF
Profiling
Warping
Peptidomics

A B S T R A C T
The goal of this study was to investigate differences in tear peptidome/proteome profiles of aqueous-deficient dry eye, Meibomian gland dysfunction and control individuals. Tears from 93 individuals were collected by capillary and subjected to solid-phase extraction followed by MALDI-TOF for profiling analysis.

Obtained spectra were aligned by variable penalty dynamic time warping (VPdtw) and the resulting data analyzed using multivariate statistics. Comparative analyses revealed good performance of VPdtw and a high discrimination of groups with a correct assignment of 89.3% using twelve informative peaks. SDS–PAGE followed by MALDI-TOF/TOF analysis allowed identification of lipocalin-1 as a biomarker candidate.

© 2014 The Authors. Published by Elsevier B.V. on behalf of European Proteomics Association (EuPA). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction
Dry eye (DE) is understood as a complex syndrome which is frequently associated with other symptomatic ocular surface disorders (SOSD), such as allergic eye disease, chronic non-dry eye conjunctivitis, infective conjunctivitis, keratoconjunctivitis and Meibomian gland dysfunction (MGD) among others [1]. Frequently, some of these diseases may coexist and interact
in the same patient. MGD and these other conditions may cause or contribute to dry eye, but exist in their own right as either symptomatic or asymptomatic disorders. In general, dry eye is classified in two subtypes, aqueous-deficient (ADDE), in which a failure of lacrimal secretion occurs, and evaporative dry eye (EDE), caused by an excessive water loss from the exposed ocular surface [2], with MGD being the main cause of this last type of dry eye [3]. MGD is a subtype of posterior blepharitis, which is defined as chronic, diffuse abnormality of the Meibomian glands, commonly characterized by terminal duct obstruction and/or qualitative/quantitative changes in glandular secretion. The alterations of the normal lipids composition in Meibomian gland secretions induce abnormalities in tear film composition and function thus inducing tear film evaporation and hyperosmolarity, the cause of evaporative dry eye [4]. In fact, MGD is likely to be the most common cause of evaporative dry eye [2], and is responsible for approximately 40–60% of dry eye cases [5]. As a final point, this condition may result in alteration of the tear film, symptoms of eye irritation, clinically apparent inflammation, and ocular surface disease [6].

The physiological process involved in both conditions is complex and represents the interaction of multiple mechanisms such as tear film hyperosmolarity, instability, and activation of an inflammatory cascade that releases inflammatory mediators into the tears and this, in turn, can damage the ocular surface epithelium [2,7]. MGD and other subtypes of dry eye diseases compromise tear film stability and ocular surface welfare, and share signs and symptoms in patients. The dry eye workshop (DWES) recommends that any diagnostic protocol should include a classification scheme that addresses the differential diagnosis of dry eye from other subtypes and/or OSDS conditions. To this regard a classification of dry eye for practical clinical use has been also proposed based on three parameters such as etiopathogenesis, glands and tissues affected and disease severity [8].

Nevertheless, similar signs and symptoms frequently make a specific discrimination between dry eye subtypes complex, and impose a difficulty in the selection of patients in basic and applied clinical research studies as well as clinical trials for evaluating the response to treatments. The correct interpretation of the results obtained in this kind of studies may often be hampered by outliers, or individuals whose behaviour is different from the rest of a specific group in which they have been clinically classified.

Clinical proteomics is a promising practice with several features that include high throughput, high sensitivity and low cost. Of all the proteomic approaches used for this purpose, peptide and protein profiling has become a technique very close to clinical application [9]. Peptide/protein profiling is based on the detection of discriminatory patterns that enable classifying and distinguishing between different study populations (e.g., patients versus healthy controls) and is currently being used as a clinical tool for the diagnosis of ovarian cancer [9], microbial typing [10] or MS imaging [11]. The peptide/protein pattern is comprised of combinations of peaks in the mass spectra plot that are defined by the intensity of each signal on the y-axis, and the mass-to-charge ratio on the x-axis. The changes in peptide/protein levels can be investigated by looking for a subset of peaks on the spectrum that vary in different study populations or at different stages of a disease.

Previous tear film proteomic studies based on mass spectrometry (MS) have been useful in mapping the tear protein profiles of patients and normal individuals. In tear fluid, the first approaches have been done by surface enhanced laser desorption/ionization (SELDI) mass spectrometry, combining the selective retention capacity of peptides/proteins by different surfaces and mass spectrometry, and exploring the diagnostic potential of this technique in patients with Sjögren’s syndrome [12], dry eye patients [13], and the effect of contact lenses on tear protein composition [14].

Given the limitations of SELDI-TOF for analytical performance [15–17], the use of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) has been increasing due to higher resolution and sensitivity and the possibility of coupling a solid phase extraction (SPE) protocol to enrich complex fluids in a certain type of analyte [18].

In this regard, MALDI-TOF profiling coupled with SPE has been used to characterize low molecular weight (LMW) proteins in the diurnal variation of tears [19]. More recently, our group published a study on the inter-day and inter-individual variability of normal tears and standardization of tear profiling based on SPE-MALDI-TOF [20].

In this type of studies, however, instrument precision can be reduced by variations in the triggering time from spectrum to spectrum, small variations in the accelerating voltage, or simply by random errors [21]. Therefore, in order to extract the maximum biologically meaningful information from a group of samples, aligning the spectra to ensure that features occur at the same time position in all spectra being analyzed is recommended [22–24]. This is especially important in multicentric studies where different MALDI-TOF spectrometers equipped with varying resolutions might be used and consequently, the application of correction of spectra phase variations enhances the reliability of results.

This study investigates the differences in tear peptide/LMW protein profiles between three groups, dry eye, MGD associated with DE and normal subjects in combination with multivariate statistical analysis.

A variable penalty dynamic time-warping (VPdtw) based alignment algorithm was implemented before using the msProcess R package for mass spectra data pre-processing in order to minimize bias due to phase variation. The comparison of the obtained profiles by different multivariate statistical methods enabled identification of a subset of specific peaks which allow differentiation of groups. The method presented here constitutes an objective tool that contributes to correct discrimination of pathologies in cases with confusing or insufficient signs; it enables the identification of potential outliers in clinical trials, and the monitoring of response to different treatments.

2. Materials and methods

2.1. Patients

A prospective case-controlled study was carried out, in which 93 patients were enrolled, 29 with dry eye (DE), 27 suffering
Meibomian gland dysfunction (MGD) and 37 healthy control (CT) subjects. All patients and normal subjects were recruited from the Cornea Unit at the Instituto Clínico Quirúrgico de Oftalmología (Bilbao, Bizkaia) and the Ophthalmology Department at Cruces Hospital (Baracaldo, Bizkaia). The diagnosis was based on clinical examinations including the Schirmer I test with anaesthesia to measure only the basal secretion, slit lamp examination of the lid margin and Meibomian glands, fluorescein staining, and subjective symptoms. Each patient answered a modified “National Eye Institute Visual Functioning Questionnaire 25” (VFQ-25) which includes some statements about problems involving their vision or feeling. Patients were classified as having MGD if they presented symptoms including eyelid inflammation, and a Schirmer I test result of >5 mm/5 min and alteration in Meibomian glands. In the control group, healthy subjects were recruited who were not suffering from any ocular disease (no allergic or atopic history), and presented Schirmer I test values of >5 mm/5 min, no corneal fluorescein staining or sensations of discomfort, and no evident eyelid inflammation. The exclusion criteria included the presence or history of any systemic or ocular disorder or condition (including ocular surgery, trauma and disease) and patients presenting Sjögren’s syndrome. Contact lens users were also excluded to avoid any possible interference with the interpretation of the results.

2.2. Tear sample collection

Tear samples were collected with 10 μL glass microcapillaries (Blaubrand; Wertheim, Germany), by placing the capillary close to the temporal lid margin of the eye and taking special care not to touch the conjunctiva. Written informed consent was obtained in accordance with the declaration of Helsinki on Biomedical Research Involving Human Subjects. The study was approved by the local ethics committee.

Tear samples were always collected in the morning, placed in 1.5 mL Protein LoBind Tubes (Eppendorf; Germany), centrifuged at 4000 × g for 1 min at 4 ºC to remove any possible cellular debris and immediately stored at −80 ºC until analysis.

2.3. Tear sample preparation by solid-phase extraction (SPE)

A purification step was prior to MALDI-TOF analyses with the aim of enriching tear samples in peptides and low molecular weight proteins. Ferromagnetic nanospheres functionalized with C18 from the MB-HIC18 Magnetic Bead Purification Kit (Bruker Daltonics; Bremen, Germany) were used to purify the tear samples following the manufacturer’s instructions. All tear fluid samples were progressively thawed on ice, and subsequently 1 μL of each sample was diluted to 10 μL in MilliQ quality water. After the binding and washing steps (following the manufacturer’s instructions), each purified sample was eluted in a single elution step with 5 μL acetonitrile (50% in deionized water).

2.4. MALDI-TOF mass spectrometry

For MALDI-TOF target preparation, each purified sample was diluted (1:10) in alpha-cyano-4-hydroxy cinnamic acid (HCCA) matrix (0.3 mg/mL in ethanol:acetone (2:1) solution). In all cases, 1 μL of each diluted sample was spotted in quadruplicate onto an MTP AnchorChip 600/584 TF plate using a standard dried droplet method. Profile analyses were performed on an Autoflex III TOF/TOF Smartbeam spectrometer with FlexControl 3.0 (Bruker Daltonics) data acquisition software. Mass spectra were acquired in linear-mode geometry, using the following settings: 1000–20,000 Th window, linear positive mode, ion source 1: 20 kV, ion source 2: 18.5 kV, lens: 9 kV, pulsed ion extraction of 120 ns, high gating ion suppression up to 500 m/z. Ionization was achieved by irradiation with a solid state 360 nm pulsed laser operating at 66 Hz (30% attenuator). All spectra were obtained randomly over the spot surface (5000 shots fired per spot) using an automatic method in a mass range between 1 and 20 kDa. Mass calibration was performed externally using the Protein Calibration Standard I (Bruker Daltonics).

2.5. Data processing and statistical analysis

After data collection, pre-processing steps were applied to each raw spectrum in order to extract putatively meaningful features before presentation to feature selection, clustering, and classification algorithms. These steps included a baseline correction based on local regression (LOESS), alignment of the spectra by VPDtw using the VPDtw package [25] from R statistics software [26]. For VPDtw implementation, a penalty based on dilation or moving local maximum over a window of 850 consecutive registered data points were applied. Peak identification and clustering by multiresolution decomposition (MRD) applying maximal overlap discrete wavelet transform with a signal to noise threshold of 10 and a m/z precision of 0.001 were performed. Finally total ion count normalization was applied. The last three pre-processing steps were performed using the mSprocess R-package [27].

Once all pre-processing steps had been completed, including a warping procedure to reduce phase variation, the data was confident enough to perform the statistical analyses required to obtain clinical relevant information.

Subsequently, the data matrix of peak intensities containing the aligned peak list was extracted and used for multivariate statistical analysis purposes. Feature subset subtraction was performed by stepwise discriminant analysis to obtain the most significant attributes in the differentiation between groups with a p-value < 0.01 cut-off.

The clustering of the samples was performed by means of a nonlinear iterative partial least squares (NIPALS) data exploratory technique. In addition, using features subtracted by stepwise discriminant analysis, machine learning approach based on multilayer perceptron neural network was performed to assess statistical model accuracy for the correct classification of problem samples. Machine learning algorithm was trained by the selected peaks and, in order to assess the performance of the learner, 10-fold random sampling was carried out with 70% of the samples as a training set and the remaining 30% were considered as test set. This test set was
2.6. **SDS–PAGE and protein identification by MALDI-TOF/TOF**

For protein peak identification, 20 μg of total tear protein was resolved in 12% sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE). After electrophoresis, proteins were stained with 0.1% Coomassie Brilliant Blue R-250. Protein bands from approximately 14–17 kDa were excised and subjected to an in-gel digestion protocol consisting of dithiothreitol reduction at 56 °C (10 mM, 20 min) followed by iodoacetamide alkylation (55 mM, 20 min in a dark environment), and finally trypsin incubation (12.5 μg/μL) overnight at 37 °C. After digestion, peptides were extracted with ammonium bicarbonate (25 mM) and trifluoroacetic acid (0.1% in water). Peptide mass fingerprinting (PMF, based on MS data) and peptide fragment fingerprinting (PFF, based on the MS/MS data from some of the peptides located in the PMF and fragmented by LIFT-TOF) analyses were carried out by spotting peptides using the dried droplet method (0.5 μL sample + 0.5 μL HCCA matrix) on a GroundSteel 384 target (Bruker Daltonics). The mass spectrometer was used in the same way as described above, but in reflectron mode for PMF and Lift mode for MS/MS. The routine included collecting 1400 scans for PMF and 2000 scans for MS/MS (400 for parent ion selection and 1600 for fragment ions). Identification was performed with Swissprot v56 as a database using a MASCOT 2.1 search algorithm with tolerances of 50 ppm for PMF and precursor ions and 0.7 Da for fragment ions. Both MS (PMF) and MS/MS (PFF) data is combined onto the same search to obtain the best identification possible.

3. **Results**

3.1. **Patients**

Analyses of subjective symptoms reported on the modified VFQ-25 questionnaire indicated that DE patients mainly complain of a gritty sensation, dryness, and a burning sensation, and referred to the use of ocular lubricants. MGD patients referred to redness and swelling, and the presence of sensory symptoms (itching, irritation and soreness). Healthy subjects did not mention any of these symptoms. In general, the DE and MGD patients presented mild to moderate forms of the diseases. The groups were gender-matched. Thus, the male vs. female relation in each group was 34.5% male vs. 65.5% female for DE, 41% male vs. 59% female for MGD, and 35% male vs. 65% female for CT. Demographical data are shown in **Table 1**.

3.2. **Solid-phase extraction (SPE) for tear sample enrichment**

All tear samples were treated with HIC18 Magnetic Beads (Bruker Daltonics) for peptide and protein enrichment in the range of 1000–20,000 Da as described previously [20]. The SPE purification allowed high quality spectrum profiles suitable for peptide/LMW protein profiling analysis.

The enriched tear samples were subsequently analyzed by MALDI-TOF, and the resulting raw spectra generally showed a high abundance of peak signals, mainly in the 1000–4000 Th range, and less presence of peak signals above 10,000 Da. Due to the complexity of the mass spectrometry outcome, a pseudo-gel view of all spectra obtained from analysis of the DE, MGD and CT groups was used to simplify data visualization and comprehension (Fig. 1). In this type of representation, the x-axis shows the mass/charge (m/z) ratio, the y-axis is the running spectrum (each with four technical replicates) and the peak intensity expressed in arbitrary units (a.u.). In general, the spectra showed a high abundance of peaks in the peptide region up to 4 kDa and several masses up to 18 kDa.

3.3. **Spectra alignment and peak detection**

Before performing peak detection and statistical analysis, the raw data obtained from MALDI-TOF assays were subjected to a spectra alignment procedure by Vpdwt to adjust any possible drifts introduced by the instrument. As result of this alignment, the peaks were synchronized in all spectra and thus were ready for peak assignment and clustering (Fig. 2).

The peaks were detected by spectra decomposition using an MRD method applying a maximal overlap discrete wavelet transform (MODWT). Finally, a peak-clustering protocol was applied to the different clinical groups, giving a total panel of 236 features.

3.4. **Feature subset and clustering**

Features that contained discriminatory information for classifying the three groups of this study were determined by selecting subsets of peaks using stepwise discriminant analysis. This analysis confirmed that the spatial features with the best ability to discriminate between DE, MGD and CT subjects were peaks 1096, 1231, 1702, 1835, 2508, 2524, 2576, 3557, 4345, 4775, 5781 and 17,467 (Fig. 3). The dimensionality
of the dataset was reduced by nonlinear iterative partial least squares (NIPALS) analysis using only the 12 most discriminant variables indicated above. The separation obtained between the groups revealed a clustering of the samples not only between controls and pathological conditions but also between both diseases, indicating that these 12 features (peaks) provide good accuracy for discrimination of DE, MGD and CT (Fig. 4). A spatial separation of DE and MGD pathological groups from the CT group was clearly observed and a differentiation between disease groups was also evident. Two outliers were identified, one belonging to the CT group and the other to the MGD group.

3.5. Machine learning and statistical model generation

The generation of predictive models for use as classifiers was explored by performing 10-fold random sampling of 70% of the whole data using machine learning approach based on artificial neural networks. The obtained classifier was subsequently validated using the remaining 30% of data. As a result of this analysis, good classification power was achieved using the 12 most significant peaks panel as confirmed ROC curves analysis (Supplementary Figure), with a percentage of correct assignment (%CA) of 89.3% including the three groups in the study. Correct assignment represents the percentage of
samples correctly classified using the panel of 12 discriminating peaks compared to clinical diagnosis.

3.6. Identification of mass spectra peak by SDS–PAGE and MALDI-TOF/TOF

SDS–PAGE of total tear protein was performed to isolate and identify discriminant peaks using MALDI-TOF/TOF mass spectrometry. The gel obtained is shown in Fig. 5. The two clearest bands below the 20 kDa marker were excised and digested with trypsin. Identification using a combination of PMF + Lift (MS/MS) resulted in lipocalin 1 (Uniprot: P31025) for the upper band indicated in Fig. 5, and lysozyme C (Uniprot: P61626) for the lower band (data not shown). These results were coherent with the mass spectra obtained in the 6000–20,000 Th window shown in Fig. 3, where the two main peaks corresponding to 17,429 and 14,664 Th correlate with the results (lipocalin 1: 17,445 Da and lysozyme C: 14,700 Da, indicated in Fig. 5C), within the instrument tolerance limits in linear mode.

Fig. 3 – Average mass spectra of tear samples processed with solid-phase extraction (SPE) protocol for dry eye (DE), Meibomian gland dysfunction (MGD) and control groups (CT) showing the informative peaks obtained by stepwise discriminant analysis.

<table>
<thead>
<tr>
<th>N°</th>
<th>Discriminant Peaks (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1096.25</td>
</tr>
<tr>
<td>2</td>
<td>1231.35</td>
</tr>
<tr>
<td>3</td>
<td>1702.88</td>
</tr>
<tr>
<td>4</td>
<td>1835.71</td>
</tr>
<tr>
<td>5</td>
<td>2508.69</td>
</tr>
<tr>
<td>6</td>
<td>2524.77</td>
</tr>
<tr>
<td>7</td>
<td>2576.45</td>
</tr>
<tr>
<td>8</td>
<td>3557.07</td>
</tr>
<tr>
<td>9</td>
<td>4345.97</td>
</tr>
<tr>
<td>10</td>
<td>4775.82</td>
</tr>
<tr>
<td>11</td>
<td>5781.48</td>
</tr>
<tr>
<td>12</td>
<td>17467.56</td>
</tr>
</tbody>
</table>

Fig. 4 – Nonlinear iterative partial least squares (NIPALS) analysis using the most discriminant variables obtained by stepwise discriminant analysis. The spatial ordering and the separation produced between the samples as a function of the values of the variables included in the analysis is apparent. In this representation, each of the points represents one individual from each group, and good separation between the three groups can be seen. The presence of one CT and one MGD outlier can also be observed and are indicated by corresponding arrows (squares: DE group; circles: MGD group; triangles: CT group).
4. Discussion

The tear film study of ocular pathologies requires methodologies sensitive enough to allow the use of a just few microliters of sample as starting material, especially in pathologies like dry eye syndrome where the tear amount may be reduced.

Tear film is a biofluid that covers the eye surface, keeping it wet, nourished and lubricated. The foundation of the profiling use for diagnostic purposes is based on the fact that basal tear peptidome and proteome reflect changes depending on the pathological stage and the abundance of peptides and/or proteins of different natures are part of pathophysiological and defense processes. This pathological stage is, in turn, induced by disturbance in any of the lacrimal function unit (LFU) component, including lacrimal system, the ocular surface and its accessory glands and their neural interconnections [2]. In fact, it has been recently reported that dry eye patients present both structural and functional alterations of subbasal corneal nerves and that those changes are found to be related to the severity of dry eye [29].

Several efforts have been pursued lately to get clinically valuable information based on tear proteome data and the role of proteins in disease [30-34]. However, those studies have mainly explored the disease of interest independently and are distant from high throughput application in clinical studies. Possibly from these studies, SELDI technology might be closer to this kind of application. In this regard, Tomosugi et al. explored the potential use of SELDI to analyze proteomic patterns in Sjögren’s syndrome [12], and Grus et al. used it for the analysis of tears from patients with dry eye [13]. However, it has been postulated that solid phase extraction in combination with MALDI-TOF analysis overcomes the limitations of SELDI analysis [9]. Consequently, this study applied a SPE-MALDI-TOF method for the analysis of tear proteomics by comparing the peptide and small proteome profiles of three different conditions, DE, MGD and healthy controls.

Given that both pathological conditions are closely related to each other, complicating the analysis of samples, high quality results in terms of peptides/protein profiles and separation of groups based on these profiles are required.
For an optimal profiling analysis, the data obtained from mass spectrometry must undergo several pre-processing stages to precisely identify the location of peaks in the mass/charge ratio. In practice, instrument precision may be reduced by variations in the triggering time from spectrum to spectrum, small variations in the accelerating voltage, or simply by random errors [21] and this may lead to incorrect peak assignment and deficient pattern recognition based on mass spectra. It has been reported that the accuracy of peak location in MALDI-TOF is within 0.15–0.3% of the mass (m/z) value [35]. Therefore, it is essential to align or synchronize spectra to ensure that features occur at the same time position in all spectra being analyzed. After this step, features of biological interest can be clearly revealed by multivariate analysis.

In this context, the alignment of spectral data has been studied extensively in the MS field [22–24]. There are various approaches to aligning MALDI-TOF data, including the use of frequent calibration [36], clustering or re-binning [37–40], cross-correlations [41], minimizing entropy [42] or, more recently, self-calibrated warping in SELDI-TOF [43]. A prerequisite of the majority of all of these methods is that the peaks must first be identified. However, as synchronization errors may appear in the peaks throughout the whole spectrum, the corrections applied to the complete time range should provide greater precision and better alignment results [21]. This study used an alignment stage prior to the detection and assignment of peaks that involved VPdtw [25]. To our knowledge this is the first work in which VPdtw alignment procedure has been applied to MALDI-TOF data and more specifically to tear film analysis. This method was used with practical purposes to improve predictive model accuracy throughout the study. In-depth studies of the best method for processing MALDI-TOF data have been extensively reported by other groups [43–45]. However, this study tested different pre-processing techniques for spectra data in order to find which ones best fitted DE, MGD and CT MALDI-TOF profiles. This means, the same MALDI-TOF raw spectra were subjected to warping and non-warping procedures. There was a remarkable improvement in the classification accuracy from 71.4% to 85.7% of correct assignment when including a warping step in the MALDI-TOF data pre-processing pipeline (data not shown).

Due to the complexity of DE and MGD diseases and the implication of different biological processes in the evolution of these pathologies, a panel of peaks obtained by feature subtraction methods can offer more precise discrimination between diseases than discrimination based on a single peak. Supervised feature subtraction performed by stepwise discriminant analysis produced a data reduction from 236 to 12 significant peaks. The use of this panel of 12 peaks in conjunction with multivariate statistical analysis allowed discriminating the study groups with a high percentage of correct assignment. A clear separation of the control group from pathological groups is easily seen. Besides, a spatial separation between the DE and MGD groups is also observed, even more discrete than in the case of the control group. In addition, it is also possible to identify possible outliers. More specifically, in our analysis we were able to identify two possible outliers as depicted in Fig. 4 by arrows. It is evident that one control subject is out of the main group and a review of the clinical record revealed that this subject did not present signs or symptoms of disease, however the shift from the main control group towards the MGD group in an intermediate position between those groups, suggests possible early onset of the disease even in an asymptomatic phase, and the necessity for follow-up of the patient. Furthermore, the profile of a MGD patient suggests that this patient behaved more as a DE patient since is located within this group. In this case, even if this patient was recruited as MGD, a coexistence with DE could explain this result, fact that often occurs in daily clinical practice. All these results delineate peptide/protein profiling as a promising tool to aid clinical practice.

Once good separation of the groups was observed by data exploratory techniques, machine learning algorithm based on neural networks was used to build predictive model for distinguishing between the DE, MGD and control groups. In any statistical modelling process validation of the model is a crucial step. The effectiveness of the model was 10-fold evaluated by random sampling using the panel of 12 peaks obtaining a 89.3% of correct assignment when comparing all three groups included in the study.

In order to identify some of the discriminant peaks, 20 μg of tear protein from a control subject was resolved in SDS gel. The resulting SDS–PAGE correlates well with the masses observed in the 10,000–20,000 Da range (Fig. 3), with two main peaks located at 17,429 and 14,664 Th, and the corresponding double charged masses (8705 and 7331 Th, respectively) (Fig. 5C). Identification by combined PMF + Lift (MS/MS) resulted in LCN1 (Uniprot: P31025, Mm 17445 Da) for the upper band, and lysozyme C (Uniprot: P61626, Mm 14,700 Da) for the lower band. The mass of 17,445 attributed to LCN1 is located in the feature subset analysis as a deregulated marker peak, where a reduction of signal peak is observed under pathological conditions. Interestingly, LCN1 is a major protein in tears and has a variety of functions, including regulation of tear viscosity, lipid binding, and antimicrobial and anti-inflammatory activities, among others. This protein has already been reported to be deregulated in tears of dry eye [13,33,46], Sjögren’s syndrome [47] and MDG patients [46]. However, LCN1 has been mentioned as a potential marker in different dry eye and non-dry eye diseases, being apparently associated with non-specific inflammation, inflammation induced by therapy, or stress conditions [48–51], thus implying that the specificity of lipocalin as a single protein for discriminating pathologies seems to be limited. It is clear that in multifactorial diseases one unique biomarker is normally not informative enough to explain all the complex signalling and functional networks that are altered in a pathological condition. In our particular study it is clear that lipocalin 1, or any other discriminative peak by itself, lacks enough diagnostic power to discriminate DE and MGD in dry eye groups.

Therefore, these very promising results should be expanded with even more samples of each condition, and further efforts to identify informative peptides by combining alternative mass spectrometry approaches such as off-line MALDI-TOF/TOF and/or nano liquid chromatography (nLC) together with on-line tandem mass spectrometry would help in consolidating and framing the data obtained in a physiological scenario.

In summary, this study applied solid phase extraction methods in combination with MALDI-TOF peptide and low
molecular weight protein analysis. In addition, the application of a warping procedure and multivariate statistical analysis of the resulting spectra allowed differentiation between DE, MGD and control, the three conditions studied, using a panel of 12 informative peaks.

To the best of our knowledge, this is the first time two related ocular surface diseases have been analyzed by tear peptide/LMW protein MALDI-TOF profiling, and effectively differentiated.

This study constitutes a sensitive, fast method that allows for the possibility of high throughput analysis and easy classification of patients into appropriate groups according to the peptide/protein profile provided by the panel of 12 peaks found being discriminative. Just like in other health areas where sample profiling is already applied for diagnostic purposes, we suggest that combining this strategy in ophthalmology with routine clinical protocols such as questionnaires and clinical tests (Schirmer I test, tear breakup time, staining of the cornea and conjunctiva, among others) would help in better diagnosis, prognosis and monitoring of aqueous-deficient dry eye and MGD.

Financial support

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness-CDTI through the INNPACTO subprogram, grant IPT-2011-1429-01000. The funding organization had no role in the design or conduct of this research.

Conflict of interest statement

The authors declare the absence of any conflict of interest related to the manuscript or work reported therein.

Acknowledgements

The authors wish to express their gratitude to Dr. Pedro López from the San José Geriatrics Residency (Erandio, Bizkaia) for assistance in selection of volunteers.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.euprot.2014.02.016.

REFERENCES

MS using high resolution peak detection and label-free alignment. Proteomics 2008;8:1530–8.


