SELDI-TOF-MS ProteinChip Array Profiling of Tears from Patients with Dry Eye

Franz H. Grus,1 Vladimir N. Podust,2 Kai Bruns,3 Karl Lackner,3 Siyu Fu,2 Enrique A. Dalmasso,2 Anton Wirthlin,4 and Norbert Pfeiffer1

PURPOSE. Protein and peptides in tears play an important role in ocular surface diseases. In previous studies, changes have been demonstrated in the electrophoretic protein profiles of patients with dry eye. The purpose of this work was to determine the usefulness of surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) ProteinChip Array (Ciphergen Biosystems, Inc., Fremont, CA) technology for the automated analysis of proteins and peptides in tear fluid.

METHODS. Patients with dry eye (DRY, n = 88) and healthy subjects (CTRL, n = 71) were examined. Their tear proteins were analyzed using SELDI-TOF-MS ProteinChip Arrays with three different chromatographic surfaces (CM10 cation exchange, Q10 anion exchange, and H50 reversed-phase) prepared by means of a laboratory liquid-handling robotic workstation. The data were analyzed by multivariate statistical techniques and artificial neural networks, and the most important biomarkers were purified and identified by tandem MS.

RESULTS. Complex patterns of tear proteins and peptides were detected. The different chromatographic surfaces revealed the selective enrichment of proteins such as lipocalin and lysozyme. Discriminant analysis demonstrated highly significant changes in the protein profiles in patients with dry eye (P < 0.001). With a seven-peptide multimarker panel, an artificial neural network could differentiate between patients with dry eye and healthy individuals with a specificity and sensitivity of 90%. The identification of biomarkers revealed an increase of some proteins that may have protective functions.

CONCLUSIONS. The SELDI-TOF-MS technology seems to be ideally suitable for the mass screening of peptides and proteins in tears. This highly sensitive approach dramatically reduces the analysis time and provides protein profiles with great mass accuracy. Thus, it may become a very useful tool in the search for potential biomarkers for diagnosis and new therapeutics in ocular diseases such as dry eye. (Invest Ophthalmol Vis Sci. 2005;46:863–876) DOI:10.1167/iovs.04-0448

The number of patients with dry eye is rapidly increasing. In the United States, >6% of the population aged >40 years have dry eye, and this prevalence increases up to 15% of people aged >60 years.1,2 Patients with dry eye typically report discomfort, burning, irritation, photophobia, and blurred vision and have an elevated risk of corneal infection and resultant irreversible tissue damage.3 In patients with dry eye, some components such as lactoferrin were found to be altered in comparison to levels in healthy subjects.4–6

Studies have demonstrated that the electrophoretic analysis of tear proteins is able to detect differences in the tear protein profiles between patients with dry eye syndrome and healthy subjects.7 The stained gel images were digitized, and the peak volumes quantified and analyzed by multivariate statistics and artificial neural networks. In these studies, the detection of dry-eye syndrome based on the specific electrophoretic patterns was demonstrated to be superior to detection by clinical standard tests (e.g., Schirmer test [BST]) and break-up time [BUT]18–19). Furthermore, altered protein profiles in tears were found, not only in dry eye, but also in some other ocular and systemic diseases (e.g., diabetes mellitus11–14) and in cases with other factors that might interfere with the ocular surface (e.g., contact lenses and smoking15,16). All these studies provide evidence that specific proteins or peptides in tear film can be used as diagnostic biomarkers for dry eye, ocular surface diseases, and even systemic disease states such as diabetes mellitus.

Previous investigation of tear film proteins has not been limited to one-dimensional (1-D) gel electrophoretic techniques. In other studies, the tear film proteins have been extensively analyzed by high-performance liquid chromatography,17 two-dimensional (2-D) gel electrophoresis,18 or other chromatographic approaches.19 2-D electrophoresis is a step forward in proteome research, because it offers an improved protein separation compared with 1-D electrophoresis. In the first dimension, the proteins migrate to the isoelectric point that is specific for each protein. In the second dimension, the isoelectrically focused proteins are separated according to their molecular masses. Using this technology, we were able to analyze and quantify single proteins in tear film and to establish a tear film protein map.11,12,18,20,21

However, the 2-D electrophoretic analysis of proteins is very time consuming and is limited by problems in reproducibility—particularly important when this technique is used for mass-screening of many samples in clinical routine. To eliminate gel-to-gel variation, very complex and tedious procedures to eliminate gel deterioration must be performed to ensure the quality of the spot-matching process. Furthermore, the sensitivity of this method limits the analysis to proteins larger than approximately 8 to 10 kDa. Considering the increasing demand for mass screening of protein profiles in tear fluid, a method is needed that overcomes the problems of 2-D electrophoretic separations.

The purpose of this study was to analyze the applicability of surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) ProteinChip Array technology (Ciphergen Biosystems, Inc., Fremont, CA) for the protein...
profiling of tears. This method was first described by Hutchens and Yip and utilizes chromatographic surfaces to retain proteins based on their physicochemical characteristics, followed by TOF-MS using a ProteinChip Reader (PBSII, PBSIIIc, or PCS Series 4000; Ciphergen Biosystems, Inc.). This technique can be used for a rapid and efficient analysis of body fluids such as sera and cell culture samples. Proteins are separated on different array surfaces (e.g., cationic and anionic exchangers, hydrophobic, and metal-ion affinity surfaces). Proteins that are retained on the chromatographic surfaces due to the chosen binding and washing conditions can be easily purified from contaminants such as buffer salts or detergents before the analysis by MS. Furthermore, the very high sensitivity of ProteinChip Readers is ideal for the analysis of small sample volumes, such as those typically available from tear fluid. An additional advantage of this technique in comparison to 1-D and 2-D electrophoresis is its ability to screen down to the molecular size of small peptides (~500 Da). As a result, SELDI-TOF-MS has been widely used in a number of proteomic studies. The first objective of this work was to demonstrate the use of SELDI-TOF-MS as a method to profile human tear proteins. The goal of this study was to find protein or peptide biomarkers for dry-eye disease by comparing the protein profiles of tears from patients with dry eye and healthy control subjects and analyzing those using univariate analysis, multivariate statistical techniques, and artificial neural networks. The objective of the analysis phases of the study is to build a multivariate panel of markers that will classify the two groups as accurately as possible.

**MATERIALS AND METHODS**

**Patients**

One hundred fifty-nine patients were included in this study, 71 of whom were categorized as control subjects (CTRL) and 88 of whom had dry-eye syndrome (DRY). Informed consent was obtained from all patients participating in the study and the protocols were approved by the institutional ethics committee and conformed to the provisions of the Declaration of Helsinki. The ophthalmic examination included subjective symptoms, visual acuity, Schirmer test with anesthesia (basis secretory test BST), biomicroscopy with careful examination of the lid margin and meibomian glands and tear BUT. To assess the subjective symptoms, the patients’ history was taken. Each patient was asked for subjective symptoms such as burning, itching, foreign body sensation, dryness, and photophobia. Patients were classified as having dry eye if they had these symptoms of dry-eye and abnormalities of test dynamics determined by Schirmer (<5 mm/5 min). Patients with Sjögren syndrome were excluded from this study. Furthermore, only tear substituents without preservatives were accepted as treatment during the month before examination. To focus mainly on aqueous-deficient dry eye in this study, we excluded patients with clear meibomian gland dysfunctions. The normal control subjects were matched in age and sex to the patients with dry eye. In all patients, the BST was performed and the Schirmer strips stored at ~80°C until use.

**SELDI-TOF-MS Analysis**

**Samples.** The first 10 mm of the Schirmer strips were eluted overnight in 500 μL phosphate-buffered saline (PBS) with 1% Triton X-100. For automatic handling of all binding and washing steps, a robotic laboratory automation station (Biomek 2000; Beckman Coulter, Fullerton, CA) was used. The station was extended by an integrated microplate mixer (Micromix 5; Diagnostic Products Company, Los Angeles, CA), which holds the ProteinChip Array BioProcessor (Ciphergen Biosystems, Inc.). This BioProcessor was equipped with 12 ProteinChip Arrays, each having eight spots. With two BioProcessors, up to 192 wells with an approximately volume of 250 to 300 μL can be handled by the robotic station per well.

We chose to test the tear fluid on three different chromatographic surfaces: a weak cation exchange surface (CM10), a strong anion exchange surface (Q10), and a reversed-phase surface (H50). All ProteinChip Arrays were pretreated according to the standard protocols of the manufacturer. Binding buffers were 5% acetonitrile/0.1% trifluoroacetic acid (H50), 20 mM sodium acetate buffer (pH 5; CM10), and 50 mM Tris (pH 8; Q10). Twenty microliters of the Schirmer strip elution and 20 μL of binding buffer were applied to each spot using the Biomek. The arrays were incubated on the DPC shaking platform for 2 hours, and the solution in each well was removed by the Biomek. All wells were washed three times with 150 μL binding buffer followed by a wash step with 20 μL distilled water for 5 minutes to remove buffer salts. After the wells were dry, the Biomek was used to apply 1 μL of a saturated sinapinic acid solution (an energy-absorbing molecule) in 50% acetonitrile and 0.5% trifluoroacetic acid to each spot. After the spots were air dried, each spot was analyzed in a ProteinChip Reader. Each sample was bound to each array surface in duplicate on separate arrays and BioProcessors.

**Data Acquisition and Preprocessing.** ProteinChip Arrays were analyzed on a PBS-IIc ProteinChip Reader equipped with a ProteinChip Array AutoLoader using the ProteinChip Software version 3.2 (Ciphergen Biosystems, Inc.). The AutoLoader is able to analyze up to 24 ProteinChip Arrays (192 spots) at one time. Each array was read at two laser intensities: low intensity optimized for lower-molecular-mass proteins and high intensity for higher-molecular-mass proteins. The high-intensity protocol averaged 195 laser shots from each spot with a laser intensity of 190, a deflector setting of 3,000 Da, a detector sensitivity of nine, and a molecular mass detection range of 2,000 to 200,000 Da. For low-intensity measurements, the laser intensity was set at 180 and the deflector set at 1500 Da. The raw data were transferred to the CiphergenExpress Data Manager Software version 2.1 (CE; Ciphergen Biosystems) for analysis. The baseline was subtracted using a setting of 12 times the expected peak width. The spectra were calibrated with external calibrants (bovine insulin, 5,735.6 Da; ubiquitin, 8,654.8 Da; cytochrome c, 12,360.2 Da; β-lactoglobulin, 18,363.3 Da; and horseradish peroxidase, 43,240 Da). The low- and high-intensity laser settings were calibrated separately. Spectral intensities were normalized by total ion current (TIC) to an external normalization coefficient of 0.2. The starting mass for the normalization procedure differed according to the laser energy and the settings in the spot protocol for data acquisition. In this study, the low mass cutoffs were 3000 m/z for arrays at low laser intensity and 8000 m/z for arrays at high laser intensity. Spectra with normalization factors that were >100% above the mean were evaluated for possible deletion. They were deleted if their normalization factors were more than double the value calculated for the replicate spectrum. In the case in which both spectra were above the cutoff, both spectra were kept. At least one replicate was retained for every sample.

Automatic peak detection was performed using the settings of five times the signal-to-noise ratio for the first pass of peak detection and two times the signal-to-noise ratio for the second pass. From these detected peaks, a list of peak clusters was created. The Biomarker Wizard feature in the CE generates consistent peak sets across multiple spectra. When comparing a given protein peak across various sample conditions, it is important to obtain an intensity value for every spectrum, even though they may not have been found with a given set of automatic peak detection settings. The Biomarker Wizard operates in two passes. The first pass uses low sensitivity to detect obvious and well-defined peaks. The second pass uses higher-sensitivity settings to search for smaller peaks, with the masses found in the first pass. In this study, a peak cluster was created if the given peak was found in 10% of all spectra for an individual condition above the first-pass cutoff. The mass window for peak clustering was defined as 0.3% of the peak mass for spectra read at low laser intensity and 2% of the peak mass for spectra read at high laser intensity. A larger-mass window is used for

**Acknowledgments**

This work was supported by the National Institutes of Health grant RO1 EY 13087 and the Research to Prevent Blindness Career Development Award to K.G. The authors thank John Shaffer for assistance in the collection of data, and the Biomek and CiphergenExpress Data Manager Software version 2.1 for analysis of the data. The authors especially thank Dr. John Shaffer for the kind assistance in the preparation of the manuscript.
the higher-molecular-mass proteins to account for broadening of peaks due to an increase in protein isofoms.

**Data Analysis.** The CE data manager software was used to normalize the spectra, to automatically detect peaks, and to create the peak cluster lists. Because of the inherent variability in the total signal for raw laser desorption/ionization mass spectrometric data, normalized peak intensities were used exclusively in the statistical analyses. The cluster lists were exported as ASCII files to a statistical analysis program (Statistica, ver. 6.2; StatSoft, Tulsa, OK). The cluster lists contained normalized peak intensity values for each sample within a group. Based on these normalized peak intensities, probabilities based on t-tests and multivariate discriminant analysis were calculated. The separation between clinical groups for diagnostic purposes can generally be enhanced by using multimarker panels of protein biomarkers rather than single biomarkers.

The statistical software package (Statistica; StatSoft) was used to perform a multivariate discriminant analysis based on combinations of multiple biomarker peaks. This discriminant analysis tests the zero hypothesis (namely, that mean biomarker peaks of the different clinical groups derive from a multivariate normal distribution) and also shows which of the various groups are statistically different. Based on this, discriminant function analysis can be used to determine which variables (peaks at particular molecular masses) generated a significant mean value comparison, or which variables can discriminate between groups. The discriminant analysis can calculate which variables are the most important to discriminate between the groups and ranks the variables according to their significance. For each condition (surface type and laser energy), a discriminant analysis was calculated. Each analysis creates a list of peaks that was found to be significantly different between the groups. All these lists were combined and a final discriminant analysis was calculated based on the most significant protein biomarkers from all conditions. This final discriminant analysis selected the seven most important protein biomarkers for best discrimination between the groups.

To assess the diagnostic power of this seven biomarker panel, a neural network was trained using these markers as input neurons. The traditional artificial neural network (ANN) model is the most widely used in today and is the multiple-layer feed-forward network (MLFN) with the back-propagation training algorithm.33 In this study, we used the neural network module provided by the software (Statistica; StatSoft). The network was trained by the seven biomarker peaks and the output neurons defined with patient was classified as dry-eye protein pattern or not. To assess the performance of the network, the data set was randomly divided in two parts: The first half of the patients and control subjects was used as training set, the second as a test set. After completion of training, the success of the algorithm was tested using the test data set, which comprises patients and control subjects to which the neural network has not previously been exposed.

To demonstrate the results, the software generates an ROC curve by plotting sensitivity against 1 – specificity. A test that perfectly discriminates between two groups would yield a “curve” that coincides with the left and top sides of the plot. A global assessment of the performance of the test (sometimes called diagnostic accuracy) is given by the area under the ROC curve (ROC area under the curve [AUC]). A perfect marker with complete group separation would have an ROC AUC value of 1.

**Alternative Sampling of Tears.** In this study, we used tear fluid which was eluted from Schirmer strips. However, the more common technique of sampling tears for biochemical studies is the use of a microcapillary. Thus, we compared SELDI-TOF-MS protein profiles from tears eluted from Schirmer strips with tears collected by micro-capillaries. A tear volume of approximately 5 μL was sampled from one eye of three selected patients with a 5-μL glass capillary without touching the lid margins and eye lashes of the patients and stored at −80°C until use.

H50 ProteinChip Arrays were used for detection of protein profiles from tear fluid samples. For microscale analysis, a sample volume of 2 μL tears in 2 μL binding buffer (5% acetonitrile/0.1% trifluoroacetic acid) was used. After incubating in a humidity chamber at room temperature for 2 hours, all spots were washed three times with 5 μL binding buffer followed by a wash step with 3 μL distilled water for 5 minutes. After air drying, 1 μL of saturated sinapinic acid (energy absorbing molecule) in 50% acetonitrile and 0.5% trifluoroacetic acid was applied. After air drying, each spot was analyzed by SELDI-TOF-MS, as described earlier.

**Protein Purification and Identification.** Eluted samples from Schirmer strips were first enriched by combination of cutoff membrane fractionation (YM30 and YM3; Millipore, Bedford, MA), anion exchange chromatography (Q HyperD F; Ciphergen Biosystems, Inc.), and/or reversed phase chromatography (RPC PolyBio; Ciphergen Biosystems, Inc.). Polypeptides smaller than 4.2 kDa were identified by direct sequencing using a tandem mass spectrometer (QTOF2; Micromass, Manchester, UK) equipped with a PCI-1000 ProteinChip Interface (Ciphergen Biosystems, Inc.). Proteins were bound from enriched fractions to CM10 or IMAC-Cu arrays. Spectra were collected from 2 to 4.5 kDa in single MS mode. After reviewing the spectra, specific ions were selected and introduced into the collision cell for CID fragmentation. The CID spectral data were submitted to the database-mining tools (Mascot; Matrix Sciences) or Protein Prospector MS-Tag (UCSF) for identification.

Proteins larger than 4.2 kDa were further purified by SDS-PAGE (16% tricine gel; Invitrogen). Polyacrylamide gels were stained (Colloidal Blue Staining Kit; Invitrogen). Selected protein bands were excised from the polyacrylamide gel. The gel pieces were washed with 200 μL of 50% methanol/10% acetic acid for 30 minutes, dehydrated with 100 μL ACN for 15 minutes, and extracted with 70 μL of 50% formic acid/25% ACN/15% isopropanol/10% water for 2 hours at room temperature with vigorous shaking. Two microliters of extracts were applied to NP20 ProteinChip arrays and reanalyzed with the ProteinChip Reader to confirm m/z values of the excised/extracted proteins.

The gel extracts were vacuum dried (SpeedVac; ThermoSavant, Holbrook, NY). Twenty microliters of ammonia solution (concentrated stock diluted 25-fold with water) was added to the tubes, and the solution vacuum dried again. Twenty microliters of 2 ng/μL modified trypsin (Roche Applied Science, Mannheim, Germany) in 50 mM ammonium bicarbonate (pH 8) were added to each tube and the samples incubated for 2 to 4 hours at 37°C. Peptide identification was performed using the tandem mass spectrometer equipped with a PCI-1000 ProteinChip Interface. MS/MS spectra were submitted to the database mining tool (Mascot; Matrix Sciences) for identification.

**RESULTS**

Tear samples were analyzed on three different ProteinChip Array chemistries as has been described. Figures 1 and 2 demonstrate several SELDI-TOF mass spectra of tear protein profiles that illustrate differences in profiles corresponding to different array surfaces and laser-intensities. Both figures show mass spectra that were obtained from arrays prepared by the automated robotic station (Biomek; Beckman Coulter). Figure 1 shows mass spectra of healthy control subjects (CTRL) and patients with dry eye (DRY), with high laser energy used on H50, CM10, and Q10 surfaces. In the gel view, the same protein patterns are shown in a gray scale presentation that simulates an electrophoretic separation. In Figure 2, representative spectra of both groups are shown on the three different surfaces for data collected at low laser energy. Both figures provide an insight into the complexity of tear protein profiles in the lower-molecular-mass range and into the variability from patient to patient.

In Figure 3, the elution profiles of two of the main tear proteins, lysozyme, and lipocalin, are demonstrated for the different chip surfaces (H50, Q10, and CM10) in the tears of a
FIGURE 1. SELDI-TOF-MS patterns of two healthy subjects (CTRL) and two patients with dry eye (DRY) produced by means of automated sample preparation. Top: grayscale images of SELDI-TOF-MS data. Bottom: protein MS patterns of tears. The intensity of proteins (U) is plotted versus the molecular mass in a range up to 35,000 Da. The tear samples were prepared on three different array surfaces (H50, CM10, and Q10) and analyzed at the high energy-laser setting.
Figure 2. SELDI-TOF-MS patterns of two healthy subjects (CTRL) and patients with dry eye (DRY) prepared by means of automated sample preparation. Top: grayscale images of SELDI-TOF-MS data. Bottom: protein MS patterns of tears. The intensity of proteins (U) is plotted versus the molecular mass in a range up to 40,000 Da. The tear samples were prepared on three different array surfaces (H50, CM10, and Q10) and analyzed with the low energy protocol.
healthy volunteer. As expected, the protein profiles are different due to the on-chip chromatographic enrichment. The peaks which correspond to lysozyme and lipocalin are marked in Figure 3. Whereas lipocalin was present on the reversed-phase (H50) and anion exchange surface (Q10), it was very weakly bound to the weak cation exchanger surface (CM10). Lysozyme, the other most prominent protein peak in tears, shows the strongest peak on the cationic exchange surface (CM10). Furthermore, Figure 3 demonstrates a heat map where the protein expressions are shown as pseudocolors for the different groups (DRY and CTRL) and conditions (array surfaces [H50, CM10, and Q10] and laser settings [low and high energy]). In contrast to the spectra plot at the top, this heat map represents the averaged data for each group at each condition.

FIGURE 3. Top: detection of the two main tear proteins lysozyme and lipocalin analyzed on three different array surfaces (H50, Q10, and CM10) at high laser energy. Lipocalin was mainly enriched on the anion exchange surface (Q10) at pH 8, whereas it was only slightly bound on the cation exchange surface (CM10). Lysozyme was highly enriched on the CM10 surface. Bottom: heat map of protein expression patterns at each condition. In this graph, the intensity of protein expression is shown as gray values. The x-axis displays the corresponding molecular mass region and the y-axis the different groups (DRY and CTRL) and conditions (array surfaces [H50, CM10, and Q10] and laser settings [low and high energy]).
molecular mass region between 11,000 and 20,000 Da ($x$-axis) and relative peak intensity ($y$-axis) for the different groups (DRY and CTRL) for the various conditions (array surfaces: H50, CM10, and Q10, and laser-settings: low energy and high energy). In addition to other complex protein expression patterns, lysozyme is most prominent on the CM10 array, whereas lipocalin is most prominent on the Q10 array. In contrast to the spectra plots at the top of Figure 3, this heat map reveals the averaged data for each group at each condition. Thus, the heat map can provide a more comprehensive overview about the patterns, lysozyme is most prominent on the CM10 array, whereas lipocalin is most prominent on the Q10 array. In contrast to the spectra plots at the top of Figure 3, this heat map reveals the averaged data for each group at each condition. Thus, the heat map can provide a more comprehensive overview about the patterns.

In agreement with previous studies, the analysis of the SEIIDI-TOF-MS spectra confirmed that lysozyme amounts are decreased in tears of patients with dry eye (Fig. 4). To evaluate day-to-day variations between the ProteinChip Array profiles, a control (pooled tears from healthy subjects) was processed in every BioProcessor. Figure 5 shows a representative measurement using the same sample in different BioProcessors on different days. Coefficients of variation (CVs) were calculated for this pooled tear sample for all peak clusters and for all surfaces. The average CVs were calculated to be between 28% and 34%. According to the manufacturer, these values are in the same range as those obtained in serum profiling.

In this study, we chose to use automated SEIIDI-TOF-MS analysis of the elution of Schirmer strips as means of a robotic workstation to improve consistency and reproducibility in the sample preparation process and to make an important step toward developing an automated tear protein test. However, we wanted to compare the protein profiles obtained from the Schirmer strips to those obtained using the conventional microcapillary technique. Therefore, we analyzed the samples obtained with microcapillaries from three patients in addition to the Schirmer elutions. Figure 6 shows the SEIIDI-TOF-MS spectra of a patient with dry eye, using microcapillary sampling compared with spectra from Schirmer strip elutions. The $x$-axis in this figure shows the molecular mass of detected proteins and peptides for two different high- and low-molecular mass ranges. Several of the main protein peaks predicted from electrophoretic analysis were detected. According to their molecular masses, these proteins are lactoferrin, albumin, lipocalin, lysozyme, and several chains of IgA. In addition to these high-molecular-mass peaks, a very complex spectrum of peptides was detected in the molecular mass region below 20 kDa in both analyses of tear samples obtained from Schirmer strips and microcapillaries. In general, both collection methods yielded similar results. However, in the protein spectra of the Schirmer strips, an increased number of proteins in the lower-molecular-mass range could be observed. In sum, the results shown in Figures 1 to 6 provide evidence that the SEIIDI-TOF-MS ProteinChip Array analysis is very useful for the analysis of tear proteins.

The primary objective of this study was to search for proteins or peptides that are differentially expressed between healthy subjects and those with the dry-eye syndrome. Samples were prepared, and then the raw spectral data were acquired and processed as described earlier. Cluster lists were generated for each condition (low- and high-energy laser settings) and for each surface (CM10, H50, and Q10) using the Biomarker Wizard (Ciphergen Biosystems, Inc.). For each of these cluster lists, a multivariate discriminant analysis was performed which defined a total of approximately 50 peaks with significant differences between the tear protein profiles of CTRL and DRY ($P < 0.01$). The objective of protein profiling is to distinguish healthy from disease states based on protein patterns. In previous studies using electrophoretic separations with subsequent digital image analysis and multivariate statistics, we were able to demonstrate that the tear protein patterns of patients with dry eye differ significantly from those of healthy subjects and can be used for diagnostic purposes. Because SEIIDI-TOF-MS is especially sensitive in the low-molecular-mass range of proteins, we were interested to differentiate between healthy and dry eye based on peptides and small proteins. Thus, for this analysis we excluded all members of the cluster list which had a molecular size greater than lysozyme (~14 kDa). An additional discriminant analysis was performed using a cluster list containing all previously significant clusters with molecular masses smaller than 14 kDa. This analysis revealed a statistical significant difference between both groups using the biomarker candidates from all conditions and surfaces simultaneously ($P < 0.0000001$). Table 1 lists the first seven peaks that had statistically the most significant effect on the separation between both groups in the discriminant analysis and were shown not to be MS artifacts such as double charged ions.
Using this multimarker panel for input, an artificial neural network was trained with a training data set. The performance of the trained net was assessed using a test data set. Figure 7 shows an ROC curve with an AUC of 0.93 and a specificity and sensitivity of approximately 90% for each.

Several of the biomarker candidates of the multimarker panel were enriched and identified using tandem MS. The biomarker at 5792 Da was identified as proline-rich protein 3 (PRP3, Fig. 8). Figure 9 shows candidate biomarkers at 4020 and 4050 Da. These biomarkers were detected on the CM10 surface at low laser energy. There was a significant decrease of both peptides in patients with dry eye ($P < 0.000001$). The 4050-Da biomarker was identified as a C-terminal fragment of PRP4 (amino acids 85-118). The 4020-Da biomarker was identified as a variant of the same peptide, the C-terminal fragment of nasopharyngeal carcinoma-associated PRP4 (NCAPP4, amino acids 85-118). The only difference between these two peptides is a single amino acid substitution from Arg to Gln.

Figure 7 shows an elevation of the 4136-Da biomarker in patients with dry eye. This biomarker was identified as the C-terminal fragment of $\alpha$1-antitrypsin (amino acids 358-394).

Finally, the 10,834-Da marker was identified as calgranulin A.

**DISCUSSION**

In recent years, the search for new biomarkers for human diseases has become more successful because of new high-throughput techniques in the fields of genomics and proteomics. Biomarkers or biomarker patterns should enable the scientific community to develop more reliable diagnostic tools and may also allow prediction of disease progression and optimization of therapeutic approaches. One such approach was recently demonstrated for the detection of ovarian cancer.$^{36,37}$

This study used one of the most promising new proteomic tools for the detection of biomarkers, the SELDI-TOF-MS technology. This technique allows for straightforward and rapid sample analysis because very small sample volumes can be directly applied to the ProteinChip Array (Ciphergen Biosystems, Inc.) surfaces and the process can easily be automated for high-throughput analysis.$^{38}$ This technique has been widely used for the analysis of blood samples and cell culture medium. In the present study, we applied this technique to the analysis of tear-fluid proteins.

Proteins and peptides play an important role in preserving the integrity and stability of the ocular surface. Furthermore, the tear film is involved in defense against bacterial and viral infections. In previous studies, we demonstrated that changes in the tear film proteins may be used as biomarkers for dry-eye disease,$^8$ but the importance of biomarkers in tear film is not limited to ocular surface diseases. They may be helpful also for studying systemic diseases, such as diabetes mellitus$^{14,39,40}$ or cancer.$^{41}$ The electrophoretic techniques used in earlier studies did not allow for the screening of smaller proteins (<9 kDa) and peptides. These peptides, as well as the larger proteins, play an important role in the physiological functions of the tear film. They mediate inflammatory processes, and act as growth factors. In a variety of systemic diseases, the study of peptides in particular may lead to the discovery of novel potential biomarkers.$^{36}$ The SELDI-TOF-MS technology allows for a sophisticated analysis of peptides, because this technique has its
sensitivity maximum in the molecular range between 500 and 25,000 Da.

In the present study, we desired a platform suitable for screening large numbers of samples. Therefore, we used an approach that is based on an automated robotic sample preparation workstation and a SELDI-TOF-MS ProteinChip Reader equipped with an AutoLoader. For the volume of the samples to be large enough to be handled by robotic workstations, we used Schirmer strips instead of the microcapillary sample acquisition technique of previous studies. To validate this change, we compared the SELDI-TOF-MS protein profiles of Schirmer strip extracts to those for microcapillary tear samples from a few patients. We detected very complex, but similar, patterns in the tears obtained by both sampling techniques. We found no indication that proteins could be lost during elution from the strips. Furthermore, there was an increase of small protein and peptide peaks in the lower-molecular-mass range detected from Schirmer strip samples compared with microcapillary samples. Because of the nature of this test, in addition to basal tears, the Schirmer strip elutions could include some proteins from impression cytology caused by the strip itself. Further studies will be necessary to show whether the protein profiles of tears obtained by microcapillaries or by Schirmer strips will better discriminate between dry-eye patient and healthy samples. It could be speculated that the inclusion of additional protein and peptide markers from the Schirmer strips might allow better distinction between the different kinds of dry-eye disease (e.g., recognizing meibomian gland

**Figure 6.** SELDI-TOF-MS patterns obtained from microscale analysis of 2 μL tear fluid sample and Schirmer strip of a patient with dry eye on a reversed-phase surface (HS0). The intensity of the peptides and proteins is plotted versus the molecular mass, which is separated in two graphs for high- and low-molecular-mass spectra. Known main protein peaks were detected: lysozyme (1), lipocalin (2), and albumin (3). Although the spectra of both tear samples that were obtained by different sampling techniques were quite similar, there were small differences, such as an increased number of very small peaks in the lower-molecular-mass range.
disease, leading to a more differentiated treatment of the dise-

Nevertheless, the use of Schirmer strips should increase the compliance of physicians. It will not be possible to establish routine clinical use of protein profiling for dry-eye disease if the physician has to use a microcapillary, which requires up to 15 minutes obtaining a few microliters of tears from patients with severe dry eye. If they only have to freeze the Schirmer strips, they are already using to determine the BST, they might consider performing this protein profiling. A possible problem using Schirmer strips could be a requirement to determine the total protein amount used in the analysis. As a step toward automation, we used the first 10 mm of the Schirmer strips without any compensation for the wetting lengths. We found that after total ion current normalization, the protein profiles do not depend significantly on the total protein amount loaded on the arrays. In pilot studies, we investigated the correlation between the total protein amount (which depends at least in part on the wetting length) and the calculated normalization factor. We found no correlation between the protein amount and the normalization factor. Therefore, in this study we used the eluted Schirmer strips without considering the wetting length.

In proteomics, high-throughput analysis has been hampered by the fact that the most common techniques for protein research (i.e., 2-D electrophoresis) are very time consuming and have poor reproducibility. Furthermore, tedious alignment procedures are necessary to ensure that multiple gels are correctly matched for comparison of protein patterns. In comparison to 2-D electrophoresis, the SELDI-TOF-MS technology provides much more precise mass information, allows on-chip protein modifications and protein-protein interaction studies, and enables protein identification. Furthermore, it allows the simple analysis of peptides and is highly sensitive, with just a few microliters of sample needed for analysis. Using our automated approach, it is possible to analyze 96 samples in approximately 4 to 5 hours. In comparison, the 2-D electrophoretic analysis runs overnight, the staining procedure taked several hours, and the quantitative analysis is an extremely time-consuming process involving digitization of gels, peak detection, and matching of the 2-D gels. Furthermore, these steps include several sources for very common systematic errors. Using the SELDI-TOF-MS method, we were able to consistently characterize complex spectra of proteins and peptides. The most complex spectra are found in the range <25 kDa, mirroring the sensitivity of the instrument. Above 30 kDa, the sensitivity of the SELDI-TOFMS is similar to Coomassie staining in electrophoretic assays.

Using different chromatographic surfaces, we have demonstrated another advantage of the SELDI-TOF-MS technology: depending on what array surface is chosen, protein and peptides can be enriched or selectively washed away from the surface due to their biochemical properties. As known from previous studies, lysozyme was significantly decreased in patients with dry eye, and herein we show that lysozyme is best detected on the CM10 arrays. In contrast, lipocalin was detected on the H50 and Q10 chips. Lipocalin is believed to be a heterogeneous protein that may reflect a genetic polymorphism of this protein. Furthermore, some minor variations in the molecular mass of lipocalin could be found in the SELDI-TOF-MS data that may be due to different amounts and to different species of lipids associated with lipocalins.

The primary objective of this study was to find differences in the protein and peptide expression in the tears of patients with dry-eye disease versus healthy control individuals. Many clinical trials in the field of dry-eye rely mainly on the subjective symptoms of patients and do not succeed because of the

### Table 1

<table>
<thead>
<tr>
<th>Size (Da)</th>
<th>CTRL</th>
<th>DRY</th>
<th>P</th>
<th>Name</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,700</td>
<td>3.64</td>
<td>1.42</td>
<td>0.000009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,916</td>
<td>0.18</td>
<td>2.50</td>
<td>0.000003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,027</td>
<td>11.98</td>
<td>4.49</td>
<td>0.000002</td>
<td>Nasopharyngeal carcinoma-associated proline-rich protein</td>
<td>NCAPP4</td>
</tr>
<tr>
<td>4,052</td>
<td>5.62</td>
<td>1.78</td>
<td>0.000048</td>
<td>Proline-rich protein 4</td>
<td>q16378</td>
</tr>
<tr>
<td>4,136</td>
<td>0.53</td>
<td>0.98</td>
<td>0.001155</td>
<td>alpha-1-Antitrypsin, C-terminal fragment</td>
<td>p01009</td>
</tr>
<tr>
<td>5,792</td>
<td>0.55</td>
<td>0.21</td>
<td>0.00014</td>
<td>Proline-rich protein 3</td>
<td>p02814</td>
</tr>
<tr>
<td>10,834</td>
<td>0.12</td>
<td>0.36</td>
<td>0.00018</td>
<td>calgranulin A</td>
<td>p05109</td>
</tr>
</tbody>
</table>

**Figure 7.** *Top:* statistical distribution for the C-terminal fragment of α-1-antitrypsin as a significant biomarker. For this biomarker, the tear sample was fractionated on 30kDa cutoff membrane (YM-30). The 4136-Da protein was found in the retentate. The protein was bound to a CM-10 array and identified by the direct sequencing by MS/MS. **Bottom:** ROC AUC showing the diagnostic power of the artificial neural network. The tear protein patterns of patients with dry eye could be differentiated from healthy controls with an AUC of 0.93.
FIGURE 8. Identification of the 5792-Da peak as PRP3. The 5792-Da protein was detected on CM10 arrays (A). Approximately 400 μL of Schirmer strip elution was passed through a 30-kDa cutoff membrane (YM-30), and the 5792-Da protein was found in the filtrate. The filtrate fraction was then passed through a 3-kDa cutoff membrane (YM-3) and the 5792-Da protein was found in the retentate. The retentate fraction was vacuum dried, redissolved in SDS-PAGE buffer, and loaded onto a 16% tricine gel (B). Coomassie stained bands were extracted and reprofiled on NP20 arrays. Band A corresponded to the 5792-Da protein according to the molecular mass (C). In-solution digestion of the 5792-Da protein yielded two unique ions. MS/MS analysis identified both ions as fragments of PRP3 (D). The theoretical molecular mass of the polypeptide is 5809.76 Da. However, the N-terminal Gln of PRP3 is known to be modified to pyrrolidone carboxylic acid (pyroGlu, 17 Da). Therefore, the predicted molecular mass of the protein is 5792.76 Da, which is very close to the observed experimental mass.
lack of objective parameters to score the disease and to determine the effect of treatment. Thus, the development of tear protein profiling could lead to more clearly defined disease-specific protein biomarkers and biomarker patterns. However, our success depends on the fact that the patients included in this study are well defined. Although we attempted to follow strict and accepted diagnostic criteria for dry eye, we are aware that this issue may be a critical point in this and all other studies searching for dry-eye disease-specific parameters.

Each sample was measured as two replicates and was profiled under a combination of three array types and two laser energies to generate a large amount of data per sample. A total of >1000 peaks per sample were detected across all conditions. Some of the proteins were detected in multiple fractions and under multiple binding conditions. The present study revealed a complex pattern of peptides in tears using the SELDI-TOF-MS ProteinChip Array technology. In the peptide range, seven candidate biomarkers were determined with the help of multivariate discriminant analysis. Using this seven-biomarker panel, an artificial network was able to separate the two groups with sensitivities and specificities of approximately 90%. Some of the biomarkers were downregulated in dry-eye disease, whereas others were upregulated. Thus, both dry-eye disease and the aqueous deficient dry-eye cannot be described as a simple reduction of the total amount of tear proteins.

Two of the seven identified biomarkers are members of the family of PRPs (PRP3 and PRP4) and are decreased in dry-eye samples. PRPs are thought to mediate protective functions in the eye, such as modulation of the microflora. PRP3 is a highly abundant protein in tears. PRP4 is abundantly expressed in lacrimal gland where it is found in the acinar cells, but not in the intralobular ducts. Also found in the submandibular gland, the parotid gland, and the sublingual glands (Swiss-PROT Database; http://www.expasy.org/ provided at no fee to academic users by the Swiss Institute of Bioinformatics, Geneva, Switzerland). In contrast to other PRPs, PRP4 is expressed in the lacrimal acinar cells and other anterior exocrine glands. Therefore, PRP4 should be a useful marker for human lacrimal gland acinar cell function. Furthermore, the members of the PRP family PRP are suggested to act as neuroprotectors and are considered to be neurotrophic factors. In addition, a variant of PRP4 was identified (NCAPP4). This variant was also decreased in patients with dry eye as PRP4 counterpart and the percentage of patients having both variants was less than half in patients with dry eye compared with controls. The biological importance underlying genetic polymorphism of PRP4 and NCAPP4 is still unknown and further studies will be needed to shed light on functions of proteins of the PRP family.

The 10.8-kDa biomarker, identified as calgranulin A, was increased in patients with dry eye. Calgranulin is expressed in...
neutrophils and is thought to be an inflammatory marker. Recent findings suggest that these molecules are important in fighting inflammation. The 4136-Da biomarker was identified as the C-terminal fragment of α-1-antitrypsin. As an inhibitor of proteases, the major physiological function of this molecule is the protection of tissues against proteolytic destruction. In our study, we found an increase of the C-terminal fragment in patients with dry eye, which could indicate an increase in antitrypsin molecules which were inactivated due to cleavage at the reactive bond by enzymes such as metalloproteinases. However, cleavage of an enzyme inhibitor does not result in only enzyme inactivation, but can also lead to new functions such as being a powerful proinflammatory activator of neutrophils. Two biomarkers at 3700 and 3916 Da remain to be identified at this time. In summary, some of the 18 biomarkers in our multimarker panel that may provide protective functions are downregulated in patients with dry eye, whereas others, including some inflammatory markers, are upregulated.

Although this was a successful proof-of-principle study, it is important to keep in mind that these biomarkers must be further tested in validation and multisite studies. In future studies, we will investigate the relevance of these peptides in dry-eye disease and their roles in the pathogenesis of the disease. It will also be investigated how these identified biomarkers are distributed in different forms of dry-eye and if these biomarkers can lead to objective follow-up parameters in clinical studies of new approaches in dry-eye treatment. In conclusion, this is the first study of which we are aware that demonstrates the use of SELDI-TOF-MS ProteinChip Array technology for high-throughput biomarker discovery in tear fluid. This approach is not limited to ocular surface diseases, but may be applicable to systemic diseases as well. Using complex computational approaches (e.g., multivariate statistics, decision trees, or artificial neural networks), protein biomarker patterns can be detected and may lead to powerful diagnostic tools, to advances in predictive medicine and a discovery of new therapeutic targets for drug development.

References