Human Cornea Proteome: Identification and Quantitation of the Proteins of the Three Main Layers Including Epithelium, Stroma, and Endothelium

Thomas F. Dyrlund,† Ebbe Toftgaard Poulsen,‡ Carsten Scavenius,†‡ Camilla Lund Nikolajsen,†‡ Ida B. Thøgersen,†‡ Henrik Vorum,§ and Jan J. Enghild*†‡

†Interdisciplinary Nanoscience Center (iNANO) and Department of Molecular Biology, Aarhus University, Denmark
‡Center for Insoluble Protein Structures (inSPIN), Aarhus University, Denmark
§Department of Ophthalmology, Aalborg Hospital, Aarhus University Hospital, Denmark

Supporting Information

ABSTRACT: Diseases of the cornea are common and refer to conditions like infections, injuries and genetic defects. Morphologically, many corneal diseases affect only certain layers of the cornea and separate analysis of the individual layers is therefore of interest to explore the basic molecular mechanisms involved in corneal health and disease. In this study, the three main layers, including, the epithelium, stroma and endothelium of healthy human corneas were isolated. Prior to analysis by LC−MS/MS the proteins from the different layers were either (i) separated by SDS-PAGE followed by in-gel trypsinization, (ii) in-solution digested without prior protein separation or, (iii) in-solution digested followed by cation exchange chromatography. A total of 3250 unique Swiss-Prot annotated proteins were identified in human corneas, 2737 in the epithelium, 1679 in the stroma, and 880 in the endothelial layer. Of these, 1787 proteins have not previously been identified in the human cornea by mass spectrometry. In total, 771 proteins were quantified, 157 based on in-solution digestion and 770 based on SDS-PAGE separation followed by in-gel digestion of excised gel pieces. Protein analysis showed that many of the identified proteins are plasma proteins involved in defense responses.

KEYWORDS: human cornea, epithelium, stroma, endothelium, proteomics, label free quantitation, TripleTOF 5600

1. INTRODUCTION

The human cornea is the transparent tissue of the eye that refracts and transmits light to the lens and retina. It has a diameter of about 11.5 mm and a thickness of about 0.6 mm. The cornea is composed of five main layers, (i) the epithelium (~50 μm thick), (ii) the Bowman’s membrane (the most superficial, modified part of stroma, ~17 μm thick), (iii) the stroma (~450 μm thick), (iv) the Descemet’s membrane (~5–20 μm thick), and (v) the endothelium (~5 μm thick).1,2 The cornea protects the eye against damaging UV radiation and toxic agents,3 accounts for most of the eye’s focusing power, and provides biomechanical stability and structural resiliency.4 Diseases of the cornea are common and together with injuries or genetic defects cause numerous visual impairments worldwide. Corneal diseases include inflammation (keratitis), deposition of insoluble and opaque macromolecules (corneal dystrophies), disruption of the epithelial layer (corneal ulcers), excessive ingrowth of blood vessels into the cornea (corneal neovascularization), loss of corneal transparency due to cornea swelling (Fuchs’ dystrophy) and structural changes and thinning (keratoconus and keratoglobus). Despite considerable research, the causes and molecular mechanisms of the listed diseases still remain unclear.

Previous studies have focused on identifying the proteins of the cornea, and a total of 138 unique Swiss-Prot proteins (corrected for duplicates) were identified in the first human cornea proteome study.5 Since then, advances in mass spectrometry and separation methods have facilitated the identification of more proteins resulting in 1555 unique Swiss-Prot proteins (corrected for duplicates) being identified in human corneas,6 341 unique proteins in rabbit corneas7 and 2173 unique proteins in mouse corneas.8 However, in order to understand the biology and the molecular basis for corneal disorders, it is of interest to perform quantitative deep proteome analyses of individual layers of the human cornea. With the high acquisition speed, resolution and sensitivity of the TripleTOF 5600 mass spectrometer (AB Sciex), it is

Received: April 13, 2012
Published: June 14, 2012
possible to identify hundreds of proteins in a single nano-HPLC run while performing extracted ion chromatogram (XIC) label-free quantification on the proteins.

The purpose of the present study was to identify and quantify the proteins of the epithelium, the stroma and the endothelium of normal human cornea using the TripleTOF 5600 mass spectrometer. The ability to perform quantitative analysis provides valuable clues regarding the protein profile required to maintain corneal homeostasis. This information may be employed as a reference to explore the basic molecular mechanisms involved in corneal health and diseases. We identified 3250 unique corneal proteins, 2737 in the epithelium, 1679 in the stroma and 880 in the endothelium layer. Of these, 771 proteins were quantified.

2. EXPERIMENTAL PROCEDURES

2.1. Human Corneas

The first group of analyses were conducted on five human donor eyes obtained post-mortem from three humans at Aarhus University Hospital. The ages ranged from 66 to 93 years (mean 81.3 years), with 3 eyes from females and 2 from males. The second group of analyses were conducted on five human donor eyes obtained post-mortem from three humans, age 77 to 84 years (mean 81.3 years), with 3 eyes from females and 2 from males. All eyes were collected within 48 h post-mortem, carefully rinsed in isotonic saline and the freshness and clearness of the cornea was controlled under a dissection microscope after which the eyes were stored at −80 °C. Permissions were obtained from the local ethical committee according to the Declaration of Helsinki.

2.2. Tissue Homogenization

Donor eyes were thawed and the corneal epithelium removed by 70% ethanol treatment and gentle scraping using a surgical blade. The remaining part of the cornea was carefully excised avoiding the corneal limbus and washed briefly using milli-Q water. The endothelium was then removed and collected in the same manner as described for the epithelium. The three different corneal layers were lyophilized and separately pooled, homogenized (IKA Werke Ultra-Turrax T8) in liquid nitrogen and transferred to 15 mL Sarstedt tubes by adding 4 mL ice cold milli-Q water. The samples were instantly frozen in liquid nitrogen and lyophilized overnight to a fine powder.

2.3. Sample Preparation for LC–MS/MS – SDS-PAGE Separation

From the first group of corneas, the epithelium, stroma and endothelium powders were boiled for 5 min in SDS sample buffer (240, 600 and 240 μL, respectively) containing 50 mM DTT, separated by SDS-PAGE using 10% (w/v) gels9 and stained using Coomassie blue (RAPIDstain, G-Biosciences). Each lane was cut into 8 equal size slices, these were washed 3 times in milli-Q water and incubated two times 15 min in 130 μL 50% acetonitrile, dehydrated in 130 μL acetonitrile for 15 min and equilibrated in 150 μL 0.1 M NH4HCO3 for 5 min before 150 μL acetonitrile was added. After 15 min the supernatants were removed and the gel pieces were lyophilized for 20 min. In-gel digestions were performed by incubating the gel pieces with 400 ng sequencing grade modified trypsin (Promega) in 50 mM NH4HCO3 at 37 °C for 16 h. The resulting peptides were desalted using C18 StageTips (Thermo Scientific) and stored at −20 °C before LC–MS/MS analysis.

The SDS insoluble pellets were washed in cold milli-Q water to remove residual SDS sample buffer, frozen in liquid nitrogen and lyophilized overnight. One milliliter of 0.66 M CNBr in 70% TFA was added for each milligram of protein to the samples before incubation overnight at 23 °C. After lyophilizing, the samples were resuspended in 0.2 M Tris-HCl pH 8.3 containing 8 M Urea, reduced in 15 mM DTT for 1 h at 23 °C and alkylated for 1 h in 30 mM iodoacetamide. The samples were then diluted 5 times with 0.1 M Tris-HCl, pH 8.3 and digested overnight with 1:50 w/w sequencing grade modified trypsin (Promega) at 37 °C for 16 h. The resulting peptides were desalted using C18 StageTips (Proxeon) and stored at −20 °C before LC–MS/MS analysis.

2.4. Sample Preparation for LC–MS/MS – In-solution Digest and Cation Exchange Chromatography

From the second group of corneas, the epithelium, stroma and endothelium powder were incubated with 1 mL 0.66 M CNBr in 70% TFA for each mg of protein overnight at 23 °C. Approximately, 200 μL of the stroma sample, and everything from the two other samples, were lyophilizing and resuspended in 8 M Urea 0.2 M Tris-HCl pH 8.3, reduced in 15 mM DTT for 1 h and alkylated in 30 mM iodoacetamide for 1 h. The samples were then diluted 5 times with 0.1 M Tris-HCl pH 8.3 and digested overnight with 1:50 w/w sequencing grade modified trypsin (Promega) at 37 °C for 16 h in a final volume of 1 mL. The samples were desalted using Poros 50 R2 reverse phase column material (PerSeptive Biosystems) packed in GE_loader Tips (Eppendorf)10 before 30 μL from each of the stroma and epithelium samples and 100 μL from the endothelium sample were removed and stored at −20 °C. The rest of the samples (3 × 50 μL from each of the stroma and epithelium samples and 3 × 300 μL from the endothelium sample) were dried in a speedvac (Savant), resuspended in buffer A (0.5% formic acid and 5% acetonitrile) and separated by strong cation exchange (SCX) using a PL-SCX 100 Å 5 μm 20 × 2.1 mm column (Higgins Analytical) connected to an Ettan LC system (Amersham Biosciences) equilibrated in buffer A. The column was operated at 23 °C at a flow rate of 100 μL/min. The peptides were eluted using a 1% B/min linear gradient from buffer A to buffer B (0.5% formic acid, 5% acetonitrile containing 1 M NaCl). A total of 13 fractions were collected and individual samples from the same layers were pooled from three cation exchange runs. The samples were concentrated in a vacuum centrifuge and desalted using Poros 50 R2 material packed in GE_loader tips, dried in a vacuum centrifuge, resuspended in 33 μL 5% formic acid and stored at −20 °C.

2.5. LC–MS/MS Analysis

LC–MS/MS analyses were performed on an EASY-nLC II system (Thermo Scientific) connected to a TripleTOF 5600 mass spectrometer (AB Sciei) equipped with a NanoSpray III source (AB Sciei) and operated under Analyst TF 1.5.1 control. The trypsin digested samples were dissolved in 0.1% formic acid, injected, trapped and desalted isocratically on a ReproSil-Pur C18-AQ column (5 μm, 2 cm × 100 μm ID; Thermo Scientific) after which the peptides were eluted from the trap column and separated on an analytical ReproSil-Pur C18-AQ capillary column (3 μm, 10 cm × 75 μm ID; Thermo Scientific) connected in-line to the mass spectrometer at 250 nL/min using a 80 min gradient from 5% to 35% phase B (0.1% formic acid and 90% acetonitrile). An Information dependent acquisition method was employed to automatically run
2.6. Protein Identification

The collected MS files were converted to Mascot generic format (MGF) using the AB SCIEX MS Data Converter beta 1.1 (AB SCIEX) and the “proteinplot MGF” parameters. The peak lists were used to interrogate the Swiss-Prot (version 2011_10, 532,792 sequences) Homo sapiens (20,331 sequences) database using Mascot 2.3.02 (Matrix Science). For in-gel digests, the search parameters were allowing one missed trypsin cleavage site and propionamide as a fixed modification. For the solution digest, a combination of CNBr and Trypsin was employed allowing one missed cleavage with carbamidomethyl as a fixed modification. Oxidation of methionine, and hydroxylation of proline residues were entered as variable modifications. The mass accuracy of the precursor and product ions were 10 ppm and 0.6 Da and the instrument setting was hydroxylation of proline residues were entered as variable.

2.7. Protein Quantitation

All raw MS files were processed using Mascot Distiller 2.4.2.0 (Matrix Science). The MS data obtained by the analysis of a single gel lane were merged into a multi file project using the default settings from the ABSciex_5600.opt file except that the MS/MS Peak Picking “Same as MS Peak Picking” was deselected and “Fit method” was set to “Single Peak”. The CNBr+Trypsin in-solution digests were processed separately but using the same settings as described above. After peak picking all scans, a Mascot search was performed using the same settings as for protein identification above except that the default average [MD] quantitation protocol was selected using a significance threshold at 0.01, number of peptides used for quantitation was 3, matched rho was 0.8, XIC threshold was 0.3 and isolated precursor threshold was set at 0.7. This label-free quantitation protocol relies on the average MS signal response for the three most intense tryptic peptides for each protein. When calculating protein amount based on total XIC area for matches to the three most intense peptide sequences, Mascot Distiller failed to recognize cases where two different modification states had the same precursor m/z and elution time and hence resolve to the same XIC. This caused double counting of XICs in the original report, leading to errors in the relative protein amounts. In our data, any such duplicates were found by manual inspection and eliminated. The average relative protein amount and standard deviation was calculated for all proteins quantified in a minimum of 3 samples. The complete quantitation method with all settings is provided in XML format as Supporting Information 8.

3. RESULTS

Human corneas were separated into the epithelial, stromal and endothelial layers. Since the stroma mostly consists of collagen, this separation reduced the amount of collagen present in the cellular layers and therefore, increased the number of proteins that could be identified and quantified. The proteins of the different layers were separated by 1D SDS-PAGE (Supporting Information 7) followed by in-gel trypsin digestion, or a combination of CnBr treatment and trypsin digestion followed by strong cation exchange separation. Proteins were then identified by MS/MS analysis and searched in Mascot against Homo sapiens sequences in Swiss-Prot. All mascot results were parsed with in-house developed software (MS Data Miner v. 1.0) and spectra for single-peptide-based protein identifications were manually verified. A total of 3250 unique proteins were identified, 2737 in the epithelium, 1679 in the stroma and 880 in the endothelial layer. With the strict criteria employed (significance threshold (p) of 0.01, ions score cutoff of 30 and the manual validation), false positive protein identifications in the data set were kept to a minimum. The lower number of unique proteins identified in the endothelium compared to the epithelium is likely caused by the lower amount of total protein extracted from this single layer of cells. Similarly, only four SDS-PAGE separations, loaded less compared to the epithelial lanes, were performed on the endothelial layer. The lists of proteins identified in the individual layers and the information reported by Mascot is provided as supplementary data (Supporting Information 1) along with assigned MS/MS spectra for all single-peptide-based protein identifications (Supporting Information 4–6).

A total of 609 unique proteins were identified in all three layers and 13, 18 and 51% were unique to the endothelium, stroma and epithelium, respectively (Figure 1). A total of 836 unique proteins were identified by all three methods and 1388 (43%), 29 (1%) and 222 (7%) were only identified by SDS-PAGE followed by in-gel trypsinization, in-solution digest without prior protein separation or in-solution digest followed by cation exchange chromatography respectively (Figure 2).

![Figure 1](https://dx.doi.org/10.1021/pr300358k.i/J. Proteome Res. 2012, 11, 4231−4239)
Corneal diseases are often restricted to a single layer and identification and quantification of proteins present in the different layers is therefore of interest to understand the molecular mechanisms of disease. In this study, healthy human corneas were divided into the three main layers: epithelium, stroma, and endothelium. Not all of the proteins of the cornea, especially the collagens of the stroma, are soluble in SDS. Thus, it is not possible to resolve all proteins by SDS-PAGE. Quantifying proteins based on the gel analysis will therefore introduce a bias in the relative amounts and to overcome this we also quantified the proteins based on an in-solution digest where no separation was performed on the protein level. The intensities of the identified proteins were calculated based on XIC using the Average Protocol in Mascot Distiller. This label-free quantification protocol relies on the average MS signal response for the three most intense peptides for each protein. A minimum of three replicate runs were performed for all samples to account for variation. The relative abundance of proteins quantified in a minimum of three MS runs was calculated as the average MS signal for all quantified proteins in the layer. A total of 771 proteins were quantified, 157 based on in-solution digest and 770 based on gel separations. Total proteins quantified for the individual layers are epithelium, 110 [663]; stroma, 44 [342] and endothelium, 59 [140].

Figure 2. Venn diagram showing the number of overlapping proteins identified in human cornea by (i) SDS-PAGE followed by in-gel trypsinization, (ii) in-solution digest without prior protein separation or (iii) in-solution digest followed by cation exchange chromatography. A total of 2958 unique proteins were identified by SDS-PAGE separation, 1010 by in-solution digest without prior protein separation and 1729 by in-solution digest followed by cation exchange chromatography.

In 2005, the first human cornea proteome study was published identifying a total of 138 unique Swiss-Prot proteins (corrected for duplicates). Only 4 of these proteins (Apolipoprotein B-100, Deoxynucleobase-1-like 2, E3 ubiquitin-protein ligase NEDD4 and Integrin alpha-1) were not found in the present study. In a more recently published study a total of 1555 unique Swiss-Prot proteins (corrected for duplicates) were identified and only 115 of these were not detected in the present study. Of the 115 proteins identified, 74 were only identified in one MS analysis in the previous study with Mascot scores ranging from 44 to 229 (mean, 71). Taken together, the present study has identified 1787 proteins not previously identified in the human cornea by mass spectrometry.

Protein Quantification

Not all of the proteins of the cornea, especially the collagens of the stroma, are soluble in SDS. Thus, it is not possible to resolve all proteins by SDS-PAGE. Quantifying proteins based on the gel analysis will therefore introduce a bias in the relative amounts and to overcome this we also quantified the proteins based on an in-solution digest where no separation was performed on the protein level. The intensities of the identified proteins were calculated based on XIC using the Average Protocol in Mascot Distiller. This label-free quantification protocol relies on the average MS signal response for the three most intense peptides for each protein. A minimum of three replicate runs were performed for all samples to account for variation. The relative abundance of proteins quantified in a minimum of three MS runs was calculated as the average MS signal for all quantified proteins in the layer. A total of 771 proteins were quantified, 157 based on in-solution digest and 770 based on gel separation. The large difference in the number of proteins quantified was due to the lower complexity of samples separated by PAGE. For the in-solution digest, a total of 110 proteins were quantified in the epithelium, 44 in the stroma and 59 in the endothelium layer. For the gel separation, the numbers are 663, 342, and 140 respectfully as illustrated (Figure 3). A total of 18 unique proteins were quantified in all three layers based on the solution digests and 95 proteins from the PAGE. In total, collagens account for about 50% of the proteins in the stroma layer, 30% in the endothelium layer and only about 2% in the epithelium layer (Table 1 and Supporting Information 2).

Genome ontology (GO) classification of cornea proteins

The GO annotations for the identified corneal proteins were compared to the quantified proteins based on a series of summary categories (GO slim categories) including, Cellular Component, Molecular Function and Biological Process. The relative amount of proteins in the summary categories were calculated based on the quantitation data and used to refine the distribution profile of the individual proteins. Since proteins are annotated with several GO terms, they can be represented in multiple summary categories, and the percentage values were therefore normalized to a total value of 100%. Only about 10% of the proteins identified in the stroma have GO terms associated to the extracellular region, however, these make up more than half of the protein amounts (Figure 4). Likewise, the identified cytoplasmic proteins accounted for 27% of the identified proteins but for only 5% based on the quantitative data. These analyses revealed that the categories with the highest number of proteins did not necessarily also contain the highest relative protein amount. When comparing diseased and control tissues, it is thus relevant to consider the amount associated to a GO category as well as the number of proteins. As expected, the gene ontology distributions are different for the three corneal layers. However, categories like protein binding, structural molecule activity and regulation of biological process are high in all three layers (Supporting Information 1 and 3).

4. DISCUSSION

Corneal diseases are often restricted to a single layer and identification and quantification of proteins present in the different layers is therefore of interest to understand the molecular mechanisms of disease. In this study, healthy human corneas were divided into the three main layers: epithelium,
Biased toward proteins that are compatible with SDS-PAGE. Protein quantitation based on SDS-PAGE and in-solution digestion is both influenced by the dynamic detection range of the mass spectrometer. However, in-solution digests are not biased toward particular properties of the proteins and provide a more accurate picture of the relative amounts. This can be observed by the large variation between the relative amounts of the most abundant proteins, especially the collagens, in the stromal and endothelial layers when comparing in-solution and in-gel digestions following SDS-PAGE (Supporting Information 2). This effect may explain that only 44 stromal proteins were quantified based on the in-solution digest compared to 342 proteins based on the SDS-PAGE separations. With the strict search criteria employed, manual verification of single-peptide-based protein identifications and reporting only Swiss-Prot hits, there should only be a very limited number of false positive protein identifications. Of the 771 proteins quantified, stra and endothelium. The extracted proteins were subsequently identified and quantified by tandem mass spectrometry following separation by SDS-PAGE or strong cation exchange chromatography. A total of 3250 unique Swiss-Prot annotated proteins were identified in the three corneal layers, 2737 proteins the epithelium, 1679 proteins in the stroma and 880 in the endothelial layer, where 609 proteins (19%) were detected in all three layers. Additionally, 1398 proteins were unique to the epithelial, 304 to the stromal and 111 to the endothelial layer. A total of 770 proteins were quantified following SDS-PAGE. Of these, 95 (12%) were quantified in all three layers. For the in-solution digests where no protein separation was performed, 157 proteins were quantified and 18 (11%) were quantified in all layers. Some proteins are lost during SDS-PAGE separation if they are insoluble in SDS sample buffer, very large or small. Quantification based on SDS-PAGE separation is therefore biased toward proteins that are compatible with SDS-PAGE.  

Table 1. Thirty Most Abundant Proteins of the Three Layers of the Human Cornea Including Relative Abundance and Standard Deviation$^a$

<table>
<thead>
<tr>
<th>Epithelium</th>
<th>%</th>
<th>Stroma</th>
<th>%</th>
<th>Endothelium</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Keratin, type I cytoskeletal 12 8.1 ± 0.7 1. Collagen-alpha-1(VI) chain 20.0 ± 2.0 1. Transforming growth factor-beta-induced protein Ig h3 36.8 ± 1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Keratin, type II cytoskeletal 5 7.3 ± 0.3 2. Transforming growth factor-beta-induced protein Ig h3 17.6 ± 3.3 2. Collagen-alpha-1(VII) chain 5.8 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Keratin, type II cytoskeletal 6A 6.0 ± 0.5 3. Collagen-alpha-2(1) chain 17.2 ± 3.1 3. Collagen-alpha-2(VII) chain 4.4 ± 0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Keratin, type II cytoskeletal 3 4.7 ± 0.5 4. Decorin 5.1 ± 0.4 4. Collagen-alpha-4(VI) chain 4.2 ± 0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Keratin, type II cytoskeletal 2 oral 4.7 ± 0.3 5. Collagen-alpha-3(VI) chain 4.7 ± 0.9 5. Collagen-alpha-3(V) chain 4.0 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Keratin, type I cytoskeletal 15 3.8 ± 0.4 6. Lumican 3.5 ± 0.5 6. Collagen-alpha-1(1) chain 3.9 ± 0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Keratin, type II cytoskeletal 4 3.3 ± 0.1 7. Serum albumin 3.5 ± 0.5 7. Keratocan 3.5 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Keratin, type I cytoskeletal 19 3.1 ± 0.8 8. Keratocan 3.4 ± 0.9 8. Collagen-alpha-5(V) chain 3.0 ± 0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Aldehyde dehydrogenase, dimeric NADP-prefering 3.1 ± 0.4 9. Collagen-alpha-1(VI) chain 2.2 ± 0.5 9. Clusterin 2.5 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Alpha-enolase 3.0 ± 0.4 10. Ig gamma-1 chain C region 1.9 ± 0.1 10. Thrombospondin-4 2.3 ± 0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Keratin, type I cytoskeletal 14 2.9 ± 0.3 11. Collagen-alpha-2(VI) chain 1.7 ± 0.2 11. Collagen-alpha-2(I) chain 2.3 ± 0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Keratin, type II cytoskeletal 8 2.5 ± 0.2 12. Collagen-alpha-1(XIII) chain 1.6 ± 0.6 12. Ig gamma-1 chain C region 2.2 ± 0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Histone H4 2.1 ± 0.4 13. Ig kappa chain C region 1.6 ± 0.3 13. Serum albumin 1.9 ± 0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Histone H2B type I-C/E/F/G/I 1.4 ± 0.1 14. Ig gamma-3 chain C region 1.5 ± 0.3 14. Ig gamma-3 chain C region 1.7 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Transketolase 1.4 ± 0.2 15. Ig gamma-2 chain C region 1.3 ± 0.2 15. Ig gamma-4 chain C region 1.4 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Actin, cytoplasmic 1 1.3 ± 0.1 16. Collagen-alpha-2(V) chain 1.3 ± 0.2 16. Ig gamma-2 chain C region 1.3 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Histone H2B type I-B 1.3 ± 0.1 17. Ig gamma-4 chain C region 1.2 ± 0.2 17. Fibulin-5 1.2 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Collagen-alpha-1(I) chain 1.1 ± 0.1 18. Collagen-alpha-1(V) chain 1.1 ± 0.2 18. Ig kappa chain C region 1.2 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Annexin A2 1.1 ± 0.6 19. Aldehyde dehydrogenase, dimeric NADP-prefering 0.9 ± 0.2 19. Apolipoprotein D 0.9 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Heat shock protein beta-1 1.1 ± 0.1 20. Biglycan 0.6 ± 0.1 20. Decorin 0.9 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Transforming growth factor-beta-induced protein Ig h3 1.0 ± 0.3 21. Alpha-enolase 0.6 ± 0.1 21. Collagen-alpha-3(VI) chain 0.9 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Serum albumin 1.0 ± 0.1 22. Collagen-alpha-1(III) chain 0.5 ± 0.1 22. Ig lambda-2 chain C regions 0.8 ± 0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Keratin, type I cytoskeletal 13 0.9 ± 0.1 23. Apolipoprotein D 0.4 ± 0.1 23. Immunoglobulin lambda-like polypeptide 5 0.7 ± 0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Glyceroldehyde-3-phosphate dehydrogenase 0.7 ± 0.0 24. Clusterin 0.4 ± 0.3 24. Lumican 0.7 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. L-lactate dehydrogenase A chain 0.7 ± 0.1 25. Alpha-1-antitrypsin 0.4 ± 0.1 25. Myocilin 0.6 ± 0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Protein S100-A4 0.7 ± 0.1 26. MAM domain-containing protein 2 0.3 ± 0.0 26. Apolipoprotein E 0.6 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Pyruvate kinase isozymes M1/M2 0.6 ± 0.0 27. Prolargin 0.2 ± 0.0 27. Serine protease HTRA1 0.6 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Retinal dehydrogenase 1 0.6 ± 0.1 28. Vimentin 0.2 ± 0.0 28. Alpha-enolase 0.5 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Elongation factor 1 alpha 1 0.6 ± 0.1 29. Ig alpha-1 chain C region 0.2 ± 0.0 29. Collagen-alpha-1(III) chain 0.5 ± 0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Ferritin heavy chain 0.6 ± 0.0 30. Alpha-1-antichymotrypsin 0.2 ± 0.0 30. Collagen-alpha-2(VI) chain 0.4 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Relative abundance of the proteins were calculated as the average MS intensity for the three most intense peptides for each protein divided by the total sum of the average signal for all quantified proteins in the layer. The full list is available as Supporting Information 2.
S73 (74%) were identified in all three digest and separation methods, 25 (3%) failed to be identified by in-solution digestion followed by cation exchange chromatography, 112 (15%) were not identified by in-solution digestion without prior protein separation and the remaining 61 (8%) were only identified by SDS-PAGE separation followed by in-gel trypsinization.

**Corneal Keratins**

In the human cornea data set, we identified a total of 36 keratins of which 30 belong to type I and type II epithelial keratins (Type I: K9, K10, K12−K19, K24, Type II: K1−K6a, K6b, K6c, K7, K8, and K71−79) and 6 belong to the type I and type II hair keratins (K31, K36, K81, K82, K83, and K84). Since keratins are naturally present in the epithelial layers of the cornea, a general exclusion of keratins from the data set cannot be performed. Recently, immunohistochemical studies of corneal keratins showed that K1, K3−K6, K8, K13−K16, K18 and K19 are present within the human cornea. Furthermore, K3 and K12 are believed to be cornea specific, and it has been shown that mutations in either of these keratins resulted in the corneal dystrophy known as Meemann’s corneal dystrophy. By comparing our data to the previous studies we are confident that the following 15 keratins are present in the human cornea: K1, K3−K6a, K6b, K6c, K8, K12−K16, K18, and K19. The other 20 keratins: K2, K7, K9, K10, K17, K24, K71−K79, K31, K36, K81, K82, K83, and K84 are uncertain. Of the confident keratins, K12, K5, K6a, K3, K76, K4, K19, K14, K8, and K13 are found to be within the top 30 quantified proteins of the epithelial layer and of the uncertain keratins, K76 account for more than 4.7% of the quantified proteins (Table 1). However, the absence of keratins within the top 30 proteins for the stromal and endothelial layers suggests that K76 may be a true corneal epithelium keratin as well.

The epithelial keratins accounts for more than 47% of the total abundance of quantified proteins in the epithelial layer (Table 1 and Supporting Information 2). The abundance of these keratins is in agreement with what has been observed using immunohistochemistry. Additionally, one-third of the protein content in the epithelial layer has been reported to be the K3/K12 cornea specific keratin pair, supporting that keratins are abundant in the corneal epithelium.

**Bowman’s and Descemet’s Membranes**

Most of the quantified extracellular or membrane proteins were derived from the stromal and endothelial layers (87% based on the solution digest). In contract only 9% of the quantified proteins were extracellular or membrane proteins in the epithelial layer (Figure 5). Similarly, type IV and VIII collagens,

![Figure 4](https://example.com/f4.png)

**Figure 4.** Gene Ontology distributions for summary categories based on cellular component of stroma proteins. The IDs bars represent the percentage of identified proteins in the summary category and the amount bars represent the percentage of quantified proteins accounting for the relative amount of each protein. Distributions for all layers and ontologies are provided in Supporting Information 3.

![Figure 5](https://example.com/f5.png)

**Figure 5.** Quantified proteins for the three different cornea layers were classified into 3 broad cellular component categories. The bars represent the percentage of quantified proteins accounting for the relative amount of each protein.

localized in the basement membranes, were present in high amounts in the endothelium together with noncollagenous basement membrane proteins such as fibulins, fibronectin, vitronectin, laminin and nidogen. It is therefore likely, that the endothelial layer was removed with the Descemet’s membrane similar to what occur during the Automated Endothelial Keratoplasty transplant procedure used for treating of Fuchs’ dystrophy. Bowman’s membrane specific proteins such as collagen type VII and XVIII and Basement membrane-specific heparin were found in both the epithelium and stroma and it is therefore not possible to determine if the Bowman’s membrane was removed and analyzed together with the epithelial or stromal layer.

**Plasma Proteins**

Although the cornea is avascular, 1050 plasma proteins were identified when compared to a reference set of 1805 human plasma proteins (only hits which could be correlated to Swiss-Prot proteins were compared). Among the most abundant plasma proteins, albumin, immunoglobulin’s, serotransferrin, alpha-1-antitrypsin, complement component 3 and 9 and haptoglobin were quantifiable indicating that they are also among the most abundant proteins in the human cornea. Hemoglobin subunit alpha, beta and zeta were also detected together with haptaglobin, a hemoglobin binding protein, and heme-binding protein 1 and 2. Hemoglobin has been shown to be expressed in developing mouse corneas and it is hypothesized that besides oxygen transport, hemoglobin may also be involved in apoptosis development, nuclear degeneration/organelle loss and stem cell differentiation. It is assumed that there is an import of plasma proteins into the cornea, but it is still unclear if the import is selective or passive and how important it is. Studies have shown that plasma proteins like alpha-1-antichymotrypsin, alpha-2-antiplasmin, alpha-1-antitrypsin, coagulation factor X and prothrombin are also synthesized in the human cornea but the function of the classical plasma proteins in the cornea remains unclear.
Angiogenin, a stimulator of new blood vessel formation, was also detected in the human cornea. Angiogenin is likely imported from the blood through the ciliary arteries and the presence of several angiogenesis inhibitors (pigment epithelium-derived factor, serpin B5, alpha-1-antitrypsin and antithrombin III) is likely the reason why no vascularization is observed as an imbalance between angiogenic and antiangiogenic factors may result in corneal neovascularization.

Tear Film and Aqueous Humor

Of the identified cornea proteins, 823 have also been identified in human tears when compared to a reference set of 990 proteins and 252 were found in human aqueous humor when compared to a reference set of 361 proteins (Figure 6). Of the above-discussed plasma proteins, serrotransferrin was not found in the reference sets of tear film or aqueous humor. Haptoglobin, angiogenin and pigment epithelium-derived factor were not found in tear film and serpin B5 was not identified in aqueous humor. Of the 990 human tear proteins, 529 (53%) are found in the plasma protein reference set and of the 361 aqueous humor proteins, 296 (82%) are in the plasma protein reference set. Among these proteins are proteins of the complement system that cause neutralization and opsonization of pathogens. C-reactive protein is an acute phase protein generally associated with inflammation and has previously been identified in the epithelium of corneas with lattice dystrophy. Here, C-reactive protein was detected in all layers of healthy human corneas.

Plasma kallikrein has not previously been identified in the cornea or discussed in relation to corneal function, but was recently identified in human tear and the aqueous humor. Together with the identified coagulation factor XII and kinogen-1, a precursor of high molecular weight kininogen, plasma kallekrein form the initial part of the kallikrein-kinin system. The system is able to activate coagulation through the contact/intrinsic pathway, promote inflammation and regulate vascular permeation. The kallikrein-kinin system has been identified in other parts of the eye and seems to be implicated in the development of retinal edema in diabetic retinopathy.

Several other components of the coagulation cascade are present in the cornea. The fibrinogen chains, pro-thrombin, coagulation factor X, coagulation factor VIII and coagulation factor XIII-A, and the two coagulation factor XIII-A related enzymes, transglutaminase-1 and -2 were all identified. Immunohistochemistry of corneal tissue has previously revealed the presence of transglutaminase-1 and -2, but failed to identify coagulation factor XIII-A. It has been shown that coagulation factor XIII-A is present in tears with a higher concentration after corneal transplantation suggesting that it may be involved in corneal wound healing. The study also indicated a possible correlation between coagulation factor XIII-A and neovascularization of the cornea. In contrast to in plasma, the regulatory coagulation factor XIII-B was not identified in our study. In plasma, this subunit is in large excess over coagulation factor XIII-A and most all coagulation factor XIII-A is found in complex with coagulation factor XIII-B, thus hindering its activation.

Previous studies have shown that coagulation with consequent fibrin deposition is involved in corneal wound healing. These studies also showed that balancing fibrin deposition and degradation is essential for maintaining corneal function. The fibrinolytic system is especially important for obtaining this balance. Several well-known regulators of coagulation including plasminogen, α2-antiplasmin, antithrombin-III and carboxypeptidase B2 were identified together with two proteins, β2-glycoprotein 1 and heparin cofactor 2, which have not previously been characterized in the cornea. β2-glycoprotein 1 is a plasma protein with anticoagulant and procoagulant functions as well as possible functions in modulation...
of angiogenesis.\textsuperscript{46,47} \(\beta_2\)-glycoprotein 1 has been identified in the vitreous fluid with higher levels associated with proliferative diabetic retinopathy.\textsuperscript{48} Heparin cofactor 2 is also a plasma protein and is able to inhibit thrombin. This inhibition is more efficient when heparin cofactor 2 is bound to the glycosaminoglycan dermatan sulfate which is one of the major glycosaminoglycans of the cornea.\textsuperscript{49,50}

Proteins with antimicrobial activities are crucial for the defense against pathogen infections in the cornea and several proteins with known antimicrobial functions were also identified including lipopolysaccharide binding protein, neutrophil defensin 1, dermicidin, lysozyme C and deleted in malignant tumors 1 protein, all of which have previously been identified in the cornea.\textsuperscript{24,51–53} Lactotransferrin, a protein with antimicrobial and possibly anti-inflammatory functions was also identified. Lactotransferrin is a major protein constituent of tears but has previously only been identified in diseased corneas.\textsuperscript{54–56}

\section*{CONCLUSION}

In conclusion, this study provides protein identification and quantitative information for the proteins of the corneal epithelium, stroma and endothelium. The separation of the corneal layers, combined with modern mass spectrometry facilitated the identification of 3250 unique proteins. Of these, 1787 proteins have not previously been reported present in the human cornea by mass spectrometry. The separation provides new insight into the proteins present in the individual layers and the relative abundance in each layer. This provides a useful reference data set when exploring basic molecular mechanisms involved in corneal diseases, many of which are restricted to a specific corneal layer. Furthermore, our analysis showed that many of the identified proteins are human plasma proteins involved in defense responses.

\section*{ASSOCIATED CONTENT}

\subsection*{Supporting Information}

This material is available free of charge via the Internet at http://pubs.acs.org. The data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche network, https://proteomecommons.org/tranche/, using the following hash: NXZDWWNSAOZovh0WBeDr8KY-Q29EoE6FQ6KwAqmImZhbW1GUFhil3QSnKXEHoMJa-NYoJt8s1wNWfjpwUMAhgObfc5UAAAAAAAAAFZCQ==.

\section*{AUTHOR INFORMATION}

Corresponding Author

*Aarhus University, Gustav Wieds Vej 10C, 8000 Aarhus C., Denmark. Tel.: +45 89425062. Fax: +45 89425063. E-mail: jjc@mb.au.dk.

Notes

The authors declare no competing financial interest.

\section*{ACKNOWLEDGMENTS}

This study was supported by the Danish Agency for Science, Technology and Innovation, the Faculty of Science at Aarhus University, the Danish Society for Eye Health and the National Eye Institute (R01 EY012712).

\section*{REFERENCES}

(19) Schlotzer-Schrehardt, U.; Bachmann, B. O.; Laaser, K.; Cursiefen, C.; Kruse, F. E. Characterization of the cleavage plane in


(41) Rughunath, M.; Cankay, R.; Kubitscheck, U.; Fauteck, J. D.; Mayne, R.; Aeschlimann, D.; Schlotzer-Schrehardt, U. Transglutami-