Protein Composition of TGFBI-R124C- and TGFBI-R555W-Associated Aggregates Suggests Multiple Mechanisms Leading to Lattice and Granular Corneal Dystrophy

David G. Courtney,[1,2] Ebbe Toftgaard Poulsen,[2,3] Susan Kennedy,[4] Johnny E. Moore,[1,5]
Sarah D. Atkinson,[1] Eleonora Maurizi,[1] M. Andrew Nesbit,[1] C. B. Tara Moore,[1]
and Jan J. Enghild[2,3]

[1]School of Biomedical Sciences, Centre for Molecular Biosciences (CMB), University of Ulster, Coleraine, Northern Ireland, United Kingdom
[2]Department of Molecular Biology and Genetics, Science Park, Aarhus University, Aarhus, Denmark
[3]Interdisciplinary Nanoscience Center (iNANO) and Center for Insoluble Protein Structures (inSPIN), Aarhus University, Aarhus, Denmark
[4]Department of Histopathology, Royal Victoria Eye and Ear Hospital, Dublin 2, Ireland
[5]Cathedral Eye Clinic, Belfast, Northern Ireland, United Kingdom

Correspondence: C. B. Tara Moore, School of Biomedical Sciences, CMB, University of Ulster, Coleraine, Northern Ireland, UK; t.moore@ulster.ac.uk.

Jan J. Enghild, Department of Molecular Biology and Genetics, Science Park, Aarhus University, Aarhus, Denmark; jje@mbg.au.dk.

DGC and ETP contributed equally to the work presented here and should therefore be regarded as equivalent authors.

Submitted: March 20, 2015
Accepted: June 10, 2015

Citation: Courtney DG, Toftgaard Poulsen E, Kennedy S, et al. Protein composition of TGFBI-R124C- and TGFBI-R555W-associated aggregates suggests multiple mechanisms leading to lattice and granular corneal dystrophy. *Invest Ophthalmol Vis Sci*. 2015;56:4653–4661. DOI:10.1167/iovs.15-16922

PURPOSE. Transforming growth factor beta-induced (TGFBI)–related dystrophies constitute the most common heritable forms of corneal dystrophy worldwide. However, other than the underlying genotypes of these conditions, a limited knowledge exists of the exact pathomechanisms of these disorders. This study expands on our previous research investigating dystrophic stromal aggregates, with the aim of better elucidating the pathomechanism of two conditions arising from the most common TGFBI mutations: granular corneal dystrophy type 1 (GCD1; R555W) and lattice corneal dystrophy type 1 (LCD1; R124C).

METHODS. Patient corneas with GCD1 and LCD1 were stained with hematoxylin and eosin and Congo red to visualize stromal nonamyloid and amyloid deposits, respectively. Laser capture microdissection was used to isolate aggregates and extracted protein was analyzed by mass spectrometry. Proteins were identified and their approximate abundances were determined. Spectra of TGFBIp peptides were also recorded and quantified.

RESULTS. In total, three proteins were found within GCD1 aggregates that were absent in the healthy control corneal tissue. In comparison, an additional 18 and 24 proteins within stromal LCD1 and Bowman's LCD1 deposits, respectively, were identified. Variances surrounding the endogenous cleavage sites of TGFBIp were also noted. An increase in the number of residues experiencing cleavage was observed in both GCD1 aggregates and LCD1 deposits.

CONCLUSIONS. The study reveals previously unknown differences between the protein composition of GCD1 and LCD1 aggregates, and confirms the presence of the HtrA1 protease in LCD1-amyloid aggregates. In addition, we find mutation-specific differences in the processing of mutant TGFBIp species, which may contribute to the variable phenotypes noted in TGFBI-related dystrophies.

Keywords: TGFBI, granular dystrophy, lattice dystrophy, proteomics, corneal stroma

---

Transforming growth factor beta-induced protein (TGFBIp; also known as *BIG-H3* or *keratoepithelin*), is a 72-kDa extracellular matrix protein consisting of four consecutive and highly conserved fascilin 1 (FAS1) domains flanked N-terminally by a small cysteine-rich domain (EMI) and C-terminally by an integrin-binding motif, Arg-Gly-Asp (RGD).[1] Mutations in *TGFBI* cause different forms of corneal dystrophy with the most common mutations arising at codons Arg124 and Arg555. The resultant corneal dystrophy is dependent upon the particular missense mutation at these codons. For example, an Arg124Cys (c.417 C>T) mutation results in lattice corneal dystrophy type 1 (LCD1), while an Arg124Ser (c.417 C>A) mutation causes granular corneal dystrophy type 1 (GCD1).[2] Five distinct corneal dystrophies noted in the recently published in the second edition of the IC3D Classification of Corneal Dystrophies[3] are attributed to mutations within *TGFBI*; Reis-Bücklers corneal dystrophy (RBCD), Thiel-Behnke corneal dystrophy (TBCD; both Bowman's layer dystrophies), LCD1, GCD1, and GCD2 (regarded as stromal dystrophies). Over 63 disease causing mutations within *TGFBI* have now been reported.[4] It remains unclear how these varying phenotypes arise from different mutations within this one gene[5,6] and why stromal TGFBIp deposits are amyloidogenic in LCD1 yet appear as nonamyloid hyaline aggregates in GCD1.[3]

Our previous research investigating the protein profile of amyloid dystrophic deposits in the cornea focused on rare genetic variants of LCD (Table 1).[7–9] In these studies, focusing on mutations within the FAS1-4 domain of TGFBIp, a number of

Table 1. Genotypes LCD1, GCD1, and GCD2 Analyzed by LCM and LC-MS/MS

Genotype	Phenotype	Characteristics	Reference
R124C	LCD1	Stromal and Bowman's amyloid deposits	Present
R124H	GCD2	Stromal amorphous aggregates	Karring et al.[7]
A546D	LCD variant	Stromal amyloid deposits	Karring et al.[8]
A546D/P551Q	LCD variant	Stromal amyloid deposits	Poulsen et al.[9]
R555W	GCD1	Stromal amorphous aggregates	Present
V624M	LCD variant	Stromal amyloid deposits	Karring et al.[7]

proteins unique to LCD deposits compared with healthy stromal tissue were reported. In particular, the serine protease HtrA1 (HtrA1) was identified as a lead candidate in the proteolytic processing of mutant TGFBIp species. These studies also identified variances in the abundance of one polypeptide region (Y571-R588) as well as increased proteolytic processing within the FAS1-4 domain in the LCD1 amyloid deposits when compared with healthy stroma. In addition, we have analyzed the protein composition of aggregates in one GCD2 case caused by a R124H mutation[7]; however, none of these analyses have been carried out on stromal aggregates of patients with the more common R555W-associated GCD1.

This study aimed to investigate further the protein profile of stromal amyloid deposits within LCD1 across multiple samples ($n = 5$ corneas) by analyzing deposits caused by the R124C FAS1-1 mutation. In addition to central stromal deposits, anterior amyloid material located just beneath Bowman's layer was identified, isolated, and analyzed in three of the LCD1 cases. This constituted an additional novel level of analysis, comparing protein and TGFBIp peptide abundances, not simply within characteristic central stromal deposits, but also within these "Bowman's deposits," with healthy stromal tissue. Furthermore, this research constitutes the first profiling of the GCD1 stromal aggregate proteome caused by the TGFBI-R555W mutation ($n = 2$ corneas). Proteins unique to each group of GCD1 aggregates and LCD1 deposits, but not found within the healthy corneal stroma, were identified. In addition, preferential cleavage sites of mutant TGFBIp species were found. The data suggest that different molecular pathways give rise to GCD1 and LCD1 phenotypes, and additional differences between LCD1 cases are dependent on the site of the TGFBI mutation.

METHODS

Patients

In total, corneal samples from 10 individuals were used in this study; 3 control cases (average age of 88 years); 2 GCD1 cases (average age of 77 years); and 5 LCD1 cases (average age 56 years). Corneas from GCD1 and LCD1 patients were collected following corneal transplantation. Healthy corneal stroma from three recently deceased individuals was collected from the donor eye bank and used as control tissue. Full ethical approval was in place for this study. Informed consent was obtained from all patients prior to tissue being collected. All work was carried out in accordance with the tenets of the Declaration of Helsinki.

Preparation of samples

All tissue samples were fixed in 4% paraformaldehyde (in PBS) for 10 minutes prior to dehydration through graded alcohol solutions before being embedded in paraffin wax. Samples were then cut into 7-µm sections and placed on glass slides and left to adhere at 37°C overnight. Deparaffinization and rehydration through graded alcohol washes was performed. Hematoxylin and eosin (H&E) staining of GCD1 samples and Congo red staining of LCD1 samples was performed[10] to visualize the TGFBIp nonamyloid and amyloid aggregates, respectively. Both were viewed under polarized light, though Congo red staining could also be visualized by fluorescent microscopy under a Texas Red filter.

Laser-Capture Microdissection

Capture of dystrophic TGFBIp aggregates was carried out as described previously.[8,9] Briefly, a laser capture microscope (Arcturus Autopix Laser Capture Microscope; Applied Biosystems, Life Technologies, Paisley, UK) was used to isolate aggregates from within the stroma of dystrophic tissue. Deposits were collected on laser capture microdissection caps (CapSure Macro LCM Caps; Applied Biosystems) and stored for processing.

Sample Preparation and LC-MS/MS Analysis of Deposits

Membranes were removed from the LCM Caps and incubated at 95°C for 90 minutes in 20 µL 10 mM Tris-HCl (pH 8) containing 0.1% surfactant reagent (RapiGest; Waters Corporation, Milford, MA, USA). Samples underwent centrifugation every 15 minutes. Samples were then chilled on ice. To generate peptides for analysis using liquid chromatography-tandem mass spectrometry (LC-MS/MS), samples were incubated with 5 µg trypsin (Sequencing Grade Modified Porcine Trypsin; Promega, Southampton, UK) for 16 hours at 37°C. Digested samples were reduced (5 mM dithiothreitol) for 15 minutes at room temperature and subsequently alkylated (15 mM iodoacetamide) for 15 minutes at room temperature, followed by micropurification using POROS R2 material packed in gel loader tips. Eluted peptides were lyophilized under vacuum before being resuspended in 0.1% formic acid and analyzed using a commercial system (Easy-nLC II; Thermo Scientific, Epsom, UK) coupled directly to a mass spectrometer (TripleTOF 5600$^+$; AB Sciex, Framingham, MA, USA).

All samples were loaded on a 0.1 × 21-mm C18 trap column and a 0.075 × 150-mm C18 analytical column (columns packed in-house). Upon elution peptides were immediately electrosprayed into the mass spectrometer using a 20-minute gradient of 5% to 40% acetonitrile in 0.1% acetic acid at a constant flow rate of 250 nL/min. Information-dependent data acquisition was performed using survey scans of 250 ms followed by up to 25 product ion scans resulting in a total cycle time of 0.925 s. Dynamic exclusion was set to 6 seconds, mass range was 350 to 1250 m/z, and precursor charge state was set to +2 to +5.

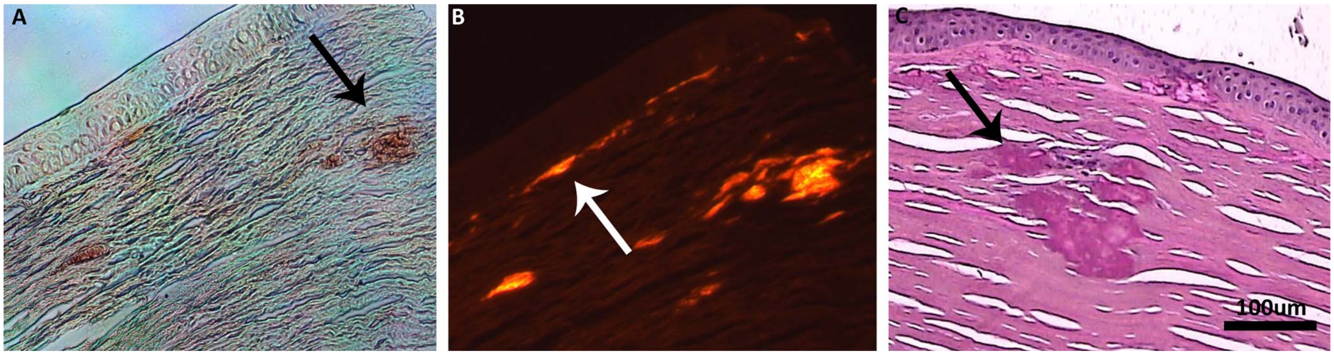

Figure 1. Visualization of plaques in LCD1 and GCD1 dystrophic stromal tissue. Tissue sections of LCD1 and GCD1 underwent staining to visualize dystrophic plaques before LCM. Congo red staining allowed for the identification of LCD1 stromal amyloid deposits ([**A**], *black arrow*) and LCD1 Bowman's amyloid deposits ([**B**], *white arrow*) using either light or fluorescent microscopy; H&E staining was employed to visualize GCD1 aggregates (**C**).

Identification and Semiquantification of Proteins

Raw data from the LC-MS/MS was converted to mascot generic format (MGF) using conversion software (MS Data Converter, beta 1.1; AB Sciex) and the "proteinpilot mgf" parameters, and peptides were searched using an inhouse search engine (Mascot, version 2.5.0; Matrix Science, London, UK) and queried against the Swiss-Prot database (version 2014_11). As search parameters, semitrypsin was specified as the enzyme allowing one miscleavage. Carbamidomethyl was set as fixed modification and oxidation of methionine and proline as variable modifications. The tolerances of precursor and product ions were set to 10 ppm and 0.2 Da, respectively, and ESI-QUAD-TOF was selected as the instrument used. All searches had a P value < 0.01 applied and an expected value of 0.005. Only protein hits identified by a minimum of two peptides with Mascot ion scores above 30 were considered for further analyses.

All cytokeratins identified within the analysis were regarded as contaminants and removed from the proteome profiles, as carried out in previous studies.[9] In addition, porcine tryptic peptides identical to human peptides were removed as porcine trypsin was used during the processing of the samples. The remaining proteins were arranged based on their exponentially modified protein abundance index (emPAI) values reflecting their abundances within the aggregates. To compile a comprehensive list of proteins found in multiple GCD1 and LCD1 aggregates, relative abundances were calculated for proteins present in at least two tissues of the same biological significance. Relative abundance was termed as the average emPAI value for each protein across biological samples divided by the sum of the emPAI values of all proteins present within the tissue specimens. These values were then recorded as a percentage for all proteins within each biological group,[11] (Supplementary Table S1).

Determining Spectral Counts for Tryptic and Semitryptic TGFBIp Peptides

Spectral counts were carried out for every TGFBIp peptide with no missed cleavages, using methods previously described.[7,8] Initially tryptic and semitryptic peptides were separated, with tryptic peptides being those identified as cleaved by trypsin during sample preparation on the C-terminal side of a Lys or Arg residue at both termini, while the list of semitryptic peptides included those cleaved on the C-terminal side of a Lys or Arg residue in just one terminus, and containing a nontryptic cleavage site at the other peptide terminus. These nontryptic cleavage site events are most likely due to the processing of TGFBIp by endogenous proteases present within the tissue samples collected. Spectral counts for all technical replicates were performed using the MS Data Miner v.1.3.0 software.[12] Spectral count ratios were calculated by dividing the count of spectra for each TGFBIp peptide by the total number of spectra for tryptic or semitryptic TGFBIp peptides. Spectral count ratios were displayed as a percentage. The spectral count ratio for each peptide from each tissue was also presented graphically and grouped by disease. Significance of the variance between counts from disease tissue compared to those from control tissue was calculated using Fisher's least significant difference (LSD) test.

In addition, the positions and frequencies of nontryptic TGFBIp peptide cleavage sites were counted and compared against previous data.[7–9]

Statistics

Three technical replicates of LC-MS/MS were run for each sample. Error bars indicate the standard deviation across samples of the same biological significance. For spectral counts, Fisher's LSD test was used to calculate significance between diseased tissue and control healthy tissue.

RESULTS

Staining and Microdissection of Corneal Sections

Central stromal amyloid deposits from five patients with LCD1 (heterozygous R124C TGFBI mutation) were identified by Congo red staining (Figs. 1A, black arrow, 1B) and isolated by LCM. Interestingly, in the corneal tissue of three individuals with LCD1 anterior stromal deposits of an amyloid nature were visible just beneath Bowman's layer (Figs. 1A, 1B, white arrow), hereafter referred to as Bowman's deposits. These Bowman's deposits were isolated separately to the central stromal deposits by LCM and were treated as separate samples throughout all subsequent analysis. Amorphous stromal aggregates from two patients with GCD1 (heterozygous TGFBI R555W mutation) were identified by standard H&E staining procedures and were visualized as dark pink regions within the corneal stroma (Fig. 1C, black arrow) and isolated by LCM. Stroma from sections of unstained healthy corneal tissue was collected by LCM and used as the control corneal stroma tissue.

Establishment of a Proteomic Profile of Corneal Deposits and Healthy Corneal Stroma

On establishing a profile of the aggregate and stromal proteomes, and following the removal of common contaminating exogenous proteins, a total of 39, 25, 45, and 46 proteins were identified in healthy stroma, GCD1 aggregates, LCD1 stromal deposits, and LCD1 Bowman's deposits, respectively. These proteins were then sorted based on their relative abundance in the tissue samples using their emPAI values (S1). The top 25 proteins found within each disease sample set, their average relative abundances and standard deviations are recorded in Table 2.

Identification of Proteins Unique to the Proteomes of GCD1 or LCD1 Deposits

Analysis of the proteins identified in each tissue type elucidated proteins present within the GCD aggregates and LCD deposits, but not present in a healthy corneal stroma. In total, three unique proteins were identified in the GCD1 aggregates, while 18 and 24 unique proteins were isolated from LCD1 stromal and Bowman's amyloid deposits, respectively (Table 3).

Comparison of the Proteolytic Processing of TGFBIp in all Corneal Samples

Spectral counts of tryptic and semitryptic peptides for TGFBIp found within GCD1 aggregates, LCD1 deposits, or control healthy stromal tissue were carried out. This demonstrated the proportion of each peptide in relation to the overall quantity of peptides within each biological sample (Fig. 2). Spectral counts for each individual tissue were also plotted graphically (tryptic, Supplementary Fig. S1; semitryptic, Supplementary Fig. S2). Upon assessment of tryptic TGFBIp peptide abundances, it was noted that the S28-R38, 470-484, and 591-596 tryptic peptides were significantly reduced in all diseased tissues. In addition, the Q43-K53 tryptic peptide was significantly underrepresented by half in all dystrophic aggregates (Fig. 2A) suggesting this region was N-terminally truncated to a higher degree in diseased corneal tissue than in the control corneal tissue. In total, the abundances of 21 tryptic peptides were found to be significantly different in GCD1 or LCD1 tissues when compared with control tissue (Supplementary Fig. S1).

The semitryptic TGFBIp peptide profile (Fig. 2B) of the GCD1 aggregates demonstrated significantly higher abundances of peptides within the polypeptide regions Q43-K53, G96-R124, V235-R257, D356-K377, T378-R396, Y448-K461, and F515-R533 when compared to the lower or nonexistent levels found in the control tissue (Supplementary Fig. S2). Three of these polypeptide regions, G96-R124, V235-R257, and T378-R396 are interdomain regions, which are likely to be more accessible to proteolytic degradation.

In Vivo Differential Cleavage of TGFBIp in Stromal Deposits

The total number of endogenous cleavage sites of TGFBIp in each biological sample was determined and averaged across specimens of similar biological significance and each observed P1 residue was calculated as a percentage of the total nontryptic P1 residues observed. Of note, it was evident that there was a reduction by approximately one-third in TGFBIp peptides cleaved after the Phe residue in GCD1 aggregates compared to the control tissue, while contrarily, an increase by more than one-third in cleavage after this residue is apparent in both stromal and Bowman's LCD1 cases. In addition to this, although with lower frequencies, cleavage within the GCD1 aggregates occurs after five additional residues, Asp, Glu, Gly, Gln, Thr, not seen in the TGFBIp fragments in either the control or LCD1 samples. There is also a decrease in cleavage after Asn residues in both stromal and Bowman's LCD1 aggregates in comparison to GCD1 and control tissues. Cleavage after Leu residues also appears to be increased in all mutant species of TGFBIp, compared with the control tissue (Fig. 3).

DISCUSSION

In this study the protein profile and TGFBIp processing within the stromal aggregates of patients with LCD1 and GCD1 was investigated. By further elucidating the proteins and molecular processes involved in these disease states it is hoped that current therapies under development for such conditions could take another step toward the clinic,[13–16] or novel therapies developed based on these latest findings.

Proteins Identified Within Dystrophic Deposits and Their Relevance in Disease

A comparison of control and disease groups (Table 2) revealed proteins present only in the diseased tissue but not in the control tissue. In total, 18 and 24 proteins were observed only in the central stromal and Bowman's amyloid deposits in LCD1, respectively, while three such proteins were solely found in the GCD1 aggregates (Table 3). In agreement with previous reports, amyloidogenic proteins serum amyloid P-component,[17] lysozyme C,[18,19] cystatin A,[20] apolipoprotein A-I, A-IV and D, HtrA1,[21] and a number of proteins from the S100 protein family,[22] constituted a large part of the LCD1 amyloid deposits, in addition to TGFBIp itself (Table 2).

We only identified HtrA1 in the stromal deposits of LCD1 specimens, not within Bowman's deposits, and appeared with a similar molar fraction to that which we demonstrated previously in similar proteomic analyses of amyloid stromal deposits arising from mutations within the FAS1-4.[8,9] This further supports our previous proposal as to its potential role in the proteolytic processing of mutant TGFBIp species. The bacterial HtrA1 homologue is involved in bacterial protein quality control machinery[23] and it could be that HtrA1 has a similar role in humans, expressed at higher levels in response to an elevation of protein misfolding. Of particular interest was the lack of HtrA1 in the amorphous GCD1 aggregates, corroborating our previous studies where HtrA1 was only observed in amyloid TGFBIp deposits.[7,8] This indicates that this protease may only be involved in the proteolytic processing of TGFBIp aggregates of an amyloid nature; HtrA1 has also been observed in the healthy human cornea,[24] indicating that HtrA1 is implicated in normal protein turnover but expression is elevated in diseases associated with amyloid formation.

An accumulation of proteins from the protein S100 family were found in both stromal and Bowman's LCD1 TGFBIp deposits, where none of these proteins are present within the healthy corneal stroma. Recently a link between amyloid precursor proteins and protein S100A9, the most highly expressed of the S100 family members, was demonstrated in

Table 2. Relative Amounts of the 25 Most Abundant Proteins in Control Tissue, GCD1 Aggregates, and LCD1 Deposits

Control Tissue				GCD1 Aggregates				LCD1 Stromal Deposits				LCD1 Bowman's Deposits			
Name	**%**		**SD**	**Name**	**%**		**SD**	**Name**	**%**		**SD**	**Name**	**%**		**SD**
1 Collagen alpha-1(I) chain	28.8	±	5.7	1 TGFBIp	57.1	±	7.9	1 Collagen alpha-1(I) chain	14.9	±	2.2	1 Collagen alpha-1(I) chain	11.6	±	5.3
2 Collagen alpha-2(I) chain	19.2	±	0.2	2 Collagen alpha-1(I) chain	14.1	±	1.2	2 TGFBIp	11.2	±	3.9	2 Collagen alpha-2(I) chain	9.1	±	4.4
3 TGFBIp	7.6	±	1.3	3 Collagen alpha-2(I) chain	9.3	±	2	3 Collagen alpha-2(I) chain	11.1	±	1.9	3 Dermcidin	8.5	±	2.5
4 Ig kappa chain C region	7.1	±	6.4	4 Hemoglobin subunit beta	2.9	±	3	4 Serum amyloid P-component	6.2	±	2.1	4 TGFBIp	7.1	±	1.5
5 Serum albumin	4.9	±	1	5 Lumican	2.2	±	0.1	5 Dermcidin	6.1	±	3.1	5 Protein S100-A9	6.8	±	2.7
6 Decorin	4	±	1	6 Dermcidin	1.6	±	0.1	6 Ig kappa chain C region	5.3	±	2.9	6 Cystatin-A	6.1	±	1.3
7 Keratocan	3.3	±	0.7	7 Keratocan	1.6	±	0.1	7 Protein S100-A9	4.9	±	1.8	7 Serum amyloid P-component	5.2	±	1.1
8 Lumican	2.9	±	0.2	8 Collagen alpha-3(VI) chain	1.5	±	0.8	8 Cystatin-A	3.9	±	0.6	8 Protein S100-A6	3.6	±	0.8
9 Collagen alpha-3(VI) chain	2.3	±	0.1	9 Aldehyde dehydrogenase, dimeric NADP-preferring	1.2	±	0.1	9 Histone H2B type 1-A	3.7	±	0.8	9 Protein S100-A4	3.4	±	0.3
10 Ig gamma-1 chain C region	1.7	±	0.3	10 Histone H2B type 1-A	1.1	±	0.3	10 Hemoglobin subunit beta	2.7	±	0.2	10 Aldehyde dehydrogenase, dimeric NADP-preferring	3	±	1.4
11 Dermcidin	1.6	±	0	11 Collagen alpha-1(VI) chain	1	±	0.4	11 Ubiquitin-40S ribosomal protein S27a	2.7	±	1.3	11 Histone H2B type 1-A	2.8	±	0.2
12 Collagen alpha-1(VI) chain	1.3	±	0.1	12 Collagen alpha-2(VI) chain	0.8	±	0.3	12 Lumican	2.6	±	0.5	12 Alpha-enolase	2.6	±	1.6
13 Histone H2B type 1-A	1.3	±	0.1	13 Decorin	0.8	±	0.3	13 Clusterin	2.3	±	1	13 Vimentin	2.4	±	1.7
14 Collagen alpha-2(V) chain	1.2	±	0.1	14 Serum albumin	0.8	±	0.5	14 Apolipoprotein A-IV	2.3	±	0.9	14 Actin, cytoplasmic 1	2.3	±	0.7
15 Collagen alpha-2(VI) chain	1.2	±	0.1	15 Alpha-enolase	0.7	±	0.1	15 Keratocan	2.1	±	0.6	15 Peroxiredoxin-1	2.2	±	0.9
16 Aldehyde dehydrogenase, dimeric NADP-preferring	1.1	±	0.2	16 Collagen alpha-1(III) chain	0.6	±	0.1	16 Serum albumin	2.1	±	0.8	16 Fatty acid-binding protein, epidermal	2.1	±	0.4
17 Collagen alpha-1(XII) chain	1	±	0.4	17 Collagen alpha-1(V) chain	0.5	±	0	17 Apolipoprotein D	2	±	0.3	17 Ubiquitin-40S ribosomal protein S27a	2.1	±	0.1
18 Collagen alpha-1(V) chain	1	±	0.1	18 Actin, cytoplasmic 1	0.5	±	0.1	18 Alpha-enolase	2	±	1	18 Serum albumin	1.9	±	0.3
19 Vimentin	0.9	±	0.2	19 Collagen alpha-2(V) chain	0.4	±	0.1	19 Collagen alpha-3(VI) chain	1.9	±	1	19 Lysozyme C	1.9	±	0.4
20 Ig gamma-2 chain C region	0.9	±	0.1	20 Clusterin	0.4	±	0	20 Ig gamma-1 chain C region	1.9	±	0.8	20 Heat shock protein beta-1	1.8	±	0.5
21 Annexin A2	0.8	±	0.4	21 Biglycan	0.4	±	0.1	21 Prostaglandin-H2 D-isomerase	1.8	±	0.6	21 Apolipoprotein A-IV	1.6	±	1.1
22 Prostaglandin-H2 D-isomerase	0.8	±	0	22 L-lactate dehydrogenase A chain	0.4	±	0.1	22 Vimentin	1.7	±	0.6	22 Ig gamma-1 chain C region	1.5	±	0.9
23 Apolipoprotein D	0.7	±	0	23 MAM domain-containing protein 2	0.2	±	0.1	23 Aldehyde dehydrogenase, dimeric NADP-preferring	1.6	±	1.5	23 Keratocan	1.5	±	0.6
24 Collagen alpha-1(III) chain	0.7	±	0.2	24 Collagen alpha-1(XII) chain	0.1	±	0	24 Decorin	1.6	±	0.6	24 Annexin A2	1.4	±	0.3
25 Dermatopontin	0.7	±	0.1	25 Desmoplakin	0	±	0	25 Ig gamma-3 chain C region	1.6	±	0.5	25 Decorin	1.3	±	0.7

Table 3. Relative Amounts of Proteins Present Only in GCD1 Aggregates and LCD1 Deposits

	GCD1 Aggregates					LCD1 Stromal Deposits					LCD1 Bowman's Deposits			
	Name	**%**		**SD**		**Name**	**%**		**SD**		**Name**	**%**		**SD**
1	Hemoglobin subunit beta	2.9	±	3	1	Serum amyloid P-component	6.2	±	2.1	1	Protein S100-A9	6.8	±	2.7
2	Actin, cytoplasmic 1	0.5	±	0.1	2	Protein S100-A9	4.9	±	1.8	2	Cystatin-A	6.1	±	1.3
3	Desmoplakin	0	±	0	3	Cystatin-A	3.9	±	0.6	3	Serum amyloid P-component	5.2	±	1.1
					4	Hemoglobin subunit beta	2.7	±	0.2	4	Protein S100-A6	3.6	±	0.8
					5	Ubiquitin-40S ribosomal protein S27a	2.7	±	1.3	5	Protein S100-A4	3.4	±	0.3
					6	Apolipoprotein A-IV	2.3	±	0.9	6	Actin, cytoplasmic 1	2.3	±	0.7
					7	Apolipoprotein D	2	±	0.3	7	Peroxiredoxin-1	2.2	±	0.9
					8	Actin, cytoplasmic 1	1.3	±	0.3	8	Fatty acid-binding protein, epidermal	2.1	±	0.4
					9	Serine protease HTRA1	1.2	±	0.6	9	Ubiquitin-40S ribosomal protein S27a	2.1	±	0.1
					10	Apolipoprotein A-I	0.9	±	0.3	10	Lysozyme C	1.9	±	0.4
					11	Actin, aortic smooth muscle	0.9	±	0.1	11	Heat shock protein beta-1	1.8	±	0.5
					12	Serpin B12	0.8	±	0	12	Apolipoprotein A-IV	1.6	±	1.1
					13	Junction plakoglobin	0.6	±	0.2	13	Pyruvate kinase PKM	1.3	±	1
					14	Retinal dehydrogenase 1	0.6	±	0.1	14	Glyceraldehyde-3-phosphate dehydrogenase	1.3	±	0.2
					15	Olfactomedin-like protein 3	0.5	±	0.1	15	Junction plakoglobin	1.2	±	0.8
					16	Keratinocyte proline-rich protein	0.4	±	0.1	16	Retinal dehydrogenase 1	1.2	±	0.9
					17	Desmoplakin	0.3	±	0.1	17	L-lactate dehydrogenase A chain	1.1	±	0.1
					18	Desmoglein-1	0.3	±	0.1	18	Elongation factor 1-alpha 1	0.8	±	0.2
										19	Transketolase	0.7	±	0.5
										20	Desmoplakin	0.6	±	0.1
										21	Keratinocyte proline-rich protein	0.5	±	0
										22	Desmoglein-1	0.4	±	0
										23	Desmocollin-1	0.3	±	0.1
										24	Filaggrin-2	0.1	±	0

vitro[25] where, in the presence of three peptides of the amyloid precursor-like protein 2, mRNA and protein levels of S100A9 rose significantly. This is in addition to previous research that discovered similar propensities for amyloid aggregate formation between all three S100 proteins discovered within these LCD1 deposits; S100A4, A6, and A9.[26]

Irregular Proteolytic Processing of TGFBIp in Dystrophic Tissue

Through analysis of the proteolytic processing of TGFBIp within dystrophic aggregates it may be possible to better understand the mechanisms by which these aggregates arise. We, along with others, have previously shown that the thermodynamic stability of mutant isoforms of TGFBIp is altered compared to the wild type.[27–29] A study by Grothe et al.[29] investigated the stability of two mutant TGFBIp isoforms (R124C and R555W). It was observed that in the presence of superoxide species, both mutant proteins readily unfolded over 48 hours, in comparison to wild-type TGFBIp. In addition, this study assessed amyloid fibril formation, quantified by a Thioflavin T (ThT) fluorescence assay, in the presence or absence of superoxide. It was observed that fibril formation was greatly increased among mutant TGFBIp species, both R124C and R555W when in the presence of superoxide, while wild type TGFBIp showed equivalent fibril formation regardless of the presence of superoxide. This is particularly significant to disorders affecting the cornea, as the cornea absorbs approximately 90% of incident UV-B radiation, which is the main trigger for superoxide species production.[30] It is interesting to note the comparable ThT levels, indicating amyloid fibril formation, in R124C and R555W TGFBIp isoforms, in support of previous research carried out by Kim et al.[31] Nevertheless, the stability of the TGFBIp within the stroma appears pertinent to the accumulation of stromal deposits.

The spectral count analysis of diseased and control tissues did not indicate accumulation of any TGFBIp polypeptide region in either GCD1 or LCD1 when compared with control tissue (Fig. 2). Korvatska et al.,[32] reported altered processing of TGFBIp in R124C corneal tissue leading to accumulation of the N-terminal part of TGFBIp within the stroma. In addition, Schmitt-Bernard and colleagues[33] showed that a polypeptide region within the first FAS1 domain of TGFBIp had a high propensity for amyloid fibrillation in vitro. Both stromal and Bowman's deposits are most abundantly represented by the semitryptic peptides within the polypeptide region L128-R172 constituting approximately 30% of all semitryptic peptides

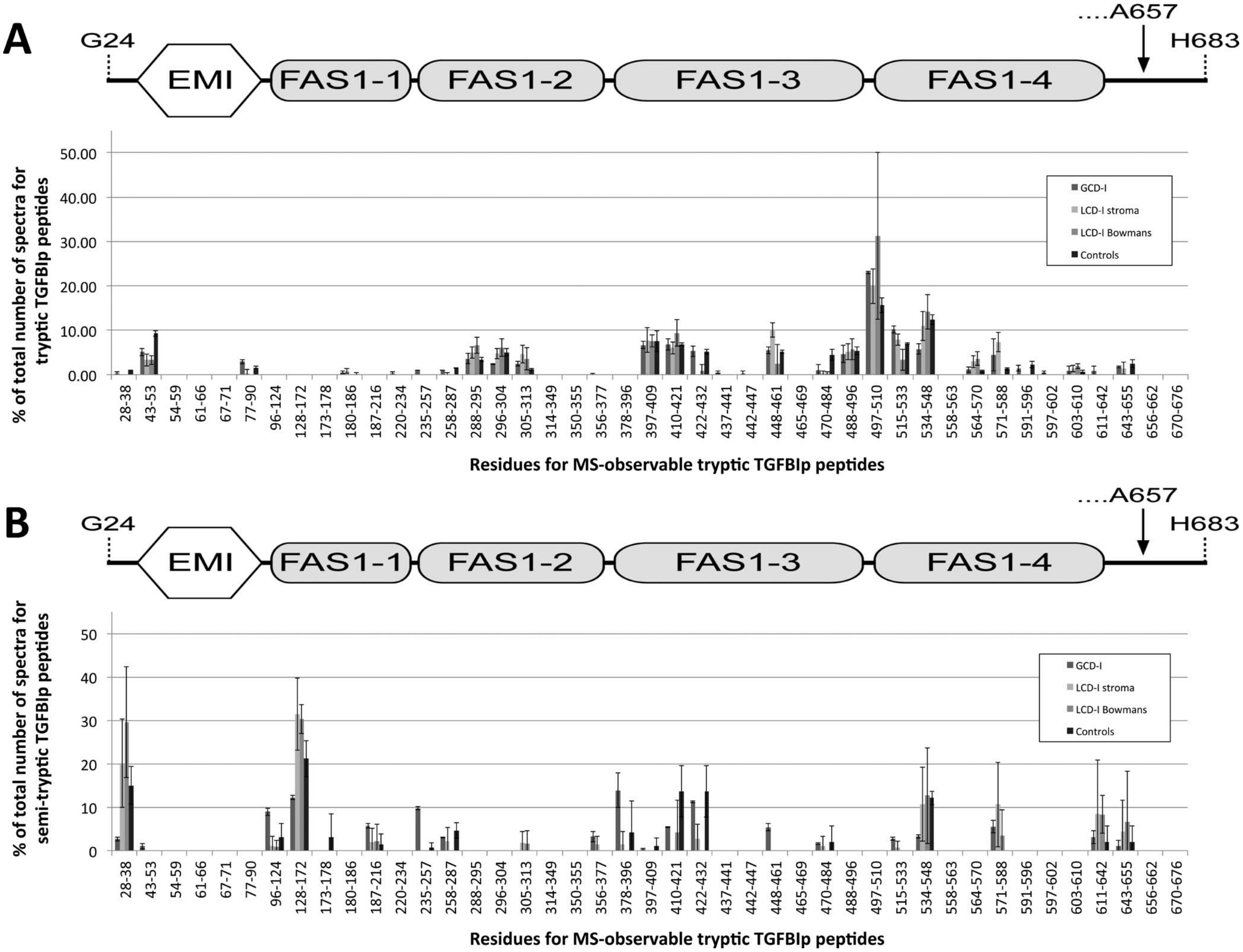

Figure 2. Spectra of TGFBIp peptides found in healthy stromal, GCD1 aggregates, and LCD1 deposits. The peptide spectra in control tissue, GCD1 aggregates, and LCD1 deposits allow for a better understanding of the proteolytic processing of TGFBIp. Peptide abundances were separated into those identified by tryptic (**A**) and semitryptic (**B**) analysis.

observed for both types of amyloid deposits. In addition a significant increase in abundance of this L128-R172 region by approximately 50% in the stromal amyloid deposits compared to the same polypeptide region in the control group was observed (Fig. 2B). The polypeptide region L128-R172 partly overlaps with the amyloid prone peptide described by Schmitt-Bernard et al.[33] supporting their conclusion that this region is involved in progression of LCD1 caused by the R124C mutation. Of note, the polypeptide region L128-R172 was only observed to a minor extent in all the previous proteomic analyses performed on LCD variants caused by mutations in the FAS1-4 domain,[7–9] which implies the exploitation of a distinctive molecular mechanism that leads to a similar phenotype to LCD1. Furthermore, distribution of endogenous cleavage sites in R124C deposits also differ in some of the P1 positions seen in previous observations,[9] further supporting this assumption of an altered mechanism leading to LCD1, depending on the site of the mutation.

The spectral count of peptides for the GCD1 cases presented a somewhat uniform distribution across the full-length TGFBIp (Fig. 2A). This suggests that TGFBIp-R555W may be processed to a lesser degree than TGFBIp-R124C in LCD1, and exists as a full-length, or close to full-length, protein within the stromal aggregates. Previously the R555W mutation has been shown to increase overall thermodynamic stability of the mutant TGFBIp.[28] This structural preservation is supported by the identification of only three proteins solely identified in the GCD1 aggregates when compared to the control group (Table 3). This may result from a lower in vivo immunological response to the presence of the TGFBI-R555W isoform or any protein accumulation.

Comparative Analysis of Stromal and Bowman's LCD1 Deposits

Throughout, both stromal and Bowman's LCD1 deposits were analyzed separately with the aim of identifying differences that may exist between both deposits. The proteome of these two deposits was found to be comparable (Table 2), though the presence of a few proteins did appear to be specific to either one of the deposits. For example, HtrA1 was present within the stromal deposits, but absent in the Bowman's deposits. This

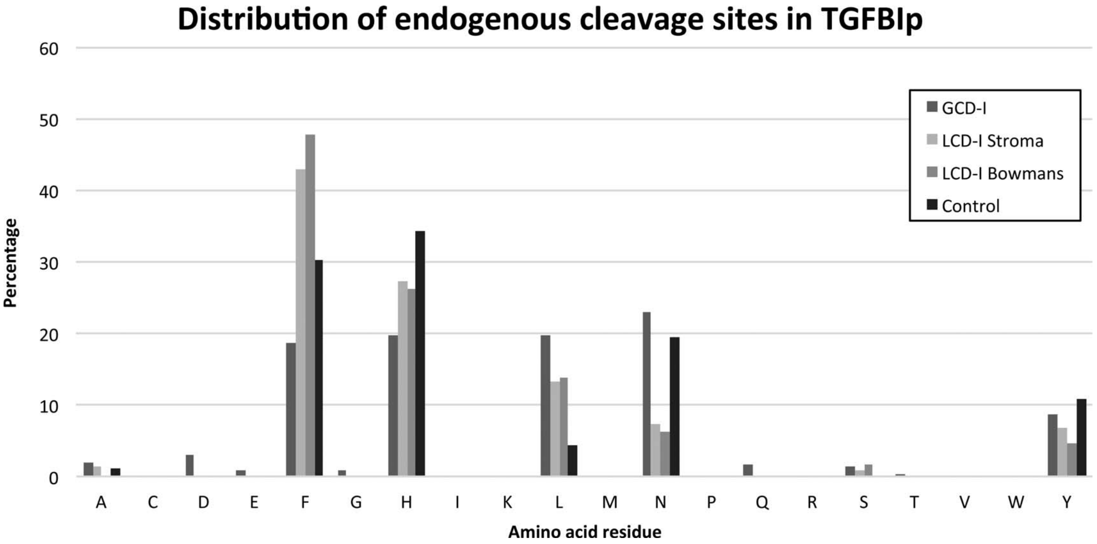

FIGURE 3. Analysis of endogenous cleavage sites of TGFBIp in healthy stromal, GCD1 aggregates, and LCD1 deposits. The number of endogenous cleavage sites for TGFBIp from healthy stroma, GCD1 aggregates and LCD1 deposits were quantified and the exact residue at each endogenous cleavage site was identified to better understand the processing of mutant TGFBIp isoforms in vivo.

could imply that more than one molecular pathway should be considered when investigating the two types of amyloid deposits observed in the R124C-associated deposits. In addition to HtrA1, the serine proteinase inhibitor Serpin B12 was likewise only found in the LCD1 stromal deposits (Table 3), suggesting a role as a potential inhibitor of HtrA1. Serpin B12 has previously been identified in LCD1 amyloid deposits caused by mutations in the FAS1-4,[9] also in presence of HtrA1. Further investigation is required to better elucidate the potential interplay between these two proteins.

The apolipoproteins Apo D and Apo A-I were likewise only observed in the stromal LCD1 deposits as also observed in our previous LCD1 studies,[8,9] whereas protein S100-A4 and protein S100-A6 were solely observed in Bowman's deposits. Interestingly, summing the molar fraction of the three members of the protein S-100 family found in Bowman's deposits exceeds the molar fraction of TGFBIp by almost double (13.8% vs. 7.1%). The capabilities of members of the protein S100 family to form amyloid on their own[26] may suggest that development of Bowman's deposits are highly dependent on coaggregation between TGFBIp and the three S100 proteins observed in Bowman's deposits.

Upon analysis of tryptic and semitryptic peptide spectra no apparent differences were observed between stromal and Bowman's deposits, most likely reflecting that both types of deposits encounter the same proteolytic machinery in the cornea. However, our previous research on LCD amyloid deposits of FAS1-4 domain mutants indicated excessive proteolytic processing in the polypeptide region F515-R533 of TGFBIp and accumulation of the polypeptide region Y571-R588.[8,9] This further supports the notion that different molecular mechanisms take place in the occurrence of a LCD1 phenotype.

CONCLUSIONS

It is clear that a different proteomic environment exists within these dystrophic aggregates and deposits, in comparison to healthy stromal tissue. A much greater abundance of amyloidogenic proteins was found within both types of LCD1 deposits. The fact that the stromal and Bowman's amyloid deposits also differentiated from each other's protein profiles in terms of amyloidogenic proteins could imply that different molecular mechanism are involved in the formation of the two types of amyloid deposits. The tryptic peptide spectra for TGFBIp do indicate unusual processing of the FAS1-4 domain in the R124C form of TGFBIp within LCD1 deposits. As this processing was found to be different from previously conducted analyses on FAS1-4 domain related TGFBIp mutations, it could well be the case that multiple pathological pathways should be considered when describing the common LCD1 phenotype characterized by amyloid deposition in the corneal stroma. The typical degradation pathways that eliminate wild-type TGFBIp from the extracellular matrix may be ineffective in removing mutant forms of TGFBIp. This incomplete proteolytic processing of TGFBIp, perhaps by serine protease HtrA1,[8] could result in the formation of amyloid deposits within the stroma, while other proteases within the extracellular matrix may be involved in the occurrence of nonamyloid deposits. These results continue to develop our understanding into the onset of these dystrophic aggregates and deposits, while further research into the exact processing of mutant TGFBIp species would continue to enhance this knowledge base.

Acknowledgments

We thank the Winston Churchill Memorial Trust for the travel fellowship that allowed for the establishment of this collaboration and for the retrieval and processing of these samples. The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium[34] via the PRIDE partner repository with the dataset identifier PXD002236 and 10.6019/PXD002236.

Supported by the Danish Council for Independent Research–Medical Sciences (DFF-4004-00471), Fight for Sight Denmark, and Fight for Sight UK.

Disclosure: **D.G. Courtney**, None; **E. Toftgaard Poulsen**, None; **S. Kennedy**, None; **J.E. Moore**, None; **S.D. Atkinson**, None; **E.**

Maurizi, None; **M.A. Nesbit**, None; **C.B.T. Moore**, None; **J.J. Enghild**, None

References

1. Runager K, Enghild JJ, Klintworth GK. Focus on molecules: Transforming growth factor beta induced protein (TGFBIp). *Exp Eye Res*. 2008;87:298–299.
2. Munier FL, Frueh BE, Othenin-Girard P, et al. BIGH3 mutation spectrum in corneal dystrophies. *Invest Ophthalmol Vis Sci*. 2002;43:949–954.
3. Weiss JS, Møller HU, Aldave AJ, et al. IC3D classification of corneal dystrophies—edition 2. *Cornea*. 2015;34:117–1159.
4. Lakshminarayanan R, Chaurasia SS, Anandalakshmi V, et al. Clinical and genetic aspects of the TGFBI-associated corneal dystrophies. *Ocul Surf*. 2014;12:234–251.
5. Liu Z, Wang Y, Gong Q, Xie L. An R124C mutation in TGFBI caused lattice corneal dystrophy type I with a variable phenotype in three Chinese families. *Mol Vis*. 2008;14:1234–1239.
6. Yoshida S, Yoshida A, Nakao S, et al. Lattice corneal dystrophy type I without typical lattice lines: role of mutational analysis. *Am J Ophthalmol*. 2004;137:586–588.
7. Karring H, Runager K, Thøgersen IB, Klintworth GK, Højrup P, Enghild JJ. Composition and proteolytic processing of corneal deposits associated with mutations in the TGFBI gene. *Exp Eye Res*. 2012;96:163–170.
8. Karring H, Poulsen ET, Runager K, et al. Serine protease HtrA1 accumulates in corneal transforming growth factor beta induced protein (TGFBIp) amyloid deposits. *Mol Vis*. 2013; 19:861–876.
9. Poulsen ET, Runager K, Risør MW, et al. Comparison of two phenotypically distinct lattice corneal dystrophies caused by mutations in the transforming growth factor beta induced (TGFBI) gene. *Proteomics Clin Appl*. 2014;8:168–77.
10. Klintworth GK. Lattice corneal dystrophy. An inherited variety of amyloidosis restricted to the cornea. *Am J Pathol*. 1967;50: 371–399.
11. Ishihama Y, Oda Y, Tabata T, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*. 2005;4:1265–1272.
12. Dyrlund TF, Poulsen ET, Scavenius C, Sanggaard KW, Enghild JJ. MS Data Miner: a web-based software tool to analyze, compare, and share mass spectrometry protein identifications. *Proteomics*. 2012;12:2792–2796.
13. Courtney DG, Atkinson SD, Moore JE, et al. Development of allele-specific gene-silencing siRNAs for TGFBI Arg124Cys in lattice corneal dystrophy type I. *Invest Ophthalmol Vis Sci*. 2014;55:977–985.
14. Yellore VS. Rayner SA, Aldave AJ. TGFB1-induced extracellular expression of TGFBIp and inhibition of TGFBIp expression by RNA interference in a human corneal epithelial cell line. *Invest Ophthalmol Vis Sci*. 2011;52:757–763.
15. Yuan C, Berscheit HL, Huang AJW. Identification of an amyloidogenic region on keratoepithelin via synthetic peptides. *FEBS Lett*. 2007;581:241–247.
16. Choi SI, Kim BY, Dadakhujaev S, et al. Inhibition of TGFBIp expression by lithium: implications for TGFBI-linked corneal dystrophy therapy. *Invest Ophthalmol Vis Sci*. 2011;52:3293–3300.
17. Coria F, Castaño E, Prelli F, et al. Isolation and characterization of amyloid P component from Alzheimer's disease and other types of cerebral amyloidosis. *Lab Invest*. 1988;58:454–458.
18. Frare E, Mossuto MF, Polverino de Laureto P, Dumoulin M, Dobson CM, Fontana A. Identification of the core structure of lysozyme amyloid fibrils by proteolysis. *J Mol Biol*. 2006;361: 551–561.
19. Pepys MB, Hawkins PN, Booth DR, et al. Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature*. 1993;362:553–557.
20. Bernstein HG, Rinne R, Kirschke H, Järvinen M, Knöfel B, Rinne A. Cystatin A-like immunoreactivity is widely distributed in human brain and accumulates in neuritic plaques of Alzheimer disease subjects. *Brain Res Bull*. 1994;33:477–481.
21. Grau S, Baldi A, Bussani R, et al. Implications of the serine protease HtrA1 in amyloid precursor protein processing. *Proc Natl Acad Sci U S A*. 2005;102:6021–6026.
22. Yanamandra K, Alexeyev O, Zamotin V, et al. Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS One*. 2009;4:e5562.
23. Krojer T, Pangerl K, Kurt J, et al. Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins. *Proc Natl Acad Sci U S A*. 2008;105:7702–7707.
24. Dyrlund TF, Poulsen ET, Scavenius C, et al. Human cornea proteome: identification and quantitation of the proteins of the three main layers including epithelium, stroma, and endothelium. *J Proteome Res*. 2012;11:4231–4239.
25. Li G, Chen H, Cheng L, Zhao R, Zhao J, Xu Y. Amyloid precursor-like protein 2 C-terminal fragments upregulate S100A9 gene and protein expression in BV2 cells. *Neural Regen Res*. 2014;9:1923–1928.
26. Carvalho SB, Botelho HM, Leal SS, Cardoso I, Fritz G, Gomes CM. Intrinsically disordered and aggregation prone regions underlie β-aggregation in S100 proteins. *PLoS One*. 2013;8: e76629.
27. Underhaug J, Koldsø H, Runager K, et al. Mutation in transforming growth factor beta induced protein associated with granular corneal dystrophy type 1 reduces the proteolytic susceptibility through local structural stabilization. *Biochim Biophys Acta*. 2013;1834:2812–2822.
28. Runager K, Basaiawmoit RV, Deva T, et al. Human phenotypically distinct TGFBI corneal dystrophies are linked to the stability of the fourth FAS1 domain of TGFBIp. *J Biol Chem*. 2011;286:4951–4958.
29. Grothe HL, Little MR, Sjogren PP, Chang AA, Nelson EF, Yuan C. Altered protein conformation and lower stability of the dystrophic transforming growth factor beta-induced protein mutants. *Mol Vis*. 2013;19:593–603.
30. Kolozsvári L, Nógrádi A, Hopp B, Bor Z. UV absorbance of the human cornea in the 240- to 400-nm range. *Invest Ophthalmol Vis Sci*. 2002;43:2165–2168.
31. Kim JE, Park RW, Choi JY, et al. Molecular properties of wild-type and mutant bIG-H3 proteins. *Invest Ophthalmol Vis Sci*. 2002;43:656–661.
32. Korvatska E, Henry H, Mashima Y, et al. Amyloid and non-amyloid forms of 5q31-linked corneal dystrophy resulting from kerato-epithelin mutations at Arg-124 are associated with abnormal turnover of the protein. *J Biol Chem*. 2000;275: 11465–11469.
33. Schmitt-Bernard CF, Chavanieu A, Derancourt J, et al. In vitro creation of amyloid fibrils from native and Arg124Cys mutated betaIGH3((110-131)) peptides, and its relevance for lattice corneal amyloid dystrophy type I. *Biochem Biophys Res Commun*. 2000;273:649–653.
34. Vizcaíno JA, Deutsch EW, Wang R, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol*. 2014;32:223–226.