A Dataset of Human Cornea Proteins Identified by Peptide Mass Fingerprinting and Tandem Mass Spectrometry*

Henrik Karring‡, Ida B. Thøgersen‡, Gordon K. Klintworth§, Torben Møller-Pedersen¶, and Jan J. Enghild‡,

Diseases of the cornea are extremely common and cause severe visual impairment worldwide. To explore the basic molecular mechanisms involved in corneal health and disease, the present study characterizes the proteome of the normal human cornea. All proteins were extracted from the central 7-mm region of 12 normal human donor corneas containing all layers: epithelium, Bowman’s layer, stroma, Descemet’s membrane, and endothelium. Proteins were fractionated and identified using different procedures: (i) two-dimensional gel electrophoresis and protein identification by MALDI-MS and (ii) strong cation exchange or one-dimensional SDS gel electrophoresis followed by LC-MS/MS. All together, 141 distinct proteins were identified of which 99 had not previously been identified in any mammalian corneas by direct protein identification methods. The characterized proteins are involved in many processes including antiangiogenesis, antimicrobial defense, protection from and transport of heme and iron, tissue protection against UV radiation and oxidative stress, cell metabolism, and maintenance of intracellular and extracellular structures and stability. This proteome study of the healthy human cornea provides a basis for further analysis of corneal diseases and the design of bioengineered corneas. Molecular & Cellular Proteomics 4:1406–1408, 2005.

The human cornea is a transparent, avascular, and highly specialized connective tissue that provides ~70% of the total refraction in the optical system of the eye. Other essential properties of the cornea include protection against noxious agents, biomechanical stability, and structural resiliency as well as the ability to filter out damaging UV light (1), thereby protecting both the crystalline lens and retina against injury. The human cornea (thickness, ~530 μm) is a multilayered tissue composed of five main layers: the epithelium (~50 μm), Bowman’s layer (~10 μm), the stroma (~450 μm), Descemet’s membrane (~5–15 μm), and the endothelium (~5 μm). In the healthy eye, these layers interact in a complex manner to strictly maintain the properties of the cornea. Increased biochemical knowledge of normal and diseased corneas is essential for the understanding of corneal homeostasis and pathophysiology.

The present study explores the proteome of the intact normal human cornea. We identified 141 distinct corneal proteins by peptide mass fingerprinting or LC-MS/MS preceded by fractionation using 2D PAGE, 1D PAGE, and strong cation exchange of peptides (supplemental experimental information). Four different protocols were used for extraction and separation of the proteins/peptides that facilitated the identification of both soluble and insoluble proteins and increased the number of identified proteins in general.

RESULTS

Proteome Study Using 2D PAGE (Protocols 1 and 2)—Proteins from the corneal powder were extracted by 5 M urea and 2 M thiourea under reducing conditions (Protocol 1) and analyzed by 2D PAGE using five different pH gradients (Supplemental Fig. 1, A–E). 2D gel spots were excised, and proteins were identified by peptide mass fingerprinting. The 165 identified spots represented only 67 distinct proteins because several proteins existed as multiple isoforms (Supplemental Table I). Especially transforming growth factor-β-induced protein (TGFBlp) (29 isoforms), serum albumin (13 isoforms), and immunoglobulin κ light chain (11 isoforms) were found in a significant number of isoforms. Several of the TGFBlp and serum albumin isoforms have molecular masses lower than the calculated molecular masses of the mature full-length proteins (TGFBlp, M_m ~ 72.4 kDa; and serum albumin, M_m ~ 66.5 kDa).

To analyze the water-soluble proteome of the human cornea, proteins were extracted from the corneal powder using 100 mM NaCl under non-reducing conditions (Protocol 2) and separated by 2D PAGE using a 4.0–7.0 pH gradient (Supple-

* The abbreviations used are: 2D, two-dimensional; 1D, one-dimensional; TGFBlp, transforming growth factor-β-induced protein; EC, extracellular; IC, intracellular.
Comparisons of the protein patterns (Supplementary Fig. 1, B and F) showed that most of the proteins and protein isoforms extracted using Protocol 1 are also present in the water-soluble fraction (Protocol 2). However, four abundant isoforms of immunoglobulin α-1 heavy chain (spots 63, 111, 112, and 136) present in the water-soluble fraction (Supplementary Fig. 1F) were not detected using extraction Protocol 1 (Supplemental Fig. 1B). In contrast, most of the TGFβI Spots and serum albumin isoforms migrating between 35 and 45 kDa (Supplemental Fig. 1B) were not detected in the water-soluble fraction (Supplemental Fig. 1F).

Proteome Study Using Cyanogen Bromide Prior to LC-MS/MS (Protocol 3)—To facilitate the identification of the insoluble proteins and proteins too large/small or too acidic/basic to be analyzed by 2D PAGE, the corneal powder was chemically fragmented using CNBr to facilitate the trypsin digestion and LC-MS/MS analysis. Using this procedure, 31 distinct proteins were identified exhibiting Mascot scores ranging from 37 to 13,162 (Supplemental Table II). The moderate number of identifications made using Protocol 3 is likely caused by the large excess of peptides derived from highly abundant proteins. To avoid this problem, another approach was used where the proteins were separated by 1D PAGE prior to the generation of peptides (Protocol 4).

Proteome Study Using 1D PAGE and LC-MS/MS (Protocol 4)—The corneal proteins were extracted in reducing SDS sample buffer and separated by 1D PAGE (Supplemental Fig. 2) (Protocol 4). Slices of the gel lane were digested with trypsin, and each sample was subjected to LC-MS/MS. From this analysis, 103 distinct proteins with total Mascot scores ranging from 31 to 10,532 were identified (Supplemental Table III). However, this procedure is not suitable for the identification of small/large proteins or proteins not soluble in SDS sample buffer such as collagens. Among the identified proteins, two hits (entries 99 and 100 in Supplemental Table III), were classified as hypothetical proteins (predicted proteins not verified by analysis of the proteins in vivo). Protein-protein BLAST searches revealed that the Unnamed protein product (Sequence 21 from Patent W00214358, accession number CAD29037, Mass Spectrometry protein sequence Database entry) has very high identity to human MAM (meprin A5 protein tyrosine phosphatase μ) domain-containing proteins 1 and 2 (accession number NP_694999), whereas Hypothetical protein FLJ20261 (accession number Q9NXG7, Mass Spectrometry protein sequence Database entry) shows high identity to human keratin 24 (accession number NP_061889) (E-value, e-180) suggesting that these proteins are expressed in the normal human cornea.

Categorization and Distribution of the Corneal Proteins—All together, the identified proteins from the normal human corneas represent 141 distinct proteins (Supplemental Table IV). Proteins not previously identified in any mammalian cornea using direct identification methods such as Edman degradation or mass spectrometry are indicated in the table. Thus, 99 proteins (70%) of the 141 proteins have not been detected previously in mammalian corneas by direct methods. The proteins are categorized according to their predominant function as reported in the current literature. A, intracellular proteins (IC); B, extracellular matrix proteins (EC); Bl, classical blood/plasma proteins excluding those involved in immune defense mechanisms (EC, 39%); Fo, protein folding and degradation (IC, 11%); Im, immune defense and inflammatory response (IC, 1%; EC, 17%); Me, metabolism (IC, 24%); Ot, other functions including cell proliferation, differentiation, apoptosis, signal transduction, ion transport, osmotic regulation, DNA replication and regulation, protein synthesis, cell adhesion and migration, endocytosis, exocytosis, etc. (IC, 21%; EC, 11%); Re, redox regulation and oxidative stress defense (IC, 13%; EC, 2%); St, structural and structural associated proteins (IC, 28%; EC, 27%); Un, unknown function (IC, 2%; EC, 4%).
associated proteins (St), nine proteins are involved in immune defense and inflammatory response (Im), one protein is involved in oxidative stress defense (Re), six proteins have other functions (Ot), and two proteins have unknown function (Un) (Fig. 2B).

**DISCUSSION**

The 2D PAGE analysis revealed that TGFBIp, serum albumin, and immunoglobulin κ chain were found in a significant number of isoforms indicating post-translational additions and fragmentations. Most of the isoforms of TGFBIp were absent in the water-soluble and non-reduced fraction (Supplemental Fig. 1, B and F, and Table I). In a recent study, we have shown that ~60% of human cornea TGFBIp (~65 kDa) is covalently associated with insoluble components of the extracellular matrix and that this insoluble fraction of TGFBIp is released after reduction of disulfides (2). The present results show that the low molecular mass isoforms of TGFBIp (35–45 kDa) in the cornea are not water-soluble indicating that some of these fragments are also associated with insoluble components of the extracellular matrix. Furthermore this finding suggests that cleavage of extracellular matrix proteins is a common event in the normal human cornea. Alternatively it cannot be out ruled that some of the degradation occurred postmortem or during sample preparation. However, the addition of a broad spectrum protease inhibitor mixture makes this less likely.

As expected, most of the intracellular proteins are involved in metabolism or have structural roles. We identified 15 different keratins (six type I and nine type II) representing at least seven particular keratin pairs (pairs K10/1, K12/3, K13/4, K14/5, K16/6A or 6C, K16/6B, and K16/6F). In addition, Hypothetical protein FLJ20261 has high identity to human keratin 24, which apparently is a type I keratin. However, because several keratins are commonly present in the environment our identification of some keratins (keratin 1, epidermal keratin 2, keratin 7, keratin 9, and keratin 10) might be contaminations from the laboratory (3).

Common plasma proteins and structural proteins dominate the extracellular matrix of the cornea. Most of the extracellular proteins are probably from the corneal stroma as this connective tissue constitutes about 90% of the corneal volume and mainly consists of extracellular space as opposed to the cell-dense endothelium and epithelium. A few of the identified proteins probably originate from the tear fluid including lysozyme, tear lipocalin, and apolipoprotein D and are synthesized by the lacrimal glands (4, 5).

The proteins identified in the cornea are involved in many processes including antiangiogenesis, antimicrobial defense, protection from and transport of heme and iron, tissue protection against UV radiation and oxidative stress, cell metabolism, and maintenance of intracellular and extracellular structures and stability. This human cornea protein dataset provides a useful reference library for further studies to define the specific roles of the identified proteins and for comparative proteomic studies of cornea disease and wound healing (6). Furthermore the identification of corneal components will assist the efforts to generate artificial corneas (7).

**Acknowledgments**—We thank the Protein Research Group at University of Southern Denmark, particularly Peter Højrup and Allan Steensballe, for helpful suggestions.

* This work was supported in part by National Institutes of Health Grant RO1-EY12712 and grants from The Danish Medical Research Council and The Danish Natural Science Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

To whom correspondence should be addressed. Tel.: 45-8942-5062; Fax: 45-8942-5063; E-mail: jje@mb.au.dk.

**REFERENCES**

1. Kolozsvári, L., Nogradi, A., Hopp, B., and Bor, Z. (2002) UV absorbance of the human cornea in the 240- to 400-nm range. Invest. Ophthalmol. Vis. Sci. 43, 2165–2168

2. Andersen, R. B., Karring, H., Møller-Pedersen, T., Valnickova, Z., Thøgersen, I. B., Hedegaard, C. J., Kristensen, T., Klintworth, G. K., and Enghild, J. J. (2004) Purification and structural characterization of transforming growth factor β induced protein (TGFβp) from porcine and human corneas. Biochemistry 43, 16374–16384

3. Porter, R. M., and Lane, E. B. (2003) Phenotypes, genotypes and their contribution to understanding keratin function. Trends Genet. 19, 278–285

4. Holzfeind, P., Mersch, P., Dieplinger, H., and Redl, B. (1995) The human lacrimal gland synthesizes apolipoprotein D mRNA in addition to tear prealbumin mRNA, both species encoding members of the lipocalin superfamily. Exp. Eye Res. 61, 495–500

5. Glasgow, B. J. (1995) Tissue expression of lipocalins in human lacrimal and von Ebner’s glands: colocalization with lysozyme. Graefes Arch. Clin. Exp. Ophthalmol. 233, 513–522

6. Karring, H., Thøgersen, I. B., Klintworth, G. K., Enghild, J. J., and Møller-Pedersen, T. (2004) Proteomic analysis of the soluble fraction from human corneal fibroblasts with reference to ocular transparency. Mol. Cell Proteomics 3, 660–674

7. Carlsson, D. J., Li, F., Shimmura, S., and Griffith, M. (2003) Bioengineered corneas: how close are we? Curr. Opin. Ophthalmol. 14, 192–197