TGFB1p-associated corneal dystrophy : exploring structure-based drug development strategies for disease prevention and treatment

Anandalakshmi, Venkatraman

2018

Anandalakshmi, V. (2018). TGFB1p-associated corneal dystrophy : exploring structure-based drug development strategies for disease prevention and treatment. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/73264>

<https://doi.org/10.32657/10356/73264>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**TGFBIP -ASSOCIATED CORNEAL DYSTROPHY:
EXPLORING STRUCTURE-BASED DRUG
DEVELOPMENT STRATEGIES FOR DISEASE
PREVENTION AND TREATMENT**

**VENKATRAMAN ANANDALAKSHMI
School of Biological Sciences
2017**

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**TGFB β -ASSOCIATED CORNEAL DYSTROPHY:
EXPLORING STRUCTURE-BASED DRUG DEVELOPMENT
STRATEGIES FOR DISEASE PREVENTION AND
TREATMENT**

By

VENKATRAMAN ANANDALAKSHMI (G1301256H)

Supervisors: Assoc. Prof. Konstantin Pervushin

Assoc. Prof. Jodhbir S. Mehta

School of Biological Sciences

A thesis submitted for the degree of Doctor of Philosophy

NANYANG TECHNOLOGICAL UNIVERSITY

2017

Acknowledgments

Firstly I would like to express my gratitude to both my main Supervisors Associate Professor Konstantin Pervushin and Associate Professor Jodhbir S. Mehta for giving me an opportunity to pursue my PhD and for their great supervision. I would like to express my sincere sincere thanks to Prof Jod for trusting me and for the wonderful support and guidance throughout the journey. I always remain a great fan of your knowledge, energy and keen enthusiasm towards corneal research. I am grateful to Prof Konstantin for accepting me as a part of his research group and patience to teach me NMR. I thank Prof Konstantin for his support and allowing me to work in my own research space and exploring my research strengths. My sincere thanks to Prof Konstantin for the various scientific discussions, brainstorming and knowledge transfer sessions. I am also grateful to Singapore Eye research Institute (SERI) for providing an outstanding research environment and SERI graduate Scholarship. I would also like to thank the leaders and Trailblazers of SERI, Professor Donald Tan, Professor Aung Tin and Professor Wong Tien Yin for their encouragement. I would also like to thank my role models and “Super Women” in research, Associate Professor Eranga Vithana and Associate Professor Tina Wong for their wonderful inspiration. I am also thankful for the guidance and motivation from my “Pillars of Strength” Associate Professor Lakshminarayanan and Assistant Professor Shyam S. Chaurasia without whom my PhD would still be a dream. I would like to express a very special thanks to Associate Professor Lakshminarayanan for all the help and guidance, patiently correcting and proofreading my manuscript drafts and discussions to carry out my key experiments.

I am also grateful to my Thesis advisory committee (TAC) members Associate Professor Newman Sze, Associate Professor Mu Yuguang for all the valuable suggestions and comments. I also thank our Collaborators Professor Sten Ohlson, Professor Newman Sze and Dr Anita Chan for their commitment and assistance towards the project and fruitful collaboration. I am thankful to my colleagues both at SERI and NTU. I thank Ms. Margaret Phillips for her incredible patience in helping me out whenever I needed an extra pair of hands. I would also like to thank my SERI colleagues Dr Elavazhagan and Dr Gary Yam and my NTU lab mates Mr Yossa Dwi Hortono, Ms Areetha D’Souza, Ms Rubing Liu, Ms Bai Yang, Dr Mangesh Joshi, Dr Harini Mohanram, Dr Mukesh Mahajan, Dr Deepak, Ms Bhuvaneshwari, Ms Bhargy Sharma and Mr Lewis for their help in the lab.

I would like to thank my colleagues from various institutes who helped me with my experiments. I thank Dr Bamaprasad Dutta, Dr Hao Piliang, Mr Chua Khi Pin, Dr Dao Minh Thi, Ms Prachi Singh, Ms Jagjeet Kaur, Mr. Tim, Ms Adeline, Dr Ravisankar, Ms Qiu Xuan and Ms Vivian, for their patience in teaching me a new technique and reciprocating equal interest in the project. I express my heartfelt thanks to Dr Toshio Yamazaki and Dr Toshio Nagashima from Riken Institute, Japan for helping me with the solid-state NMR experiments and their patience in helping me learn about solid-state NMR. I would also like to extend my thanks to Dr Andreas Larsson for listening to me patiently and for the many long discussions on the drug discovery project. I thank Dr Larsson for sharing his ideas and personal experiences in designing experiments to screen and validate compounds.

I thank my family members for their kind understanding and support throughout my life especially during the past 4 years where I tried to juggle between research, studies and family. I also thank my husband and loving sons for their unconditional love, care and concern towards me. I express my heartfelt thanks to my husband, Mr Venkatraman for being there with me always.

I express my gratitude to the numerous researchers and clinicians who paved an excellent path by their valuable contributions and discoveries in the field of Corneal Dystrophy research. “If I have seen further than the others, it is merely by standing on the shoulders of giants and pioneers who paved the way for me”. I have named only a few here and would like to take this opportunity to thank all the kind souls who have helped me either directly or indirectly to reach where I am now.

Table of Contents

1	Introduction	6
1.1	Structure of healthy human cornea.....	6
1.2	Layers of the human cornea	6
2	Dystrophies of the Cornea	9
2.1	Classification of Dystrophies based on anatomical locations of cornea	10
2.2	History of <i>TGFBI</i> associated Corneal Dystrophies.....	11
2.3	International Committee for Classification of Corneal Dystrophies (IC3D).....	13
2.4	Survey on the epidemiology of <i>TGFBI</i> -corneal dystrophies	16
2.5	Possible Bias on the reported epidemiological studies.....	21
2.6	Classification and characteristics of <i>TGFBI</i> associated corneal dystrophy	26
2.7	Phenotype-Genotype correlation in <i>TGFBI</i> associated mutations	33
3	Structure and function of TGFBIp	40
3.1	Gene and protein structure	40
3.2	Expression and Regulation of TGFBIp	42
3.3	Structural Organisation of TGFBIp	42
3.4	Functional properties of TGFBIp	46
4	Alteration of structural conformation and stability of TGFBIp	53
4.1	Cornea specific aggregation of mutated TGFBIp	54
4.2	Localization of protein deposits in various layers of cornea	55
4.3	Granular deposition of mutant protein in the cornea	56
4.4	Heterogeneity in the proteolytic processing of the TGFBIp-mutant proteins ...	56
4.5	Treatment Strategies of Corneal Dystrophies	59
5	Protein composition analysis of amyloid corneal aggregates from patients	69
5.1	Introduction	69
5.2	Materials and Methods.....	71
5.2.1	Corneal dystrophy patient samples	71
5.2.2	Preparation of corneal tissue sections and laser capture microdissection (LCM).72	
5.2.3	Sample preparation for mass spectrometry LC-MS/MS analysis	72
5.2.4	TGFBIp synthetic peptides.....	74
5.2.5	Circular dichroism spectroscopy	74
5.2.6	Thioflavin T (ThT) assays.....	75
5.2.7	Transmission electron microscopy (TEM).....	75
5.3	Results.....	76
5.3.1	Patient sample identification and laser capture microdissection.....	76

5.3.2	Protein composition of amyloid deposits in corneal dystrophy patients.....	76
5.3.3	Identification of proteins enriched in amyloid deposits from LCD patients.....	80
5.3.4	Mutation-specific enrichment of TGFB1p peptides in Lattice amyloid deposits..	86
5.3.5	Enrichment of endogenous peptide cleavage sites in amyloid corneal deposits...	87
5.3.6	Change in conformation and secondary structure of synthetic TGFB1p peptides	92
5.3.7	Favourable aggregation kinetics of peptides enriched in LCD amyloid deposits.	93
5.4	Discussion	97
6	Bio-physical and structural characterization of TGFB1p peptides	105
6.1	Introduction	105
6.2	Materials and Methods	108
6.2.1	Materials	108
6.2.2	Spectral counts of tryptic peptides from TGFB1p using LC MS/MS	108
6.2.3	Circular Dichroism Spectroscopy Secondary structure of peptides.....	110
6.2.4	Thioflavin T assay.....	111
6.2.5	Transmission Electron Microscopy.....	112
6.2.6	LC-MS/MS analyses	112
6.2.7	Peptide Aggregation by 1D Solution NMR	113
6.2.8	Solid State NMR for the amyloid fibrils.....	114
6.2.9	Cytotoxicity of amyloids on Cultured Human Corneal Fibroblast (HCSF).....	115
6.3	Results.....	115
7.3.1	Identification of peptides enriched in the amyloid deposits of patients	115
7.3.2	Time-dependant transition of secondary structure of model peptides	118
7.3.3	Determination of Kinetics of amyloid formation by ThT assays.....	120
7.3.4	Aggregation of peptides measured by 1-D NMR	122
7.3.5	Verification of amyloid fibril morphology by TEM	126
7.3.6	Stability of the amyloid fibrils derived from peptides	129
7.3.7	Solid State NMR analysis of the most stable amyloid fibrils	132
7.3.8	Cytotoxicity of amyloid fibrils.....	139
6.4	Discussion	141
7	Screening of compounds to prevent protein aggregation	146
7.1	Introduction.....	146
7.2	Materials and methods.....	152
7.2.1	Expression and Purification of WT and mutant TGFB1p	152
7.2.2	F-Pocket analysis on WT-TGFB1p and R555W mutant.....	155
7.2.3	Screening of compounds using Weak affinity Chromatography (WAC)	156
7.2.4	Screening of Maybridge library using WAC	158
7.2.5	Sample preparation for NMR and NMR experimentation.....	160
7.2.6	Limited proteolysis by Trypsin	161
7.2.7	Peptide Mass fingerprinting (PMF) of the tryptic peptides.....	162
7.3	Results.....	163

7.3.1	Expression and Purification of 4th FAS1_WT and mutant TGFB1p	163
7.3.2	F-Pocket analysis of 4th FAS1 domain of TGFB1p-WT and mutant R555W	168
7.3.3	Screening of compounds using Weak Affinity Chromatography	170
7.3.4	Selection of compounds for validation screening by NMR	175
7.3.5	Limited proteolysis of 4th FAS-1 domain of TGFB1p-WT, R555W and H572 .	193
7.4	Discussion.....	197
8	Future Work and conclusion	202
8.1	Protein composition analysis of amyloid corneal aggregates from patients.....	202
8.2	Bio-physical and structural characterization of TGFB1p peptides.....	203
8.3	Screening of compounds to prevent protein aggregation.....	207

List Of Figures:

Figure 1-1: Schematic representation of cornea cross-section.....	7
Figure 2-1: Time line of TGFBI associated Corneal Dystrophies	14
Figure 2-2: Corneal dystrophies and mutations in genes responsible for the dystrophy	16
Figure 2-3: Distribution of TGFBI associated mutations worldwide.	17
Figure 2-4: Slit lamp photographs of clinical phenotype of TGFBI-related corneal dystrophy ..	25
Figure 2-5: Pathological classification of <i>TGFBI</i> -associated corneal dystrophies	27
Figure 2-6: Age of onset of mutations associated with TGFBI-associated corneal dystrophy....	28
Figure 3-1: The domain structure of TGFBIp.....	41
Figure 3-2: Distribution of Cys residues in TGFBIp..	42
Figure 3-3: Sequence alignment of TGFBIp and the proteolytic cleavage sites of TGFBIp	43
Figure 3-4: Available structure of TGFBIp from the PDB website.....	45
Figure 5-1: Corneal images showing lattice lines and aggregated proteins in patients.	77
Figure 5-2: Isolation of amyloid corneal deposits by laser capture microdissection..	77
Figure 5-3: Protein profiles of amyloid deposits and surrounding tissues in LCD patients..	78
Figure 5-4: Protein profiles of corneal deposits and surrounding tissues grouped in patients.....	79
Figure 5-5: Protein profile of healthy control cornea.	79
Figure 5-6: Spectral counts of peptides derived from TGFβIp by trypsin digestion..	88
Figure 5-7: Normalized logarithmic area of peptide enrichment in individual patient samples..	91
Figure 5-8: Endogenous cleavage sites in corneal deposits, adjacent tissues, and controls..	92
Figure 5-9: Circular dichroism spectra of TGFBIp peptides.	94
Figure 5-10: Aggregation kinetics of TGFBIp peptides 515-533 and 571-588.....	94
Figure 5-11: TEM images of amyloid fibrils derived from TGFBIp peptides.	96
Figure 5-12: Aggregation propensity of 4th FAS1 human TGFBIp algorithms.....	101
Figure 6-1: The spectral counts of relative abundance of peptides that are highly enriched in dystrophy patients compared to the healthy control cornea.....	117
Figure 6-2: Circular Dichroism spectra of the peptides at different time points.....	119
Figure 6-3: Aggregation kinetics of peptides as measured by Thioflavin T assays.....	121
Figure 6-4: Aggregation kinetics of peptides followed by 1-D NMR	125
Figure 6-5: Rate Constant (K) of peptides based on 1-D NMR.....	126
Figure 6-6: Morphology of the amyloid fibrils.....	128
Figure 6-7: Thermal denaturation of Amyloid fibrils:	131
Figure 6-8: TEM of amyloid fibrils after heat-cool cycle:.....	132
Figure 6-9: Cp MAS experiment for the amyloid fibrils derived from peptides.	133
Figure 6-10: DARR spectra of the most stable P5 peptide:.....	135
Figure 6-11: NCA and NCO spectra for P5.....	136
Figure 6-12: C-alpha and C-beta chemical shifts of P5	136
Figure 6-13: Aliphatic region of the DARR spectrum of the amyloid fibrils from P1	137
Figure 6-14: NCA experiment for fibrils derived from P1 peptide.	138
Figure 6-15: Cytotoxicity and inhibition of cell proliferation on corneal fibroblasts.....	140
Figure 7-1: Size exclusion Chromatography (SEC) elution profile for TGFBIp-WT protein...	164
Figure 7-2: SEC elution profile for TGFBIp-R555W and H572R protein.	165
Figure 7-3: MS analysis of 4 th FAS-1 domain of TGFBIp-WT, R555W and H572R proteins..	166
Figure 7-4: A. SDS-PAGE analysis of 4 th FAS1 domain proteins.	167
Figure 7-5: F-pocket analysis of 4thFAS-1 domain of TGFBIp-WT and R555W proteins.	169

Figure 7-6: Ligand Efficiency (LE) values of interactions between the final hits and targets R555W, H572R and WT.....	174
Figure 7-7: List of top 10 compounds that showed higher affinities towards mutants.....	174
Figure 7-8: Retention time profiles of compounds and DMSO with the column immobilised with proteins of the 4 th FAS-1 domain of TGFBIp H572R, WT and R555W	177
Figure 7-9: ¹⁵ N-HSQC spectra of TGFBIp-WT and the mutant H572R protein.	180
Figure 7-10: Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with MO07617 compound..	184
Figure 7-11: Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with RJF00203 compound.....	185
Figure 7-12: (A) Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with BTB05094 compound..	186
Figure 7-13: Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with compounds CC03813 and CC07101.	187
Figure 7-14: Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with compounds BTB07813 and BTB10358.	188
Figure 7-15: Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with compounds SEW04816 and RDR00814.	189
Figure 7-16: Chemical shift perturbations of individual amino acid of 4 th FAS-1 domain of TGFBIp with compounds BTB07813 and BTB10358.	190
Figure 7-17: RMSD values for the first half of the 132 amino acid of TGFBIp-R555W mutant protein to top 10 compound binding validated by ¹⁵ N-HSQC based NMR assays.....	191
Figure 7-18: RMSD values for the second half of the 132 amino acid of TGFBIp-R555W mutant protein to top 10 compound binding validated by ¹⁵ N-HSQC based NMR assays.....	192
Figure 8-1: Vector map of pET 28a (+) vector and the sequence of TGFBI used in the study .	204
Figure 8-2: Expression optimization of target protein with Histidine tag in E-coli BL21 (DE3).	205
Figure 8-3: SDS-PAGE analysis of 2ug purified protein with Histidine tag.....	206
Figure 8-4: Western blot analysis of the purified protein using anti-His tag antibody.	206

List of Tables:

Table 2-1: Classification of corneal dystrophies.....	10
Table 2-2: List of <i>TGFBI</i> mutations reported so far and the corresponding clinical phenotypes	18
Table 2-3: Distribution of p.Arg124His mutation worldwide.....	21
Table 2-4: Distribution of p.Arg124cys worldwide.....	22
Table 2-5: Distribution of p.Arg555Trp mutation worldwide	24
Table 2-6: List of two simultaneous mutations in the <i>TGFBI</i> -associated corneal dystrophies....	36
Table 5-1: Patient genotype and disease phenotypes used in this study	71
Table 5-2: Number of proteins identified in amyloid corneal deposits and adjacent tissues.....	80
Table 5-3: Relative abundance of the top 25 proteins identified in amyloid corneal deposits. ...	82
Table 5-4: List of proteins enriched in LCD deposits relative to healthy control cornea	85
Table 5-5: List of protein enriched in the tissue surrounding the amyloid deposits	89
Table 5-6: Aggregation kinetics of TGFBIp peptides 511-533 and 571-588	95
Table 6-1: Change in overall net charge in clinically relevant mutations.....	108
Table 6-2: Peptide sequence of TGFBIp-4th FAS-1 domain (amino acids 611 to 633) WT and various disease-associated mutants.....	113
Table 6-3: Peptide sequence of TGFBIp-4th FAS-1 domain	117
Table 6-4: Aggregation kinetics of the peptides.	120
Table 6-5: Aggregation kinetics of the peptides.	126
Table 7-1: composition of the LB-broth used for culture	154
Table 7-2: Characteristics of target proteins (4th FAS-1 domain of TGFBIp)	156
Table 7-3: Properties of the compounds included in the R03 Maybridge library	157
Table 7-4: Details of immobilizations of target proteins and obtained columns	158
Table 7-5: Output from the F-Pocket analysis for 4 th FAS-1 domain of TGFBIp-WT and mutant R555W	168
Table 7-6: Immobilization results and affinity column characteristics.....	170
Table 7-7: SDs of retention times of fragments by duplicate mixture injections and detection of mixture screening on affinity columns.....	171
Table 7-8: Hits from the primary screening in mixtures.....	171
Table 7-9: Characteristic of final hits after confirmation by analyzing single solution.....	172
Table 7-10: List of compounds used for secondary NMR screening.....	176
Table 7-11: Peptide fragments generated by limited proteolysis of R555W with trypsin and analysis by MALDI mass fingerprinting	195
Table 7-12: Peptide fragments generated by limited proteolysis of H572R with trypsin and analysis by MALDI mass fingerprinting	196

Abbreviations

1D	One-Dimensional
2D	two-Dimensional
3D	three-Dimensional
¹ H	Protons
AA	Amino Acid
ACN	Acetonitrile
Ala	Alanine
ALK	Anterior Lamellar Keratoplasty
Arg	Arginine
AUC	Area under the curve
CD	Circular Dichroism
CI	Cell Index
CP-MAS	Cross-Polarization magic-angle spinning
CR	Congo red
CSCD	Congenital stromal dystrophy
DARR	Dipolar Assisted Rotational Resonance
DALK	Deep Anterior Lamellar Keratoplasty (DALK)
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DSS	4, 4-dimethyl-4-silapentane-1-sulfonic acid
DTT	Dithiothreitol
ECM	Extracellular Matrix
EIC	Extracted Ion Chromatogram
EMI	EMILIN- like domain
emPAI	Exponentially Modified Protein Abundance Index
ERED	Epithelial recurrent erosion dystrophy
FA	Formic Acid
FAS-1	Fasciclin like Domain-1
FCD	Fleck corneal dystrophy
FDR	False Discovery Rate
FFPE	Formalin-Fixed, Paraffin-Embedded
FPLC	Fast Protein Liquid Chromatography
GAGs	Glycosaminoglycans
GCD	Granular Corneal Dystrophy
GDLD	Gelatinous drop like corneal dystrophy
Gly	Glycine
HCSF	Cultured Human Corneal Fibroblast
HFIP	Hexafluoro-2-propanol
His-tag	Histidine tag
HSQC	Heteronuclear Single Quantum Coherence Spectroscopy
HTRA1	Human high temperature requirement serine protease A1
IAA	Indole 3- Acetic acid
IC3D	International Committee for Classification of Corneal Dystrophies
Ile	Isoleucine
IPTG	Isopropyl-beta-D-thiogalactopyranoside

KDa	Kilodalton
KE	Keratoepithelin
LASIK	Laser-assisted in situ keratomileusis
LB	Luria Bertani
LC	Liquid Chromatography
LCD	Lattice Corneal Dystrophy
LCM	Laser Capture microdissection
LC-MS	Liquid chromatography-Mass Spectrometry
LE	Ligand Efficiency
LECD	Lisch epithelial dystrophy
MAS	Magic angle spinning
MALDI	Matrix Assisted Laser Desorption/Ionization
MCD	Macular corneal dystrophy
MECD	Meesmann dystrophy
MGF	Mascot Generic Format
MHz	Mega hertz
MMC	Mitomycin C
MMP 9	Matrix Metalloprotease-9
MRW	Mean Residual Weight
ms	Millisecond
MT	Masson Trichrome
MW	Molecular weight
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PACD	Posterior amorphous corneal dystrophy
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEN	Polyethylene Naphthalate
PK	Penetrating Keratoplasty
PMF	Peptide Mass Fingerprinting
PPCD	Posterior polymorphous dystrophy
PTA	Phosphotungstic Acid
PTK	Phototherapeutic Keratectomy
RBCD	Reis-Bücklers Corneal Dystrophy
RMSD	Root Mean Square Deviation
RP-HPLC	Reverse Phase- High Performance Liquid Chromatography
RTCA	Real-Time Cell Analysis
SAXS	Small Angle X-ray Scattering
SCD	Schnyder corneal dystrophy
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulphate-polyacrylamide gel electrophoresis
SEC	Size Exclusion Chromatography
SMCD	Subepithelial mucinous corneal dystrophy
ssNMR	Solid State Nuclear Magnetic Resonance
TBCD	Thiel-Behnke Corneal Dystrophy
TEM	Transmission Electron Microscopy
<i>TGFB1</i>	Transforming Growth Factor Beta Induced Gene
TGFB1p	Transforming Growth Factor Beta Induced protein

ThT	Thioflavin T
T _m	Melting temperature
TNF- α	Tumor Necrosis Factor-alpha
T _t	Transition temperature
UV	Ultraviolet
VT-CD	Variable Temperature-Circular Dichroism
WAC	Weak Affinity Chromatography
WT	Wild Type
XIC	Extracted ion chromatogram
XECD	X-linked endothelial corneal dystrophy

Abstract

Corneal Dystrophies are a group of inherited disorders localized to various layers of the cornea that affect corneal transparency and visual acuity. The deposition of insoluble protein materials in the form of extracellular amyloid fibrils or intracellular cysts is pathognomonic. The gene associated with the stromal and Bowman's related Corneal Dystrophy is *TGFBI*. The protein product of the gene is known as Transforming Growth Factor-induced protein (TGFBIP). This protein is a 68 KDa protein with 683 amino acids with 4 Fasciclin-like domains. In pathological conditions, the protein accumulates as insoluble deposits in various forms. The severity, clinicopathologic variations, age of the onset, and location of the deposits depend on the type of amino acid alterations in the protein. Until 2006, 38 different pathogenic mutants were reported in the *TGFBI* gene associated with various phenotypes of corneal dystrophy. To date, there are more than 65 reported mutations in the gene, and 84% of the mutations are found in the 4th FAS1 domain, making it a mutational hotspot. Hence our research focused mainly on the 4th FAS1 domain and mutations involved in the region. There is no effective treatment to prevent, halt, or reverse the deposition of TGFBIP.

The main aim of this project is to understand the aggregation mechanism of mutant TGFBIP in the cornea. We have sought different approaches to understand the disease pathology and mechanism of protein aggregation and deposition. The first was a proteomics-based approach where we aimed to identify the composition of various proteins present in the amyloid corneal deposits of the patients with different mutations reported in *TGFBI*. The study aimed to understand the clearance mechanism of the mutant protein and compare it with the wild-type (WT) protein. The study also aimed to identify specific proteases and proteolytic mechanism involved in clearing the mutant protein in the cornea. The proteolytic processing of mutant TGFBIP protein compared to the WT protein may provide insight to the disease pathology. The proteomics data was vital to understand the role of specific proteases and the proteolytic clearance activity in WT and mutant protein. We identified enrichment of serine protease HTRA1 in the amyloid deposits of corneal dystrophy patients compared to the normal healthy

controls. The corneal amyloid deposits are also enriched with amyloid-associated proteins, non-fibrillar amyloid proteins and TGFBIp itself. The study also showed that the patient corneal amyloid deposit samples showed enrichment of certain tryptic peptide fragments like TGFBIp 515-533, TGFBIp 571-588 and TGFBIp 611-642, which are either not present or present only in low amounts in healthy controls. The study showed the possibility of mutant protein processing by serine protease HTRA1, which may give rise to peptides that potentially act as amyloid fibril seeds and enhance the aggregation process.

The second approach was to understand the mechanism of fibril formation using a peptide model. Based on the mass spectrometry analysis, we identified many short peptide residues (TGFBIp 515-533, TGFBIp 571-588 and TGFBIp 611-642) enriched in the amyloid corneal deposits of patients compared to the healthy control cornea. The main aim of the study was to understand the molecular-functional relationship of TGFBIp-derived mutant peptides where the amino acid substitutions decreased the overall net charge and their effects on the propensity of the peptides to form amyloid fibrils. We modified the 23 residue peptide (TGFBIp E⁶¹¹PVAEPDIMATNGVVHVITNVLQ⁶³³) with clinically relevant cationic charged amino acid substitution and investigated the amyloid-forming properties of the mutant peptides using various biophysical approaches. The peptide sequences were predicted to have a very high propensity to form amyloid fibres. The study also aimed to investigate the structure of amyloid fibrils derived from the peptides by Solid-state NMR. The study may be used to understand the fibrillation process, intermediate states of fibril formation, and will be useful for drug discovery and treatment of *TGFBI*-associated Corneal Dystrophy. The study provides an insight into the effects of substitution of individual amino acid on the rate of amyloid formation in the peptides, the structure, biological and biophysical properties of the amyloid fibrils formed by the peptides. We also explored how a position specific and type of amino acid substitution hugely contributes to the amyloidogenic property of the peptides.

The third was a drug discovery approach to identify small fragments or compounds that can bind to the TGFBIp and modulate the protein, thereby delaying the aggregation process or changing the proteolytic processing of the mutant protein and preventing the formation of highly amyloidogenic peptide seeds. The lead compound identified will be tested to see if it can alter the thermodynamic stability of the protein and thereby stabilize the protein or delay the formation of toxic oligomers or fibrils.

The compounds may possibly play a role in altering the proteolytic processing of the mutant TGFBIp. Thus the screening of 2500 compounds from the Maybridge library using initial Weak affinity chromatography and secondary ^{15}N -HSQC-based assays have identified 3 lead compounds that bind to the mutant protein near a predicted binding pocket. The lead compounds were added to the proteins, and limited proteolysis by trypsin showed that the addition of the compounds stabilised the secondary structure of the protein and prevented the formation of highly amyloidogenic peptide seeds. The compounds, when bound to the mutant proteins, may alter the proteolytic processing for efficient clearance or to prevent the formation of smaller peptides. The lead compounds are to be taken to the next stage of the drug discovery process.

A combination of computational, proteomics, structural and biophysical studies along with the clinical data provided an insight into the mechanism of cornea-specific protein aggregation and identification of lead compounds that may play a role in protein processing and proteolysis.

Even the smallest progress made in each of the different approaches will provide additional information to the basic understanding of the mechanism of protein aggregation, disease progression, and development of novel therapeutic strategies to treat *TGFBI*-associated Corneal Dystrophies.

Chapter 1

“Of all the senses, sight must be the most delightful.”-Helen Keller

1 Introduction

1.1 Structure of healthy human cornea

The cornea is a highly transparent avascular tissue. The cornea is situated in the anterior part of the eye. This unique tissue accounts for focusing the light on the optical centre of the retina and protects the ocular tissue from oxidative damage from UV light (Land and Fernald 1992). A thin layer of tear film and the aqueous humour covers the anterior part of the cornea. Both the tear film layer and aqueous humour are very crucial for normal corneal function. Any changes in the components of the aqueous humour or the tear fluid may result in a disease.

In an adult, the shape of the cornea is oval. The diameter of the cornea is approximately 11 mm and the thickness is 0.5 mm (Mishima 1968). The cornea requires a fine balance of cellular and molecular processes to retain structural and functional integrity. The transparency, convexity, surface smoothness and refractive index contribute to the optical property of the cornea. Any small changes in the specifications will affect the corneal function and leave the corneal transparency compromised.

The cornea is a densely innervated tissue, and it is estimated that per square mm it has 7000 nerve terminals (Muller, Marfurt et al. 2003). The nerves enter as unmyelinated bundles into the periphery of central stroma and the nerves branch out to smaller nerve fibres towards the anterior surface of the cornea (Rozsa and Beuerman 1982). The nerve fibres are able to penetrate the Bowman's layer easily, and they spread outward towards the epithelium in between the epithelial cells (Muller, Vrensen et al. 1997).

1.2 Layers of the human cornea

The human cornea is composed of five main layers: the corneal epithelium, Bowman's membrane, stroma, Descemet's membrane and the corneal endothelium (**Figure 1-1**). Each of the corneal layers has a distinct role and each layer is highly important to keep the cornea functional.

Figure 1-1: Schematic representation of cornea cross-section showing various layers of the cornea (Image adapted from discovery.lifemapsc.com and reproduced with permission from publishers).

The Corneal Epithelium

The anterior layer, the corneal epithelium is made up of 5 to 7 layers of cells. The epithelium is about 50 μm thick, which is approximately 10% of the entire thickness of the human cornea. The corneal epithelium has two major functions: to maintain a smooth surface of cells in order to prevent scattering of the incoming light and the second function is to prevent any harmful substance or agents from entering the cornea (Lu, Reinach Ps Fau - Kao et al. 2001). The corneal epithelium is constituted of cells that are five to six layers thick and three non-identical layers of epithelial cells namely superficial cells, wing cells and basal cells. Among the different cell types, the basal cells have the ability to proliferate. The basal cells synthesise laminin and collagen Type IV. Laminin and collagen are the eminent constituents that form the basement membrane.

Bowman's membrane

The thickness of the Bowman's membrane is about 10-12 μm and is an acellular layer. This layer is made up of non-specifically arranged collagen fibres (mainly of type I, III, IV and V) and different proteoglycans. The major function of this layer is not clear but this layer does not regenerate after a corneal injury and it is the strongest layer of the cornea.

The Corneal Stroma

The stromal layer of the cornea is about 500 μm thick and contributes to more than 90% of the corneal thickness (Bettelheim and Plessy 1975). The corneal stroma is mainly constituted by collagen and glycosaminoglycans. The main collagen fibres present in the stroma are types I, V and VI (Michelacci 2003). The arrangement of collagen in the stroma is very important, and their important function is to maintain the biomechanical and optical properties of the stroma (Freund, McCally R1 Fau - Farrell et al. , Jue and Maurice 1986). The glycosaminoglycans function to subsume and conserve huge amounts of water. The stroma has only about 2% to 3% of cells (keratocytes), which synthesise enzymes like matrix metalloproteinase. The hydration of the cornea is accomplished by a membrane pump located at the endothelium. The hydration is also regulated by surface evaporation and intraocular pressure.

Descemet's membrane

The Descemet's membrane is a thick membrane and is about 8-10 μm in thickness. This membrane consists of collagen types IV, VIII and XVIII (Sawada, Konomi et al. 1990, Leung, Rife et al. 2000, Kato, Chang et al. 2003) as well as laminin, dermatan sulphate, fibronectin, heparin sulphate and nidogen (Schittny 1988, Schittny, Kresse et al. 1995). The corneal endothelial cells continuously secrete the various constituents of the Descemet's membrane. This explains the increased thickness of the Descemet's membrane with age (Murphy, Alvarado et al. 1984). The Descemet's membrane also reflects any change in the shape of the corneal stroma. The Descemet's membrane tends to fold as a result of any swelling in the stroma. The function of the Descemet's membrane is to maintain structural integrity (Hopfer, Fukai et al. 2005) and accurate alignment and focusing of the light (Jakus 1956). The Descemet's membrane

is also involved in the differentiation and proliferation of the corneal endothelial cells (Biswas 2001, Hopfer, Fukai et al. 2005).

The Corneal Endothelium

The endothelial layer is about 5 μm thick and is made up of single layer of cells. The cells are polygonal, and each of them is 5 μm in thickness with a width of 20 μm and have a property of very slow rate of cell division. The main function of the corneal endothelial cells is the regulation of corneal hydration (Maurice 1972, Sherrard 1976).

2 Dystrophies of the Cornea

Dystrophies related to the Cornea are genetic disorders that progress gradually, symmetric/bilateral and are localised to various layers of the cornea. The disease does not have any relationship to environmental or systemic factors. Most corneal dystrophies are characterised by deposition of protein aggregates that cause photophobia, corneal erosions and pain and eventually lead to decreased corneal transparency and reduced visual acuity. Based on clinical observation, dystrophies of the cornea can be classified into three groups (**Table 2-1**) based predominantly on the location of protein deposits in the cornea (Klintworth 2009). The first group of corneal dystrophies (anterior corneal dystrophies) is the ones that affect the epithelial layer, the Bowman's membrane or superficial stromal layer of the cornea. The second group of dystrophies (stromal corneal dystrophies) is associated with the deep corneal stroma, and the third group (posterior corneal dystrophy) is where the protein deposits originate in the Descemet's membrane or in the corneal endothelium. Majority of the corneal dystrophies reported so far follow a Mendelian pattern of inheritance (autosomal dominant, recessive or X-linked recessive) (Klintworth 2009). A few of the corneal dystrophies are genetically inherited and identified to portray developmental abnormality.

It should also be noted that since the knowledge about the dystrophies has increased, there are certain dystrophies that do not fit with the exact definitions of 'corneal dystrophy'. A few examples are as follows: most of the patients diagnosed with epithelial dystrophy of the basement membrane do not follow a genetic pattern of inheritance. In some of the other patients with Posterior polymorphous dystrophy (PPCD), the changes are manifested unilaterally. In the case of macular dystrophy, the

phenotypes of the disease directly correlate with the level of antigenic serum keratin sulphate. This indicates that corneal changes play an integral role in systemic abnormalities in some corneal dystrophies (Weiss 2010).

2.1 Classification of Corneal Dystrophies based on anatomical locations of cornea

The classification system of dystrophies related to the cornea is solely dependent on the layer of cornea involved in the pathology.

Table 2-1: Classification of corneal dystrophies

	MODE OF INHERITANCE
SUPERFICIAL CORNEAL DYSTROPHIES	
Meesmann dystrophy (MECD)	Autosomal Dominant
Stocker-Holt dystrophy	Autosomal Dominant
Granular corneal dystrophy type III (Reis-Bücklers dystrophy)	Autosomal Dominant
Thiel-Behnke dystrophy	Autosomal Dominant
Gelatinous drop like corneal dystrophy (familial sub epithelial corneal amyloidosis) (GD CD)	Autosomal Resistant
Sub epithelial mucinous corneal dystrophy (SMCD)	Autosomal Dominant
Lisch epithelial dystrophy (LECD)	X-linked, Resistant
Epithelial recurrent erosion dystrophy (ERED)	Autosomal Dominant
CORNEAL STROMAL DYSTROPHIES	
Macular corneal dystrophy (MCD)	Autosomal Dominant
Granular corneal dystrophy type 1 (GCD 1)	Autosomal Dominant
Granular corneal dystrophy type 2 ((Avellino dystrophy, combined lattice-granular dystrophy) (GCD 2)	Autosomal Dominant
Lattice corneal dystrophy type I and variants (LCD I)	Autosomal Dominant
Lattice corneal dystrophy type II (LCD II)	Autosomal Dominant
Fleck corneal dystrophy (FCD)	Autosomal Dominant
Schnyder corneal dystrophy (SCD)	Autosomal Dominant
Posterior amorphous corneal dystrophy (PACD)	Autosomal Dominant
Congenital stromal dystrophy (CSCD)	Autosomal Dominant
POSTERIOR DYSTROPHIES	
Fuchs dystrophy (early onset) (FECD)	Autosomal Dominant
Fuchs dystrophy (late onset) (FECD)	Autosomal Dominant
Posterior polymorphous dystrophy type 1 (PPCD)	Autosomal Dominant
Posterior polymorphous dystrophy type 2 (PPCD)	Autosomal Dominant
Posterior polymorphous dystrophy type 3 (PPCD)	Autosomal Dominant
Congenital hereditary endothelial dystrophy type 1	Autosomal Dominant
Congenital hereditary endothelial dystrophy type 2	Autosomal Resistant
X-linked endothelial corneal dystrophy (XECD)	X-linked, Resistant

2.2 History of *TGFBI* associated Corneal Dystrophies

The term *dystrophy* is from the Greek word dys+trophe, which means *difficult nourishment* (Boutboul, Black et al. 2006), and the first time the word was instigated into the literature of medical observation was by Wilhelm Erb in 1884. Arthur Groenouw in 1890 described the “noduli corneae” in 2 patients: one of the patients had corneal dystrophy with granular protein deposits and the second patient had corneal dystrophy with macular corneal dystrophy. In 1899, Biber published his detailed research on Lattice Corneal Dystrophy (LCD) (Corden, Swensson et al. 2000). He described his findings as a corneal dystrophy, although he did not examine the phenotype and familial predisposition.

In 1917, a report was published on unknown superficial dystrophy by Reis. He described and published on the amorphous powdery aggregates, on the superficial layers of the cornea. In 1967, Thiel-Behnke was the first to report “honey-comb”-shaped deposits on the cornea. The deposits were also on the superficial layers of the cornea. Since the publication from Reis and Thiel-Behnke, there have been several reports to describe the clinical, histological and ultrastructure of various corneal dystrophies that are associated with Bowman’s membrane and also with the superficial stroma. Later with the improvement in technology, clinicians and scientists began to report the genetics and also included the surgical aspect of corneal dystrophy. The literature was confusing and misleading on whether both the dystrophies were the same or distinct entities.

It was in 1967, Klintworth published the association of Lattice lines observed in Lattice corneal dystrophy to amyloid fibrils. He demonstrated the association by staining the Lattice dystrophy samples with congo red and ThT, the dye that binds to amyloid fibrils (Klintworth 1967). Klintworth’s report was also the first report to provide staining evidence that associated amyloid deposits with lattice corneal dystrophy. The report also provided histological evidence that the cardinal factor in the disorder was collagen degeneration in the stroma (Klintworth 1967).

In 1949, the first-ever classification of corneal dystrophies was published by Bucklers. He had classified the disease based on the layer of the cornea that was involved in the pathology and position of the tissue affected based on anatomy. In the publication by Bucklers, the classification was purely based on anatomy. The genetic

pattern of inheritance, disease severity or characterisation based on biochemical and histopathology were not included.

In the 1960s and early 1970s, the classification of corneal dystrophy and the description of the disease were based on reports from a single family. In a number of cases reported, the already described pathology of a corneal dystrophy would often be reported as a new dystrophy or a variant of the current existing form due to small changes in the disease presentation. Thus there were innumerable atypical, unreported and indeterminate corneal dystrophies. For understanding the disease better, a proper system of classification was needed (Weiss, Moller et al. 2008).

In 1988, Folberg reported the Lattice-granular dystrophy that was combined in 4 of his patients. According to the reports published on the combined dystrophy, the patients had several fusiform deposits supplementary to the granular protein deposits. The presence of the acuminous deposits was confirmed to be amyloid deposits by electron microscopy and histochemical staining. The amyloid deposits were similar to the fusiform deposits identified in the lattice corneal dystrophy (Folberg, Alfonso et al. 1988).

In 1992, the gene *TGFBI* (Transforming Growth Factor Induced) was identified in a research study of genes that were instigated by TGF- β signalling in human adenocarcinoma cell lines originating from the lungs (Skonier, Neubauer et al. 1992). Since then, there are several publications and reports that showed manifestation of the protein in various organs including the cornea. The studies also showed over-expression of the protein in disease conditions like Granular corneal dystrophy (Klintworth 1999).

In 1994, the *TGFBI* gene was associated with corneal dystrophy. Mutations present in the *TGFBI* gene, which can be mapped on to chromosome 5 within 10-centimorgan region, were identified and associated with corneal dystrophy (Stone, Mathers et al. 1994). There were additional reports that prognosticate the involvement of more than one gene in the disease and took allelic effects into consideration. The report was a major contribution to understanding the disease better. The report highlighted that disparate mutations in a single gene might result in different phenotypic expressions. The phenotype may vary due to the effect of many genes on a single mutation. This linkage analysis reported in 1994 was a quantum leap that was able to identify the gene that could be associated with Lattice/Granular dystrophies.

In 1997, Munier and his co-workers took a major effort to further refine chromosome 5 to identify the exact genetic location of the *TGFBI* gene. The group also reported for the first time on four missense mutations at amino acids 124 and 555 of the TGFBIp by DNA sequence analysis (Munier, Korvatska et al. 1997). Since 1997, there have been several reported pathogenic modification in the *TGFBI* gene and phenotypic expression related to corneal dystrophies. A graphical representation of various major findings and reporting on corneal dystrophies is summarised in [Figure 2-1](#).

The advanced tools for genotypic investigations in recent years have contributed to understanding the defects in the gene and the consequences resulting from a mutant protein. The genetic analysis of corneal dystrophy has brought about a revolution in the knowledge of corneal dystrophies and was able to clarify inaccurate information in the earlier classifications of corneal dystrophies.

Stromal Corneal dystrophies lead to impaired vision, irritation in the eye, photophobia and recurrent corneal erosions in the dystrophic patients. These symptoms may vary significantly from patient to patient based on the type of amino acid change, the nature of protein deposits and the anatomical location of the protein deposits (Klintworth 2009).

There was a demand for a modern classification system that could take into account the clinical and genetic nature as well as pathology of the disease. The modern nomenclature of the corneal dystrophy was essential to traverse the disconnection between the terminology used by clinicians and researchers. The International Committee for Classification of Corneal Dystrophies (IC3D) was formed in April 2005 (Lisch and Seitz 2008). In 2007, the new classification system was published, and this system has been widely followed by clinicians and basic scientists to report on corneal dystrophies and new genotypic changes in the *TGFBI* gene.

2.3 International Committee for Classification of Corneal Dystrophies (IC3D)

According to the recommendations by the IC3D committee, corneal dystrophies were classified under four major categories based on the mode of inheritance, gene loci,

Figure 2-1: Time line of TGFBI associated Corneal Dystrophies

and the genes involved in the pathogenesis as: the epithelial and sub epithelial dystrophy, Bowman's layer dystrophy, stromal dystrophy and endothelial dystrophy.

Based on the supporting evidences the IC3D committee came up with a method of classifying the corneal dystrophies.

Categories

Category 1: This group contains all well-characterized dystrophy and the gene responsible for the dystrophy has been identified and mutations in the gene are linked to the disease.

Category 2: This group belongs to the corneal dystrophies that are mapped to very specific loci of a chromosomal. The gene present in or around this locus that may be responsible for the disease remains to be mapped.

Category 3: This group contains dystrophy that is well characterized phenotypically but the chromosomal locus still remains un-mapped.

Category 4: This category is kept for a newly suspected dystrophy that may be well documented. (Weiss, Moller et al. 2008).

The categorization of CDs (Corneal Dystrophies) had provisions to accommodate the change in category over time of a particular class when the knowledge progressively advances with the identification and association of the genetic

mutation responsible for the particular type of CD. The mutations in genes responsible for the dystrophies were compiled and made available. Mutational analysis became essential in identification and classification of CDs. The epithelial dystrophies are typically symptomatic with recurrent corneal erosions and/or the presence of irregular corneal astigmatism resulting from abnormalities in the epithelium, the corneal basement membrane, or Bowman's layer. The stromal dystrophies cause opacification from deposition of generated abnormal material. The endothelial dystrophies are characterized by the development of edema due to dysfunctional endothelial hydration regulation (Maurice 1972).

In 2015, IC3D published an updated version of classification of corneal dystrophies and a revision of the historical and traditionally accepted version of classification based on "anatomic level" was made (Weiss, Moller Hu Fau - Aldave et al. 2015). Some of the corneal dystrophies like Reis-Bücklers dystrophy and Thiel-Behnke dystrophy affected more than one layer of cornea and did not fit in the conventional classifications. The revised classification system was intended to accurately demonstrate which cellular layer of the cornea was involved. This system classified corneal dystrophies as epithelial and sub-epithelial dystrophies, epithelial-stromal *TGFBI* dystrophies, stromal dystrophies and endothelial dystrophies. The classification of corneal dystrophies and the genes associated with dystrophies are shown in **Figure 2-2**.

Figure 2-2: Corneal dystrophies and mutations in genes responsible for the dystrophy.

2.4 Survey on the epidemiology of *TGFBI*-corneal dystrophies

There are only very few studies conducted and reported in the prevalence of genetic mutation in *TGFBI* in a true epidemiological setting (Mashima, Yamamoto et al. 2000, Fujiki, Nakayasu et al. 2001, Lee, Cristol et al. 2010, Yang, Han et al. 2010). Most of the studies reported so far were based on clinical observations. There are more than 100 reports on the *TGFBI* associated clinical phenotypes that report dystrophies in the Bowman's membrane or corneal stroma. The web search results of various mutations reported so far in the *TGFBI* gene cite the original article published by Munier et al in 1997 (Munier, Korvatska et al. 1997). Since then more than 65 mutations have been reported and published that include more than 30 countries, involves over 1800 corneal dystrophy patients worldwide (Figure 2-3).

Figure 2-3: Distribution of TGFBI associated mutations worldwide. Figure adapted from (Lakshminarayanan, Chaurasia et al. 2014)

The number of reported mutations, the phenotype associated number of patients and demography are listed down in (Table 2-2). Most of the reported mutations are localised to the 1st and 4th Fasciclin 1 (FAS1) domain of the TGFBIp. Mutations in the amino acid residues 124 and 555 are considered to be mutational hotspots. Mutations p.Arg124Cys, p.Arg124His and p.Arg555Trp have been reported in different ethnic population around the globe. It has been identified that p.Arg124His is one of the highest reported mutation in the world accounting for about 26% of reported patients and is also one of the most frequently observed mutation in South-east Asian population accounting for about 88% of mutation.

An epidemiological study from Korea (Lee, Cristol et al. 2010) reported the presence of p.Arg124His (Table 2-3) mutation as one of the most prevalent genotype in the Korean population. This accounts for 11.5/10,000 individuals in Korea with this particular mutation. Studies from Japan reported that p.Arg124His was most prevalent in the population and accounts for up to 72% of cases with corneal dystrophy.

Table 2-2: List of *TGFBI* mutations reported so far and the corresponding clinical phenotypes

Number	Mutation	Phenotype	Number of reported cases	Countries
1.	p.Val113Ile	GCD I ^a	3	Mexico
2.	p.Asp123His	Atypical GCD	1	Vietnam
3.	p.Arg124Cys	LCD I ^b TBCD ^c RBCD ^d GCD II ^e LCD I LCD I	372 18 6 1 3 5	See Table 5-2 and 5-4. China Korea
4.	p.[Arg124Cys];[Gly470X]*	LCD I	1	Japan
5.	p.Arg124His	GCD II GCD	563 9	See Table 5-3
6.	p.[Arg124His];[Asn544Ser]*	GCD II- LCD	3	Japan
7.	p.[Arg124His;307-308delCT]	GCD	1	Hong Kong
8.	p.Arg124Leu	RBCD Atypical GCD	63 2	Brazil, USA, India, Czech Republic USA, France China
9.	p.[Arg124Leu;Thr125_Glu126del]	Atypical GCD	10	France
10.	p.Arg124Ser	GCD I	2	UK
11.	p.Asp299Asn	Polymorphic LCD	1	USA
12.	p.Arg496Trp	LCD IV	1	Japan
13.	p.Pro501Thr	LCD IIIA	19	Japan, China
14.	p.Met502Val	Unknown	1	Mexico
15.	p.[Met502Val];[Arg555Gln]	Atypical TBCD	2	France
16.	p.Val505Asp	LCD I	10	China
17.	p.Leu509Pro	GCD II LCD I	2 4	Germany France
18.	p.Leu509Arg	EBMD ^f LCD	4 7	France France
19.	p.[Arg514Pro];[Phe515Leu]	LCD I	5	China
20.	p.Ser516Arg	Atypical GCD I	1	India
21.	p.Leu518Pro	LCD I	12	Japan
22.	p.Leu518Arg	LCD I/IIIA	1	Italy
23.	p.Ile522Asn	LCD I	4	China
24.	p.Leu527Arg	LCD IIIA	29	Japan, Korea
25.	p.Thr538Arg	LCD I/IIIA	3	USA, Ukraine
26.	p.Thr538Pro	LCD I	4	China, India
27.	p.Val539Asp	LCD I	2	India
28.	p.Phe540Ser	LCD I/IIIA	4	Germany

Number	Mutation	Phenotype	Number of reported cases	Countries
29.	p.phe540del	LCD I/IIIA	1	Sardinia
30.	p.Pro542Arg	Variant LCD	1	Korea
31.	p.Asn544Ser	LCD IIIA	1	Japan
32.	p.Ala546Asp	GCD LCD Polymorphic LCD	10 8 10	China Mexico, India USA, Germany
33.	p.[Ala546Asp];[Pro551Gln]	LCD I	8	USA
34.	p.Ala546Thr	LCD IIIA	7	Brazil, China, France
35.	p.Phe547Ser	Atypical LCD	1	Hungary
36.	p.[A549T];[R555W]*	GCD I (early onset)	1	Germany
37.	p.Leu550Pro	GCD II	4	Mexico and Singapore
38.	p.[Leu550Pro];[His626Arg]	Atypical GCD	1	Mexico
39.	p.Arg555Trp	GCD I GCD II RBCD	375 1 1	See Table 2-5
40.	p.Arg555Gln	TBCD RBCD	66 6	Brazil, China, Czech Republic, France, Hungary, Japan, Singapore, Switzerland, Ukraine, USA,
41.	p.Leu558Pro	LCD III	8	Ukraine
42.	p.Leu559Val	Atypical GCD	3	India
43.	p.Leu569Arg	LCD I	4	USA
44.	p.His572Arg	LCD I/IIIA	38	Thailand, Singapore, Chile, China
45.	p.His572del	Unilateral LCD	1	USA
46.	p.Gly594Val	LCD IIIA	2	India
47.	p.Val613Gly	LCD IIIA	2	France
48.	Del613-616VAEP	Variant LCD	2	China
49.	p.Met619Lys	Variant GCD II	7	USA
50.	p.Ala620Asp	LCD I/IIIA	1	Singapore
51.	p.Asn622Lys	LCD IIIA	2	Italy, South America
52.	p.Asn622His	LCD I/IIIA	1	USA
53.	p.Gly623Asp	LCD RBCD LCD/Bowman LCD	2 11 4 4	Switzerland China, USA USA Germany
54.	p.Gly623Arg	LCD I/IIIA	11 CR+/KE2+	Germany

Granular corneal dystrophy; ^b Lattice corneal dystrophy; ^c Thiel-Behnke corneal

Number	Mutation	Phenotype	Number of reported cases	Countries
55.	p.Val624Met**	Atypical LCD	2 CR+	USA
56.	p.Val625Asp	LCD I	2	China
57.	p.Val624-625del	Atypical LCD	3	India
58.	p.His626Pro	RBCD/TBCD LCD	6 4	New Zealand Czech Republic
59.	p.His626Arg	Asymmetric LCD I/IIIA	122	Mexico, Germany, India, France, UK, Ukraine, China, Singapore, Vietnam
60.	p.Val627Sx	LCD IIIA	1	Italy
61.	T ⁶²⁹ N ⁶³⁰ ins NVP	LCDI/IIIA	2	France
62.	p.Val631Asp	LCD IIIA	15	Italy
63.	p.Arg666Ser	EBMD	2	Ireland
64.	p.Ile200Phe	Not mentioned	Not mentioned	Not mentioned
65.	p.Leu269Phe	Not mentioned	Not mentioned	Not mentioned
66.	p.Arg496Gly	Not mentioned	Not mentioned	Not mentioned

dystrophy; ^d Reis-Buckler corneal dystrophy; ^e Avellino corneal dystrophy; ^f Epithelial basement membrane dystrophy. *Compound heterozygous mutations. ** Phenotype was observed only in the homozygous state.

The mutations responsible for Lattice Corneal Dystrophy type 1 (LCD I) and Granular Corneal Dystrophy type 1 (GCD 1) are p.Arg124Cys (**Table 2-4**) and p.Arg555Trp (**Table 2-5**) respectively. They are the second most commonly prevalent mutation and the occurrence of mutation has been reported in 22 countries worldwide. The percentage of occurrence of each of the mutation is about 21%. The mutation p.Arg124Cys is the most common form of mutation among the Asian population than in the western Caucasian population. P.Arg555Trp has been reported in Western population and European population but according to the statistical reports, the highest number of reports from patients was from Asia. **Figure 2-4** shows the slit lamp images of patient's cornea with corneal dystrophy reported in our corneal clinic at Singapore National Eye Centre. The images show various disease phenotypes and the nature of protein aggregates associated with corneal dystrophy.

The next mutation that is most prevalent is p.His626Arg, the mutation responsible for LCDIIIA or variant LCD. The phenotype associated with this mutation is asymmetric nature and late onset of the disease. The H626R mutation was first

identified and reported in the UK and has a pervasive nature in geographic distribution in the Asian and European population. This mutation is also the common form of mutation in the Vietnamese population (Chau 2003). The two highly prevalent mutations associated with corneal dystrophy of the Bowman's membrane namely Reis-Bücklers dystrophy and Thiel-Behnke dystrophy with p.Arg124Leu mutation and p.Arg555Gln mutation are known to be reported in diverse Asian and Caucasian population (Aldave 2007). Reports from Russia suggests that LCD is a rare pathology in the Russian population and the estimated incidence is 1 in 92,000 (Zinchenko, Elchinova et al. 2001). The reported table and the recent studies on the various reports on *TGFBI*-associated corneal dystrophy suggest that the mutations reported have almost doubled in the recent years due to the increased awareness in screening patients and also due to the improved genetic and immunohistological screening tests available to detect and classify .

2.5 Possible Bias on the reported epidemiological studies

The prevalence of mutations and the list of commonly occurring mutations as listed in the [Table 2-2](#) is purely collected from the reports of studies published so far, and there are certain limitations to the conclusions of the reports. There are no reports available on detailed studies published from countries like Australia. The lack of access to medical diagnostic facilities may have led to no reports from Africa. The previous publications before the formation of IC3D classification could have reported the mutations and associated it with a different phenotype causing a significant bias in the reported data.

There could be significant bias introduced to the data reported in [Table 2-2](#), for some of the large population based studies, the assessment of genetic screening would have been done on exons with the “common mutations” and not the whole gene. This approach would have possibly missed the mutations in other exons in the gene and perhaps some novel mutations relating to that particular ethnicity. Notwithstanding the limitations, the detailed report of mutation in the *TGFBI* gene causing various clinical phenotypes shows prevalence of the disease in various populations.

Table 2-3: Distribution of p.Arg124His mutation worldwide

Region/ Country	Phenotype	Number of reported cases
-----------------	-----------	--------------------------

Asia		
- China	GCD II	66
- Hong Kong SAR	GCD II	12
- India	GCD II	12
- Japan	GCD II	270
- Korea	GCD II	123
- Singapore	GCD II	6
- Vietnam	GCD II	3
Middle East		
- Iran	GCD II	20
Europe		
- France	GCD II	3
- Hungary	GCD II	3
- Germany	GCD II	14
- Switzerland	GCD II	1
- Spain	GCD II	6
- UK	GCD II	22
Americas		
- USA	GCD II	5
Various/unspecified	GCD II	6

Table 2-4: Distribution of p.Arg124cys worldwide

Region/ Country	Phenotype	Number
Asia		

- China	LCD	92
- China	TBCD	18
- China	GCD3/RBCD	6
- Japan	LCD	59
- Korea	LCD	7
- Hong Kong SAR	LCD	2
- India	LCD	27
- Singapore	LCD	6
- Vietnam	LCD	3
Americas		
- USA	GCD II/ACD	7
- Brazil	LCD	1
- Chile	LCD	12
- Canada	LCD	2
Europe		
- Austria	LCD	2
- Bulgaria	LCD	1
- Czech Republic	LCD	2
- France	LCD	10
- Greece	LCD	11
- Hungary	LCD	18
- Germany	LCD	22
- Spain	LCD	10
- Switzerland	LCD	14
- UK	LCD	8
- Ukraine	LCD	65
Various/ info not available	LCD	11

Table 2-5: Distribution of p.Arg555Trp mutation worldwide

Region/ Country	Phenotype	Number of reported cases
Asia/ Australia		
- China	GCD	94
- Hong Kong SAR	GCD	2
- India	GCD	66
- Japan	GCD	24
- New Zealand	GCD	5
- Singapore	GCD	8
- Turkey	GCD	37
- Taiwan	GCD	4
- Vietnam	GCD	5
Europe		
- Czech Republic	GCD	2
- France	GCD	10
- Hungary	GCD	5
- Germany	GCD	23
- Spain	GCD	2
- Poland	GCD	4
- Switzerland	GCD	24
- UK	GCD	26
- Ukraine	GCD	15
Americas		
- Brazil	GCD	1
- Brazil	GCD II	1
- Mexico	GCD	4
- Mexico	RBCD	1
- USA	GCD	3
Various/unspecified	GCD	11

Figure 2-4: Slit lamp photographs of clinical phenotype of TGFBI-related corneal dystrophy (Figure adapted from (Lakshminarayanan, Chaurasia et al. 2014))

A. Anterior segment slit lamp picture of a Lattice Corneal Dystrophy patient with R124C mutation with visible multiple refractile lines of amyloid fibrils. B. Anterior segment image of a granular dystrophy patient with R555W mutation, showing corneal opacities with granular protein deposits. C. slit lamp image of a patient with Avellino dystrophy or Granular dystrophy type 2 sharing features of both Lattice and granular dystrophy. D. Retro illumination of image C showing both fibrils and granular protein deposits. E. Anterior segment picture of a patient with R124L mutation associated with Reis-Bucklers corneal dystrophy or granular corneal dystrophy type 3. The image shows opacities in the cornea with granular protein deposits. F. Photograph of anterior segment of cornea from a patient with R555Q mutation associated with Thiel-Behnke corneal dystrophy. The characteristic “honeycomb” pattern with multiple deposits in the sub epithelial layer of the cornea.

2.6 Classification and characteristics of *TGFBI* associated corneal dystrophy

As described earlier, the Bowman's layer of the adult human cornea is made of collagens, mainly type I, type III, type V and type VII. *TGFBI* associated mutations in the Bowman's layer causes two different kinds of granular deposits namely Reis-Bucklers corneal dystrophy (RBCD) and Theil-Behnke corneal dystrophy (TBCD) (Figure 2-5). In the clinical setting, distinguishing the two corneal dystrophies of the Bowman's membrane is very difficult because the deposits have a similar pattern and the depth of the cornea in which the protein deposits are found are similar.

2.6.1 Bowman's Layer Corneal Dystrophy

Reis-Bucklers corneal dystrophy (RBCD)

RBCD is better known as Bowman's corneal dystrophy type 1. The clinical phenotype of this dystrophy is presence of geographical opacities with sub epithelial deposits which are mostly grey or white in colour and begin to appear in the Bowman's layer which then slowly extends to the anterior part of the stroma. The deposits start to develop from infancy and based on age of onset from the previous publications, this dystrophy is reported in the early decades of life, mostly in the first and the second decade. The patients with this type of dystrophy are reported to have frequent corneal erosions and a progressive impaired vision in the early stages of childhood. The affected patients seek early corneal transplant, and the protein deposits tend to recur with 6 years of corneal transplant (Ellies, Renard et al. 2002, Ellies, Bejjani et al. 2003).

Both the dystrophies related to the Bowman's layer (RBCD and TBCD) may not be easily distinguished only based on slit-lamp examination (Weiss, Moller et al. 2008). When the cornea is examined by light microscopy, a very thin and irregular epithelial layer with the absence of Bowman's layer is seen. Sub epithelial protein deposits replaces the Bowman's layer which when stained with Masson trichrome shows a positive red colour. The protein deposits also stain positive on haematoxylin and eosin (H&E) and shows characteristics of eosinophilic. When the protein deposits were viewed under Transmission Electron Microscope (TEM), protein deposits that were either rod-shaped or trapezoidal shaped with high electron-density were observed. The protein deposits are encompassed with fibrillar structures and proteoglycan filaments.

The granular protein deposits shows huge positive reactivity when stained with KE-2 and K-15 antibodies which is used to detect the C-terminal amino acid residues 426-682 and N-terminal residues of 69-364 of TGFBIp (Konishi, Yamada et al. 2000). The major mutation associated with this clinical phenotype is p.Arg124Leu.

TGFBI - Associated Corneal Dystrophies

Figure 2-5: Pathological classification of *TGFBI*-associated corneal dystrophies

Thiel-Behnke corneal dystrophy (TBCD)

TBCD is identified as corneal dystrophy that affects the Bowman's membrane. The disease is depicted by the advent of irregular sub epithelial reticular corneal deposits in a "honeycomb" pattern. The age of onset of the disease is very early (first or second decade of one's life) (Figure 2-6). The characteristic of the disease is the presence of the protein deposits that spread into the deep stroma with time. As the disease progresses, the protein deposits integrate and the cornea become opaque. For this disease phenotype, the protein deposits tend to recur more than 10 years after corneal transplantation (Ellies, Bejjani et al. 2003). When examined by histopathology, the protein aggregates stain negative with Masson trichrome dye. The mutation p.Arg555Gln is usually associated with the disease phenotype. The protein deposits when viewed under transmission electron microscopy (TEM) appear as twisted frizzy

fibres with a mean diameter of 9-15 nm (Kuchle, Green et al. 1995, Ridgway, Akhtar et al. 2000).

In the clinic setting, it is difficult to distinguish RBCD from TBCD as the features are not distinct in patients. Based on the age of disease onset, extremity of opacity of the cornea and successive loss of vision are the features that help distinguish TBCD and RBCD (Kuchle, Green et al. 1995, Weiss, Moller et al. 2008).

Figure 2-6: Age of onset of some of the known mutations associated with TGFBI-associated corneal dystrophy. Modified from (Lakshminarayanan, Chaurasia et al. 2014).

2.6.2 Stromal Corneal Dystrophy

TGFBI-mutations are associated with three different types of corneal dystrophy of the stromal region. They are Lattice Corneal Dystrophy (LCD), GCD type 1 (GCD1) and GCD type 2 (GCD2). Each of the corneal dystrophies of the corneal stroma can be distinguished by physical examination using slit lamp biomicroscopy and electron

microscopy, by histopathological staining of tissue and genetic screening to identify the specific mutation associated with the disease pathology.

Lattice Corneal Dystrophy (LCD)

LCD is one of the most commonly encountered and reported corneal dystrophy. LCD is associated to autosomal dominant form of genetic inheritance. The characteristic feature of the disease is the linear and branching refractile lines (amyloid fibrils) combined with or without whitish granular opacities in the stroma. LCD is associated with bilateral form of the disease (except for pHis572Del and p.His626Arg mutations). LCD can be further classified based on the clinical characteristics as Classic LCD (LCD type 1) and the variant LCDs (type IIIA, I/IIIA, IV and atypical).

The classical form of LCD is usually inherited as autosomal dominant attribute. The most common mutation associated with this phenotype is p.Arg124Cys and linked to the mutation in the 1st FAS-1 domain of the protein. The protein amyloid deposits in the form of translucent amyloid lines appear within the first ten years of patient's life in the anterior portion of the corneal stroma. The amyloid fibrils disperse centrifugally into the deeper layers of cornea and also to the periphery of cornea with time. The presence of amyloid deposits in the anterior portion of the cornea leads to frequent corneal erosion and the corneal surface becomes highly irregular. During the initial stages of the disease, the space between the white dots and the fibrillar lines are clearly visible but with time, the deposits tend to appear as dense network. This develops into corneal haze and leads to progressive loss of vision. The corneal deposits from LCD stain orange with Congo red and present dichroism and bright apple green birefringence under plane polarized light. The amyloid fibres associated with this phenotype shows highly ordered stacks with a mean diameter of 9 nm (Francois and Feher 1972). When the amyloid deposits from the patient's cornea is subjected to staining with antibodies raised against C-terminal part of TGFBIp (amino acid residues 426-682, KE2), the deposits show positive staining with the antibody.

The characteristic features of the most common mutation associated with LCD phenotypes, p.Arg124Cys is also associated to patients with RBCD-like phenotype (Yang, Zhao et al. 2011) or TBCD- like phenotype (Chang, Zhiquan et al. 2009). The reported compound heterozygous mutation p.[Arg124Cys];[Gly470Ter] which is associated with LCD is more drastic than the p.Arg124Cys mutation alone. The mutation that was heterozygous for only mutation p.Gly470Ter, had a clear cornea

(Sakimoto, Kanno et al. 2003). The other common mutations associated with LCD phenotype are p.Val505Asp, p.Leu518Pro, p.Ile522Asn, p.Thr538Pro, p.Val539Asp, p.Leu569Arg, p.Val625Asp, p.[Arg514Pro];[Phe515Leu] and p.[Ala546Asp];[Pro551Gln] (Fujiki, Hotta et al. 2000, Hirano, Hotta et al. 2000, Sakimoto, Kanno et al. 2003, Tian, Fujiki et al. 2005) (Warren, Abbott et al. 2003, Chakravarthi, Kannabiran et al. 2005, Yu, Gu et al. 2006, Zhang, Zeng et al. 2009, Zhong, Chen et al. 2010).

There were variations identified from the classical LCD in terms of the severity of the disease phenotype, the age of phenotypic expression of the disease and visible appearance of the protein deposits. This has led the scientist to further classify LCD into four types (LCD type IIIA, I/IIIA, IV and atypical). LCD IIIA is characterised by the appearance of thick lattice amyloid line and the age of onset of disease is late (5th to 7th decade of life). The occurrence of corneal erosion is very rare and since the disease onset is late, visual acuity does not decrease until later stage of the disease. This phenotype is characterised by low penetrance and autosomal dominance mode of genetic inheritance compared to the autosomal recessive inheritance pattern of LCD III. The most common mutation associated with LCD IIIA is p.Pro501Thr (Mashima, Yamamoto et al. 2000). LCD III shows thick lattice amyloid deposits which extend from one side of the limbus to the other side.

LCD IV is associated with the protein deposits mostly amyloid in nature that are found deep in the corneal stroma and the age of onset is normally around 7th to 9th decade of life based on reported cases. The most common mutation representing this phenotype is p.Leu527Arg (Fujiki, Hotta et al. 1998). The disease is presented with a less severe phenotype compared to classical LCD. Many of the mutations that are present in the 4th FAS-1 domain of TGFBIp express phenotype that are intermediate form of LCDI and IIIA and hence classified as LCDI/IIIA or LCDi. The most common mutation associated with this phenotype is p.Leu518Arg (Munier, Frueh et al. 2002). The other mutation p.His626Arg which is the most common form of LCD in south-east Asia is associated with a characteristic thick lattice amyloid fibrillar deposit that extends from the anterior to mid-stroma. The age of onset for this phenotype is typically between the 30 to 40 years of life (Stewart, Black et al. 1999, Schmitt-Bernard, Guittard et al. 2000, Dighiero, Niel et al. 2001, Munier, Frueh et al. 2002). But according to the recent guidelines from IC3D classification, it is proposed to remove all sub

classifications and just mention them as variant LCDs (Weiss, Moller Hu Fau - Aldave et al. 2015).

2.6.3 Granular Corneal Dystrophy

Granular Corneal Dystrophy type 1

Granular Corneal Dystrophy type 1 is passed down genetically as an autosomal dominant form and the phenotype is characterised by bilateral, discrete protein deposits in the form of granular opacities. The protein deposits are normally found in the anterior part of the corneal stroma saving the periphery and the age of onset is during the first twenty years of life. The most common mutation associated with the disease is p.Arg555Trp. The granules of protein deposits are found to be of different shapes and are classified as ring, crumb and drop-shaped. At the primitive stages, the cornea is usually clear and visual acuity is not compromised. The disease is characterised by severe photophobia and glared vision with severe corneal erosion that are recurrent is common in patients. When the initial protein granular deposits become into bigger protein aggregates with more confluent protein deposits, visual acuity of the patient decreases rapidly.

When the protein deposits are stained with Masson trichrome, they stain deep red colour. When the granular protein deposits are stained with polyclonal antibodies raised against TGFBIp (amino acid 69-364, KE15) and antibodies raised against the c-terminus of TGFBIp (amino acid 426-682, KE2) these deposits show a strong positive interaction. When the protein deposits are viewed under the microscope, they can be visualised as rod-shaped trapezoids with a diameter of 100-500 nm in diameter. They appear to be more filamentous, a ‘moth-eaten’, morphology that is homogenous in the inner structure of the deposits (Sornson 1965, Akiya and Brown 1970).

The patients with a homozygous form of gene inheritance of GCD1 experience a more severe phenotype and typically have a placoid type of corneal deposits and even after corneal transplant, the protein deposits tend to recur early. There have been reports of three patients from Japan with homozygous pattern of inheritance and the age of onset of the disease is usually in the early years of age (around 6 years) with a very severe phenotype.

The other mutations linked with CGD1 are p.Val113Ile (Zenteno, Ramirez-Miranda et al. 2006), p.Asp123His (Ha, Cung et al. 2003), p.Arg124Ser, p.Ser516Arg,

p.Leu559Val (Stewart, Black et al. 1999, Munier, Frueh et al. 2002, Chakravarthi, Kannabiran et al. 2005, Paliwal, Gupta et al. 2009).

Granular Corneal Dystrophy type 2

GCD2 was first reported in a large family from a district called Avellino in Italy (Holland, Daya et al. 1992) and the dystrophy has also been known as “Avellino Corneal Dystrophy”. The features of the disease in the clinic, is characterised by both Lattice lines and also granular protein deposits. When patients are first examined in the clinic by slit lamp, identifies small and faint superficial protein deposits. The age of onset of the disease is around first or second decade of life. The granular protein opacities occur much earlier in the disease and the Lattice amyloid fibrils becomes apparent in the later stage of life. In patients the amyloid lattice lines appear deeper in the stroma with superficial protein granular deposits. When the superficial protein granular deposit reaches deeper to the Bowman’s membrane, it is usually associated with severe erosion of the corneal epithelium. The amounts of granular protein deposits are fewer in GCD2 than GCD1.

The histopathological characterisation of the disease is by the red stain with Masson trichrome dye and Congo red dye. The sections of the cornea with the protein deposits are also strongly reactive to antibodies raised against TGFBIp (KE-2 and KE-15). The mid-stromal and deep stromal deposits stain positive only with KE-2 antibody. Patients who are identified to have homozygous mutations have early age of onset and experience severe phenotypes of the disease with a much clouded cornea. The patients also require corneal transplant at a much earlier stage of life around 25 years of age.

The mutation associated with this disease is p.Arg124His and two phenotypic difference of the disease is reported in Japanese populations (Watanabe, Hashida et al. 2001). The type 1 patients had spot shaped opacities did not share a common birthplace, but patient with type 2 disease had reticular opacities and shared common birthplace in Japan and traced to Tottori prefecture in western part of Japan. If the patient is heterozygous for the mutation, they experience less severity and they do not approach the clinic for any treatment as patients do not experience loss of visual acuity during the early stages of the disease.

There has been another variant of GCD2 reported in Hispanic patients with p.Met619Lys mutation (Aldave 2008). There has been another report in Mexican patients carrying p.Leu550Pro mutation with histopathological features similar to

GCD2. The clinical feature is characterised by amyloid Lattice lines in the peripheral part of the cornea and granular protein deposits in the mid or deep part of the stroma. When the corneal deposits are subjected to histopathological analysis, the deposits stain red with Masson trichrome and stain orange red with Congo red. The histopathological examination confirms the GCD2 disease phenotype.

The other mutations related to GCD2 are p.[Arg124His(;c.307_308delCT)](Yam, Wang et al. 2012), p.[Arg124His];[Asn544Ser] (Yamada, Kawamoto et al. 2009), p.Leu550Pro (Zenteno, Correa-Gomez et al. 2009, Lakshminarayanan, Vithana et al. 2011) and p.Met619Lys (Aldave 2008).

2.7 Phenotype-Genotype correlation in *TGFBI* associated mutations

In general, there is a good correlation observed between the clinical phenotype and the reported gene mutation with a few exceptions. A few examples of the variation in the correlation are as follows: The p.Arg124His mutation is related to amyloid-non-amyloid phenotype (combined GCD-LCD) because 99% (491 patients) of cases reported is known to have a good correlation while the 1% of reported did not show a good correlation. The correlation in heterozygous p.Arg124His mutation is associated with complex incomplete penetrance and it strongly resembles the classic p.Arg555Trp associated granular corneal dystrophy (Konishi, Yamada et al. 1999, Kim, Kim et al. 2008, Cao, Ge et al. 2009).

The other mutation associated with LCDI phenotype is p.Arg124Cys, there have been variations reported in the associated phenotype in patients from Greece, Japan, Bangladesh and China (El-Ashry, Abd El-Aziz et al. 2004, Yoshida, Yoshida et al. 2004, Ma, Liu et al. 2010). In the reports from Japan, the patients had opacities in the superficial cornea with absence of amyloid deposits confirmed by the histopathological staining and electron microscopy investigation.

There has been another report from Vietnam on the variant form of p.Arg124Cys that closely resembled GCD2 phenotype. The patients' cornea had granular protein deposits that were identified in the sub-epithelial layer and presence of amyloid deposits in the stroma (Patel, Chang et al. 2010). When the deposits were

stained with Masson trichrome, the deposits stain red and reacted positive to the antibody raised against full length TGFBIp (Morishige, Chikama et al. 2004). In the reports from Bangladesh and China, the patients did not have the traditional lattice lines but had grey-white opacities in the anterior and mid-stroma. However, there was no staining or histopathological data available for the reports from Bangladesh and China (El-Ashry, Abd El-Aziz et al. 2004, Ma, Liu et al. 2010).

The other mutation which is common in South-East Asia and associated with late onset LCD is p.His626Arg. This mutation is reported to have varied phenotypes in various ethnic groups. In most of the reported cases with this phenotype, the protein deposits are localised to anterior and mid-stroma. The protein deposits also stain red or dark brown with Congo red indicating the presence of corneal amyloid deposits. The variant form of the mutation was reported by (Schmitt-Bernard, Schneider et al. 2002) where the authors show the deposits between the corneal stromal lamellae and interrupts the collagen fibril organisation that is required for normal vision.

There have been several mutations reported in the *TGFBI* gene in a diverse population across the world (Table 1). There has also been disease variants of clinical phenotypes reported but the variants constitute to less than 20% of the total mutations reported. It is also worthy to note that the majority of the reported variants are associated with 4th FAS-1 domain of the TGFBIp and most of the variations are reported in Asian population. The possibility of the variation may be linked to founder effects (Fukuoka, Kawasaki et al. 2010). It is also interesting to note that different LCD mutations associated to the 4th FAS-1 domain of the protein have a large variation to the reported phenotype compared to the classic LCDI variant.

Among the most commonly reported mutation in the 4th FAS1 domain is p.Arg555Trp mutation and this mutation is also reported to have 2 cases that had a varied phenotype. One was from Brazil and the patient had p.Arg555Trp mutation but the phenotype resembled more towards GCD2 (Solari, Ventura et al. 2007). The second report was from Mexico and the phenotype was more related to RBCD than the classical GCD (Zenteno, Correa-Gomez et al. 2009). The variants reported had only genotype data and did not have histology or electron microscopy data.

The other mutation which displays significant variation in phenotype among the various ethnicities worldwide is p.Gly623Asp. The first report on the mutation was by Afshari et al and linked the mutation to the RBCD phenotype. The reports observed the presence of protein hyaline deposits in the Bowman's membrane of the cornea and

histopathological analysis of the protein deposits failed to stain positive with Masson trichrome dye and also stained negative with Congo red. The age of onset of the disease was much earlier and recurrent corneal erosion was common among the patients (Afshari, Mullally et al. 2001).

A report from China observed the same phenotype in the patient who also carried the p.Gly623Asp (Li, Qi et al. 2008) but the mutation was linked to LCDI/IIIA in other reported patients (Munier, Frueh et al. 2002, Aldave, Rayner et al. 2005, Auw-Haedrich, Agostini et al. 2009). The report from the German patient's corneal deposits showed positive staining with Congo red stain and also showed positive reactivity for KE2 antibody which is the pathological indicator of amyloidogenic phenotype (Auw-Haedrich, Agostini et al. 2009).

Another interesting phenotypic variability is seen in p.Leu509Pro mutation. In reports from Germany, the patients showed cloudy and geographic-like protein deposits in the sub-epithelial layer and the anterior part of the stroma (Gruenauer-Kloevekorn, Clausen et al. 2009). The corneal deposits showed strong reactivity for KE-2 and KE-15 antibodies and were grouped under GCD2 variant corneal dystrophy. The French patients with the same mutation showed geographic like protein deposits combined together with fine lattice amyloid protein deposits in the deep stroma. When the protein deposits were subjected to histopathological analysis, the deposits stained negative with Masson trichrome but stained positive with congo red with a bright apple green birefringence under plane polarised light (Niel-Butschi, Kantelip et al. 2011).

There has been another report from New Zealand describing a LCD variant that showed coinciding features of RBCD and TBCD for p.His626Pro mutation (Wheeldon, de Karolyi et al. 2008). The patient had clinical phenotypes similar to RBCD with early onset of the disease, visual loss and speedy recurrence of the protein deposits after the corneal transplant. When the corneal deposits were subjected to histological analysis, they stained negative with both Congo red and Masson trichrome. When the protein deposits were viewed under the microscope, they showed curly fibrils that could be associated with TBCD.

In the spectrum of mutations reported so far, to add on to the phenotype-genotype variations, the occurrence of multiple mutations in the gene further makes the clear classification difficult due to the multiple dose effect. There are about eight multiple gene mutations reported in the *TGFBI*-associated corneal dystrophy. In all these multiple mutations the substitution occurs in the same or different alleles ([Table](#)

2-6). The reported mutations may be pathogenic or non-pathogenic, this may also lead to two mutations that are pathogenic or one mutation that is pathogenic and the other mutation which is non-pathogenic in the patients reported so far.

Table 2-6: List of two simultaneous mutations in the *TGFBI*-associated corneal dystrophies

Mutation	Inheritance Pattern	Phenotype of single mutation	Phenotype of double mutation	Location of the alleles
p.[Arg124Cys];p.[G470X] ^a	AD	LCD I/ Non-pathogenic	LCD I	Different
p.[Arg124His;Asn544Ser] ^b	AD	GCD II/LCD IIIA	GCD II-LCD	Different
p.[Ala549Thr;Arg555Trp] ^c	AD	Non-pathogenic/GCDI	Severe GCD I	Different
p.[Ala546Asp;Pro551Gln] ^d	AD	PolymorphicLCD/ not known	LCD I/IIIA	Same
p.[Met502Val;Arg555Gln] ^e	AD	Unclassified/TBCD	Atypical TBCD	Same
p.[Arg124Leu;Thr125_Glu126del] ^f	AD	RBCD/not known	Atypical GCD (French Variant)	Same
p.[Leu550Pro;His626Arg] ^g	AD	GCD II/LCD	Atypical GCD	Not Known
p.[Arg514Pro;Phe515Leu] ^h	AD	Not known	LCD I/IIIA	Not Known

^a Sakimoto et al., 2003; ^bYamada et al., 2009; ^cFrising et al., 2006. ^dKlintworth et al., 2004; ^eNiel-Butschi et al., 2011; ^fDighiero et al., 2000; ^gZenteno et al., 2009; ^hZhong et al., 2010.

In the case of compound heterozygous mutations, the protein product carries both the mutations and for all the reported compound heterozygous mutations, the mutations are located in the 1st and 4th FAS1 domain or both mutations to either the 1st FAS1 domain or 4th FAS1 domain. When there are two mutations reported, the clinical phenotype is observed to be severe than the presence of single mutation or they may have a neutral effect.

One compound heterozygous mutation is p.[Arg124C;Gly470X] which is a nonsense mutation and the protein product results in the replacement of Cysteine for Arginine in amino acid 124 and removal of a major portion (~ 32%) of C-terminal part of TGFBIp (Sakimoto, Kanno et al. 2003). The severity of the disease showed resemblance to p.Arg124Cys phenotype and the patient showed dense protein deposits in the sub-epithelial layer of the cornea. The lattice lines were not clearly visible in the reported patients. When the patient's daughter was screened, she was identified to have p.Gly470X mutation, did not show any symptoms even after the 5th decade of life. In this case of a compound heterozygous mutation, the p.Arg124Cys mutation deploys a dominant effect on the phenotypic expression. It is highly possible that the change in the TGFBIp postpones or avoids the formation of amyloid fibrils. This may be possible

due to the absence of C-terminal fragment that is susceptible to proteolytic processing and these proteolytic fragments of protein are important component of LCDI phenotype.

There is another compound heterozygous mutation reported by Frising et al p.[Ala549Thr;Arg555Trp] that cause a severe form GCD1. This phenotype is matching closely to homozygotic p.Arg555Trp and the offspring of the patient did not show any clinical phenotype (Martine Frising 2006). In this reported compound heterozygous mutation, the polymorphic mutant intensifies the severity of the p.Arg555Trp phenotype related to the disease. In the case of another compound heterozygous mutation p.[Arg124H;Asn544Ser] the two mutations are localised in different chromosomes, the clinical presentation of the phenotype may be the additive effect of the two phenotypes (Yamada, Kawamoto et al. 2009). The phenotypic variations reported so far and the reasons behind the variations are not very clear.

Chapter 2

3 Structure and function of TGFBIp

3.1 Gene and protein structure

The *TGFBI* gene is 34,924- nucleotide transcript which has 17 exons, and after the transcriptional process results in a 2,052-nucleotide protein coding sequence. The gene encodes the TGFBIp also identified as keratoepithelin, Big-h3, and RGD-CAP. The protein has 683 amino acid residues which after the cleavage of N-terminal signal peptide (peptide 1-23) is secreted from the cell. The secreted TGFBIp contains a Cys-rich domain called as the EMI domain, which was identified in a group of proteins called EMILINS, four consecutive Fasciclin 1 (FAS1) domains, and a c-terminal RGD integrin recognition motif (residues 642-644) as shown in [Figure 3-1](#) . The protein TGFBIp has shown to exist in the monomeric form when purified from the corneal extracts with no post-translational modification (Andersen, Karring et al. 2004). The report also found that the majority of the TGFBIp is truncated after alanine at amino acid number 657 in the human cornea. There are few other reports that identified the carboxylation of certain glutamate residues in the protein expressed by bone marrow derived mesenchymal cells (Coutu, Wu et al. 2008) and others have reported phosphorylation of TGFBIp in Corneal endothelial cells (Srivastava and Srivastava 1999). This indicated that the post-translation modification identified in TGFBIp is tissue specific or it is possible that both unmodified and modified forms of TGFBIp co-exist in the same tissue.

The EMI domain is located near the N-terminus of TGFBIp and it got its name from the EMILIN (Elastin Microfibril Interface Located ProteIN), which are a group of multi domain glycoproteins of the extracellular matrix (Doliana 2000). The EMI domain is very common in several human proteins including EMILIN-1 and EMILIN-2, Stabilin-1 and Stabilin-2, Multimerin-1, Multimerin-2, Periostin and TGFBIp. There is some extent of sequence similarity found in the EMI domains, the pattern of Cys-residue alignment and the presence of the EMI domain always towards the N-terminus of the protein are the common features observed.

TGFBIp contains 11-Cys residues and six of the Cys residues are present in the EMI domain of the protein. These Cys residues are highly conserved throughout the evolution and they are expected to form three-intra domain disulphide bridges based on the stequence homology (Callebaut, Mignotte et al. 2003). The remaining 5 Cys residues are distributed with one residue in the 1st FAS-1 domain, two in both the 2nd

and 3rd FAS-1 domains as shown in [Figure 3-2](#). The main function of the EMI domain is difficult to define although the EMILINs is known to bind with Elastin and Fibulin-5 proteins suggesting that the group of protein may have a potential role in the elastic fibre formation (Zanetti, Braghetta et al. 2004).

Figure 3-1: The domain structure of TGFBIp. The individual FAS-1 domains with the relevant amino acid sequence is marked in the figure.

The FAS-1 domain identifies to a cell adhesion domain which is homologous to the glycoprotein Fasciclin 1 from *Drosophila*. The FAS-1 domains are found in secreted or membrane bound proteins which is not typical characteristics for extracellular domains. The FAS-1 domains are found in organism from all phyla, including mammals, insects, plant, and bacteria (Kawamoto, Noshiro et al. 1998). When the FAS-1 domain from all phyla are aligned by pair wise sequence, the identity between them is rather low but there is a presence of two highly conserved sequences that is present in all phyla.

The presence of FAS-1 domain was identified by Zinn, et al in 1988. According to the report, the main function of the FAS-1 domain was to mediate cell-cell or cell-protein contact (Zinn, McAllister et al. 1988). There are other reports which identified several integrin binding sites in the FAS-1 domain (Park, Park et al. 2004). The reports also suggest that the FAS-1 domain mediates cell-cell interaction in one of the extracellular protein Stabilin-2 (Park, Jung et al. 2009).

Three other proteins are known to have FAS-1 domains apart from TGFBIp, they are periostin, stabilin-1 and stabilin-2. These proteins have several common factors as all of them are extracellular proteins and they all have multiple FAS-1 domains. Of all the three proteins Periostin is known to be more homologous to TGFBIp as it has a domain composition similar to TGFBIp with an EMI domain followed by four FAS-1 domains. But TGFBIp is distinct among the other three proteins because only TGFBIp has an RGD signal motif. This RGD-domain may play a major role in the binding and interaction of TGFBIp with other proteins especially with integrins although no direct binding evidence has been reported.

Figure 3-2: Distribution of Cys residues in TGFBIp. The six Cys residues in the EMI domain are hypothesised to form intradomain bridges.

3.2 Expression and Regulation of TGFBIp

TGFBIp is a most enriched protein in the human cornea and is an extracellular component of the many developing organs of the human body like nuchal and periodontal ligaments. The protein is also found expressed in mature cornea, aorta, lung, kidney, skin, bladder, bone and blood (Gibson, Kumaratilake et al. 1997, Thapa, Lee et al. 2007, Karring, Runager et al. 2010). TGFBIp h is known to be a major component of elastin-rich fibres found in close association with microfibrillar proteins like Fibrilin-1, Fibrilin-2 and Microfibril-associated glycoprotein (Gibson, Hatzinikolas et al. 1996). In the human cornea TGFBIp is expressed in the corneal epithelial layer, stroma, Descemet's membrane and endothelium. The expression of TGFBIp is very high during the early years of life in accordance to the maturation of cornea after birth. The wet weight of corneal tissue is about 0.45-0.55 $\mu\text{g}/\text{mg}$ in a child which gradually increases to about 0.7-0.8 $\mu\text{g}/\text{mg}$ in an adult and maintained around the same weight throughout the rest of the life time (Karring, Runager et al. 2010). In both the human and porcine corneas, the form that is most abundant is the C-terminus truncated isoforms. This may be due to the proteolytic processing of TGFBIp into numerous small fragments of different molecular weights by yet unidentified protease(s) (Korvatska 2000).

3.3 Structural Organisation of TGFBIp

3.3.1 Primary Structure of TGFBIp

The fully mature form of the protein TGFBIp is about 683 amino acid residues with 11 cysteine residues and the protein shares greater than 90% sequence similarity with both porcine and mouse TGFBIp (Kawamoto, Noshiro et al. 1998). The overall structure of TGFBIp contains an N-terminal domain rich in cysteine residues and four FAS-1 domain and a cell –attachment RGD domain. The EMI domain in TGFBIp has 56 amino acid residues with 6 Cys residues. Only in human TGFBIp there is an even number of Cys residues in the EMI domain and lacks the suppositional hydrophobic

amino acid residues between the 3rd and 4th Cys. Each of the FAS-1 domains contain 110-150 amino acids with low sequence similarities between each of the domains. In each of the FAS-1 domains there is presence of two sets of amino acids sequences (10 to 11 amino acids) that are highly conserved. **Figure 3-3** shows the organisation of TGFBIp and the reported proteolytic sites indicated in the normal protein.

Figure 3-3: Sequence alignment of TGFBIp with its domain organisation and the proteolytic cleavage sites reported in WT-TGFBIp

3.3.2 Secondary and Tertiary Structure of TGFBIp

The full length TGFBIp has been subjected to tertiary structure determination but no successful reports on the 3D structure are available in any of the databases. There are preliminary X-ray diffraction data available for the full length TGFBIp (Runager, García-Castellanos et al. 2009). There are reports from Clout and Hohenester have reported homology models of FAS-1 domains reconstructed from crystal structure of *Drosophila* Fasciclin I (Clout, Tisi et al. 2003). The four FAS-1 domains present in TGFBIp showed a mixed α + β folds, the first half of the protein is rich in α helix and the second half of the protein is predominated by β -sheet under the physiological pH

(Grothe, Little et al. 2009). When the protein was analysed using intrinsic tryptophan fluorescence spectroscopy, an emission maximum spectra was observed around 330 nm. This result from spectra suggests that TGFBIp is compact and the reduced accessibility of the fluorophores used in the tryptophan emission spectroscopy experiments is not accessible to the quencher (I-) indicating that the hydrophobic amino acids like tryptophan is well buried inside the protein.

There has been a report on the existence of TGFBIp as a homotrimer in the periodontal ligament under non-denaturing conditions (Ohno, Doi et al. 2002). There has been another report from Basaiwamoit et al on the low-resolution structure of recombinant human TGFBIp using small-angle X-ray scattering (SAXS). The report suggests that the protein self-assembled and aggregated to form homo-trimers in a time dependant manner. The reports suggested that in the aggregated TGFBIp, the protein monomer is organised in a linear fashion forming rod-like structures (Basaiawmoit, Oliveira et al. 2011).

3.3.3 Secondary and Tertiary Structure of 4th FAS-1 domain of TGFBIp

There has been 4 reports so far in the Protein Database (PDB) on the structure of 4th FAS-1 domain of TGFBIp. Three Solution NMR structures (PDB IDs: 13XB, 2LTB and 2LTC) and one X-ray crystallography structure (PDB ID: 2VXP) has been reported (Yoneyama, Tomizawa et al. 2005, Yoo and Cho 2009). All available structures of the 4th FAS-1 domain of TGFBIp from the Protein databank (PDB) is shown in [Figure 3-4](#). The crystal structure is reported to have 6 α -helices in the first half of the protein and 9 β -strands at the C-terminal part of the protein. There were a few difference observed between the experimentally determined structure and the homology model. The differences being the residues P552-G560 form a stable α -helix and amino acid residue S580-L586 form a hydrogen-bonded turn instead of α -helix as predicted by the model (Clout, Tisi et al. 2003). There has been another report from Runager et al on a model of 4th FAS-1 domain of TGFBIp using Circular Dichroism spectra with intact C-terminus isoform of the protein (M501-A657) and a compact structure of α -helix and β -sheet (Runager, Basaiawmoit et al. 2011). Our group has reported the secondary structure of intact 4th FAS-1 domain of TGFBIp (M501-L632) and it is similar to the secondary structure elements reported by Runager et al (Elavazhagan, Lakshminarayanan et al. 2012).

Figure 3-4: Available structure of TGFBIp from the PDB website.

3.4 Functional properties of TGFB1p

3.4.1 Interaction of TGFB1p with extracellular matrix proteins

There are several proteins and proteoglycans present in the Extracellular matrix of cornea. These proteins are highly essential for maintaining the physiological function and transparency of the cornea. These proteins interact with the cells and also interact with each other covalently and non-covalently with different degrees of affinity and varying specificity that is very important for the tissue function. TGFB1p is found to co-localise with various elastic and non-elastic micro fibrils. This suggests that TGFB1p has a strong interaction with many extracellular matrix proteins (Gibson, Kumaratilake et al. 1997, Thapa, Lee et al. 2007).

Interaction of TGFB1p with Fibronectin

TGFB1p is shown to be purified using Size Exclusion Chromatography (SEC) from the culture medium extracted from primary cells derived from human bladder fibroblast. This indicates that the protein TGFB1p can be purified as a fraction with high molecular weight (>200 KDa fraction) (Billings 2002). When the fraction was subjected to further purification by ion-exchange chromatography, it was identified that TGFB1p was associated with fibronectin and in order to separate the complex to individual protein components a high salt concentration was required (~ 0.5M NaCl) (Billings, Herrick et al. 2000). TGFB1p is shown to have moderate affinity interaction with fibronectin ($K_d=72$ nM) and very weak interaction with collagen I (Billings 2002). The interaction between TGFB1p and fibronectin may possibly be stabilised by hydrophobic bonds and interactions as the complex could not be disintegrated even with very high concentration of NaCl. Fibronectin is reported to have a heparin-binding domain and a collagen/gelatin-binding domain. TGFB1p binds with greater affinity to the collagen/gelatin-binding domain of fibronectin ($K_d =50$ nM) than the heparin-binding domain.

Interaction of TGFB1p with Periostin

Periostin is an ECM protein which is around 93 KDa and it has a greater sequence similarity to TGFB1p in the domain organisation (around 43% identity). Periostin has an additional C-terminal extension of highly hydrophilic amino acid

which is not present in TGFBIp (Hoersch and Andrade-Navarro 2010). Periostin is evaluated as a paralogue of TGFBIp, and periostin is expressed in various layers of cornea like epithelium and fibroblast as a shorter form of the protein (Kim, Jeong et al. 2002). Periostin is found expressed in trans-Golgi network and fibroblasts derived from cornea co-localized with TGFBIp. There has been reports showing that TGFBIp interacts with periostin and the interaction is mediated by EMI domain of TGFBIp (Kim, Olzmann et al. 2009).

Interaction of TGFBIp with Collagens

TGFBIp is found to be co-localised with collagen fibrils in the mature cornea and also in a number of developing tissues (Gibson, Kumaratilake et al. 1997). In later reports Hansen et al showed that TGFBIp can be co-eluted with collagen VI from nuchal ligament tissue from the insoluble collagenase-digested tissue extract using by using size exclusion chromatography (Hanssen, Reinboth et al. 2003). The reports also showed that for TGFBIp binding, the triple helical structure of collagen is highly essential. The triple helical structure also enhances the rate of assembly of native TGFBIp and also pepsin-digested collagen VI (Hanssen, Reinboth et al. 2003, Reinboth, Thomas et al. 2006). A majority of TGFBIp protein (~60%) can be extracted from human cornea as an insoluble fraction by reducing conditions. This results confirms that TGFBIp and collagen VI forms a covalent interaction through a disulphide bond (Andersen, Karring et al. 2004).

There has been another report from Hashimoto et al on the co-localisation of TGFBIp with collagen I. The proteins interact through a strong non-covalent interaction and the complex is difficult to separate the complex using a buffer with high salt concentration (1.5 M NaCl) but the complex can be disintegrated with 7M urea. This shows that there is intense hydrophobic interaction between TGFBIp and collagen I (Hashimoto, Noshiro et al. 1997). Kim et al reported the strong interaction of recombinant human TGFBIp (amino acids 69 to 653) which is truncated N-terminal construct towards chicken collagens type I and type II and also with bovine collagen (type IV and type VI) (Kim, Park et al. 2002). From the reports it can be understood that TGFBIp has strongest interaction with collagen type I followed by collagen type II, type VI and type IV.

Interaction of TGFBIp with Proteoglycans

In the human cornea there is a presence of many glycosaminoglycans (GAGs) among which keratin sulphate and chondroitin/dermatan sulphate are dominant proteoglycans. These GAGs are linked to numerous protein cores and form lumican, keratocan, decorin, biglycan and mimecan (Knupp, Pinali et al. 2009). There has been reports on the interactions of decorin and biglycan with collagen VI with a similar site of interaction to TGFBIp (Wiberg 2001). There is a report from Reinboth et al on the role of TGFBIp interaction with decorin and biglycan and how the complex influences collagen assembly in the cornea (Reinboth, Thomas et al. 2006). Their reports also measured the affinity of the interaction between TGFBIp and biglycan ($K_d=58.8$ nM) and TGFBIp and decorin ($K_d=102$ nM). The interaction also showed that both biglycan and decorin bind to TGFBIp at sites that are in close proximity and the binding increases the affinity of the complex to bind to proteoglycans and collagen VI to form a large ternary complex. There are reports on the competitive binding assays of TGFBIp proteoglycans and GAGs, the results predict that TGFBIp binds to the core proteins of proteoglycans and does not show any binding to GAGs. The reports from Reinboth et al have shown that TGFBIp has a role to regulate the size and rate of decorin/biglycan and collagen VI complex assembly. The report also suggested that TGFBIp is highly needed for controlling the shape and organisation of these collagen and GAG complex (Reinboth, Thomas et al. 2006).

3.4.2 Role of TGFBIp in cell Adhesion and migration

TGFBIp and Epithelial cells

TGFBIp is known to interact with ECM but interaction of TGFBIp with various cell types has also been studied extensively. There is a report from Kin et al on the role of TGFBIp on cell adhesion and migration. According the reports WT-TGFBIp and C-terminal truncated TGFBIp that does not have RGF motifs were also effective in mediating cell adhesion and migration. The interaction is mediated by binding to integrin $\alpha_3\beta_1$ and metal ions that are predominantly divalent in the order $Mn^{2+}>Mg^{2+}>Ca^{2+}$ (Kim, Kim et al. 2000). The cell adhesion property of TGFBIp is mapped to 2 pentapeptides, one present in the 2nd FAS-1 ($N^{354}KDIL^{358}$) and the other in the 4th FAS-1 domain ($E^{615}PDIM^{619}$). When the interaction was studied by solution

NMR it was shown that the peptide **EPDIM** showed maximum interaction in a membrane-mimicking environment and this peptide showed maximum conformational changes in the secondary and tertiary structure upon binding to integrins (Park, Park et al. 2004). Corneal epithelial cell adhesion and migration is greatly enhanced by TGFBIp with a double-repeat of the RGD signal sequence which is mediated by integrin $\alpha_3\beta_1$.

TGFBIp and Fibroblasts

There are reports on the interaction of TGFBIp and fibroblasts through integrin $\alpha_1\beta_1$ and enhancing the migration of fibroblasts. This process requires the presence of divalent metal ions but does not require the presence of RGD signal motif (Ohno, Noshiro et al. 1999). Kim et al reported on the role of TGFBIp in mediating fibroblast migration. Full-length TGFBIp and 4 FAS-1 domains interact strongly with MRC-5 fibroblast cells through $\alpha_v\beta_5$ integrins but not by $\alpha_1\beta_1$ (Kim, Kim et al. 2000). Assays on cell inhibition of cell adhesion shows that the 4th FAS-1 domain of TGFBIp has a greater contribution ($EC_{50}=5-10\ \mu M$) of the inhibition of full length TGFBIp on cell adhesion. The report identified a conserved stretch of peptides each around 18 amino acids along the 4th FAS-1 domain ($I^{165}-T^{182}$, $K^{301}-A^{318}$, $D^{442}-G^{459}$ and $K^{563}-S^{580}$) that could inhibit adhesion of MRC-5 towards full length TGFBIp which is dose dependent with comparable efficacy ($EC_{50}=100-200\ \mu M$). When some of the amino acids from the 4th FAS-1 domain are simultaneously replaced by other amino acids, the efficiency to mediate cell adhesion decreased by 4 folds. The results suggests that these amino acid residues play an important role and may mediate cell adhesion through interaction of 4th-FAS-1 domain or full-length protein with integrin $\alpha_v\beta_5$.

TGFBIp and Endothelial Cells

The full-length TGFBIp and also all the 4 FAS-1 domains is shown to interact with vascular endothelial cells derived from human umbilical vein. The interaction of TGFBIp and endothelial cells is reported to be mediated by integrin $\alpha_v\beta_3$ (Nam, Kim et al. 2003). When amino acid residues (L^{565} , I^{568} , L^{569} , I^{577} and L^{578} with serine and Y^{571} and H^{572} with alanine) along the 4th FAS-1 domain are replaced simultaneously, an 8 fold decrease was observed in the cell adhesion properties of TGFBIp. When cell inhibition assays were performed, the presence of 18-residue peptides that contribute

to cell adhesion along all the 4 domains were identified. The peptides K³⁰¹-A³¹⁸ had a lower efficacy (EC₅₀=200-300 µM) compared to the other three peptides derived from various domains of TGFBIp. The peptide derived from 4th FAS-1 domain of TGFBIp (K⁵⁶³-S⁵⁸⁰) is also shown to possess anti-angiogenic activity both in-vivo and in-vitro.

3.4.3 Role of TGFBIp in Wound Healing

In response to any injury in the corneal stroma, the corneal keratocytes release TGFβ, which in turn causes overexpression of TGFBIp in the cornea near the areas of wound healing (Aldave, Sonmez et al. 2007). The activation of TGFBIp is further supported by the fact that when cornea of patients are treated by PTK or LASIK, the treatment tend to aggravate the protein corneal deposits and elevates the release of the protein TGFBIp in cornea of patients with GCD (Aldave, Sonmez et al. 2007, Kim, Roh et al. 2008, Lee, Stulting et al. 2008). In the previous sections there are several reports that shows interaction of TGFBIp with other matrix proteins like collagen, integrin and fibronectin. The role of TGFBIp in the above interaction is cell adhesion and migration of fibroblast in the cornea to synthesise the ECM. In the process of repair, increased levels of TGFBIp may influence the cells of the corneal stroma to turn them to fibroblasts and tend to synthesise ECM which mostly results in opaque cornea.

3.4.4 Role of TGFBIp in inflammation

The protein TGFBIp is observed to be over expressed in a number of inflammatory diseases like rheumatoid arthritis (Nam, Sa et al. 2006), atherosclerosis (O'Brien, Bennett et al. 1996) and cyclosporine nephropathy (Li, Lim et al. 2005). From the reports on TGFBIp overexpression, it can be understood that TGFBIp is upregulated in inflammatory diseases, both acute and chronic conditions (Gratchev, Guillot et al. 2001). A report from Choi et al observed elevated oxidative stress with decreased levels of antioxidant enzymes like catalase in patient derived corneal fibroblast with GCD2 (Choi, Kim et al. 2009). There is no direct correlation observed between TGFBIp over expression and decrease in catalase enzyme in these patient corneas. However, it can be inferred that a decrease in antioxidant levels and over expression of cytokine markers such as TNFα and interleukin-1β which is observed in inflammatory diseases may be related to simulate a primary role in the progression of TGFBIp-related corneal dystrophies (Nam, Sa et al. 2006).

3.4.5 Role of TGFB β 1 in Morphogenesis

TGFB β 1 is found to be expressed in different tissue in the development of an embryo in mouse models (Schorderet, Menasche et al. 2000) and in zebrafish (Hirate, Okamoto et al. 2003) the gene is expressed throughout embryogenesis. The expression of TGFB β 1 like all other matricellular proteins, is well regulated in the life of an adult except in pathological conditions like inflammation and wound healing. The expression of *TGFB β 1* in mouse during the cushion formation of the endocardium and the spatiotemporal pattern of expression was reported by Lindsley et al (Lindsley, Li et al. 2005). *TGFB β 1* is also found to be expressed in chick embryo during cardiac development (Norris, Kern et al. 2005) and also known to mediate TGF- β 1's effect during mouse embryonic lung development (Lu, Qian et al. 2004). Although the expression of *TGFB β 1* is found in vertebrate embryonic development in various tissues, the exact functional contribution of *TGFB β 1* towards morphogenesis of specific tissue type is still unknown.

3.4.6 Role of TGFB β 1 in Nephropathy

TGFB β 1 expression is eminent in vertebrate embryogenesis in several organ systems and tissues including the kidney. The direct role of TGFB β 1 is still largely under investigation, the induction of TGFB β 1 in response to TGF- β 1 is investigated as a potential biomarker for TGF- β 1 bioactivity. The abnormal expression of TGF- β 1 is reported to have an important role in many kidney related disorders (Kitamura and Suto 1997). There is a reported correlation between TGF- β 1 and *TGFB β 1* expression in Sprague Dawley rats with diabetes induced by streptozotocin. In the experimental model, Immunohistochemical studies detected *TGFB β 1* in juxtaglomerular apparatus and also in parts of proximal tubules that lie in the outer S3 segment. The study also reported increased levels of TGFB β 1 in the urine of the experimental animal model (Gilbert, Wilkinson-Berka et al. 1998, Lee, Bae et al. 2003). It was observed that a combination of excretion of albumin and TGFB β 1 in the urine could directly forecast the extremity of diabetic nephropathy (Ha, Kim et al. 2004).

The other disease where TGF- β 1 and *TGFB β 1* is over expressed is cyclosporine-induced nephropathy. The regions in the kidney that shows elevated presence of TGFB β 1 upon treatment are parietal glomerular epithelium, basement membrane and distal convoluted tubular epithelium (Kim, Lim et al. 2005). The exact mechanism of TGFB β 1 in the pathogenesis remains to be explored but involvement of TGF- β 1 in

tubulointerstitial fibrosis in both humans and animal models through ECM synthesis is observed in cyclosporine-induced nephropathy (Shihab, Andoh et al. 1996). It may be postulated that the increased expression of TGFBIp promotes enlistment of inflammatory cells present in the interstitium and accelerates proliferation of mesangial renal cells. In vitro studies on the epithelial cells derived from renal proximal tubules report the same phenomenon of TGFBIp which supports adhesion, cell migration and proliferation (Park, Bae et al. 2004).

3.4.7 Role of TGFBIp in Tumorigenesis

The association of TGFBIp and Tumorigenesis was first observed by Skonier et al. According to the reports, when ovary cells from Chinese hamster were transfected with *TGFBI* cDNA in a xenograft model of nude mouse, there was a significant reduction in the tumorigenicity (Skonier, Neubauer et al. 1992). Another report from Zhao et al showed significant reduction in cell growth and tumorigenesis in H522 lung cancer cells lacking endogenous *TGFBI* (Zhao, Piao et al. 2002, Zhao, El-Gabry et al. 2006). There are a number of reports on down regulated expression of *TGFBI* in tumor cells and lung carcinoma cells with two possible mechanism of action, hyper methylation of *TGFBI* promoter region and damage to chromosome 5q by deletion or loss of chromosomal region could possible explain the phenomenon of down regulation of *TGFBI* in tumours (Shao, Berenguer et al. 2006, Zhao, El-Gabry et al. 2006).

There are few other reports that observed increased expression of *TGFBI* in certain types of tumors. The exact significance and role of variation in expression of TGFBIp in tumorigenesis is yet to be fully defined or validated but in most tumors, the change in expression of genes including *TGFBI* is due to the change in the growth pattern of the tumor cells and the genes may directly not be involved in tumorigenesis. It may be plausible reason that interaction of TGFBIp with other ECM proteins and integrins may be involved in the aspects of disease progression like metastasis angiogenesis and tumor progression. A detailed investigation on various types of tumors, the type of micro environment in which the cells are present may provide answers to the role of TGFBIp in modulating the events that are directly linked to its pro or anti tumorigenic properties.

4 Alteration of structural conformation and stability of TGFBIp

Mutation induced in genes and the resulting structural changes in the protein products are proposed to be the major cause of amyloid fibril formation in *TGFBI* associated corneal dystrophies (Zerovnik 2002, Clout, Tisi et al. 2003). Protein misfolding due to mutations and amyloid fibril formation is also linked to other diseases like Parkinson's disease, Alzheimer's disease and Huntington's disease (Zerovnik 2002, Clout, Tisi et al. 2003). In TGFBIp the common mutations occur in 2 amino acids (Arg 124 and Arg 555), which are located in the 1st FAS-1 and the 4th FAS-1 domains of the protein and are known to be the mutational hotspots. Many reports suggest that mutation in either of the Arginine residue is less likely to lead to protein misfolding because Arg 124 is located between 2 α helices (α 1 and α 2) in the 1st FAS-1 domain and Arg 555 is located in a linker between 2 α helices (α 3 and α 4) in the 4th FAS-1 domain and both residues are thought to be exposed to the solvent. Changes in the amino acids in the form of mutation in the linker regions are less likely to lead to changes in the secondary structure or cause protein misfolding (Clout, Tisi et al. 2003).

There are other amino acids in the 4th FAS-1 domain that may cause protein misfolding and induce changes in the protein conformation which in turn may affect the protein function and stability. Some of the reported mutations like Thr538, Leu518, Leu527, Asn544, Gly623, Asn622, His626 and Val631 are hydrophobic and are hidden inside the protein and surrounded by hydrophobic environment, mutation in any of these amino acids is likely to disrupt the protein structure and cause the protein to unfold (Clout, Tisi et al. 2003).

There has been reports on the conformational stability of the proteins and the effect of position specific and the location of mutation on the protein sequence. The reports show that mutations in the 1st FAS-1 domain (p.Arg124His, p.Arg124Cys and p.Arg124Leu) of the protein did not greatly affect the stability of TGFBIp. There are other mutations in the 4th FAS-1 domain (p.Arg555Trp, p.Ala546Thr and p.Arg555Gln) of the protein which have large impact on the structure and stability of TGFBIp. The mutations cause both increased (p.Arg555Trp) and decreased stability (p.Ala546Thr and p.Arg555Gln) of the protein which may be different for different mutations and different phenotypes. This suggest that there may be also different

mechanism involved in the protein aggregation and protein deposition in the cornea (Poulsen, Runager et al. 2014).

The mutant p.Arg555Trp responsible for GCD1 had increased stability, increased aggregation propensity and decreased proteolysis susceptibility than the WT protein (Runager, Basaiawmoit et al. 2011). When the recombinant WT and mutant protein was subjected to proteolysis by thermolysin and trypsin, the amino acid residues that were more exposed to proteolysis in the WT protein was located between Arg557 and Leu558 and both amino acids are in the 4th FAS-1 domain and the mutant protein did not respond to the enzymatic cleavage. When the solution state NMR structure of both WT and R555W mutant was published, it was reported that the Arg555 residue is exposed to the solvent in the WT whereas in the mutant protein, the tryptophan residue which replaced the arginine residue was found to be buried in one of the hydrophobic cavity in the 4th FAS-1 domain (Runager, Basaiawmoit et al. 2011). Molecular dynamic simulations of both WT and R555W mutant showed that the $\alpha 3$ helix located around the C-terminus of the protein which contains the cleavage site is reported to be more rigid in the mutant protein, offering less flexibility for proteolysis resulting in the resistance to the proteolytic enzymes (Underhaug, Koldso et al. 2013).

4.1 Cornea specific aggregation of mutated TGFBIp

In *TGFBI* associated corneal dystrophy patients, the only organ affected is the cornea. There is a report from El Kochairi et al on the autopsy of a 79 year-old patient who had both LCD I and Alzheimer's disease. The report showed that the protein deposits are found only in the cornea but not in 17 other organs or tissue samples that were examined from the patient. The organs examined include lung, brain, lymph nodes, kidney, liver, heart, adrenal gland, bone marrow, urinary bladder, parathyroid gland, muscle and prostate (El Kochairi, Letovanec et al. 2006). The reason for the detection of TGFBIp in the cornea may be possible due to abundance of TGFBIp in the cornea (Dyrlund, Poulsen et al. 2012). There are more than 1679 proteins identified in the human cornea and TGFBIp is reported to be the second most abundant protein in the cornea. Since mutation in the TGFBIp leads to protein deposition specific to cornea, it is hypothesised that there may be involvement of cornea specific factors that leads to protein aggregation and deposition in the cornea.

This hypothesis was validated by observing the cultured primary corneal fibroblast cells from patients with p.Ala124Cys did not produce amyloid fibrils when cultured in-vitro (Korvatska 2000). In another report, when the ECM proteoglycan, heparin sulphate was added to the mutant protein p.Ala546Thr, the fibrillation process of the mutant protein is observed to be much rapid in-vitro (Andreasen, Nielsen et al. 2012). Physiological factors present in various layers of cornea may cause mutant TGFBIp to self-aggregate spontaneously or the process may lead to the production of degraded peptides by the action of proteases that may serve as amyloid seeds of aggregation.

Normally in the corneal dystrophy patients, there is a clear zone observed between the limbus and the aggregated protein deposits. A clear zone is also observed near the areas of blood flow and the blood vessels in patients with GCD (Lee, Himmel et al. 2007). In other reports, the protein deposits do not appear at perilimbal cataract incision, and there is a severe aggravation of the deposits in the central region of cornea (Jun 2004, Lee, Chung et al. 2006, Kim, Roh et al. 2008). The protein deposits are much denser near the central region of the cornea than the peripheral region of incision (Feizi, Pakravan et al. 2007). All the reports suggest that the blood vessels may play a major role in inhibiting the abnormal deposition of mutant TGFBIp.

4.2 Localization of protein deposits in various layers of cornea

The protein responsible for corneal dystrophy is the mutant TGFBIp. The protein is a 68 KDa protein and is finally degraded within the tissue. The position or exact location of the cornea in which these protein deposits are found depends on the degradation condition of the mutant TGFBIp and also the proteolytic products of the mutant protein. When primary cultured corneal fibroblast derived from GCD2 patient were subjected to microarray to understand the genes that are significantly up/down regulated, the cells show that 555 genes show at least two-fold difference between the WT and homozygous GCD2 cells. Out of the 555 genes that showed change, 319 of the genes showed up regulation and 236 showed down regulation in the primary corneal fibroblast derived from homozygous GCD2 fibroblasts (Choi, Yoo et al. 2010). The

location of corneal region of the protein deposits may be dictated by different ECM conditions of the mutant protein and their interaction with each other.

4.3 Granular deposition of mutant protein in the cornea

The primary source of mutant protein TGFBIp in the powdery granular protein deposits is from the epithelial cells and the keratocytes. When the corneal granular deposition pattern in the cornea was followed in a GCD2 heterozygous patient, the presence of clear single isolated protein aggregate surrounded by a clear zone and then surrounded by hazy protein deposit that may be present in the sub-epithelial layer of the cornea. This result suggests that the small central sub-epithelial deposits may be the nucleation core for further deposition of mutant proteins.

In most of the patients with corneal dystrophy, once a small protein deposit is formed, the deposit becomes bigger in size and forms amorphous powdery deposits. There has been observations made in some patients where the protein deposits suddenly disappear (Han, Chung et al. 2013). The protein deposits in GCD1 and GCD2 patients were observed over a period of time and reported that some of the superficial protein deposits tend to disappear, suggesting that the protein deposits that may be attached loosely and the friction generated by the opening and closing of the eye lid cause the protein deposits to disappear. However, no clear explanation on how the epithelial deposits disappear has not been investigated in detail (Han, Choi et al. 2016).

4.4 Heterogeneity in the proteolytic processing of the TGFBIp-mutant proteins

It has been previously reported by Fandrich et al and Munishkina et al that proteins associated with disease phenotypes are degraded abnormally under normal physiological conditions (Fändrich, Fletcher et al. 2001, Munishkina, Fink et al. 2004). In Corneal dystrophy, there has been various reports on the increased proteolytic processing and alteration in the turnover of the mutant TGFBIp. There has been several studies that reported the protein composition of the amyloid deposits from patients using 1D, 2D gel electrophoresis and mass spectroscopy methods. The process to analyse the composition of the amyloid deposits from patients is rather complicated

given the complex nature of the protein deposits and the difficulty involved in purification of the components involved in the amyloid deposits. In some methods, the harsh chemical conditions and physical treatment involved in the extraction of the proteins (usage of high salt concentration, harsh denaturants, detergents or proteases) may modify the composition or interact competitively with the aggregation-competent proteins (Tennent 1999). Despite the challenges, proteomics and mass spectrometric methods report on the processing of the various mutant proteins both from the 1st and 4th FAS-1 domains. In a normal cornea, TGFB1p is processed by various proteases into smaller fragments to facilitate clearance. In the case of mutant TGFB1p the proteolytic processing greatly differs which may be the main cause of the disease phenotype.

4.4.1 Proteolytic processing of LCD (Amyloidogenic) mutants

Korvatska et al were the first to report the proteolytic processing of the mutant TGFB1p. They used antibodies that were directed against two regions of TGFB1p that were non-overlapping and reported the proteolysis of TGFB1p by 1D and 2D gel electrophoresis combined with mass spectrometry analysis. In LCD type 1 patients carrying R124C mutation, they did not see a significant change in the full-length TGFB1p in patients compared to the healthy controls. The group reported on the presence of several proteolytic fragments of 17, 26, 39 and 44 KDa molecular weight were observed in the patient samples and they did not find these fragments in the normal control samples (Korvatska 2000). There was about 2-fold increase observed in the 44 KDa proteolytic fragment (which lacks the C-terminal part) in the patient's cornea was also observed. There is another study by Takacs et al reported the presence of shorter 42 KDa and 30 KDa fragment accumulating in the cornea of Lattice dystrophy patients (Takács, Boross et al. 1998). Stix et al reported the composition and proteolytic processing of F540S deposits from patients using electrophoresis. The mutation is responsible for the late onset LCD type IIIA and is located on the 4th FAS-1 domain of the protein. The group observed a major enrichment of 68 KDa protein fragment and also smaller fragments with molecular weights 6.5, 6.9, 14, 17 and 21 KDa (Stix, Leber et al. 2005). In another study involving the V624M mutation, there was no significant change observed in the proteolytic shorter fragments or the TGFB1p content in the patients compared to the normal controls (Afshari, Bahadur et al. 2008). In a recent study on the amyloid composition of V624M mutation, the authors reported on the

accumulation of short peptide TGFBIp-571-588 in patients compared to the healthy control cornea (Karring, Runager et al. 2012, Poulsen, Runager et al. 2014).

Small peptide fragments were also observed in cell lysates when the HEK293T cells were transfected with recombinant TGFBIp (Han, Sim et al. 2011, Han, Sim et al. 2012). The authors used different antibodies that could detect the 47 KDa fragments relating to the L518P and another 43 KDa fragment of R124C and L527R in the cell lysates. These fragments were not identified in the WT protein and hence unique to the mutant protein. These peptide fragments were considered as possible amyloidogenic peptide seeds that are responsible for mutant protein aggregation in the patients.

4.4.2 Proteolytic processing of GCD1 and GCD3 mutants

Klintworth et al, reported the composition of corneal extracts from GCD type 1 patients with R555W mutation using Western blot analysis. The group reported that the full length TGFBIp is overexpressed in the patients compared to the normal cornea (Klintworth, Valnickova et al. 1998). According to the report the authors suggested that the pathogenesis may be due to the accumulation of mutant TGFBIp and its association with collagen VI and other interacting partners of TGFBIp. In granular corneal dystrophy with R124L mutation, there is superficial granular deposits and does not show distinct amyloid fibrils. Korvatska et al reported the proteolytic processing of the mutant TGFBIp with R124L mutation using 1D, 2D and mass spectrometry analysis. According to the reports there is an observed 2.5-fold increase in the 68 KDa TGFBIp fragment in the patient samples compared to the normal controls (Korvatska 2000). In another patient with the same R124L mutation, there was a presence of 63 KDa protein fragment lacking the C-terminus and another TGFBIp fragment lacking both N and C terminus (40 KDa) were observed (Hedegaard, Thøgersen et al. 2003). The 63 KDa fragment was found to be enriched about 77-fold in the patient corneal extract compared to the healthy control cornea.

4.4.3 Proteolytic processing of GCD2 mutants

Analysis of the composition of the corneal extracts from patients with R124H mutation showed abundance of a 68 KDa mutant TGFBIp fragment and another 66

KDa fragment lacking the C-terminal region (Korvatska 2000). The corneal phenotypic analysis showed that the patient's cornea contained non-amyloid deposits.

All the above mentioned reports demonstrates that the phenotypic differences observed in TGFBIp-associated corneal dystrophies is a clear manifestation of the mutant protein status in the patients. Each of the mutant TGFBIp is processed differently compared to the WT protein. The reports clearly demonstrate that the mutant TGFBIp is proteolytically processed differently compared to the WT protein and the proteolytic processing of the mutant protein by various proteases may result in fragments that may potentially act as amyloid fibril seeds and enhance the propensity of mutant TGFBIp aggregation in the cornea. Inhibition of abnormal enrichment of TGFBIp derived peptides may hence present a potential target for treating TGFBIp-associated corneal dystrophies.

4.5 Treatment Strategies of Corneal Dystrophies

Patients associated with *TGFBI*-associated corneal dystrophies may face severe visual impairment based on the type of mutation, the severity of the phenotype and the age of onset of the disease. The only option available right now for these patients is surgical interventions which include tissue replacement or destruction. The major limitation with the surgical procedure is the disease recurrence and the high dependency on the donor corneal graft (Ing, Ing et al. 1998, Terry and Ousley 2001). The surgical method to treat corneal dystrophy depends on the characteristics of the protein deposits, the location of the protein deposits in the cornea, the morphology of the protein deposits and the type of mutation associated with the protein deposits.

Corneal Transplantation

Corneal transplantation or tissue replacement has been the only available option to treat corneal dystrophies. The surgeon replaces the affected cornea with a clear graft from the donor. Penetrating Keratoplasty (PK) is a technique whereby the affected corneal tissue from the patient is fully removed and replaced with a graft. The Anterior Lamellar Keratoplasty (ALK) is a technique where only a part of the affected cornea is removed by surgery and replaced with a donor cornea. Both the surgical procedures results in a clear cornea and delays the recurrence of the disease and slows down the

natural process of protein aggregation (Carrwik and Stenevi 2009). Despite the surgical interventions, the chances of the protein deposits to recur is high depending on the type of mutation (Ellies, Bejjani et al. 2003). Based on the reports so far patients with p.Arg124Leu and p.[R124L;T125_E126 Del] which is linked to superficial corneal dystrophy had significant recurrence of the protein deposits within 5 to 8 years of the corneal transplant. Patients associated with LCD phenotype (p.Arg124Cys, p.His626Arg and p.Ala546Thr) did not have early recurrence of the protein deposits. For the patients associated with LCD, the recurrence of the protein deposits were between 5 to 7 years after the first transplant and a significant amount of the protein deposits are found after 12 years. The patients associated with TBCD phenotype with p.Arg555Gln mutation also had a late recurrence after 16 years (Ellies, Bejjani et al. 2003, Sorour, Yee et al. 2005).

For patients with severe GCD, PK was used to improve visual acuity and the outcome of the surgery was clear cornea in patients but visible protein deposits along the suture tracts and incision lines were observed between 12 to 24 months (Mashima, Konishi et al. 1998, Okada, Yamamoto et al. 1998, Diaper, Schorderet et al. 2005). The advancement of newer surgical procedures like Deep Anterior Lamellar Keratoplasty (DALK) allows replacement of corneal tissue in patients still preserving the corneal endothelium and the Descemet's membrane of patients (Tan and Mehta 2007). The technique of selectively replacing the patient corneal tissue while leaving the endothelium intact helps in preserving the integrity of the corneal tissue and reduces the chances of graft-rejection. Although this surgical technique is successful in treating patients with corneal dystrophy, the surgery is challenging and demanding when the protein deposits are found deep in the stroma or near the Descemet's membrane (Moon, Kim et al. 2007, Noble, Agrawal et al. 2007, Park, Ki et al. 2007).

Corneal Surface Ablation

In patients with stromal corneal dystrophy, Phototherapeutic keratectomy (PTK) induced by excimer laser provides relief from symptoms and restores the vision (Rapuano 2001, Ayres and Rapuano 2006, Han, Choi et al. 2012). The technique is useful to treat different types of GCD and LCD (Sher 1991, Dogru, Katakami et al. 2000, Das, Langenbucher et al. 2005), the results are reported to be favourable in restoring the vision of patients but the recurrence of the protein deposits pose a major

problem (Dinh, Rapuano et al. 1999, Dogru, Katakami et al. 2000). The recurrence rate of the protein deposits is based on the type of genetic mutation with a homozygous or heterozygous pattern of expression in the patients (Inoue, Watanabe et al. 2002, Gruenauer-Kloevekor, Braeutigam et al. 2009). When the protein deposits tend to recur, the deposits are denser and start are observed as early as 7 months after treatment (Dogru, Katakami et al. 2000) with protein deposits starting from the laser ablation margins.

When the efficiency of PTK was examined by Ellies et al in reported patients with different types of corneal dystrophy, it was identified that the procedure improved the vision in patients up to two or more lines of corrected visual acuity in a significant number of patients (85% of patients treated with PTK) (Ellies, Bejjani et al. 2003). Based on the reports, the patients who achieved the most improvement in visual acuity are the ones with p.Arg124Leu and p.[Arg124Leu;ΔT125ΔE126] mutations. The next group of patients with improvement are the ones with p.Arg555Trp and p.Arg124Cys mutations. The patients did not complain of any post-operative complications like corneal ulcers or epithelial defects. There are some observed disadvantages of PTK like corneal scarring, pain after the surgery, discomfort, hyperopic shift and delayed re-epithelisation of corneal tissue. PTK is not suited for patients with protein deposits in the deep stroma and usually recommended for patients with more superficial corneal deposits.

For patients who complained of haze in the cornea after PTK, application of mitomycin-C (MMC) is shown to have reduced the haze and aids in wound healing (Majmudar, Forstot et al. 2000, Shah and Wilson 2010). MMC is known to be an alkylating agent that is capable of inhibiting DNA synthesis. There has been reports that suggest the possible role of MMC to induce a time-dose-dependent apoptosis of the stromal keratocytes in patients with GCD phenotype (Jun 2004, Kim, Choi et al. 2008). There is another report which identified no substantial effect of MMC after PTK with reference to disease recurrence (Ha, Kim et al. 2010). So the long term effect of MMC and its role to delay the corneal aggregates remains to be explored.

There has been another technique to treat protein deposits to have less discomfort and complications after surgery with minimal hazing of the corneal tissue and improved wound healing potential (Anderson, Beran et al. 2002, Lee, Stulting et al. 2008). This technique is known as Laser Epithelial Keratomeileusis (LASEK) but the protein deposits tend to recur is a major problem of the technique. For patients with

latent and underlying TGFB1p deposits, Laser-assisted in situ Keratomeileusis (LASIK) is shown to have adverse effect on the disease (Jun 2004, Lee, Himmel et al. 2007). LASIK technique has recently shown to aggravate the conditions of GCD2 patients if they underwent a LASIK treatment for vision correction with the latent TGFB1p deposits in the cornea (Awwad, Di Pascuale et al. 2008).

Electrolysis

The technique of electrolysis is not a commonly used form of treatment to treat protein deposits in corneal dystrophy. There has been reports from Japan on the use of electrolysis to treat corneal dystrophy. The technique uses the electrodes from an electrolyser and addition of normal saline which reacts with water to produce sodium hydroxide and this can dissolve the protein deposits. For the whole procedure, the patient is kept under topical anaesthesia. The time take for each cycle is about 5 minutes and the whole process is repeated until the protein deposits are removed completely. This procedure is not very effective in removing the corneal lesions which are present in the deep anterior part of the cornea. Although the protein deposits tend to disappear after electrolysis, they recur very fast and the long-term management of patients who reported recurrence is not documented or reported.

Study Aims and Hypothesis

Current treatment options for *TGFB1*-associated corneal dystrophies are transient due to the recurrence of the deposits after surgery. Since protein aggregation is strongly linked with the disease, treatment strategies that would halt the protein deposition and clear the insoluble deposits may improve visual acuity. The aim of the project is to identify a lead compound that can be used to alter/prevent Corneal Dystrophy. This study also aims to investigate the functions of TGFB1p in the cornea and understand more on the cornea-specific deposition of mutant TGFB1p.

Analysis of protein composition and proteolysis of corneal aggregates with mutations in the *TGFBI* gene

TGFBIp is known to be the second most enriched protein in the corneal stroma. To date there are more than 65 mutations reported in the *TGFBI* gene. The amino acid alterations are responsible for the altered proteolytic processing and formation of insoluble aggregates that are deposited in various layers of the cornea. The protein deposits in the cornea may appear as long lattice lines or amyloid fibrils in Lattice Corneal Dystrophy (LCD), amorphous protein aggregates in Granular Corneal Dystrophy I (GCD I) or a combination of amyloid fibrils and amorphous aggregates in GCD II. The deposition of insoluble aggregates is responsible for poor visual acuity and corneal opacity in the affected patients. Depending on the nature of mutation, the phenotype, age of the onset, severity of the disease and the location of deposits in the cornea may vary making the disease unique.

TGFBIp is found to be present in various other tissues. However, mutations found in *TGFBI* is associated with corneal specific disorders only. This explains the possibility of a unique mechanism of protein aggregation in the cornea and the interactions with other tissue-specific factors may play an eminent role in protein aggregation and deposition in the cornea. Biophysical and Molecular Dynamic studies indicated that some mutant forms of TGFBIp (R555W and G623D mutant proteins) are more stable than WT and some mutations destabilize the protein (A546T). It is likely that changes in the proteolytic stability, proteolytic processing and increased turnover of the mutant protein may eventually contribute to the pathogenicity of Corneal Dystrophies.

The main objectives of the study is to identify the composition of amyloid protein deposits in the affected cornea of the dystrophy patients with different types of mutation and protein deposits and to have an insight on the interaction of TGFBIp with other proteins in the cornea. By analysing the composition of proteins present in the amyloid corneal deposits of patients with different phenotypes and genotypes, then comparing the protein composition against the protein composition of the healthy control cornea will help us to understand the mechanism of proteolytic processing of mutant proteins in the cornea. This will also help us to understand if there are any differences in the clearing mechanism of amyloid protein deposits and proteolytic processing of different mutant proteins. This will help to postulate new mechanism

involved in the proteolytic processing of mutant protein and clearing mechanism of amyloid protein deposits in the cornea.

Bio-physical and structural characterization of peptide fragments derived from 4th FAS-1 domain of TGFB1p

Mutations in TGFB1p are hypothesised to induce structural changes in the mutant protein leading to exposure of hydrophobic amino acids and oligomerization of proteins. Very few studies have been carried out to understand how the mutant proteins aggregate and form insoluble deposits. Various reports show that mutant peptides with a decrease in charge, increase in hydrophobicity, and increased propensity to form β -sheet structures, accelerated the rate of aggregation compared to the wild type peptide. Apart from the intrinsic properties (hydrophobicity and overall net charge) and extrinsic factors (pH, temperature, salts etc.), the propensity to form fibrillar aggregates also depends upon the size of the protein. It has been postulated that larger proteins have a higher ability to form aggregates, but low predisposition to form amyloid fibrils. Mutant peptides from the 1st and 4th FAS-1 domain of TGFB1p have increased aggregation propensity in contrast to the wild-type protein.

In Corneal Stromal Dystrophies we observe aggregated TGFB1p protein deposits in the cornea and depending upon the type of mutation and location of the deposits, their physical and chemical properties show immense variations. All the amyloid fibrils from mutant proteins appear with a similar morphology when examined under Electron Microscopes, but their bio-physical properties and the phenomena or the trigger that causes the fibrillar deposits may vary based on the location of the protein and the pathology it causes. Many proteins are triggered by nucleation and elongation of self-assembled peptides to form amyloid fibrils. Therefore, elucidation of the molecular architecture of amyloid fibrils and characterization of the peptides capable of acting as seeds and facilitating nucleation and elongation of amyloid fibrils is important to understand the mechanism of amyloid fibril formation.

The aim of the study is to understand the molecular-functional relationship of TGFB1p derived peptides and their propensity to form amyloid fibrils. LC-MS/MS based protein profiling of Lattice corneal amyloid deposits (*TGFB1*- H626R and R124C mutations) derived from dystrophic patients showed abundance of TGFB1p in the deposits. Spectral counts of peptides derived from TGFB1p upon tryptic digestion

showed enrichment of several peptides along the 4th FAS-1 domain compared to the normal control cornea. One such peptide was TGFBIp 611-642. This peptide was also shown to be highly amyloidogenic based on four different prediction algorithms. We modified the 23 AA residue peptide (TGFBIp611-633) with cationic charged amino acid residues and investigated the amyloid forming properties of the mutant peptides. We chose mutant variations of the peptides that decreased the overall net charge of the peptides to further investigate its amyloid propensity and aggregation kinetics. Interestingly, the mutations within this region are associated with clinically relevant amyloidogenic phenotypes and 10 out of 12 mutations are associated with alteration of overall net charge.

We also investigated the secondary structures of the amyloid fibrils derived from wild type and mutant peptides which differ considerably in thermal stability by solid state nuclear magnetic resonance (ssNMR).

Modulating the stability of TGFBIp by small molecules/fragments

The main objective of our study is to understand mechanism of amyloid fibril formation in *TGFBI* associated CD and aim to find small fragment compounds that can bind and modulate the TGFBIp. This compound could possibly stabilize the protein and delay the aggregation process or modulate the oligomeric intermediates and prevent further aggregation or play a role in proteolysis and protein clearance mechanism.

It has been shown recently that aberrantly folded and aggregated proteins are the causes of almost 30 different disorders in humans that have no form of cure currently. These disorders and diseases affect different organs in different parts of the body and one such disorder is corneal dystrophy which causes protein deposits in the cornea. The targets in these kinds of protein misfolding diseases are the proteins themselves and very specifically, the whole process of mutant protein aggregation.

In general, proteins with high amyloidogenic property can be divided into two main categories. The first category consists of proteins that are highly structured naturally but when there is a mutation in the gene, the protein product from the gene becomes highly unstructured and form amyloid fibrils which aggregate in the tissue due to a clearance mechanism that has completely failed (Liu et al. 2012) or due to the lack of certain factors that prevent mutant protein aggregation. The second category of proteins are natively unstructured (A β and tau). To date there has been different

mechanisms proposed to describe aggregation and amyloid fibril formation. The most widely accepted model is nucleation-dependent polymerization model proposed by Jarrett et al in 1993. According to this model the monomers self-associate into oligomers in a slow, rate-determining step, the oligomers then act as nucleation seeds for further aggregation by addition of monomers which can be very rapid (Jarrett, Berger et al. 1993).

There are other pathways that involve self-association of monomers into small oligomers, which later assemble to form larger oligomers and amyloid fibres (Urbanc, Betnel et al. 2010). The mechanism of aggregation may vary from protein to protein or there might be an involvement of both pathways simultaneously for some proteins. It becomes difficult to spot which species of the aggregation pathway are more toxic or target the most crucial stage in the aggregation pathway. The main aim of the project is to screen for small fragments/chemical compounds that can bind and modulate TGFBIp and prevent/delay the aggregation process. This may result in a much stable protein that is not susceptible to faster aggregation.

Our aim of the study is to screen chemical fragments or compounds that can bind to the mutant protein and in a way delay the aggregation process by either stabilizing the mutant protein or modulate the proteolytic processing and clearance of the mutant proteins. The lead compound identified may be used as a therapeutic option for cornea dystrophies.

Chapter 3

Graphical Abstract of the Analysis of amyloid Corneal Aggregates from Corneal Dystrophy patients

5 Protein composition analysis of amyloid fibrils from corneal dystrophy patients

5.1 Introduction

TGFBI associated corneal dystrophies are classified into different subgroups based on genetics, nature of protein deposits and the layer of cornea in which the protein deposits are found (Weiss, Moller Hu Fau - Aldave et al. 2015). In Lattice corneal dystrophy (LCD), the protein deposits appear as long lattice lines or amyloid fibrils, whereas in granular corneal dystrophy I (GCD I) patients exhibit amorphous protein aggregates, and in GCD II both amyloid fibrils and amorphous aggregates are evident (Klintworth 2009). Deposition of these insoluble protein aggregates leads to reduced corneal clarity and profound visual impairment (Surguchev and Surguchov , Weiss 2010). Since the protein deposits play a major role in the pathology, a detailed analysis of the composition of the amyloid deposits may provide improved understanding of the mechanism of disease progression and may pave the way to therapeutic development.

Different type of amino acid substitution exert a variable impact on *TGFBI*p amyloid phenotype, age of onset, disease severity and processing of the protein (Lakshminarayanan, Chaurasia et al. 2014). Indeed, in a previous study using 1D, 2D gel electrophoresis combined with mass spectrometry analysis of protein deposits from variants p.Arg124Cys, p.Arg124His and p.Arg124Leu, Korvatska *et al.* reported abnormal turnover and increased aggregation of the amyloidogenic mutant protein compared to WT. In addition, the authors detected several truncated proteolytic fragments (17-KDa, 26-KDa, 39-KDa and 44-KDa) with a 2 fold increase in the 44-KDa fragment, which lacked the C-terminus residues, compared to the normal 68-KDa protein (Korvatska 2000). In the case of R124L and R124H, both from non-amyloid patients the authors did not report any smaller proteolytic products. Hence abnormal proteolytic processing and accumulation of these short peptide fragments in patients with amyloid related mutations as opposed to non-amyloid patients, represent mutation-specific amyloid fibril seeds that accelerate aggregation.

There have been several previous attempts (Karring, Runager et al. 2012, Karring, Poulsen et al. 2013, Poulsen, Runager et al. 2014, Courtney, Toftgaard Poulsen et al. 2015) to investigate the composition of amyloid deposits and amorphous

aggregates from the corneas of dystrophic patients, mostly of Caucasian origin. The composition of patient corneal deposits reported so far are from Lattice corneal dystrophy (R124C(Courtney, Toftgaard Poulsen et al. 2015) mutation), Granular corneal dystrophy (R555W(Courtney, Toftgaard Poulsen et al. 2015) and R124H(Karring, Runager et al. 2012) mutation) and variant LCD dystrophy (A546D(Karring, Poulsen et al. 2013), V624M(Karring, Runager et al. 2012) and A546D/P551Q(Poulsen, Runager et al. 2014) mutations). These studies revealed that the serine protease HTRA1 was substantially enriched in corneal deposits from patients with lattice dystrophy but not in patients with granular dystrophy(Courtney, Toftgaard Poulsen et al. 2015). The results also suggested that HTRA1 preferentially mediated the proteolytic processing of TGFBIp amyloid mutants, and the mutants displayed increased sensitivity to proteolytic cleavage of residues in the 4th FAS-1 domain, resulting in enrichment of peptides in corneal deposits.

In the present study, we performed in-depth composition analyses of amyloid fibrils from Lattice corneal dystrophic patients with H626R and R124C mutations with the aim of identifying peptides capable of seeding protein aggregation. The composition of amyloid deposits for patients with H626R mutation which is highly prevalent in the South-East Asia, accounting for >75% of cases in the Vietnamese population has not been previously reported (Chau 2003). While the amyloid profile of patients with R124C mutations have been reported before (patient of Caucasian origin)(Courtney, Toftgaard Poulsen et al. 2015), it is unknown whether patient ethnicity influences peptide truncation and protein aggregation. We therefore used laser capture Microdissection (LCM) to isolate amyloid deposits from the cornea of dystrophic patients and healthy controls, then subjected the aggregates to tryptic digestion and LC-MS/MS to identify proteins that were specifically enriched that could potentially act as amyloid fibril seeds.

We observed that corneal amyloid fibrils from dystrophic patients were enriched in both ‘amyloidogenic’ and non-fibrillar proteins including apolipoprotein A-IV, lysozyme, IgG chain C, serum amyloid-P, apolipoprotein E, decorin and biglycan. The enzyme HTRA1 showed 7-fold increase in the amyloid corneal deposits from dystrophic patients compared to the wild type and the enrichment of shorter TGFBIp peptides in the patient corneal deposits may likely play a key role in disease progression. Furthermore, we demonstrated that the TGFBIp peptide sequences found in the patient deposits, (F⁵¹⁵SMLVAAIQSAGLTETLNR⁵³³ and

Y⁵⁷¹HIGDEILVSGGIGALVR⁵⁸⁸) exhibited increased propensity to transform into amyloid fibrils as assessed by circular dichroism assays, Thioflavin T binding, and transmission electron microscopy. Together our data suggests that the amyloid deposits from dystrophic patients are composed of several amyloidogenic proteins, non-fibrillar amyloid associated proteins and TGFBIp. The enrichment of serine protease, HTRA1 in the corneal amyloid deposit suggests the involvement of the protease in processing the mutant protein and altering the enzymatic degradation pattern of the mutant TGFBIp. This may induce the formation of amyloid seeds and may be involved in amyloid fibril formation seen in the dystrophic patients.

5.2 Materials and Methods

5.2.1 Corneal dystrophy patient samples

Corneal tissues were collected from three patients with lattice corneal dystrophy (*TGFBI* mutations H626R [n=2] and R124C [n=1]) who were undergoing corneal transplantation at the Singapore National Eye Centre. Written informed consent was obtained from all patients prior to surgery (patient details are provided in [Table 5-1](#)). Transplant-grade Control tissues, n=3 from deceased individuals were obtained from Lion's Eye Institute, Tampa, USA. Ethical approval for the collection of patient corneas was granted by the Singhealth Institutional Review Board.

Table 5-1: Patient genotype and disease phenotypes used in this study

S. No	Mutational status	Age of patient	Disease Phenotype
1	H626R	53	Variant LCD
2	H626R	55	Variant LCD
3	R124C	61	LCD type 1

S. No	Gender	Age of control	Phenotype
1	Female	56	Normal controls
2	Male	59	Normal controls
3	Male	67	Normal controls

5.2.2 Preparation of corneal tissue sections and laser capture microdissection (LCM)

Corneal tissue samples were fixed in paraformaldehyde and dehydrated by immersion in alcohol solution before being embedded in paraffin wax. The samples were allowed to set, then a microtome was used to cut 7µm tissue sections for collection onto PEN-coated membrane slides (#LCM 0522, Life Technologies), suitable for Laser Capture Microdissection (LCM). A maximum of 6 sections were placed onto each slide and allowed to dry before being deparaffinised and rehydrated by washing in graded acetone solution. TGFβ1p amyloid deposits were located by staining with Congo red according to the manufacturer's protocol (#860-026, Ventana Medical Systems).

LCM of corneal sections was performed manually using an Arcturus Laser Capture system (Arcturus Engineering, Inc., Mountain View, CA). Once isolated, the Congo red-stained amyloid deposits were collected onto Cap Sure LCM caps (Arcturus Engineering, Inc., Mountain View, CA). The surrounding amyloid-free tissues that stains negative with Congo red (referred to peri-amyloid tissue in this manuscript) were also collected onto separate LCM caps to be used as patient-matched controls. All LCM caps were stored at -80°C until analysis by mass spectrometry.

5.2.3 Sample preparation for mass spectrometry LC-MS/MS analysis

Films were stripped from LCM caps and placed into Eppendorf tubes containing 40µl of 10mM Tris buffer pH 7.8 containing 0.1% RapiGest SF Surfactant (#186001860, Waters Corporation, Milford, MA). The tubes were then centrifuged at 10,000 X g for 60s and heated in a water bath at 95°C for 90min total duration. The centrifugation process was repeated every 20min and at the end of incubation prior to cooling the samples on ice.

Samples were reduced with 10mM DTT at 60°C for 30min, followed by alkylation with 55mM IAA at room temperature for 45min in the dark. Overnight digestion was performed using 10µg trypsin at 37°C. Digested peptides were desalted using Sep-Pak C18 Vac cartridges (Waters, Milford, MA). Desalted tryptic peptides were reconstituted in 0.1% FA for LC-MS/MS analysis.

Normal human cornea samples were collected from healthy deceased donors and the epithelium was removed by gentle peeling. The remaining corneal tissue was then fractionated into small pieces using a sharp scalpel and processed to extract stromal proteins.

LC-MS/MS analysis of corneal tissues

LC-MS/MS analysis was performed using a Q Exactive mass spectrometer coupled with online Dionex Ultimate 3000 RSLC nanoLC system (Thermo Scientific; MA). Samples were run on a Dionex EASY-spray column (PepMap® C18, 3µm, 100Å) using an 120min gradient of mobile phase A (0.1% FA in 5% ACN) and mobile phase B (0.1% FA in ACN) with a flow rate of 300 nl/min. Separated peptides were passed through an EASY nanospray source with a source voltage of 1.5kV. A full MS scan was set at 350–1600 m/z with resolution of 70,000 at m/z 200. The maximum ion accumulation time and dynamic exclusion time were set as 100 ms and 30 s respectively. The 10 most intense ions detected above a threshold of 1000 counts were selected for higher-energy collisional dissociation (HCD) fragmentation using 28% normalized collision energy with a maximum ion accumulation time of 120ms. Other parameters used were Automatic Gain Control $1e + 06$ for the full MS scan and $2e + 05$ for the MS2 scan, isolation width of 2Da for the MS2 scan, and exclusion of single and unassigned charged ions from the MS/MS scan (with underfill ratio set to 0.1%). Data acquisition was performed using Xcaliber 2.2 software (Thermo Scientific; USA).

Data analysis

Raw data were processed using two different approaches to achieve high-confidence of label-free quantitation. Firstly, all raw data were converted into mascot generic format (MGF) using Proteome Discoverer 1.4.1.14 (Thermo Fisher, MA) and searches were performed using an in-house search engine (Mascot, version 2.4.1; Matrix Science, London, UK). The UniProt Knowledgebase (UniProtKB) of human proteins (downloaded on July 25, 2016, including 70,849 sequences and 23,964,784 residues) was used as the search database. The following settings were applied; static modification = Carbamidomethyl at cysteine; dynamic modifications = methionine oxidation, asparagine/glutamine deamidation; digestion parameters = full trypsin with maximum 2 missed cleavages, semi-trypsin with one non-specific cleavage; search

parameters = #13C is 2, precursor mass is 10ppm, fragment mass tolerance is 0.02Da. Selection of proteins for final analysis was performed using a target-decoy search strategy with false discovery rate (FDR) $\leq 1\%$ considering only those proteins identified with multiple peptides. An additional search was performed using semi-tryptic digestion as the enzyme parameter in order to identify non-tryptic peptide cleavage sites. The emPAI value reported by Mascot and spectral counting were used to perform quantification of proteins and peptides, respectively (Ishihama, Oda et al. 2005, Shinoda, Tomita et al. 2010).

The raw data files were also analysed using the Proteome Discoverer 1.4 software package. MS/MS spectra were searched against the same UniProt human database using the Mascot search engine with identical parameters to those described above. The peptides-to-spectrum matches were filtered by application of $<1\%$ FDR threshold using Percolator. Only peptides identified with high confidence as assigned by the software were selected for further analysis. Precursor ion area detection node was enabled in order to obtain the area under the curve (AUC, or extracted ion chromatogram / XIC) of the LC elution profiles of the detected peptides. The AUCs of the same peptides obtained from both patient and control samples were used for quantitative analysis to determine the fold-changes and extent of enrichment in patient tissues. Peptide/protein lists were exported to Excel and processed using in-house analytical scripts. All statistical calculations were performed using three technical replicates for each sample.

5.2.4 TGFBIp synthetic peptides

TGFBIp synthetic peptides 515-533 and 571-588 (purity $>95\%$) were purchased from Synpeptides (Synpeptide Co Ltd, Shanghai, China). The peptides were dissolved in HFIP to remove pre-formed aggregates prior to aliquotting and lyophilization for storage as a dry powder until further use. The peptides were reconstituted in PBS to a final concentration of 0.6 mg/ml and incubated at 37°C in a shaking incubator for up to 4 weeks for use in biophysical assays.

5.2.5 Circular dichroism spectroscopy

A total of 100 μL peptide sample was diluted with 100 μL PBS and time-dependent changes in conformation were tracked by circular dichroism spectroscopy.

Far UV spectra of the peptides were collected using a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD) and Chirscan-plus (Applied Photophysics, UK) using a quartz cuvette with a path length of 0.1cm (Hellma, Müllheim, Germany/ Starna). Spectra were recorded from 190-260nm with a step size of 0.1nm and the final spectrum was calculated as an average of three scans. The final spectrum was calculated as the average of three scans. Circular dichroism data were expressed as mean residual weight ellipticity (deg cm² dmol⁻¹). Mean residual weight (MRW) ellipticity ([θ] MRW) was estimated using the following equation:

$$[\theta]MRW = [\theta]\lambda \times MRW / 10 \times l \times c$$

where [θ] λ is the observed ellipticity, MRW is mean residual weight (defined as M/N-1, where M is the molecular mass and N is number of amino acid residues), l is the path length of the cuvette, and c is concentration in mg/ml. Data were plotted using Graphpad PRISM 7 software.

5.2.6 Thioflavin T (ThT) assays

Synthetic peptides (0.6 mg/ml) were treated with 30 μ M ThT in PBS buffer in a Greiner 96-well flat-bottom microplate (Greiner, Frickenhausen Germany). The plate reader was maintained at 37°C under constant shaking to mimic physiological conditions for fibril formation. Samples were excited at 445 nm and fluorescence at 485 nm was measured at regular intervals over a total period of 17 h using a microplate reader (Tecan infinite M200 pro, Zanker Road, San Jose, USA). Graphpad PRISM 7 software was used to plot amyloid fibril formation data and calculate single exponential functional kinetics of fibrillation and peptide t_{1/2}.

5.2.7 Transmission electron microscopy (TEM)

TEM images of the peptide fibrils were acquired with a JEOL JEM-1010 transmission electron microscope using Digital Micrograph™ 1.81.78 for GMS 1.8.0 (Gatan, Pleasanton, CA). The fibrillated peptide sample (100 μ l total volume) was centrifuged and a 10 μ l volume was applied onto a Formvar carbon-coated nickel grid (mesh size 300) using bacitracin as the binding agent. Excess sample was blotted away from the filter paper and stained with 10% Phosphotungstic acid (PTA). The grids were dried and visualized at magnification 8000-50,000X (80kV).

5.3 Results

5.3.1 Patient sample identification and laser capture microdissection

Corneal tissues were collected from three patients with lattice corneal dystrophy (*TGFBI* mutations p.His626Arg [n=2] and p.Arg124Cys [n=1]) who were undergoing corneal transplantation. Protein deposits were clearly visible in corneal photographs of all three patients, and lattice fibrils were evident in two of the three donors (**Figure 5-1**). The corneal sections were stained with Congo red to locate amyloid deposits which were then collected using laser capture microdissection (LCM) as shown in **Figure 5-2**. Peri-amyloid tissue adjacent to the amyloid deposits was also collected to be used as patient-matched controls.

5.3.2 Protein composition of amyloid deposits in corneal dystrophy patients

Label-free protein quantitation was calculated based on Mascot-reported emPAI values and area under curve (AUC) values using proteome discoverer software were calculated as detailed in the materials and methods and used for analysis. The number of proteins identified within the amyloid deposits and peri-amyloid tissues from each patient are shown as Venn diagrams in **Figure 5-3** and listed in **Table 5-2**. A total of 52 amyloid deposit proteins were successfully identified in all three dystrophic patient samples (**Figure 5-4A**), whereas 62 proteins were detected in each of the peri-amyloid surrounding tissue samples from the same donors (**Figure 5-4B**). Analysis of healthy control corneal stroma samples (n=3) yielded ~126 overlapping proteins (**Figure 5-5**).

Figure 5-1: Corneal images showing lattice lines and aggregated proteins with opaque cornea in dystrophic patients with H626R mutation (A) and (B) R124C mutation (C).

Figure 5-2: Isolation of amyloid corneal deposits by laser capture microdissection. A) FFPE-fixed corneal sections were stained with Congo red and tissues staining dark brown were identified as amyloid deposits. B) Tissue section after laser capture microdissection.

Figure 5-3: Protein profiles of amyloid corneal deposits and surrounding tissues in individual Lattice dystrophy patients. Venn diagrams show the distribution of proteins identified in the corneal deposits and surrounding tissues sampled from patient 1 (A), patient 2 (B), or patient 3 (C).

Figure 5-4: Protein profiles of corneal deposits and surrounding tissues grouped in patients. A) Venn diagram showing the distribution of proteins identified in corneal deposits from all three patients (~52 proteins were common to all three donors). B) Venn diagram showing the distribution of proteins identified in unaffected tissues from all three patients (62 proteins were common to all 3 donors).

Figure 5-5: Protein profile of healthy control cornea. Venn diagram showing the distribution of proteins identified in control cornea samples (~126 proteins were common to all 3 control donors). An average of ~206 proteins was detected in control tissue)

Table 5-2: Number of proteins identified in amyloid corneal deposits and adjacent tissues from patients (or healthy control cornea). Numbers represent average values obtained from triplicate mass spectrometry runs.

Healthy control cornea proteins	Amyloid corneal deposit proteins	Surrounding tissue proteins
Control 1 =199	Patient 1= 76	Patient 1=101
Control 2=217	Patient 2=98	Patient 2=106
Control 3=203	Patient 3=105	Patient 3=77

5.3.3 Identification of proteins enriched in amyloid deposits from Lattice dystrophy patients

When the proteins identified were ranked according to their relative abundance (Table 7-3), we observed that TGFBIp was the most abundant protein in both patient amyloid deposits and corneal tissues from healthy controls. Amyloid deposits were also enriched in proteins known to be associated with systemic amyloidosis (apolipoprotein A-IV, lysozyme) or non-fibrillar amyloid formation (serum amyloid P-component, clusterin, decorin, biglycan, apolipoprotein D and E) (Tennent, Lovat et al. 1995, Botto, Hawkins et al. 1997, Stokes, Holler et al. 2000, DeMattos, O'Dell M et al. 2002, Holtzman 2004). Proteins such as apolipoprotein A IV, biglycan, clusterin, decorin, aldehyde dehydrogenase, annexin, pyruvate kinase, histone H4, histone 1, keratocan, lumican, actin and vimentin were found in amyloid deposits from all three dystrophic patient samples. In addition, Amyloid deposits from both patients with H626R mutation were found to be enriched with apolipoprotein D, apolipoprotein E, alpha enolase, kaliocin-1 and serum albumin compared to the amyloid deposits from the patient with R124C mutation. Proteins such as angiopoietin related protein 7, clusterin beta, Ig Kappa, peroxiredoxon, prolargin and ubiquitin were detected only in the dystrophic patient with R124C mutation. Proteoglycans such as biglycan, decorin, keratocan and lumican are found to be abundant in amyloid deposits from all three dystrophic patients.

Proteins such as Actin, Aldehyde dehydrogenase, alpha enolase, annexin, apolipoprotein A IV apolipoprotein E, clusterin, collagen alpha-3 (VI) chain, kaliocin

and pyruvate kinase were unique to the amyloid deposits but absent in the normal control samples. Proteins such as annexin A2, Collagen alpha-1(XII) chain, galectin-1, Ig gamma-1 chain C region, Ig gamma-4 chain C region, Ig lambda-1 chain C region, Ig lambda-7 chain C region, serum albumin, MAM domain-containing protein 2 and Thrombospondin-4 were unique to the healthy control samples and were absent in the amyloid deposits.

Specific enrichment of proteins in patients with Lattice amyloid deposits

We next sought to identify proteins that were specifically enriched in amyloid deposits compared with peri-amyloid deposits from corneal tissues obtained from the same patients. As shown in [Table 5-4](#), we observed that all the three corneal deposits from patients were significantly enriched in retinal dehydrogenase, clusterin, elongation factor 1- α 1, transketolase, actin and ubiquitin. However, the most abundant proteins detected in deposits were apolipoprotein A-IV, apolipoprotein E, and serine protease HtrA1 (HTRA1). We also identified the proteins enriched in the peri-amyloid tissue and compared it with the normal control samples ([Table 5-5](#)). Proteins such as Apolipoprotein E, actin and clusterin were found enriched in the peri-amyloid tissue but the most enriched proteins were found to be elongation factor 1- α 1, apolipoprotein A-IV and serine protease HtrA1 (HTRA1). Together, these data suggest that HtrA1 was present in higher amounts in amyloid fibrils containing mutant TGFBIp and may be involved in the clearance of the amyloid fibrils (Karring, Poulsen et al. 2013).

Table 5-3: Relative abundance of the top 25 proteins identified in amyloid corneal deposits from patients and healthy control donors (ranked by molar percentage calculation). Amyloid-associated proteins are highlighted in blue and non-fibrillar amyloid-associated proteins are highlighted in red. Average protein abundance and standard deviation were calculated from 3 individual mass spectrometry runs.

	Sample 1 amyloid (H626R) mutation	Sample 2 amyloid (H626R) mutation	Sample 3 amyloid (R124C) mutation	Top 25 proteins in control
Protein	mol %	mol %	mol %	mol %
Transforming growth factor-beta-induced protein ig-h3	11.08 ± 2.03	14.91 ± 2.14	32.98 ± 3.94	10.27 ± 2.61
Actin, cytoplasmic 1	0.87 ± 0.25	1.33 ± 0.08	0.68 ± 0.3	
Aldehyde dehydrogenase, dimeric NADP-preferring	1.33 ± 0.78	1.09 ± 0.08	0.56 ± 0.09	
Alpha-enolase	0.70 ± 0.23	1.07 ± 0.20		
Annexin (Fragment)	0.82 ± 0.47	1.42 ± 0.43	0.71 ± 0.41	
Apolipoprotein A-IV	0.98 ± 0.69	1.68 ± 0.30	1.16 ± 0.4	
Apolipoprotein D (Fragment)	1.05 ± 0.05	1.79 ± 0.38		0.78 ± 0.2
Apolipoprotein E	1.04 ± 0.20	3.53 ± 1.17		
Biglycan	1.28 ± 0.34	1.46 ± 0.59	0.8 ± 0.36	2.5 ± 0.51
Clusterin	2.12 ± 0.34	2.90 ± 0.35	1.62 ± 0.57	
Collagen alpha-1(VI) chain	1.20 ± 0.05	0.82 ± 0.13	0.56 ± 0.06	1.51 ± 0.08
Collagen alpha-2(VI) chain	0.95 ± 0.14			1.27 ± 0.12

Collagen alpha-3(VI) chain (molecular mass 323455)	1.76 ± 1.02	1.05 ± 0.61		
Collagen alpha-3(VI) chain (molecular mass 345167)	1.53 ± 0.90	0.75 ± 0.46		1.78 ± 0.24
Decorin	2.03 ± 0.71	0.94 ± 0.17	0.53 ± 0.14	3.21 ± 0.21
Histone 1, H2bn, isoform CRA_b	1.43 ± 0.07	1.19 ± 0.19	1.01 ± 0.07	
Histone H2A type 1	0.88 ± 0.63		0.59 ± 0.34	0.83 ± 0.24
Histone H4	1.54 ± 0.89	2.37 ± 0.45	2.11 ± 0.83	0.98 ± 0.42
Kaliocin-1 (Fragment)	1.06 ± 0.38	0.81 ± 0.07		
Keratocan	3.31 ± 0.24	1.45 ± 0.11	1.8 ± 0.2	1.96 ± 0.39
Lumican	1.99 ± 0.63	1.00 ± 0.37	0.75 ± 0.3	4.85 ± 0.2
Periostin	1.06 ± 0.19			
Pyruvate kinase PKM	0.82 ± 0.54	0.83 ± 0.19	0.56 ± 0.07	
Serum albumin	0.88 ± 0.24	0.92 ± 0.15		
Vimentin	1.04 ± 0.62	2.22 ± 0.54	0.52 ± 0.31	1.39 ± 0.24
Fatty acid-binding protein, epidermal		0.68 ± 0.40		
Lysozyme C		1.36 ± 0.30		
Serum amyloid P-component		1.56 ± 0.25	0.76 ± 0.21	
Truncated apolipoprotein A-I		1.78 ± 0.84	0.55 ± 0.07	
Angiopoietin-related protein 7			0.89 ± 0.13	
Clusterin beta chain (Fragment)			2.32 ± 0.78	
Ig kappa chain C region			1.08 ± 0.33	9.46 ± 1.55

Peroxiredoxin-1			0.81 ± 0.06	
Prolargin			0.5 ± 0.05	0.93 ± 0.19
Ubiquitin (Fragment)			0.9 ± 0.74	
Annexin A2				2.26 ± 0.41
Collagen alpha-1(XII) chain				1.06 ± 0.1
Collagen alpha-1(XII) chain (Fragment)				0.79 ± 0.02
Galectin-1				0.77 ± 0.28
Ig gamma-1 chain C region				1.81 ± 0.18
Ig gamma-4 chain C region				0.8 ± 0.1
Ig lambda-1 chain C regions				1.22 ± 0.34
Ig lambda-7 chain C region				1.17 ± 0.2
MAM domain-containing protein 2				1.09 ± 0.16
Serum albumin				0.75 ± 0.25
Thrombospondin-4				1.15 ± 0.4

Table 5-4: List of proteins enriched in Lattice amyloid deposits relative to healthy control cornea

Protein	Control (mol %)	Patient corneal Deposit (mol %)	Fold increase
Apolipoprotein A-IV	0.03	1.24	32.7
Apolipoprotein E	0.1	1.90	14.9
Serine protease HTRA1	0.04	0.30	7.1
Retinal dehydrogenase 1	0.02	0.13	5.1
Clusterin beta chain (Fragment)	0.49	2.37	4.8
Elongation factor 1-alpha 1	0.02	0.11	4.3
Isoform 2 of Transketolase	0.05	0.18	3.5
Actin, cytoplasmic 1	0.24	0.81	3.3
Ubiquitin (Fragment)	0.23	0.75	3.2
Serum amyloid P-component	0.19	0.62	3.1
Isoform 2 of Clusterin	0.44	1.24	2.7
Fibulin-7	0.03	0.09	2.5
Transforming growth factor-beta-induced protein ig-h3	14.2	33.63	2.3
Histone H4	0.69	1.41	2.0
Pyruvate kinase PKM	0.30	0.61	1.9
Complement component C8 gamma chain	0.06	0.13	1.9
Histone H2B	0.27	0.53	1.9
Tubulin beta-4B chain	0.11	0.21	1.8
Desmin	0.08	0.16	1.8
C-type lectin domain family 11 member A	0.13	0.23	1.7
Angiopoietin-related protein 7	0.32	0.55	1.7
Sushi repeat-containing protein SRPX2	0.04	0.07	1.7
Elongation factor 2	0.01	0.01	1.6
Vitronectin	0.07	0.12	1.6
Heat shock protein beta-1	0.21	0.33	1.6
Alpha-enolase	0.30	0.48	1.5
Peroxiredoxin-1 (Fragment)	0.34	0.52	1.5

5.3.4 Mutation-specific enrichment of TGFBIp peptides in Lattice amyloid deposits

We next used two different label-free quantitation methods to accurately identify the specific proteolytic fragments that were enriched in amyloid fibrils. Using combined spectral counting and AUC analysis to compare amyloid deposits from all three Lattice corneal dystrophy patients with tissues from healthy control donors, we observed significant enrichment of TGFBIp peptides 515-533 (G⁵¹⁵DNRF SMLVAAIQSAGLTETLNR⁵³³) and 571-588 (Y⁵⁷¹HIGDEILVSGGIGALV R⁵⁸⁸) (Figure 5-6). The list of all peptides that were enriched (Table 5-6) in amyloid deposits are TGFBIp 235-257 (V²³⁵ISTITNNIQIIEIEDTFETLR²⁵⁷), 258-295 (A²⁵⁸AVAASGLNTMLEGNGQYTLLAPTNEAFEKIPSETLNR²⁹⁵), 288-295 (I²⁸⁸PSET LNR²⁹⁵), 296-304 (I²⁹⁶LGDPEALR³⁰⁴), 356-377 (D³⁵⁶ILATNGVIHYIDELLIPDSAK³⁷⁷), 378-396 (T³⁷⁸LFELAAESDVSTAIDLFR³⁹⁶), 410-432 (L⁴¹⁰TLLAPLNSVFKDGTPPID AHT R⁴³²), 488-496 (Y⁴⁸⁸GTLFTMDR⁴⁹⁶), 497-510 (V⁴⁹⁷LTPPMGTVMDVLK⁵¹⁰), 497-514 (V⁴⁹⁷LTPPMGTVMDVLKGDNR⁵¹⁴), 515-533 (G⁵¹⁵DNRF SMLVAAIQSAGLTETL N R⁵³³), and 564-570 (E⁵⁶⁴LANILK⁵⁷⁰). There were many peptides identified that were very unique to the patient samples and not found in the control samples. A list of the peptides that were observed only in patient 3 with R124C mutation and were found to be absent in both control and the patient sample with H626R mutation are 43-59 (Q⁴³HGPNVCAVQKVIGTNR⁵⁹), 61-71 (Y⁶¹FTNCKQWYQR⁷¹), 187-216 (H¹⁸⁷GMTLTSMYQNSNIQIHYPNGIVTVNCAR²¹⁶), 314-349 (S³¹⁴AMCAEAIVA GLSVETLEGTTLEVGCSDMLTINGK³⁴⁹) and 558-570 (L⁵⁵⁸LGDAKELANILK⁵⁷⁰). Conversely, a significant number of TGFBIp-derived peptides were decreased in amyloid fibril deposits from patients compared to the normal controls, including 28-38 (S²⁸PYQLVLQHSR³⁸), 77-95 (S⁷⁷TVISYECCPGYEKVPGEK⁹⁵), 305-313 (D³⁰⁵LLNNHILK³¹³), 470-485 (N⁴⁷⁰SLCIENSCIAAHDKR⁴⁸⁵) and 591-602 (S⁵⁹¹LQGDKLEVSLK⁶⁰²). These data suggested that mutant TGFBIp undergoes a specific pattern of proteolytic processing distinct from that which degrades WT TGFBIp.

We next assessed if the mutant TGFBIp peptides could be detected in individual patient samples (3 mutants compared with healthy control donor). A protein sequence database containing the WT TGFBIp and mutant variants (R124C and H626R) was used to search the MS/MS spectra of each sample, the AUC of each identified peptide was calculated by Proteome Discoverer software, and the results were log converted for visualization of peptide distribution. In the patient amyloid deposit with R124C mutation, we were able to identify the mutant peptides TGFBIp-R124C 96-127 and 91-127 with altered cysteine residue (G⁹⁶CPAALPLSNLYETLGVVGSTTTQLYTDCTEK¹²⁷ and V⁹¹PGEKGC PAALPLSNLYETLGVVGSTTTQLYTDCTEK¹²⁷). While we were unable to detect mutant TGFBIp in one patient with H626R mutation, amyloid sampled from a second patient with the same mutation did contain mutant peptides TGFBIp-H626R 603-626 and 611-626 (N⁶⁰³NVVSVNKEPVAEPDIMATNGVVR⁶²⁶ and E⁶¹¹PVAEPDIMATNGVVR⁶²⁶). The differential peptide enrichment in individual patient samples compared with normal controls is shown in **Figure 5-7**. Thus we were able to confirm the specific mutational status in each patient by detecting the presence of the mutant peptides. The results also showed enrichment of a number of tryptic peptides from all the 3 patient samples that were completely absent in the control cornea when the database was searched with the WT (**Table 5-6**).

5.3.5 Enrichment of endogenous peptide cleavage sites in amyloid corneal deposits

We next performed a digestion database search to identify whether semi-tryptic peptides indicative of endogenous cleavage sites were enriched in dystrophic patient tissues. This approach revealed significant enrichment of semi-tryptic peptides with an Asp, Phe, His, Leu, Asn or Tyr cleavage edge in both amyloid corneal deposits and peri-amyloid tissue from patients compared with control cornea (**Figure 5-8**). We also identified a novel endogenous cleavage site at a Trp residue that was present in corneal deposits and surrounding tissues from patients but absent in healthy control samples. Enrichment of semi-tryptic peptides with endogenous cleavage sites in patient samples may indicate active proteolytic processing and attempted clearance of mutant proteins (Karring, Runager et al. 2012). The Increased abundance of endogenous cleavage sites at highly hydrophobic amino acid residues (including Phe, Trp and Tyr) in the amyloid

deposits from patients and peri-amyloid surrounding tissue compared to the normal healthy controls further suggests that mutant TGFBIp is a potential substrate for proteolytic activity by proteases including serine protease HTRA1.

Figure 5-6: Spectral counts of peptides derived from TGFBIp by trypsin digestion. Spectra show relative abundance and indicate high enrichment of 2 peptides (TGFBIp 515-533 and 571-588) in amyloid corneal deposits from patients compared with healthy control cornea. Both of the peptides enriched in dystrophic patients were found to be located in the 4th FAS-1 domain of TGFBIp which is the site of >80% of all reported mutations.

Table 5-5: List of protein enriched in the tissue surrounding the amyloid deposits

Protein	Control (mol %)	Surrounding (mol %)	surrounding/ control
Elongation factor 1-alpha 1	0.03	0.17	6.51
Apolipoprotein A-IV	0.04	0.20	5.20
Serine protease HTRA1	0.04	0.19	4.37
Apolipoprotein E	0.13	0.54	4.19
Actin, cytoplasmic 1	0.24	1.00	4.12
Clusterin beta chain (Fragment)	0.49	1.56	3.18
Sushi repeat-containing protein SRPX2	0.04	0.13	3.03
Histone H2B	0.28	0.71	2.58
Isoform 2 of Transketolase	0.05	0.13	2.47
Ubiquitin (Fragment)	0.23	0.51	2.20
Angiopoietin-related protein 7	0.32	0.70	2.19
Histone H4	0.69	1.48	2.13
Serum amyloid P-component	0.20	0.35	1.75
Transforming growth factor-beta-induced protein ig-h3	14.24	22.11	1.55
Alpha-enolase	0.30	0.46	1.51
Desmin	0.09	0.13	1.51
Isoform 2 of Clusterin	0.45	0.62	1.37
Pyruvate kinase PKM	0.31	0.36	1.16
Vitronectin	0.08	0.08	1.04
Peroxiredoxin-1 (Fragment)	0.35	0.25	0.72
Retinal dehydrogenase 1	0.03	-	-
Fibulin-7	0.04	-	-
Complement component C8 gamma chain	0.07	-	-
Tubulin beta-4B chain	0.11	-	-
C-type lectin domain family 11 member A	0.13	-	-
Elongation factor 2	0.01	-	-
Heat shock protein beta-1	0.21	-	-

Table 5-6: List of peptides enriched in the amyloid patient samples compared to the control samples

List of peptides enriched in the amyloid deposits compared to the control	
TGFB1p 235-257	VISTITNNIQQIIIEDTFETLR
TGFB1p 258-295	AASGLNTMLEGNGQYTLLAPTNEAF EKIPSETLNR
TGFB1p 288-295	IPSETLNR
TGFB1p 296-304	ILGDPEALR
TGFB1p 356-377	DILATNGVIHYIDELLIPDSAK
TGFB1p 378-396	TLFELAAESDVST AIDLFR
TGFB1p 410-432	LTLLAPLNSVFKDGTTPIDA HTR
TGFB1p 488-496	YGTLFTMDR
TGFB1p 497-510	VLTPPMGTVM DVLK
TGFB1p 497-514	VLTPPMGTVM DVLKGDNR
TGFB1p 515-533	FSMLVAAIQSAGLTETLNR
TGFB1p 564-570	ELANILK
List of peptides that are present only in the amyloid deposits but absent in controls	
TGFB1p 296-313	ILGDPEALRDLLNNHILK
TGFB1p 350-377	AIISNKDILATNGVIHYIDELLIPDSAK
TGFB1p 497-533	VLTPPMGTVM DVLKGDNRFSMLVAAIQSAGLTETLNR
TGFB1p 511-533	GDNRFSMLVAAIQSAGLTETLNR
TGFB1p 571-588	YHIGDEILVSGGIGALVR
TGFB1p 603-642	NNVSVNKEPVAEPDIMATNGVVHVITNVLQPPANRPQER
TGFB1p 611-642	EPVAEPDIMATNGVVHVITNVLQPPANRPQER
Peptides that are specific to R124C mutation	
TGFB1p 43-59	QHGPNVCAVQKVIGTNR
TGFB1p 61-71	YFTNCKQWYQR
TGFB1p 187-216	HGMTLTSMYQNSNIQIHHYPNGIVTVNCAR
TGFB1p 314-349	SAMCAEAIVAGLSVETLEGTTLEVGCSDMLTINGK
TGFB1p 558-570	LLGDAKELANILK

Figure 5-7: Normalized logarithmic area of peptide enrichment in individual patient samples. Purple dots indicate patient 1 (H626R mutation), black dots indicate patient 2 (H626R mutation) and blue dots indicate patient 3 (R124C mutation). The search considered WT protein (A), R124C mutant (B), and H626R mutant (C). Diamond symbol indicates mutant peptide detection.

Figure 5-8: Analysis of endogenous cleavage sites in corneal deposits, adjacent unaffected tissues, and healthy control cornea. This analysis revealed increased spectral counts for amino acids Asp, His, Leu, Phe, Thr and Tyr in patient samples (both deposits and adjacent tissues) indicating enrichment of endogenous cleavage sites compared with control cornea. These data suggest that mutant TGFBIp exhibits increased susceptibility to proteolysis relative to the WT protein.

5.3.6 Change in conformation and secondary structure of synthetic TGFBIp peptides

Having identified the most abundant peptides that could possibly form the core region of corneal amyloid fibrils, we next proceeded to investigate change in secondary structure and amyloid-forming tendency of 2 representative peptides (TGFBIp 515-533 and 571-587) by circular dichroism analysis. The two peptides were chosen for the study because they were the top two peptides that were highly enriched in the patient corneal amyloid deposits compared to the control cornea and were from the 4th FAS-1 domain which is a mutational hotspot. These two peptides were also predicted to have a higher aggregation propensity by various prediction algorithms. In circular dichroism spectra of β -structured amyloid fibrils, the amplitude of the positive peak at ~195-200nm was increased to levels far higher than are detected in the native β -structured protein(Matsumura, Uemura et al. 2004). Here we observed that under physiological

pH, both representative peptides (TGFBIP 515-533 and 571-587) displayed substantial differences in secondary structure; peptide 515-533 displayed negative minimum at ~205nm and a broad peak in the n- Π^* region, but after 4 week period the negative minimum was instead detected at ~218nm, indicating a clear time-dependent conformational transition to β -sheet structure (**Figure 5-9 A and B**). For TGFBIP peptide 515-533, we observed a broad negative minimum between 210-220nm, potentially indicating a mixed fibrillar content. For TGFBIP 571-587, the intensity of the broad negative peak at ~220-230nm progressively increased over time and a positive maximum peak was detected at ~198-218nm after 1 week. Taken together, these data indicate that mutant peptides TGFBIP 515-533 and 571-587 have high aggregation tendency under physiological conditions and readily form ordered β -sheet amyloid fibrils, as confirmed by transmission electron microscopy (**Figure 5-11**). While TGFBIP peptides 515-533 and 571-587 both clearly formed ordered β -sheet secondary structured amyloid fibrils within 4 weeks, longitudinal analysis by DichroWeb(Lobley, Whitmore et al. 2002) failed to shed further light on changes in secondary structure content over time

5.3.7 Favourable aggregation kinetics of T peptides enriched in LCD amyloid deposits

Peptide aggregation kinetics were assessed over a period of 17 hours by recording fluorescence emission of ThT dye, which increased with progressive binding into amyloid fibrils. **Figure 5-10A and 5-10B** showed average time-dependent changes in ThT fluorescence for each of the model TGFBIP peptides. This analysis revealed an immediate and progressive increase in fluorescence intensity without an initial lag phase, indicating that the peptides may form amyloid fibrils even during the early stages of incubation. For TGFBIP 515-533, the fluorescent dye time response was readily fit to a single exponential function. The corresponding kinetic parameters are summarized in **Table 5-6**. For TGFBIP 571-588, fluorescent dye binding was better fit to a two-phase exponential function. These data suggested that TGFBIP 571-588 forms rapid pre-nucleation clusters with K_{agg} of 50.4 sec⁻¹, contributing approximately 20-30% of total kinetics, with a subsequent slow-growth phase accounting for the remaining 70% of overall kinetics (K_{agg} 8.1 sec⁻¹). These results indicate that the TGFBIP 515-533 and 571-587 peptides identified in the core regions of corneal amyloid deposits exhibit

distinct patterns of aggregation and thus follow different pathways of amyloid formation.

Figure 5-9: Circular dichroism spectra showing TGFBIp aggregation and amyloid fibril formation over time by mutant 515-533 (A) or 571-588 (B). The peptides were incubated in PBS at 37°C for up to 4 weeks and were able to form amyloid fibrils as early as day 4.

Figure 5-10: Aggregation kinetics of TGFBIp peptides 515-533 and 571-588 as assessed by incorporation of fluorescent ThT into developing fibrils. Fluorescence intensity is expressed as % of maximum. Both peptides displayed increased tendency to form amyloid fibrils without any initial lag phase. Data shown are average values from two experiments performed in triplicates.

Table 5-6: Aggregation kinetics of TGFBIp peptides 511-533 and 571-588

Sample	Kinetic Parameters		
	k_1, sec^{-1}	$t_{50\%}, \text{sec}$	R^2
Repeat 1	26.9×10^{-5}	2575	0.822
Repeat 2	17.2×10^{-5}	4034	0.884
Repeat 3	14.1×10^{-5}	4930	0.820
Repeat 4	4.7×10^{-5}	14852	0.905
Repeat 5	6.9×10^{-5}	10006	0.961
Repeat 6	8.5×10^{-5}	8196	0.977
Mean	$13.1 \pm 8.2 \times 10^{-5}$	7432 ± 4553	

TGFBIp 571-588 aggregation kinetics

Sample	Kinetic Parameters		
	k_1, sec^{-1}	k_2, sec^{-1}	R^2
Repeat 1	43.9×10^{-5}	2.27×10^{-6}	0.997
Repeat 2	46.1×10^{-5}	6.4×10^{-6}	0.997
Repeat 3	43.0×10^{-5}	10.3×10^{-6}	0.997
Repeat 4	58.6×10^{-5}	6.4×10^{-6}	0.998
Repeat 5	62.19×10^{-5}	12×10^{-6}	0.998
Repeat 6	48.8×10^{-5}	10.91×10^{-6}	0.997
Mean	$50.4 \pm 8.1 \times 10^{-5}$	$8.1 \pm 3.7 \times 10^{-6}$	

A Peptide 515-533

B Peptide 571-588

Figure 5-11: Transmission electron microscopy images of amyloid fibrils derived from TGF β Ip peptides 511-533 and 571-588 after 5 weeks incubation. Images were acquired with a JEOL JEM-1010 transmission electron microscope using Digital Micrograph™ 1.81.78 for GMS 1.8.0 (50,000X magnification) at 80kV. Both peptides displayed increased tendency to form amyloid fibrils showing that the native protein sequence itself is highly amyloidogenic.

5.4 Discussion

In the current report, we assessed the molecular composition of amyloid deposits from the cornea of Lattice corneal dystrophy patients, identified a likely role for serine-protease HTRA1 in cleaving mutant TGFBIp, and confirmed the rapid aggregation kinetics of the peptide fragments generated. While all proteins and polypeptide chains are likely to form aggregates under certain conditions (Bemporad, Calloni et al. 2006), in amyloid-related disorders, this process is typically enhanced by specific mutations in protein-coding genes. Often these mutations result in amino acid substitutions that alter protein charge, hydrophobicity or tendency to form β -sheet secondary structures (Chiti, Taddei et al. 2002). Indeed, in *TGFBI*-linked corneal dystrophy, the disease-associated mutations encode proteins that are either thermodynamically unstable (Køldso, Andersen et al. 2015) (A546T) or overly stable (R555W [PDB 2LTC] and G623D) relative to the native protein (Runager, Basaiawmoit et al. 2011, Underhaug, Køldso et al. 2013, Murugan, Venkatraman et al. 2016, Stenvang, Christiansen et al. 2016). Where unstable mutant proteins are produced, excessive proteolysis may generate highly amyloidogenic peptide fragments. Conversely, enhanced protein stability may instead prevent effective protein clearance as required to maintain transparency of the cornea (Underhaug, Køldso et al. 2013). We identified that the proteolytic products of two TGFBIp mutants which are prevalent in Asian populations, with high aggregation propensity.

In our analysis of protein deposits from the patient with R124C mutation, we observed a higher abundance of apolipoprotein A-IV, D, E, and A-I, as well as clusterin, decorin, biglycan, lysozyme, thrombospondin-4 and kaliocin which are linked to systemic amyloidosis. We also detected the presence of mutant TGFBIp protein in amyloid deposits from patients with known disease-associated mutations H626R and R124C. Our data are consistent with a previous study on the composition of amyloid deposits from R124C-linked lattice dystrophy (Courtney, Toftgaard Poulsen et al. 2015). However, our analysis further revealed that the corneal deposits contained other proteins including angiopoietin related protein, clusterin beta, Ig Kappa, peroxiredoxin, prolargin and ubiquitin are a specific feature of R124C-linked amyloid deposit compared to the healthy cornea. Apolipoprotein A-IV enrichment has previously been linked with systemic amyloidosis and shown to accumulate in

TGFBIp-associated lattice corneal dystrophies as well (Bergström, Murphy et al. 2001, Bergström, Murphy et al. 2004). These non-fibrillar proteins interact with amyloid fibrils formed by other proteins and co-assemble to form longer amyloid fibrils (Coria, Prelli et al. 1988, Namba, Tomonaga et al. 1991, Gallo, Wisniewski et al. 1994, MacRaid, Stewart et al. 2004, Yerbury, Poon et al. 2007). Our mass spectrometry data also indicated that Lattice amyloid deposits from all three patients were enriched in proteoglycans implicated in amyloid fibril formation in human renal glomeruli and corneal tissues, namely decorin, biglycan and collagen type 1 (Moss, Shore et al. 1998, Stokes, Holler et al. 2000, Baldwin, Knowles et al. 2011, Basaiawmoit, Oliveira et al. 2011).

All three amyloid corneal deposit samples from patients analysed here were found to be enriched in actin, pyruvate kinase and histone H4 which have not been previously identified as components of amyloid deposits in the cornea. These proteins are reported to be involved in other protein misfolding disorders like Alzheimer's disease (Penzes and Vanleeuwen 2011, Newington, Rappon et al. 2012, Wang, Takata et al. 2012, Plagg, Ehrlich et al. 2015, Roque, Sortino et al. 2015, Hwang, Shim et al. 2016). The proteins may form amyloid, or incorporate themselves to the amyloid seeds, assist in the recruitment of key components into fibrils or perhaps promote fibrillation of other amyloid-associated proteins (Karring, Runager et al. 2012).

Some of the proteins listed like aldehyde dehydrogenase and Apolipoproteins were found to be co-localized with the amyloid deposits and have been reported before (Bergstrom et al. 2001 and Bergstrom et al. 2004). The exact role of these proteins and their involvement in amyloid fibrillation remain unknown. The proteins like aldehyde dehydrogenase and α -enolase are associated with oxidative stress pathway found in Alzheimer's disease (Murakami et al. 2009 and Castegna et al. 2002). The enzyme peroxiredoxin-1 was also found in the patient sample which is involved in oxidative stress.

The protein Clusterin is also reported to be involved in the early stages of amyloid fibril formation (Desikan et al. 2014). Clusterin is also reported to be associated with the oxidative stress in amyloid plaques in Alzheimer's disease (Yerbury et al. 2007). Since cornea is directly exposed to the UV rays from the sun, there might be an alternative oxidative stress pathway that may trigger the amyloid fibril formation. Collagen VI is reported to protect neurons and also prevents oligomerization of A β binding in Alzheimer's disease (Cheng et al. 2009). Annexin A1 is known to be anti-

inflammatory and reported to aid in the clearance of amyloid- β Peptide (Ries M et al. 2016). There is also a possibility of some proteins to have been recruited by the cornea to aid in the clearance of aggregated proteins and prevent further oligomerization of mutant peptides/proteins.

Our finding that serine protease HtrA1 was enriched in LCD deposits resembles previous reports that this enzyme co-localizes with β -amyloid deposits in the human brain in Alzheimer's disease (Grau, Baldi et al. 2005). Previous studies have also reported that stromal amyloid deposits from LCD 1 patients are enriched in serine protease HtrA1, whereas the enzyme is absent from bowman's deposits and amorphous GCD1 aggregates (Poulsen, Runager et al. 2014, Courtney, Toftgaard Poulsen et al. 2015). The serine protease (HTRA) family processes a wide range of different proteins and plays a critical role in regulation of protein folding (Clausen, Southan et al. 2002, Ehrmann and Clausen 2004), (Spiess, Beil et al. 1999, Krojer, Garrido-Franco et al. 2002, Walsh, Alba et al. 2003, Wilken, Kitzing et al. 2004, Grau, Baldi et al. 2005). The absence of HTRA1 in corneal deposits from patients with granular dystrophy (*TGFBI*-R555W) suggests that this protease may be involved in processing of mutant protein in lattice corneal dystrophy, thereby promoting the formation of amyloid deposits (David G. Courtney and Enghild 2015). In our study we detected that HTRA1 was increased by ~7-fold in the amyloid deposits compared with normal healthy control cornea, suggesting a major role for this enzyme in the pathogenesis and/or disease progression.

Spectral counts and AUCs of the *TGFBI*p-derived peptides identified in amyloid deposits and normal cornea may reflect patterns of proteolytic processing in the human eye. In the majority of amyloid-related disorders, proteolysis of precursor proteins generates small fragments that enhance fibril formation (Korvatska 2000). It has been reported that the cornea of patients with R124H mutations exhibit accumulation of 68 kDa protein fragments (Korvatska 2000), whereas R124C mutant cornea instead displayed accumulation of 44 and 26 kDa fragments (Korvatska 2000). These data support the concept that mutant *TGFBI*p is processed in a mutation-specific manner that differs from proteolysis of the WT protein. Indeed, we observed that each of the dystrophic patient samples was enriched in mutation-specific *TGFBI*p-derived peptides (I²⁹⁶LGDPEALRDLLNNHILK³¹³, L⁴¹⁰TLLAPLNSVFKDGTTPPIDAHTR⁴³², V⁴⁹⁷LTPPMGTVMDVLKGDNRFSMLVAAIQSAGLTETLNR⁵³³,

F⁵¹⁵SMLVAAIQSAGLTETLN^{R533}, Y⁵⁷¹HIGDEILVSGGIGALVR⁵⁸⁸) that were absent from normal healthy control cornea. The sequences of these amyloid patient-specific peptides corresponded to aggregation hotspots predicted by various computational algorithms **Figure 5-12** Previous study by Yuan et al., indicated that the peptide **G⁵¹⁵DNRF SMLVAAIQSAGLTETLN⁵³²** readily forms amyloid fibrils in vitro (Yuan, Berscheid et al. 2007). We also detected an increased spectral count for several other peptides from the 4th FAS-1 domain when comparing dystrophic patient cornea with normal healthy controls, suggesting that the 4th FAS-1 domain is likely to play a major role in the pathology of protein aggregation related to corneal dystrophy. Our observation that peptide **Y⁵⁷¹HIGDEILVSGGIGALVR⁵⁸⁸** was enriched in LCD amyloid deposits from R124C corroborates with previous study by Sorensen et al.(Sørensen, Runager et al. 2015) which indicated that TGFBIp 571-588 forms the core of amyloid fibrils in A546T-linked lattice corneal dystrophy. When considered together with the results of our ThT and TEM analyses, these data strongly indicate that the **Y⁵⁷¹HIGDEILVSGGIGALVR⁵⁸⁸** peptide exhibits potent amyloidogenic properties.

Our analysis of ThT aggregation kinetics suggested that both TGFBIp 515-533 and 571-588 peptides readily aggregate under physiological conditions. Kinetic analysis suggested that the peptide fragment TGFBIp 571-588 followed a biphasic aggregation profile, indicating continuous addition of monomers or oligomers to the fibril assembly. In contrast, TGFBIp 515-533 displayed monophasic aggregation kinetics, consistent with a high amyloidogenic tendency of this peptide fragment. Morphological analyses of the amyloid fibrils generated by these two peptides after 4 weeks incubation revealed that TGFBIp 571-588 formed fibril stacks that are denser than those generated by TGFBIp 515-533. The highly amyloidogenic properties of the peptides suggested that each of these peptides could possibly function as amyloid fibrillar seeds that can trigger the aggregation of other amyloidogenic proteins including un-cleaved mutant TGFBIp.

We observed that serine protease HTRA1 was associated with LCD fibrils that were also enriched in proteolytic cleavage products of TGFBIp. We therefore propose that HTRA1 could be the primary mediator of amyloidogenic peptide generation in dystrophic patients. While the exact HTRA1 cleavage sites in TGFBIp are currently unknown, the potential target sites of other serine proteases such as elastase-2, cathepsin G and glutamyl peptidase can be predicted by Protease Specificity Prediction

502 PMGTVM DVLK GDNRF SMLVA AIQSAGLTET LNREGVY TVF APTNEAFRAL PPRER SRLLG DAKELANILK
YHIGDEILVS GGIGALVRLK SLQGD KLEVS LKNNVSVNK **EPVAEPDIMA TNGVVHVIT N VLQ** 633

Figure 5-12: Aggregation propensity of 4th FAS1 human TGFBIp using empirical (A) Aggrescan and (B) TANGO and structure based (C) Waltz and (D) Fold amyloid algorithms. The red coloured region in A, C and D and blue coloured regions in B show the hotspots of aggregation in the 4th Fas-1 domain. The protein sequence of the 4th Fas-1 domain is shown and the amino acids highlighted in blue is shown to be the highly aggregation prone regions as predicted by various algorithms.

Server (PROSPER) [<https://prosper.erc.monash.edu.au>]. This analysis predicted multiple likely cleavage sites in the 4th Fas-1 domain of TGFBIp (after amino acids V519, A525, E534, L559, Y571, V579, V606, M619, T621, V625 and T629). The majority of these potential cleavage sites are located near the borders of highly amyloidogenic regions as predicted by various prediction algorithms and tools (Song, Tan H Fau - Perry et al.), suggesting that thermodynamic destabilization of this domain due to a mutation could expose HTRA1 target residues for excessive proteolysis, thereby generating peptides with high amyloidogenic propensity.

The exact mechanism by which HTRA1 processes TGFBIp will be the subject of follow-up studies using mass spectrometry-based assays (Poepsel, Sprengel et al. 2015). A possible strategy for counteracting the effect of destabilizing mutations on TGFBIp integrity could be the development of low/medium-affinity small molecular ligands that bind to and increase overall thermodynamic stability of the target protein

in biological tissues and cells(Roca, Althaus et al. 2013). An alternative approach might involve partial inhibition of the serine protease HTRA1 via topical application of small molecular ligands with high specificity and low cellular toxicity.

We have reported the protein constitution of the corneal amyloid deposits from patients of Asian origin with the most commonly prevalent mutations H626R and R124C. Moreover, the study is limited by the small amount of tissue extracted by laser capture microdissection and its sensitivity to detect low abundance proteins. Notwithstanding these limitations, we still obtained adequate TGFB1p coverage and identified peptides containing the expected mutation sites in the samples of Lattice amyloid deposits. Furthermore, two of the peptides derived from mutant TGFB1p were found to be enriched in corneal amyloid deposits which were confirmed to display increased tendency to aggregate and form amyloid fibrils in biochemical assays.

Chapter 4

Graphical abstract of characterization of peptide fragments derived from 4th FAS-1 domain of TGFBIp

Decrease in overall net charge,
high β -sheet propensity and
nature of amino acid substitution

Peptides derived from 4th FAS-1 domain of TGF β Ip

Increased propensity to form amyloid fibrils

6 Bio-physical and structural characterization of peptide fragments derived from 4th FAS-1 domain of TGFBIp

6.1 Introduction

As mentioned in the previous chapters, there are more than 65 mutations reported in the *TGFBI* gene. The nature and the type of amino acid substitution is directly related to when the disease presents itself in the patients, the type of protein deposits and the layer of cornea in which the aggregated proteins are deposited. The eye is known for its sophisticated mechanism of clearing aggregated proteins and providing a clear and transparent cornea and protecting the eye against microbes. Despite the clearance mechanism which is robust to provide a clear cornea, the protein deposits from the mutant TGFBIp remains aggregated in the cornea. There may be alterations in the protein clearance mechanism and other contributing factors that are compromised in the case of corneal dystrophy patients and mechanism that trigger protein aggregation in patients is still under research and is worth investing efforts to understand the processing of the mutant protein.

Most of the *TGFBI*-corneal dystrophies are associated with the Bowman's membrane and the corneal stroma. It has been shown that the processing of the mutant protein is different from the native WT protein. It has also been postulated that mutations in the *TGFBI* gene vary the overall protein turnover rate in the human cornea (Korvatska 2000). The proteolytic processing and the process of clearing aggregated proteins are reported to be different in the cornea compared to other tissues in the body. Alterations in proteolytic degradation of the mutant proteins also results in retention of C-terminal fragments of TGFBIp protein in the cornea (Karring, Thogersen et al. 2004) and likely further promotes amyloidosis in the affected eye. There are few reports on the proteolytic processing of the mutant proteins in the corneal dystrophy patients. According the reports, the mutant TGFBIp is processed into many shorter fragments which may be the seeds of aggregation of other larger aggregation prone proteins.

Mutations in TGFBIp are hypothesised to induce changes in the secondary structure at the mutation sites which may cause exposure of hydrophobic amino acids

and oligomerization of proteins. Dobson and Chiti et al, have proposed an empirical formula, based on differences in the physiochemical properties such as hydrophobicity, secondary structure propensity and net charge between wild type (WT) and mutant peptides that may be correlated to rates of amyloid formation. In their study, they demonstrated that mutant peptides with a decrease of charge, increase in hydrophobicity, and increased propensity to β -sheet structures, accelerated the rate of aggregation compared to the wild type peptide. Apart from the intrinsic properties (hydrophobicity and overall net charge) and extrinsic factors (pH, temperature, salts etc.), the propensity to form fibrillar aggregates also depends upon the size of the protein (Dobson 2003, Dobson 2006). It has been postulated that larger proteins have a greater propensity to form toxic amorphous aggregates, but have a lower propensity to align themselves and form ordered amyloid fibrils (Ramshini, Parrini et al. 2011). The TGFBIp protein is known to be subjected to proteolytic activity by proteases e.g. HtrA1.

TGFBIp and its mutants are known to form a number of proteolytic fragments in the cornea (Karring, Runager et al. 2012, Karring, Poulsen et al. 2013, Courtney, Toftgaard Poulsen et al. 2015). It has been suggested that the proteolytic products with high aggregation propensity may trigger an aggregation cascade (Karring, Runager et al. 2012). Studies have reported an increased aggregation propensity of mutant peptide fragments derived from the 1st and 4th FAS-1 domain of TGFBIp compared to the wild-type protein (Schmitt-Bernard, Chavanieu et al. 2002, Yuan, Berscheid et al. 2007). In the studies reported by our group (Lakshminarayanan, Vithana et al. 2011, Elavazhagan, Lakshminarayanan et al. 2012, Murugan, Venkatraman et al. 2016), we have expressed the 4th FAS-1 domain of TGFBIp and identified aggregation prone regions and the ability of the mutant 4th FAS-1 domains to form β -oligomers.

Recently, we analysed the protein profiles of corneal amyloid deposits derived from dystrophic patients and compared with the normal control cornea. Tryptic digestion of the deposits showed enrichment of several TGFBIp peptide fragments, predominantly derived from the 4th FAS-1 domain when compared to the control cornea. This may be due to the altered proteolytic processing between the WT and mutant TGFBIp. One of the peptides enriched in the patient deposit but absent in control cornea was TGFBIp (E⁶¹¹PVAEPDIMATNGVVHVITNVLQPPANRPQER⁶⁴²).

We have previously shown that, a 23 residue peptide within this region (E⁶¹¹PVAEPDIMATNGVVHVITNVLQ⁶³³), readily formed amyloid fibrils in vitro

under physiological conditions (Lakshminarayanan, Vithana et al. 2011). Interestingly, mutations within this region are associated with amyloidogenic phenotypes in the patients (Stewart, Black et al. 1999, Munier, Frueh et al. 2002, Aldave 2008, Gruenauer-Kloevekor, Braeutigam et al. 2009) and majority of the mutations are associated with alteration of overall net charge of the sequence. There are about 11 different mutation reported in patients along the amino acid residues of TGFBIp 611 to 632 and with the altered net charge. The higher tendency of these peptides to form amyloid fibrils is interesting and clinically relevant and hence we decided to explore more on the bio-physical properties of the peptides derived from the 4th FAS-1 domain of TGFBIp.

The main aim of the study was to understand the molecular-functional relationship of TGFBIp derived mutant peptides where the amino acid substitutions decreased the overall net charge and their effects on the propensity of the peptides to form amyloid fibrils. We modified the 23 residue peptide (TGFBIp E611PVAEPDIMATNGVVHVITNVLQ633) with clinically relevant cationic charged amino acid substitution and investigated the amyloid forming properties of the mutant peptides. The list of clinically relevant mutations are listed in [\(Table 6-1\)](#). We investigated the effect of decreased charge on secondary structure, amyloid forming propensities and cytotoxicity on corneal fibroblasts. We also performed solid state NMR (ssNMR) experiments on the amyloid fibrils derived from wild type and the most stable mutant peptide. Finally, the cytotoxicity of peptides in both monomeric and amyloidogenic forms were investigated by label-free xCELLigence cell impedance measurements. The study provides an insight into the effects of substitution of individual amino acid on the rate of amyloid formation in the peptides, the structure, biological and biophysical properties of the amyloid fibrils formed by the peptides. We also explored how a position specific and nature of amino acid substitution hugely contributes to the amyloidogenic property of the peptides.

Table 6-1: Change in overall net charge in the clinically relevant mutations and their phenotypes.

S.No	Mutation	Change in overall net charge	Phenotype
1	p.Met619Lys	+1	Amyloid
2	p.Ala620Asp	-1	Amyloid
3	p.Asn622His	+1	Amyloid
4	p.Asn622Lys	+1	Amyloid
5	p.Gly623Arg	+1	Amyloid
6	p.Gly623Asp	-1	Amyloid
7	p.Val624Met	0	Amyloid
8	p.Val625Asp	-1	Amyloid
9	p.His626Pro	0	Amyloid
10	p.His626Arg	+1 (at pH 7.0)	Amyloid
11	p.Val631Asp	-1	Amyloid

6.2 Materials and Methods

6.2.1 Materials

The TGFBIp 611-633 WT and mutant synthetic peptides were synthesized from Synpeptides (Synpeptide Co Ltd, Shanghai, China). The purity of peptides was greater than 95% when synthesised. Formvar-carbon coated nickel grids were bought from EMS (Electron Microscopy Sciences, PA). Chemicals were bought from Sigma-Aldrich (Sigma-Aldrich Inc., MO). Acetonitrile (ACN), formic acid, water for Mass Spectrometric studies were obtained from Sigma-Aldrich (St. Louis, MO, USA). Ultra Micro Spin column was bought from The Nest Group (Southborough, MA, USA).

6.2.2 Spectral counts of tryptic peptides from TGFBIp using LC MS/MS

Corneal sections were collected from 3 corneal dystrophy patients with amyloid Lattice corneal dystrophy (2 patients with H626R mutation and 1 patient with R124C mutation in *TGFBI* gene) who were undergoing corneal transplantation at the Singapore National Eye Centre. Written informed consent was obtained from all patients prior to

surgery. Ethical approval for the collection of patient corneas was granted by the Singhealth Institutional Review Board. Control tissues (n=3), transplant-grade corneal samples from deceased individuals were obtained from Lion's Eye Institute, Tampa, USA. The patient corneal sections were stained with Congo red to locate the amyloid fibril deposits. Positive Congo red sections were collected using Laser Capture Microdissection (LCM) and proteins were extracted from the LCM cap. 3 normal control corneas without epithelium were dissected into smaller pieces using a sharp scalpel and they were processed to extract corneal stromal proteins. The proteins from patients and controls were digested overnight with trypsin and used for Mass Spectrometry analysis.

LC-MS/MS analysis was performed using a Q Exactive mass spectrometer coupled with online Dionex Ultimate 3000 RSLC nanoLC system (Thermo Scientific; MA). Samples were run on a Dionex EASY-spray column (PepMap® C18, 3µm, 100Å) using an 120min gradient of mobile phase A (0.1% FA in 5% ACN) and mobile phase B (0.1% FA in ACN) with a flow rate of 300nl/min. Separated peptides were passed through an EASY nanospray source with a source voltage of 1.5kV. A full MS scan was set at 350–1600m/z with resolution of 70,000 at m/z 200. The maximum ion accumulation time and dynamic exclusion time were set as 100ms and 30s respectively. The 10 most intense ions detected above a threshold of 1000 counts were selected for higher-energy collisional dissociation (HCD) fragmentation using 28% normalized collision energy with a maximum ion accumulation time of 120ms. Other parameters used were Automatic Gain Control 1e + 06 for the full MS scan and 2e + 05 for the MS2 scan, isolation width of 2Da for the MS2 scan, and exclusion of single and unassigned charged ions from the MS/MS scan (with underfill ratio set to 0.1%). Data acquisition was performed using Xcaliber 2.2 software (Thermo Scientific; USA).

Data analysis

Raw data were processed using two different approaches to achieve high-confidence of label-free quantitation. Firstly, all raw data were converted into mascot generic format (MGF) using Proteome Discoverer 1.4.1.14 (Thermo Fisher, MA) and searches were performed using an in-house search engine (Mascot, version 2.4.1; Matrix Science, London, UK). The UniProt Knowledgebase (UniProtKB) of human proteins (downloaded on July 25, 2016, including 70,849 sequences and 23,964,784

residues) was used as the search database. The following settings were applied; static modification = carbamidomethyl at cysteine; dynamic modifications = methionine oxidation, asparagine/glutamine deamidation; digestion parameters = full trypsin with maximum 2 missed cleavages, semi-trypsin with one non-specific cleavage; search parameters = #13C is 2, precursor mass is 10ppm, fragment mass tolerance is 0.02Da. Selection of proteins for final analysis was performed using a target-decoy search strategy with false discovery rate (FDR) $\leq 1\%$ considering only those proteins identified with multiple peptides. An additional search was performed using semi-tryptic digestion as the enzyme parameter in order to identify non-tryptic peptide cleavage sites. The emPAI value reported by Mascot and spectral counting were used to perform quantification of proteins and peptides, respectively.

The raw data files were also analysed using the Proteome Discoverer 1.4 software package. MS/MS spectra were searched against the same UniProt human database using the Mascot search engine with identical parameters to those described above. The peptides-to-spectrum matches were filtered by application of $<1\%$ FDR threshold using Percolator. Only peptides identified with high confidence as assigned by the software were selected for further analysis. Precursor ion area detection node was enabled in order to obtain the area under the curve (AUC, or extracted ion chromatogram / XIC) of the LC elution profiles of the detected peptides. The AUCs of the same peptides obtained from both patient and control samples were used for quantitative analysis to determine the fold-changes and extent of enrichment in patient tissues. Peptide/protein lists were exported to Excel and processed using in-house analytical scripts. All statistical calculations were performed using three technical replicates for each sample.

6.2.3 Circular Dichroism Spectroscopy Secondary structure of peptides

Far UV-CD spectra of the peptides were collected using a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD) and Chirascan-plus (Applied Photophysics, UK) using a quartz cuvette with a path length of 0.1 cm (Hellma, Müllheim, Germany/ Starna). The TGFBIp 611-633 WT and mutant synthetic peptides were reconstituted in PBS (pH 7.2) to a final concentration of 0.6 mg/ml and their CD spectra were recorded.

For time-dependent conformational transition of the model peptides, the peptides were shaken at 180 rpm at 37°C up to 5 weeks in a shaking incubator. The spectra were recorded from 260 nm to 190 nm with a step-size of 0.1 nm, at a scan rate of 50 nm/min for every 1°C. Initial and final spectra before and after heating and after cooling back to 20°C were recorded. The final spectrum was the average of three scans. The CD data were expressed as mean residual weight ellipticity (deg cm² dmol⁻¹). The mean residual weight (MRW) ellipticity ([θ]MRW) was estimated using the equation:

$$[\theta]MRW = [\theta]\lambda \times MRW/10 \times l \times c$$

where [θ]λ is the observed ellipticity, MRW is mean residual weight and defined as M/N-1 where M is the molecular mass and N is number of amino acid residues, l is the path length of the cuvette and c is concentration in mg/ml. The data obtained were plotted using Origin 8.0 software.

To infer if the amyloid fibrils formed by the TGFBIp mutant peptides possess similar or varying thermal properties, we monitored the melting and reassembly of the preformed amyloid fibrils by variable temperature CD spectropolarimetry. To achieve the homogeneous amyloid fibrils preparations, the peptides were incubated in PBS buffer at 37 °C for 5 weeks (10 weeks for P6) until only amyloid fibrils were observed in the TEM. For thermal denaturation experiments, 0.6 mg/ml of the peptide fibrils in PBS (pH 7.2) were heated using the inbuilt Peltier heating system (Applied Photophysics, UK). The CD spectra were recorded as described earlier at various time intervals. Variable temperature scanning was done by heating the sample from 20°C to 90°C at a rate of 1°C/min. For studying the effects of ramping rates on the P5 peptide fibrils, the sample was analysed by heating the fibrils to 90 °C at various heating rates (0.5 °C/min, 1 °C/min and 2 °C/min).

6.2.4 Thioflavin T assay

The TGFBIp⁶¹¹⁻⁶³³ WT and mutant synthetic peptides (0.6 mg/ml) were shaken at 180 rpm in a 37 °C incubator and samples taken at various time points up to 36 h and were treated with 30 µM ThT in PBS buffer (pH 7.2) in a Greiner 96-well flat bottom polystyrol microplate (Greiner, Frickenhausen Germany). The samples were excited at 445 nm and the resulting emission fluorescence at 485 nm was measured using a microplate reader (Tecan infinite M200 pro, Zanker Road, San Jose, USA).

6.2.5 Transmission Electron Microscopy

Transmission electron microscopy images of the peptide fibrils were acquired with a JEOL JEM-1010 transmission electron microscope using Digital Micrograph™ 1.81.78 for GMS 1.8.0 (Gatan, Pleasanton, CA) at the National University of Singapore Electron Microscopy facility. Aliquots of amyloid fibrils were obtained by shaking 0.6 mg/ml of each of the peptide at 37° C in a shaking incubator. About 500 µL of amyloid fibrils samples were withdrawn at different time points, centrifuged and 10µl of amyloid fibril samples were applied onto Formvar-carbon coated 300-mesh-size nickel grids with bacitracin as the binding agent. Excess samples were blotted off the filter paper and negatively stained with 10% phosphotungstic acid (PTA) and incubated for 5 -10 min and washed with water. The grids were dried and observed at magnifications 8000 – 50000X at 80kV. For thermal denaturation studies on P5 derived fibrils, the fibrils were heated to 90°C and cooled back, centrifuged and examined using TEM as described above. The length and mean diameter of the fibrils were calculated by the measurements obtained using Image J software (National Institute of Mental Health, Bethesda, MA, USA.) using at least 25 samples lying within the frame per micrograph image. For enzyme-assisted proteolysis studies, the trypsin digested fibril samples were examined using TEM as described.

6.2.6 LC-MS/MS analyses

Amyloid fibrils collected from P5 (300 µg) were digested with trypsin (Trypsin: Fibril= 1:25 ratio) for 12 h. The digested mixture were analysed by LC–MS/MS [Ultimate 3000 nano-LC (Thermo Fisher Scientific/Dionex, Sunnyvale, CA), coupled with AB SCIEX Triple TOF TM 5600 mass spectrometer] and monitored for the short fragment (E611R623) in the sequence. The peptide mixture was first desalted and pre-concentrated in a trap column (Acclaim PepMap 100 C18, 75 µm × 3 µm, 100 Å from Thermo Fisher Scientific/Dionex) for 3 min at a flow rate of 5 µl/min. After desalting, the system was switched into line with the reversed phase analytical capillary column (25 cm × 75 µm i.d. Acclaim PepMap RSLC C18, 2 µm, 100 Å, Thermo Fisher Scientific/Dionex and Sunnyvale, CA). A 35 min gradient was used at 300 nl/min. All data was acquired using information-dependent acquisition (IDA) mode with Analyst

TF 1.5.1 software (AB Sciex, USA). Protein Pilot software (version 4.01, AB Sciex) was used to analyse the MS/MS data.

6.2.7 Peptide Aggregation by 1D Solution NMR

For all NMR experiments the peptides P1 to P6 were selectively, ^{13}C and ^{15}N labeled for selected amino acids (Glycine, Alanine, Valine and Isoleucine) and were synthesized from Synpeptides (Synpeptide Co Ltd, Shanghai, China) with purity greater than 95%. The amino acid residues that were ^{13}C and ^{15}N labelled for each peptide is highlighted in blue and given in **Table 6-2**.

Peptide	Sequence	Number of Amino acids ^{13}C , ^{15}N labeled	Reference in this manuscript
TGFBIP ⁶¹¹⁻⁶³³ WT	EPVAEPDIMA ^T NGV ^V HVITN ^V LQ	10	P1
TGFBIP ⁶¹¹⁻⁶³³ M619K	EPVAEPDI ^K ATNGV ^V HVITN ^V LQ	10	P2
TGFBIP ⁶¹¹⁻⁶³³ N622H	EPVAEPDIMA ^T ATHGV ^V HVITN ^V LQ	10	P3
TGFBIP ⁶¹¹⁻⁶³³ N622K	EPVAEPDIMA ^T TKGV ^V HVITN ^V LQ	10	P4
TGFBIP ⁶¹¹⁻⁶³³ G623R	EPVAEPDIMA ^T NRV ^V HVITN ^V LQ	9	P5
TGFBIP ⁶¹¹⁻⁶³³ H626R	EPVAEPDIMA ^T NGV ^V RVITN ^V LQ	10	P6

Table 6-2: Peptide sequence of TGFBIP-4th FAS-1 domain (amino acids 611 to 633) WT and various disease-associated mutants. The blue coloured residues are labelled with ^{13}C and ^{15}N isotopes. The total number of amino acids labelled for each residue is also tabulated.

To follow the aggregation profiles of the peptides using 1D solution NMR, 0.6 mg/ml of peptides were dissolved in 1X PBS buffer along with 10% D₂O and 5% DSS. Samples were transferred to 5mm NMR tubes and experiments were recorded at 37° C. Aggregation profiles were studied using 600 MHz Bruker Avance II spectrometer by recording series of 1D NMR spectra. Spectra were collected every 30 minutes until a clear drop in amide peak intensity was observed in the 1D NMR spectra that served as an indirect readout of aggregation. The NMR data were processed using TopSpin 2.0 (www.bruker-biospin.com) and analyzed using CARA (cara.nmr.ch).

1D NMR intensities between 6.8 to 8.8 ppm for each time point for each peptide was obtained by using the integral function of the Topspin software and normalized to the integral values of the internal standard, DSS. The normalized intensity values were fit using a single-phase decay exponential function. The aggregation rate (K) values for each of the peptides were tabulated.

Packing of fibrils in the rotor for Solid State NMR

About 6 mg of ^{13}C and ^{15}N labeled peptides were dissolved in PBS to attain a final concentration of 0.6 mg/ml. The peptides were incubated at 37° C in a shaking incubator in PBS buffer for amyloid fibril formation. The fibrils were collected after 5 weeks. The presence of amyloid fibrils was confirmed by TEM. The fibrils were spun gently and the pellet was collected and sent to RIKEN institute for Solid State NMR experiments. Pellet sample was filled into the NMR sample rotor by centrifugation in a swing rotor and a pipette tip as a funnel.

6.2.8 Solid State NMR for the amyloid fibrils

One-dimensional and two-dimensional ^{15}N – ^{13}C , correlation experiments were performed on a 700 MHz Bruker (Bruker BioSpin GmbH, Rheinstetten, Germany) Avance III spectrometer with a MAS rate of 15.75 KHZ. Cross-Polarization magic-angle spinning (CP-MAS) ^{13}C SS NMR experiments were performed at 25 °C. Two-dimensional ^{13}C - ^{13}C dipolar assisted rotational resonance (DARR) experiments for the P5 and P1 peptide were performed at 25 °C and 10 °C, with a mixing time of 20 milli second (ms), 60 ms, 200ms and 600 ms. The spectral width (f2 and f1) was set to 270 ppm and center (f2 and f1) was 97.7 ppm. The number of time domain data points (TD) along f2 and f1 dimensions were 1024 and 768 respectively. The recycling delay was set to 1.5 seconds with 64 scans. For the NCO and NCA experiments, the spectral width (f2 and f1) was set to 270, 74 ppm and center (f2 and f1) was 97.7, 115 ppm. The number of time domain data points (TD) along f2 and f1 dimensions were 1024 and 96 respectively. The recycling delay was set to 2 seconds with 384 scans. For all ^{13}C and ^{15}N experiments, chemical shifts were calibrated from gyromagnetic ratio based on the

H₂O signal. The data was processed using TopSpin 2.0 (www.bruker-biospin.com) and analyzed using CARA (cara.nmr.ch).

6.2.9 Cytotoxicity of amyloids on Cultured Human Corneal Fibroblast (HCSF) cells by xCELLigence

The xCELLigence system (ACEA, San Diego, CA, USA) was used to investigate the effects of the peptide fibrils of TGFBIp on HCSF proliferation and cell death. The E-plate 96 is incorporated with sensor arrays at the bottom of each well, which detects and translates cell attachment in the form of electronic impedance. A parameter termed cell index (CI) is derived which corresponds to the relative density and adherence strength of cells in each well [69]. HCSF were seeded 7000 cells per well on an E-plate 96, in fibroblast media (DMEM with 10% FBS). The cells were allowed to attach onto the E-plate till the impedance curves reach a plateau. The matured fibrils were formed from 0.6 mg/ml of monomeric peptides dissolved in PBS and shaken at 37°C at 180 rpm for 4 weeks, as described previously. These peptide fibrils were centrifuged and diluted in culture media and treated with the cells. Culture was maintained for 48 h in the xCELLigence RTCA SP Station, placed in an incubator at 37°C with 5% CO₂. No media change was done to reduce the fluctuations generated from removal of plate from RTCA system.

6.3 Results

6.3.1 Identification of peptides enriched in the amyloid deposits of patients

The corneal amyloid deposits were isolated from the patient's cornea using a technique called Laser captured Microdissection (LCM). The corneal amyloid deposits were digested with trypsin and the protein composition and the enrichment of various tryptic peptides corresponding to TGFBIp were analysed. The spectral count from the mass spectrometry data showed a list of peptides from the 4th-FAS-1 domain of TGFBIp (**Figure 6-1**) with high abundance in the corneal deposits of patients compared to the

normal control cornea. The peptides samples that were enriched in the patient samples are H¹⁸⁷GMTLTSMYQNSNIQIHYPNGIVTVNCAR²¹⁶, I²⁹⁶LGDPEALRDLLNNH ILK³¹³, L⁴⁹⁷TPPMGTVMDVLKGDNRFSMLVAAIQSAGLTETLNR⁵³³, A³⁵⁰IISNKD ILATNGVIHYIDELLIPDSAK³⁷⁷, Y⁵⁷¹HIGDEILVSGGIGALVR⁵⁸⁸, E⁶¹¹PVAEPDIM ATNGVVHVITNVLQPPANRPQER⁶⁴² and G⁵¹¹DNRFSMLVAAIQSAGLTET LNR⁵³³. This result suggest that the mutant TGFBIp is processed differently compared to the normal control cornea. This may be a due to the structural modification of the mutant protein and the partial unfolding of proteins due to the amino acid modification. The partial unfolded nature of the mutant protein may expose residues for enzymatic reactions compared to the well-folded WT native protein. We also verified if the peptides contained the modified amino acid sequence in the corneal amyloid samples from the patients.

The majority of the peptides that are enriched in the patient samples are from the 4th FAS-1 domain of the protein and it is possible that the 4th FAS-1 domain plays a major role in the aggregation of these mutant proteins. So based on the peptide results we decided to further investigate the 23 amino acid long peptide along the region of E⁶¹¹PVAEPDIMATNGVVHVITNVLQPPANRPQER⁶⁴² in this study and further study the modification of this peptides and amyloid forming properties of this peptide. The peptides used in this study and net charge are listed in **Table 6-3**. These peptide residues are also clinically important as 16 different mutations are reported around the amino acid region and 11 different mutations are associated with an amyloidogenic phenotype in patients. These 11 mutations have an amino acid alteration that decrease the nett charge of the protein.

Figure 6-1: The spectral counts of relative abundance of peptides that are highly enriched in the amyloid corneal deposits from corneal dystrophy patients compared to the healthy control cornea. Most peptides enriched in the dystrophy patients are located in the 4th FAS-1 domain which harbours more than 80% of all clinically reported mutation.

Table 6-3: Peptide sequence of TGFB1p-4th FAS-1 domain (amino acids 611 to 633) WT and various disease-associated mutants. The net charge of each peptide is also tabulated.

Peptide	Sequence	Net Charge	Reference in this manuscript
TGFB1p ⁶¹¹⁻⁶³³ WT	EPVAEPDIMATNGVVHVITNVLQ	-3	P1
TGFB1p ⁻⁶³³ M619K	EPVAEPDIKATNGVVHVITNVLQ	-2	P2
TGFB1p ⁶¹¹⁻⁶³³ N622H	EPVAEPDIMATHGVVHVITNVLQ	-3	P3
TGFB1p ⁶¹¹⁻⁶³³ N622K	EPVAEPDIMATKGVVHVITNVLQ	- 2	P4
TGFB1p ⁶¹¹⁻⁶³³ G623R	EPVAEPDIMATNRVVHVITNVLQ	-2	P5
TGFB1p ⁶¹¹⁻⁶³³ H626R	EPVAEPDIMATNGVVRVITNVLQ	-2	P6

6.3.2 Time-dependant transition of secondary structure of model peptides

The conversion of highly random coil structures to β -sheet structure is the characteristics of amyloid fibril formation (Dobson 2003, Chiti and Dobson 2006, Monsellier and Chiti 2007). Circular Dichroism assays were performed at various time points to probe for changes in secondary structure of the peptide shown in [Figure 6-2](#). The peptides were allowed to shake at 37°C incubator and form amyloid fibrils without the presence of any enhancers of amyloid fibril formation. At time zero, TGFBIP⁶¹¹⁻⁶³³WT showed an initial 200 nm peak, typical of unfolded protein and gradually shifted to show a peak around 216 nm which is typical for β sheet conformation after 5 weeks. TGFBIP⁶¹¹⁻⁶³³M619K showed a strong maximum at 200 nm at time zero with a rapid progression to β sheet structures as early as 1 week with a maximum intensity at 216 nm even after 5 weeks.

TGFBIP⁶¹¹⁻⁶³³N622H peptide was also found to have 216 nm peaks around 1 week and continued to show maximum intensity for up to 5 weeks. For both TGFBIP⁶¹¹⁻⁶³³N622K and TGFBIP⁶¹¹⁻⁶³³G623R peptides, a weak positive peak around 198 nm and a broad peak at 216 nm were observed at time zero which is very typical of β sheet confirmation. TGFBIP⁶¹¹⁻⁶³³N622K peptide, showed increased amplitude of maximum intensity peaks around 198 nm and 218 nm with progression in time. With increase in time TGFBIP⁶¹¹⁻⁶³³G623R peptide showed a slight shift in the peaks and had a positive peak around 205 nm and negative peak around 198 nm. The amplitude of these peaks was observed to be much higher than the time zero peaks. TGFBIP⁶¹¹⁻⁶³³H626R did not show any significant change in the secondary structure even after monitoring for 8 weeks.

So among the peptides, TGFBIP⁶¹¹⁻⁶³³G623R was the peptide to aggregate the fastest and form amyloid fibrils while TGFBIP⁶¹¹⁻⁶³³WT and H626R peptide were the slowest to aggregate. Thus the substitution of charged amino acids in the peptides had a greater impact in their aggregation properties and their secondary structures and how fast they converted to a β -sheet assembly. The propensity to form amyloid fibrils also depends on the position and the type of amino acid change in the form of mutation. To understand the conformational changes in the peptides, we tried to quantify the secondary structural elements in each of these peptides at different time points using two different predictive methods (Dichroweb and CDNN software), but the prediction

methods were inconsistent and not very conclusive on the secondary structural elements of these samples.

The reason may be because the peptide samples were followed for a long time and there are chances of smaller oligomers in the mixture and also precipitated protein that may have reduced the intensity of the secondary structure CD spectra and hence could not directly contribute to the individual structural elements and the transition of a random coil structure to a greatly structured β -sheet elements. The presence of ordered β -sheets was easily verified by TEM.

Figure 6-2: Circular Dichroism spectra of the peptides at different time points.

Circular Dichroism (CD) spectra of TGFBIp⁶¹¹⁻⁶³³WT (P1) and mutant synthetic peptides (P2 – P6) in PBS at 37 °C were recorded at various time points. The results show that TGFBIp⁶¹¹⁻⁶³³WT took about 5 weeks to be converted to a β -sheet structure, typical for amyloid fibrils from a random coil configuration. P3, P4 and P5 peptides showed β -sheet transitions as early as week 1. P6 peptide did not show a significant change in the secondary structure even after 8 weeks of monitoring.

6.3.3 Determination of Kinetics of amyloid formation by ThT assays

We analysed the effect of position specific amino acid changes in the peptides, the peptide samples at different time points were used and mixed with Thioflavin T (ThT) dye. The dye binds to amyloid fibrils and oligomers and the increase in the fluorescent intensity is the direct measurement of amyloid fibril formation. With time, the intensity of fluorescence increases in all samples indicating that the peptides are forming amyloid fibrils as shown in **Figure 6-3**. All peptide samples did not have a lag phase indicating that amyloid fibril formation was spontaneous and progressed with time. Among all peptides, peptide P5 was the fastest to aggregate with peptide p1 and P6 the slowest to aggregate.

The aggregation propensities for each of the peptides were calculated and tabulated in **Table 6-4**. Peptides P2, P3 and P5 had greater values of K_{agg} and took lesser time to aggregate. This result is in agreement with the CD results.

Table 6-4: Aggregation kinetics of the peptides.

Aggregation rate (k_{agg}) and Half-life ($t_{1/2}$) calculation for the synthetic peptides (P1-P6) to reach half-maximal intensity to form amyloid fibrils based on the observed ThT fluorescence intensity and Aggregation rate (k) calculation for the synthetic peptides (P1-P6) to reach half-maximal intensity to form amyloid fibrils based on normalised intensities calculated from 1-D

Peptides	k_{agg} , sec-1 (ThT assays)	$t_{1/2}$, h (ThT assays)	Rate constant (K), sec-1 by 1D-NMR
P1	$3.6 \pm 0.3 \times 10^{-5}$	5.3 ± 0.9	-
P2	$5.7 \pm 0.8 \times 10^{-5}$	3.5 ± 1.2	6.94×10^{-5}
P3	$4.7 \pm 0.35 \times 10^{-5}$	4.1 ± 0.6	4.47×10^{-5}
P4	$3.4 \pm 0.38 \times 10^{-5}$	5.9 ± 1.4	3.42×10^{-5}
P5	$6.9 \pm 1.0 \times 10^{-5}$	3.0 ± 1.0	-
P6	$3.0 \pm 0.24 \times 10^{-5}$	6.5 ± 1.1	2.16×10^{-5}

NMR analysis.

Figure 6-3: Aggregation kinetics of peptides as measured by Thioflavin T assays. (A-F) Aggregation kinetics of TGFBIp611-633WT (P1) and mutant peptides (P2 – P6) probed by ThT fluorescence. Fluorescent intensity is plotted in the Y-axis against the observed time in seconds in the X-axis. All the mutant peptides show an increased propensity to form amyloid fibrils compared to the TGFBIp 611-633WT peptide. ThT intensity data from two independent experiments were averaged for the determination of kinetic parameters.

6.3.4 Aggregation of peptides measured by 1-D NMR

To understand the conformational transition of amyloid fibril formation during the early stages of amyloidosis, we used 1-D proton NMR to follow the aggregation process of the peptides. The peptide solutions were subjected to 1-D NMR analysis at regular intervals of time to follow the aggregation process. There was a clear decrease in intensity or the broadening of peaks. This was used as an indirect measurement to follow the aggregation of peptides (**Figure 6-4**). The actual decrease in peak intensity is due to formation of aggregates and larger oligomers. The larger oligomers are species that are difficult to detect by NMR due to the slow tumbling of the larger peptide oligomers.

Comparing the aggregation process of the peptides, P1 and P6 were the slowest to aggregate and took more than 24 hours to show a clear decrease in the peak intensity. P3 took around 5 hours to aggregate while P4 aggregated around 6 hours. Peptide P6 took about 12 hours to form larger aggregates detected by NMR spectra. P5 was the fastest to aggregate and formed oligomers as early as the peptides were dissolved in PBS solution. The amide peaks of this peptides were not visible even at time zero. For peptides P2, P3 and P4 there was a clear decrease in amide peak intensity as the aggregation process was rapid forming larger order oligomers.

The 1-D NMR curves were used to calculate the Rate constant of aggregation (K) for each of the peptides. The integral values of each 1-D NMR peak between 6.5 to 8.5 ppm which showed maximum changes upon peptide aggregation were collected and normalised with integral values derived from DSS. The normalised intensities were fitted on a single phase decay curve (**Figure 6-5**). The rate constant values are tabulated as in **Table 6-4**. The rate constant values were not available for peptide P1 and P5 as there were not enough 1-D NMR curves to be used for the curve fit.

The results from the 1-D NMR and the aggregation constant calculations complement each other and they agree well with each other. There is no lag phase observed in the 1-D NMR spectrum suggest that the aggregation of these peptides are very spontaneous. The results clearly predict that P5 was the fastest to aggregate and took less than few minutes to form ordered oligomers and formed higher ordered secondary structures within few hours.

P1

After 24 hours

Time 0

P2

After 4 hours

After 3 hours

After 2 hours

After 1 hour

Time 0

Aggregation Kinetics of the selectively ^{13}C and ^{15}N labelled peptides as observed by 1D-Solution NMR. 1D NMR of all peptides at Time 0 and aggregation kinetics monitored at various time points.

P3

P4

Aggregation Kinetics of the selectively ^{13}C and ^{15}N labelled peptides as observed by 1D-Solution NMR. 1D NMR of all peptides at Time 0 and aggregation kinetics monitored at various time points.

P5

P6

Figure 6-4: Aggregation kinetics of peptides followed by 1-D NMR

Aggregation Kinetics of the selectively ^{13}C and ^{15}N labelled peptides as observed by 1D-Solution NMR. 1D NMR of all peptides at Time 0 and aggregation kinetics monitored at various time points. P1 peptide took almost 24 h to aggregate while Peptides P2, P3 and P4 aggregated within 6 h. P5 was the fastest to aggregate as soon as it was dissolved in the buffer with no major amide peaks observed in the spectra. P6 took about 12 h to aggregate.

Figure 6-5: Rate Constant (K) of peptides based on 1-D NMR

The normalised intensities of peptide aggregation profiles obtained from 1D-NMR at various time points for each of the peptides were fit to a single phase decay exponential function and the resulting aggregation constant (K) values were calculated. Out of the 4 peptides studied, peptide P2 was the fastest to aggregate and P6 was the slowest to form amyloid fibrils. This result is in agreement with the ThT kinetics.

Table 6-5: Aggregation kinetics of the peptides.

Peptides	k _{agg} , sec ⁻¹ (ThT assays)	t _{1/2} , h (ThT assays)	Rate constant (K), sec ⁻¹ by 1D-NMR
P1	3.6±0.3 × 10 ⁻⁵	5.3±0.9	-
P2	5.7±0.8 × 10 ⁻⁵	3.5±1.2	6.94 × 10 ⁻⁵
P3	4.7±0.35 × 10 ⁻⁵	4.1±0.6	4.47 × 10 ⁻⁵
P4	3.4±0.38 × 10 ⁻⁵	5.9±1.4	3.42 × 10 ⁻⁵
P5	6.9±1.0 × 10 ⁻⁵	3.0±1.0	-
P6	3.0±0.24 × 10 ⁻⁵	6.5±1.1	2.16 × 10 ⁻⁵

6.3.5 Verification of amyloid fibril morphology by TEM

The amyloid fibrils were subjected to Transmission electron microscopy at various stages of amyloid fibril formation. When the peptides were visually examined after 2 weeks of amyloid fibril formation, a clear gel like deposits was found in the falcon tubes at the buffer surface. The transparency of the solution varied greatly between the amyloid fibrils. The amyloid fibril deposits of P5 were more translucent and can be seen as early as 1 week. P3 and P4 formed transparent depositions on the

tube. All the peptides started to form visible fibrils as early as 2 weeks (**Figure 6-6**). There was a difference in the length, morphology and stacking densities of amyloid fibrils derived from different peptides. All peptides were visualised for any visible fibrillation at day 4. For most of the peptides, clear oligomers were visible with short amyloid fibrils. As the amyloid fibril formation process continued, there were clear amyloid fibrils and oligomeric species were reduced. This clearly shows that all peptides form a highly ordered amyloid fibril.

The fibril morphology of P1 was thin and long fibrils with some fibrils extending up to 1 to 1.2 μm . For P2, mature fibrils that were shorter (100-200 nm) were observed during the early stages of amyloid fibril formation and these fibrils converted to long and unbranched fibrils. For P3 and P4 a significant amount of matured fibrils were visible during the early weeks. P5 was the fastest to aggregate with dense fibril stacks in accordance to kinetics of aggregation measured by both CD and 1-D NMR. P6 also displayed fibril morphology at week 2 but the amyloid fibrils had a mixed population of oligomers, truncated and short fibrils during the early stages of amyloid fibril formation.

The diameter of the amyloid fibrils derived from each of the peptides were calculated and it did not show any significant changes or differences among the different fibrillar assemblies from the peptides. These results indicate that with a single amino acid change, the rate of amyloid fibril formation and the morphology of the amyloid fibrils during early stages of amyloidosis greatly vary with the type of amino acid substitution the peptide undergoes. The mature fibrils however did not differ greatly in their morphology or the amyloid fibril diameter properties regardless of the differences in the kinetics of aggregation.

Figure 6-6: Morphology of the amyloid fibrils

TEM images of the TGFBIp611-633WT (P1) and mutant peptides (P2 – P6) after 4 days and 2 weeks, acquired with a JEOL JEM-1010 transmission electron microscope using Digital Micrograph™ 1.81.78 for GMS 1.8.0. Stacking densities, distribution and their mean diameters vary between the peptides. It is very clear that the proto-fibrils of the different mutants display different morphologies, lengths and diameters. While, the fibrils are almost fully formed for P5 peptide, for all the other mutants they are in the process of elongation. Interestingly, the length of the fibrils and the density is in perfect correlation with the aggregation kinetics

6.3.6 Stability of the amyloid fibrils derived from peptides

There are a number of studies to show that once amyloid fibrils are formed, they are stable to thermal challenges when they are compared to the monomeric form of the protein. The amyloid fibrils were subjected to variable temperature-CD experiments (Figure 6-7) to examine thermal stability of the fibrils and the amyloid fibrils were viewed under TEM to observe the morphology of the amyloid fibrils (Figure 6-8). When the amyloid fibrils derived from P1 was subjected to variable temperature-CD experiments, the ellipticity at 218 nm [θ_{218}] which corresponds to the β -sheets was observed by subjecting the amyloid fibrils to a heat-cool cycle. The secondary structures of the amyloid fibrils after the heat cool cycle were confirmed by CD, ThT assays and TEM. The amyloid fibrils were observed for secondary structural changes when heated from 20 ° C to 90° C. For P1 the negative maxima was found to increase with temperature and found to decrease after 70 ° C. when the amyloid fibrils were cooled, the transition to β -sheet was not clear, and the curve did not follow a linear decrease to reach its original value. When the fibrils were subjected to a second heat cool cycle, the traces overlapped with the traces obtained from the cooling cycle.

When the same sample was subjected to far UV-CD, it showed a clear transition from a regular β -sheet configuration to a conformation that is highly distorted of any secondary structure. The changes in intensity of the ThT dye which was bound to amyloid fibrils upon the heating and cooling procedure were also monitored by ThT assays. The results shows a sigmoidal transition of ThT intensity and a maximum loss of intensity was observed around 70 ° C. When the amyloid fibrils were subjected to cooling, there was no reversible transition observed. These changes show that the amyloid fibrils were converted back to its monomeric forms. This result is consistent with the VT-CD experiment. A similar trend was observed for P2, P3 and P6. This indicates that amyloid fibrils derived from P1, P2, P3 and P6 did not show good thermal stability and the amyloid fibrils were distorted when the fibrils were heated and cooled back.

Peptides P4 had remarkable reversible thermal properties that clearly shows that the amyloid fibrils derived from the peptide are thermally stable. For P4 the initial temperature increase caused an increase in the [θ_{218}] intensity values but gradually reduced at a very high temperature. When the fibrils were subjected to a cool cycle, the values obtained were similar to the heating trend and reheating of the fibrils did not

alter the $[\theta_{218}]$ intensity. CD spectra also indicated the presence of β -sheets which was also confirmed by TEM analysis after a heat-cool cycle.

The amyloid fibrils from P5 showed a very good thermal stability property, better than P4. When the amyloid fibrils derived from P5 were subjected to heating, there was no change observed in $[\theta_{218}]$ intensity and a slight decrease in intensity was observed upon cooling the fibrils. The CD spectra displayed clear presence of β -sheet structure. Reheating of fibrils from P5 for about 5 heat-cool cycles showed that the intensities remained the same. ThT assays performed after the heat-cool cycle showed remarkable decrease in the ThT fluorescent intensities.

The TEM results of P1, P2, P3 and P6 after the heating and cooling cycles which did not give a typical secondary structure, did not show the presence of any amyloid fibrils. The amyloid fibrils derived from these peptides could not withstand the heat-cool cycle as the fibrils were completely disassembled. However, the amyloid fibrils derived from P4 and P5 the fibril structures remained well intact showing that the fibrils are extremely stable (**Figure 6-8**). For peptide P5 a number of the amyloid fibrils with each other along the whole length of the fibrils and showed a two fold increase in the average diameter of the fibrils (18.1 ± 1.52 to 31.7 ± 4.01 nm). For P4 the fibrils did not found to be coalesced as their diameters remain unchanged (17.5 ± 1.9 nm before heating to 18.8 ± 1.4 nm)

All the above results suggest that even with only one amino acid change, the aggregation potential of these peptides vary significantly. The type and the position of amino acids play a substantial part in the aggregation propensities and also the stability of the amyloid fibrils formed. The position of amino acid change and the type of amino acid change for P5 has the greatest possibility to propel the amyloidogenic properties of the peptides. A possibility of introducing electrostatic interactions that could stabilise the amyloid fibrils formed could also be another reason for enhanced amyloid fibril formation for P4 and P5.

The most stable peptide P5 has its amino acid glycine replaced by a more bulky and positively charged Arginine residue. The native peptide has the much smaller amino acid that lacks any side chains and may offer a greater conformational stability by keeping the protein away from any hydrophobic or electrostatic interactions. The hydrogen bonds are the key interactions which may help in the inter-residue hydrogen bond which aids in the stacked β -sheet structures in the amyloid fibril formation. The

highly charged Arginine residue may aid in the fibril formation with its hydrogen bonding ability.

Figure 6-7: Thermal denaturation of Amyloid fibrils:

Thermal denaturation studies of amyloid fibrils P1, P4 and P5 using Circular Dichroism Spectropolarimetry. Reversible thermal changes in the CD ellipticity at 218 nm ($[\theta]_{218}$) were monitored by heating the amyloid fibrils from 20 to 90 °C (Fig 8-7A, 8-7D and 8-7G). Far UV-CD spectra of P1, P4 and P5 peptide fibrils (Fig 8-7B, 8-7E and 8-7H) before heating and after heat cool cycle were measured to study the change in secondary structure. While, P1 shows almost no fibrils after heating, P4 and P5 clearly show the presence of β -sheet fibrils. Thermal denaturation using ThT fluorescence studies for P1, P4 and P5 peptides (Fig 8-7C, 8-7F and 8-7I) and inset showing both P4 and P5 peptides to be more thermally stable compared to P1 peptide.

Figure 6-8: TEM of amyloid fibrils after heat-cool cycle:

TEM images of the peptide fibrils of P1, P4 and P5 TGFBIp mutants before, after heating to 90 °C and cooling to 20 °C were acquired. There were no fibrils found on the grid for the P1 peptide after heating. Clear fibrils with almost no difference in their stacking densities were observed for both P4 and P5 peptide fibrils.

6.3.7 Solid State NMR analysis of the most stable amyloid fibrils

¹³C Cross polarisation-MAS experiments were recorded for all peptides to ensure the amyloid fibrils formed were of good quality (Figure 6-9). The results from the MAS spectra showed that the amyloid fibrils were uniform and observed line widths

were smaller than 100 Hz, which helped us to further our studies using the 2D-experiments.

Figure 6-9: Cross polarisation MAS experiment for the amyloid fibrils derived from all peptides.

Among all the peptide fibrils, P5 displayed higher stability and greater propensity to form amyloid fibrils. A further detailed investigation on the structure of amyloid fibrils formed were examined by solid-state NMR experiment. The DARR experiment with short mixing time and long mixing time were recorded to identify all intra and inter molecular carbon-carbon correlation. The DARR spectrum is able to identify 2 bond carbon-carbon correlations. Additional 2D experiments like NCO and NCA spectra were recorded to aid in assignment of the DARR spectrum. Using the NCA and NCO experiments the ^1H magnetisation is transferred to ^{15}N of the amide

nitrogen by cross –polarisation and magnetisation is selectively transferred to $^{13}\text{C}\alpha$ using carbon specific polarisation. A pair of sequential isotopic labelled VA, GV, VV and VI residues were assigned first. The chemical shift of the DARR (**Figure 6-10**) spectrum was useful in VV and VI residue pair assignment. The rest of the NCA (**Figure 6-11**) and the DARR peaks were assigned based on the classical pattern of some of the amino acids. Out of the 9 amino acids that were labeled with the isotope, only 7 of the amino acids were visible for assignment. 1 alanine amino acid (A614) and 1 valine amino acid (V613) were highly mobile and hence could not be visualised in the DARR spectra. The other 4 valine residues, 2 isoleucine and 1 Alanine were assigned on the DARR spectra.

A plot of the chemical shifts (C-alpha and C-beta) from the DARR assignment (**Figure 6-12**) that came from the rigid part of the peptide when subtracted from the chemical shifts of the random coil values showed clear β -sheet secondary structure. A DARR spectra which had a longer mixing time was also analysed and it did not give any extra cross peaks and connectivity was observed for V627 and I628. This inter connectivity shows that the β -sheet arrangement of the amyloid fibrils could be in a parallel arrangement. Thus the solid state NMR results complement the Cd and ThT results that the amyloid fibrils form β -sheet arrangement. The solid state NMR was also done for the WT-P1 peptide, however the NCA (**Figure 6-14**) experiments for the peptide showed multiple set of peaks. The multiple set of peaks may be due the presence of polymorphs and heterogeneity in the sample. It was hence very difficult to make assignment to the DARR spectra (**Figure 6-13**). The SS-NMR experiment also clearly points out that the position of the amino acid change and the charge it creates for the peptide plays an important role in the amyloidogenicity.

A

Figure 6-10: DARR spectra of the most stable P5 peptide:

Expansion of the aliphatic 2D ^{13}C - ^{13}C chemical shift correlation DARR spectra for fibrils derived from P5 peptide recorded with 200 ms mixing time. Out of the 9 amino acids that were uniformly ^{13}C and ^{15}N labelled, peaks were visible for 7 amino acids. 4 valine, 2 Isoleucine and 1 Alanine residues were identified readily based on their characteristic fingerprint region and intra residue cross peaks. 1 valine (V613) and 1 alanine residue (A614) were not visible for assignment as they might be highly dynamic.

Figure 6-11: NCA and NCO spectra for P5

NCO/NCA plot for peptide P5 used to assign nitrogen chemical shifts. In an NCA experiment magnetization is transferred from ^1H to ^{15}N via cross polarization and then selectively to the $^{13}\text{C}\alpha$ using specific cross polarization.

Figure 6-12: C-alpha and C-beta chemical shifts of P5

A plot of the C-alpha and C-beta chemical shifts from the rigid part of DARR assignment of P5 peptide subtracted from the random coil shifts ($\Delta C\alpha$ or $\Delta C\beta = [C\alpha \text{ or } C\beta - (C\alpha \text{ or } C\beta)_{\text{Random coil shifts}}]$). The plot shows the presence of a beta sheeted secondary structure.

Figure 6-13: Aliphatic region of the DARR spectrum of the amyloid fibrils from P1

DARR spectra for fibrils derived from P1 peptide recorded with 20 ms mixing time. The DARR spectrum showed fewer well resolved peaks which may be due to the presence of fibrils that are not homogenous and with different conformation of amino acids.

Figure 6-14: NCA experiment for fibrils derived from P1 peptide showing different conformation of amino acids which is evident from broad peaks.

6.3.8 Cytotoxicity of amyloid fibrils

A study on the cytotoxicity of the peptides in the soluble monomeric form and the fibrillar form on Cultured Human Corneal Fibroblast cells were assessed. Freshly prepared amyloid fibrils from the peptide solutions and monomeric peptides were added on to the fibroblasts and their proliferation was monitored using xCELLigence. None of the amyloid fibrils or the peptide solution displayed deleterious cytotoxic properties on the fibroblasts but some of the amyloid fibrils had a significant effect on the cell proliferation rates (**Figure 6-15**). For P1 peptide solution there was no change in the cell proliferation and the values obtained were comparable with the control PBS. The amyloid fibrillar form of P1 showed inhibition of cell proliferation after 24 hours of incubation ($p \leq 0.05$). The inhibition of cell proliferation was more pronounced after 48 hours and 72 hours of incubation ($p \leq 0.0001$) compared to the PBS control. The amyloid fibrils from P4 ($p \leq 0.01$), P3 and P6 ($p \leq 0.05$) had significant effect on cell proliferation compared to the control after long incubation time. The pre-fibrillar monomer forms of P2, P4 and P6, the inhibitory effect was evident soon addition of peptide solution to the fibroblast cells.

For the monomeric form of P3, the inhibitory effect was evident only after 72 hours of addition of the peptide. The monomeric form of P6 which has the slowest aggregation kinetics as followed by ThT and 1-D NMR had the strongest and maximum inhibitory effect on the corneal fibroblasts. The soluble form of P5 which showed maximum aggregation kinetics and the fastest peptide to aggregate augmented the proliferation rates of corneal fibroblasts. It is interesting to note that the fastest aggregating P5 has been shown by different experimental methods to form amyloid fibrils very rapidly and upon addition of buffer, starts to form protofibrils. The results predict that each of the soluble as well as the amyloid fibrillar form of the peptides have variable interactions with the cell proliferation of corneal fibroblasts.

Figure 6-15: Cytotoxicity and inhibition of cell proliferation effects of amyloid fibrils on corneal fibroblasts

6.4 Discussion and Conclusion

The objectives of this study were to ascertain the changes in aggregation kinetics and the concomitant physio-chemical properties of amyloid fibrils derived from the 4th FAS 1 domain of TGFBIp (E⁶¹¹-Q⁶³³). Computational and experimental observations demonstrated that the native peptide located in this region possessed high amyloid aggregation propensities. The effect of substitution of positively charged residues had direct effects on the secondary structures and rate of amyloid formation. It was clear that substitution of residues which possess a low intrinsic β -sheet propensity with residues that have higher propensity to form β -sheets caused a substantial rise in the amyloid fibril formation rate. This effect was clearly evident for P5 where substitution of glycine by arginine caused considerable increase in the β -sheet propensity and higher rate of amyloid formation. Glycine residue lacks side chains, thus improving the backbone flexibility and conformational flexibility as well as decrease the β -sheet propensity whereas arginine residue could easily be accommodated in a β -sheet structure owing to its ability to form salt bridges with other amino acid residues (Gerling, Brandenburg et al. 2011). Consistent with these observations, a good correlation was obtained between overall β -sheet propensity estimated by TANGO algorithm and the experimentally determined k_{agg} values. These results suggest that localized perturbation of the secondary structure together with decrease of overall net charge and the nature of amino acid substitution would possibly accelerate the amyloid fibril formation in P5.

TEM studies indicated that all the peptides formed unbranched amyloid assemblies with increasing incubation time. Morphologically, no evident changes in the amyloid fibril assemblies were detected among the fibrils derived from various peptides. However, significant changes in the thermal denaturation profiles were observed as peptides P4 and P5 displayed remarkable thermal stability. Their fibrillar structures remained intact and retained the β -sheet structures even after 5 heating/cooling cycles. The amyloid fibrils formed by other peptides dissociated irreversibly to the disordered monomeric forms. Morel et al., showed a similar phenomenon in the amyloid fibrils derived from N47A from the SH3 domain of α -spectrin dissociated upon heating to 90 °C and re-associated with identical morphology when cooled to around 65 °C (Morel, Varela et al. 2010). In our study, there was no

dissociation of the fibrils or hysteresis observed for P5 even after 5 cycles of heating and cooling. Furthermore, we have shown that, the thermal denaturation was not affected by the rate of heating suggesting that the process is not kinetically controlled and appears to occur under equilibrium, suggesting remarkable stability achieved by the amyloid fibrils from P5

ssNMR of the amyloid assembly formed by P5 gave insights into the secondary structure of the amyloid fibril formation. Since all the amino acids were not ^{13}C and ^{15}N labelled a complete structure of the peptide could not be determined. Results obtained suggested that P5 peptide formed very uniform amyloid fibrils as evident from the DARR spectrum and the amyloid fibrils derived from P5 peptide were more stable to various conditions used in ssNMR. The secondary structures as elucidated from the secondary chemical shifts were in agreement with the other bio-physical experiments. P1 peptide did not give a very good DARR spectrum because of the sample inhomogeneity and therefore could not be assigned for secondary structure determination.

Analysis of the cytotoxicity of peptides for the corneal fibroblasts by label-free xCELLigence in both the amyloidogenic as well as in the prefibrillar forms indicated considerable differences in the inhibition of cell proliferation between the two states, although no cytotoxic effect was observed. Interestingly, amyloid fibrils formed by mutant peptides displayed identical increase in proliferation rates when compared to PBS whereas amyloid fibrils from native peptide P1 showed a strong inhibitory effect on the proliferation rates. However, in the prefibrillar forms, all the mutant peptides displayed inhibition of cell proliferation when compared to the native P1 peptide, indicating that the oligomers of P1 did not affect the cell proliferation. Among the mutant peptides in their prefibrillar form, P6 displayed the strongest inhibitory effect. It is likely due to low aggregation rate and the mixed populations of oligomers may have caused the inhibition of fibroblasts cells.

Various studies on the TGF β Ip corneal dystrophy have shown the presence of smaller peptide fragments of the protein accumulating in the cornea possibly due to proteolytic actions (Sørensen, Runager et al. 2015). In view of this, Schmitt Bernard et al studied the amyloid forming properties of a L110-Q131 peptide of the WT and R124C and R124H mutants of 1st FAS1 domain of TGFBIp. It was found that the mutations accelerated the amyloid fibrillation in comparison to WT peptide. In our study the WT also formed fibrils presumably due to high amyloid forming propensity

of the sequence examined, as suggested by various predictive algorithms. Schmitt et al also found a decrease in overall net charge of a peptide or an increase in β -sheet propensity accelerated amyloid fibril formation. Kato et al., analysed fibrillation of TGF β Ip peptides in the presence of additives e.g. benzalkonium chloride and sodium dodecyl sulphate coupled with sonication, accelerated the amyloid formation of TGF β Ip peptide derived from 1st FAS1 domain (L¹¹⁰-E¹³¹) of TGF β Ip and R124H mutant peptides (Kato, Yagi et al. 2013). However, the WT and mutant peptides reported in our work did not require surfactants to accelerate amyloid fibrils since the peptide sequence chosen possess high intrinsic aggregation potential as predicted by the in-silico methods.

Formation of amyloid fibrils from native peptides is a characteristic feature in the mechanism of fibrillation in many protein aggregation disorders (Bitan, Kirkitadze et al. 2002, Ventura, Zurdo et al. 2004, Bemporad, Calloni et al. 2006, Monsellier, Ramazzotti et al. 2007, Breydo, Wu et al. 2012, Dikiy and Eliezer 2012, Seeliger, Estel et al. 2013). It has been shown that in many proteins e.g. heart Apo myoglobin, some regions of the sequence are more aggregation prone and can be used as determinants of aggregation of the whole protein compared to other regions (Ventura, Zurdo et al. 2004, Bemporad, Calloni et al. 2006, Monsellier and Chiti 2007). However, limitations in the study has been in the creation of fibrils that are highly homogenous. Attempts to generate amyloid fibrils using various peptides and proteins like IAPP (Seeliger, Estel et al. 2013), A β (Bitan, Kirkitadze et al. 2002), α -synuclein (Breydo, Wu et al. 2012) have employed the use of additives like metals, charged ions, proteases, biophysical methods like sonication and molecular methods like gene truncation. These maybe considered quite far from mimicking physiological conditions. Also, it has been seen in the case of α -synuclein, lipids and fatty acids have been found to accelerate α -synuclein fibrillation (Perrin 2001, Lee, Choi et al. 2002, Dikiy and Eliezer 2012). In T7endonuclease, the T7EI peptide forms amyloid fibrils at pH 6.8, but does not form stable fibrils at pH 6.0 and pH 8.0, and is predicted to dictate its transition to fibrils or amorphous aggregates (Guo and Eisenberg 2007). In our study, the region encompassing 611 to 633 amino acid residues in TGF β Ip was predicted to be aggregation prone using in-silico methods. The study was able to show that the native and the mutant peptides in this region clearly formed well-defined amyloid fibrils even in the absence of any external additives or other harsh biophysical conditions e.g. sonication.

Our group has recently shown that the R555W and H572R mutants of 4th FAS1 domains of TGFBIp display aggregation under the influence of acidic pH (Murugan, Venkatraman et al. 2016). While the critical length of peptides/proteins that constitute amyloid deposits in the dystrophic patients remain unknown, our results suggest that subtle decrease in the overall net charge, nature of amino acid substitution and the ability to form β -sheet structure could confer enhanced propensity to form amyloid deposits or may accelerate the fibrillation of TGFBIp proteins/peptides.

Conclusion

We have shown the sensitivity of a short region spanning the E⁶¹¹-Q⁶³³ of the full length protein TGFBIp, to cationic amino acid substitution. We showed that a decrease in overall net charge of the peptide altered the secondary structure and accelerated the amyloid formation, depending on the nature of cationic residues and position of the substitution. In terms of morphology the matured amyloid fibrils formed by the various peptides were indistinguishable, although considerable differences were observed in the thermal denaturation scans. Amyloid fibrils formed by P4 and P5 displayed remarkable thermal stability, indicating that the reversal or dissolution of the fibrils could be difficult and future therapeutic interventions must take into consideration this aspect.

Chapter 5

7 Screening of compounds to prevent protein aggregation by modulating protein stability and prevent formation of amyloidogenic peptide fragments

7.1 Introduction

Protein misfolding and aggregation are the most commonly followed and studied phenomena to understand the basic molecular and biological macromolecules in all living things. Aggregation of proteins is hypothesised to play a significant role in many diseases, where a protein is triggered towards the aggregation pathway by different stimulants (Stefani 2004). The main factors involved in these protein aggregation disorders are a mutation or change that is caused by the physical or chemical environment of the proteins (Koo, Lansbury et al. 1999). There are different kinds of aggregation of proteins. The most common observation is physical aggregation, where there is a physical association of one protein molecule with another without a change in the protein structure. The second process of aggregation is referred to as chemical aggregation where new chemical bonds are formed. These altered chemical bonds in the protein either aid the protein to bind to other interacting partners or form chemical bonds with other proteins, which leads to aggregation or alter the aggregation properties of the protein (Finke, Roy et al. 2000). Both aggregation mechanisms may occur simultaneously in protein, depending on the chemical environment, which may produce either a soluble or insoluble protein aggregate (Wang 2005).

TGFBI-associated corneal dystrophy is a protein aggregation disorder where mutations in the *TGFBI* gene produce a mutant protein (TGFBIP) that is prone to aggregation and the protein aggregates are deposited in various layers of the cornea (Korvatska 2000, Karring, Runager et al. 2010). The deposition of the protein aggregation leads to corneal clouding and subsequent decrease in visual acuity (Weiss, Moller et al. 2008, Surguchev and Surguchov 2010). There are more than 65 different mutations reported in the *TGFBI* gene and the type and nature of amino acid substitution dictates the type of protein deposits (amyloid fibrils, amorphous powdery aggregates or a mixed form), the layer of cornea in which these proteins are deposited or the age of

onset of the disease (Lakshminarayanan, Chaurasia et al. 2014, Anandalakshmi, Murugan et al. 2017). TGFBIp is a 68 KDa extracellular protein and is the second most abundant protein in the cornea (Klintworth, Bao et al. 2004, Dyrland, Poulsen et al. 2012). The other interesting factor in corneal dystrophies is that the protein deposits are found only in the cornea although mutations of *TGFBI* are observed in other tissues of the body. This explains the possibility of a unique mechanism of aggregation in the cornea and the interactions with other tissue-specific factors that may play a major role in aggregation (Thapa, Lee et al. 2007). It may also be possible for the presence of a corneal specific factor that may trigger the aggregation cascade or the absence of effective clearance mechanism in the cornea to process mutant proteins or clear aggregated proteins.

Normally in a folded protein, the hydrophobic residues are well buried on the inner sides of the protein folds, making the protein more stable and not exposing the hydrophobic residues for any kind of chemical interaction (Uversky, Karnoup et al. 1999). In most protein aggregation disorders, it is the partially unfolded or the misfolded protein and its intermediates that serve as precursors of protein aggregation (Fink 1998) even though the un/misfolded proteins' confirmation is not very stable and these conformers are not present in high concentrations. When a protein becomes partially unfolded, the hydrophobic residues present in close proximity are the ones responsible for initiating the aggregation process. In most of the amyloid disorders, it has been postulated that the proteins follow a nucleation-dependant aggregation process (Jarrett and Lansbury 1993) and the nucleation process is the initial and rate-determining step of the aggregation and amyloid fibril formation. After the nucleation step, it may be the addition of a monomer-cluster aggregation or a cluster-cluster aggregation (Speed, King et al. 1997), and the molecules are built up to form insoluble macro-molecules of amyloid fibrils.

There have been several carefully modelled computer simulation models to study the thermodynamics involved in the protein aggregation process (Patro and Przybycien 1994). During the initial process of protein aggregation, there is a considerable increase in the overall free energy of the protein due to the loss of well-folded monomeric conformations and translational states (decrease in entropy). This may be associated with the initial nucleation step of aggregation, which is the rate-limiting step and thermodynamically unfavourable step. After the initial process, further addition of molecules takes place, and there is a significant decrease in the

overall energy turning the aggregation process to be thermodynamically favourable (Patro and Przybycien 1994). This is mainly due to the exposed hydrophobic amino acids, which are attracted towards each other in the process of aggregation and they decrease unfavourable interaction of the protein-solvent surface.

In the case of *TGFBI*-associated corneal dystrophies, the amino acid substitution in the mutant protein may lead to changes in charge, hydrophobicity or β -sheet forming propensity (Chiti, Stefani et al. 2003), thus turning on the mutant protein aggregation cascade. The protein stability influences the protein processing, turn-over and clearance of the mutant TGFBIp, which plays a major role in aggregation. In *TGFBI* associated corneal dystrophy, mutations induce the production of either thermodynamically destabilized forms (A546T) (Køldso, Andersen et al. 2015) of proteins or alternatively over stabilized form (R555W (PDB ID 2LTC)) and G623D mutant proteins than the native protein (PDB ID 2LTB) (Runager, Basaiawmoit et al. 2011, Underhaug, Køldso et al. 2013, Lakshminarayanan, Chaurasia et al. 2015, Murugan, Venkatraman et al. 2016, Stenvang, Christiansen et al. 2016). The presence of highly unstable mutant protein may lead to protein aggregation by forming strong cross-beta amyloid fibrillar structures that eventually become difficult to clear from the cornea (Uversky, Karnoup et al. 1999, Uversky and Fink 2004). The presence of overly stable mutant protein may increase the turnover time of the protein, decrease proteolytic susceptibility and affect the efficiency of protein clearance by the protein clearance machinery of the tissue as well as trigger protein aggregation (Underhaug, Køldso et al. 2013).

There have been several previous studies and publications that analysed the differences in the proteolytic processing of the mutant and the WT-TGFBIp. The first-ever study was by Korvatska et al. where the authors reported the presence of either full length or shorter truncated TGFBIp in the amyloid deposits of corneal dystrophy patients (Korvatska 2000). Using mass spectrometry and gel electrophoresis (1D and 2D) analysis, Korvatska et al. reported the presence of small proteolytic fragments of 17-KDa, 26-KDa, 39-KDa and 44-KDa in amyloid corneal deposits of patients with the R124C mutation, and the smaller fragments were not identified in the healthy controls. They also reported the presence of a 2-fold increase in the amount of 44-KDa mutant protein, which showed the absence of the C-terminal region compared to the WT protein (Korvatska 2000). In another study by Stix et al. on the analysis of the composition of amyloid deposits from patients with the F540S mutation, the authors

observed several small TGFBIP-peptide fragments, which were 6.5 KDa, 6.9 KDa, 14KDa, 17KDa and 21KDa apart from the full-length TGFBIP (Stix, Leber et al. 2005).

Proteomic analysis of the composition of GCD corneal deposits of patients with R124L mutation did not identify smaller peptide fragments but reported a 2.5-fold abundance of the 68 KDa mutant protein (Korvatska 2000). In another patient with R124L mutation, the corneal deposits were composed of a 63 KDa TGFBIP (lacking C-terminus) and another fragment that did not have both N and C-terminus was identified (Hedegaard, Thøgersen et al. 2003). Proteomic analysis of another granular corneal deposit of patient with R124H mutation reported the increased presence of 68 and 66 KDa mutant protein (Korvatska 2000). Recently, proteomic analysis using mass spectrometry of several amyloid and no amyloid mutation identified the abundance of 2 peptide sequences TGFBIP 515-533 and TGFBIP 571-588 (Karring, Runager et al. 2012, Karring, Poulsen et al. 2013, Poulsen, Runager et al. 2014, Courtney, Toftgaard Poulsen et al. 2015, Poulsen, Nielsen et al. 2016).

In our recent mass spectrometry analysis of corneal deposits from patients with H626R and R124C mutations, we have also identified peptides TGFBIP 515-533, TGFBIP 571-588 and TGFBIP 611-642 to be highly enriched in the patient samples compared to the control cornea (Anandalakshmi, Murugan et al. 2017). The studies also reported the increased abundance of serine protease HTRA1 (HtrA1) in the amyloid deposits of lattice corneal dystrophy patients and proposed the proteolytic processing of mutant protein by HTRA1. All these studies confirm that the proteolytic processing of the mutant protein is different from the WT-TGFBIP. The exact mechanism of action of HTRA1 on TGFBIP will be the subject of follow-up research based on mass spectrometry assays (Grau, Baldi et al. 2005, Simon, Niphakis et al. 2013). It has also been proposed that the generation of the smaller peptide fragments may act as fibrillation seeds and trigger the aggregation pathway. The possible mechanism of seeded fibrillation by amyloidogenic peptide 571-588 on A546T mutant peptide was reported by Andreassen et al. (Andreassen, Nielsen et al. 2012). One of the possible strategies to counteract the destabilizing effect of the specific mutations might be the development of the low-to-medium-affinity small molecular ligand binding, which might increase the overall thermodynamic stability of the target proteins in biological tissues and cells (Martinez Molina, Jafari et al. 2013).

The main aim of the study is to look for small chemical modulators that can bind to the mutant protein (unstable mutants) and thermodynamically modulate the

protein to make it stable or increase stability closer towards WT-TGFB1p and also influence the proteolytic processing of the mutant protein. The binding of the chemical modulator to the mutant protein may prevent it from being proteolytically processed into smaller amyloidogenic peptide fragment seeds and hence possibly delay the aggregation process. The mutations or the amino acid change affects the mutant protein thermodynamically in the same order as the binding of a chemical compound. If the chemical compound can effectively bind to the mutant protein and shift the thermodynamic equilibrium of the protein towards the WT-TGFB1p, the aggregation process may be delayed, or the protein processing pathway may be reversed. The chemical modulator need not be a strong binder to cause a big shift in the equilibrium between the fully folded and the unfolded protein as it may hinder in other biological processes and cause potential side reactions that may be deleterious. The chemical modular could bind to the mutant protein in a transient manner with a weaker binding and faster kinetics to aid in the thermodynamic stabilisation and protein processing pattern of the mutant protein.

The general archetype in the drug discovery process is to screen for the tightest binding compound, with high affinity to the target protein, to have the best efficacy and selectivity, but on the other hand, such compounds are also the ones with many off-target effects (Ohlson 2008). The new concept of transient interaction has been a useful tool to identify lead candidates for a number of diseases. Transient interactions may serve as important tools to screen for weak interactions to targets that have versatile biological functions, and complete inhibition of targets may lead to shutting down or shifting the balance of different biological functions. Transient interactions, on the other hand, are effective to create the necessary biological effect and can be used to aim for several targets at the same time (Frantz 2005, Hopkins, Mason et al. 2006, Morphy and Rankovic 2007) for maximum efficiency. The main challenge for screening weak binders is the difficulty in detecting them and analysing the binding efficiency. The weak binding chemical could be easily washed out in the procedures of most routinely used detection techniques. To overcome all the difficulties in screening for a transient interaction between the compound and TGFB1p, we used affinity chromatography combined with mass spectrometry analysis adapted to high throughput to screen (Ohlson, Shoravi et al. 2006) for the binding of compounds (2500 fragments from Maybridge library) to WT and mutant TGFB1p.

The secondary validation assays on the potential lead compounds were performed based on a ^1H - ^{15}N HSQC NMR-based approach. When a compound binds to the protein, the chemical shifts of the protein change because they are very sensitive to the electronic environment of a nucleus. The changes in chemical shifts are sensitive even to the small changes caused by non-covalent interactions with the solvent molecules or to chemical or biological binding partners. Usually, these changes in the chemical shifts of a protein upon the addition of a compound are monitored by a ^1H - ^{15}N HSQC spectrum. Chemical shift mapping is also useful to monitor weak interactions (in mM range) of compound binding giving rise to a fast exchange. The chemical shifts in these weak interactions change continuously if the binding partner is added in excess. The interaction between the protein and the compound is so weak that the protein-compound complex exists only in the part of the time taken to record the chemical shift changes NMR experiment. The observed chemical shift for the weak interaction is given as a weighted average of the chemical shifts for the free protein and the complex formed between the protein and the compound until the compound is added in excess to saturate all binding sites of the protein molecule. An overlay of the recorded spectrum of the free protein and the spectrum recorded with an excess of the compound allow the identification of the chemical shift perturbations pertaining to individual amino acid. This mapping also provides information on the exact location of compound binding and the strength of the protein-compound interaction.

Limited proteolysis by various proteases can be very useful to understand the conformational features of the protein (Fontana, de Laureto et al. 2004). The proteolytic analysis helps to identify areas of intensified backbone flexibility and regions of the protein that are locally unfolded. These experiments are unique and easily reveal the molecular features of proteins. TGFBIp was subjected to limited proteolysis using trypsin with and without the compounds to identify differences in the proteolytic processing mechanism of the mutant protein by the lead compounds. By screening about 2500 compounds from the Maybridge library that can bind and modulate the mutant and WT-TGFBIp, we have identified more than 50 compounds that could be validated by secondary assays. The secondary validation by NMR analysis of top 10 compounds that had good affinity to R555W mutant identified 3 lead compounds that can bind to the R555W mutant protein and modulate the trypsin proteolysis upon compound binding.

7.2 Materials and methods

7.2.1 Expression and Purification of WT and mutant TGFBIp

The reagents and methodology for expression and purification of 4th FAS-1 domain of WT-TGFBIp, mutants R555W and H572R were followed as reported by our group (Elavazhagan, Lakshminarayanan et al. 2012). We worked with WT-TGFBIp, non-amyloidogenic mutant R555W and amyloidogenic mutant H572R.

PCR reagents, PCR and Gel purification kit, polymerases and restriction enzymes were purchased from Fermentas (Fermentas Inc., Glen Burnie, MD) and Kapa biosystems (Kapa Biosystems, Inc., Woburn, MA). *Escherichia coli* BL21 (DE3), is a lysogen of *E. coli* BL21 wherein the prophage contains the RNA polymerase gene of T7 bacteriophage under the control of lac UV5 promoter. pCDF2 vector system with the Ek/LIC cloning kit was purchased from Novagen (Novagen (EMD), Philadelphia, PA). Ni-Sepharose resin was purchased from GE Healthcare (GE healthcare Life Sciences, Piscataway, NJ). Ampicillin, streptomycin and isopropyl β-D-thiogalactopyranoside (IPTG) were purchased from Sigma- Aldrich (Sigma-Aldrich Inc., MO, USA).

Design and synthesis of wild-type and mutant constructs

The gene fragments coding for the 4th FAS1 domain of wild type TGFBIp and the mutants R555W and H572R were designed and constructed in silico using DNAMAN2 (Lynnon corporation, Quebec, Canada). The codons of these constructs were redesigned and optimized for high expression in *E. coli* using the codon optimization module provided by GenScript USA Inc. (Piscataway, NJ, USA). The modified segments were also synthesized commercially from GenScript (Piscataway).

Cloning and construction of recombinant plasmids

The constructs were obtained in pUC57 vectors (GenScript) and amplified by PCR using Kapa High Fidelity DNA polymerase (Kapa Biosystems). The specific 5' and 3' overhangs were incorporated by PCR to make the genes compatible for Ek/LIC cloning, (primer sequences: forward – 5' GACGACGACAAGATGGGCACGGTGATGG 3' and reverse – 5' GAGGAGAAGCCCGGTTACAGCACATT GGTG 3'). The 'TTA' sequence in the

reverse primer codes for the ochre stop codon, 'TAA' in the generated construct due to reverse complementarity. All the gene fragments were purified using the GeneJET PCR purification kit (Fermentas) and cloned into pCDF-2 vector by the optimized Ek/LIC cloning protocol (Novagen). The vectors cloned with the inserts were transformed into Novablue Gigasingles (Novagen) sub cloning competent cells by heat shock at 42° C for 45 s. The transformed colonies were selected by growing in Luria Bertani (LB) agar plates in the presence of streptomycin. Two to four colonies were selected for each construct and grown overnight in 5 ml of LB media with streptomycin. Plasmids were extracted from these cultures using GeneJET plasmid miniprep kit (Fermentas) and examined by restriction analysis with the NdeI and XhoI restriction enzymes. The integrity of the inserts within the plasmids was also confirmed by gene sequencing (1st BASE, Singapore). The validated plasmids were then transformed into BL21 (DE3) (Novagen) competent cells for protein expression. The transformed clones were selected based on their resistance to streptomycin.

Expression of 4th FAS1 WT and mutant TGFB1p proteins

The clones of the WT and mutant proteins were grown in 50 ml LB media at 37° C till an OD of 0.6–0.8 was reached and induced with 0.5 mM IPTG at 16° C. Following induction, the cultures were grown for 16 h. For screening the protein expression, 1 ml of culture was taken and examined. The culture was centrifuged and cell pellet was heated at 95° C for 5 min with 5X sample loading buffer that contains sodium dodecyl sulfate (SDS), b-mercaptoethanol and glycerol. The samples were loaded on to 4–20% precast SDS–PAGE gels (Bio-Rad, Hercules, CA, USA). Following electrophoresis, the protein bands were visualized by Coomassie staining. Large scale cultures were done in 2 or 4 l of LB broth containing streptomycin till an OD of 0.6–0.8 was reached and induced at 16° C with 0.5 mM IPTG. The expression of the constructs was examined using SDS–PAGE after growing for 16 h. The cultures were centrifuged at 8000 rpm and the cell pellets were frozen overnight at -80° C. The pellets were then resuspended in a cold buffer containing 50 mM Tris, 400 mM NaCl and 5% glycerol at pH 7.4. The resuspended pellets were sonicated or lysed using a French press at 1000 psi or a combination of both. The lysed cells were then centrifuged at 18,000 rpm for 45 min.

Table 7-1: composition of the LB-broth used for culture

Name	Material	Amount	Additional procedure
LB medium	Tryptone	10 grams	pH adjusted to 7.0 using sodium hydroxide
	Yeast extract	10 grams	
	NaCl	5 grams	

Purification of WT and mutant TGFBIp proteins

The supernatants containing the soluble proteins were mixed with the Ni-Sepharose (GE healthcare) resin and incubated for an hour at 4° C. The supernatant-resin mixtures were loaded on to 15 ml purification columns and washed with a cold buffer containing 50 mM Tris, 400 mM NaCl, 5% glycerol and 20 mM Imidazole at pH 7.4. Further washing and elution was done with increasing imidazole gradients up to 200 mM imidazole. The eluted fractions were concentrated by centrifuging at 3000 rpm for 1–2 h in 20 min intervals using Amicon ultra filters (Merck Millipore, Billerica, MA, USA). The samples were validated using SDS PAGE. The purified fractions were loaded on to a Fast Protein Liquid Chromatography (FPLC) system, gel filtration columns and eluted against a linear gradient of PBS. The purified fractions were stored at -80° C until further use.

¹⁵N labelled Protein Preparation for NMR experiments

The proteins TGFBIp-WT and mutants R555W and H572R were prepared and processed in the same way as the un-labelled samples but grown in M9 minimal media with ¹⁵N labelled ammonium chloride. For ¹³C labelled proteins both ammonium chloride and glucose were replaced with isotope labelled salts. All buffers were prepared in 1-liter stock, unless otherwise stated and filtered using 2 µm nylon filter. The antibiotics were prepared at 1M stock solutions and kept at -20° C. The growth media used was prepared according to the compositions stated below, pH adjusted to 7 and autoclaved prior to use.

Preparation of ^{15}N labelled M9 minimal salt media (preparation for 1 liter)

Compound	Amount
Na_2HPO_4	12.8 grams
KH_2PO_4	3.0 grams
NaCl	0.5 grams

The compounds were added in a flask with approximately 950 ml water and pH of the solution was adjusted to 7.4 with sodium hydroxide and autoclaved. The following salts and compounds (preparation for 1 liter) were dissolved in 50 ml of water and cold filtered using 30 ml syringe and 0.22 micron filter, into the autoclaved media.

Compound	Amount
NH_4Cl (labeled isotope)	1.0 g
Glucose	2.0 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0.4940 g
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.0100 g
Thiamine	0.0100 g
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	0.0100 g

7.2.2 F-Pocket analysis on WT-TGFBIp and R555W mutant

F-pocket is an open source pocket detection package based on Voronoi tessellation and alpha spheres built on top of the publicly available package Qhull. The modular source code is organized around a central library of functions, a basis for three main programs: (i) F-pocket, to perform pocket identification, (ii) T-pocket, to organize pocket detection benchmarking on a set of known protein-ligand complexes, and (iii) D-pocket, to collect pocket descriptor values on a set of proteins (Le Guilloux,

Schmidtke et al. 2009). We performed F-pocket analysis with the previously published NMR structure of the 4th domain FAS-1 domain of TGFBIp WT and R555W mutant (2LTB, 2LTC) as inputs. The f-pocket analysis was used to identify all available and potential binding pockets/cavities in both the WT and the mutant proteins. The predictions reflect the putative capacity of the pocket to bind a small molecule that might be drug-like, but might also be a sugar, cofactor or co activator.

7.2.3 Screening of compounds using Weak affinity Chromatography (WAC)

Protein targets

Protein targets were the 4th FAS1-like domain of TGFBIp as wild type (WT), R555W and H572R mutants. The wild type domain consisted of 131 amino acids, of which 7 were lysines, 7 were arginines; no cysteine nor tryptophan present. Detail characteristics of the proteins calculated by GPMaw lite online tool (http://www.alphalyse.com/gpmaw_lite.html) are described in **Table 7-2**. The protein solutions were stored in phosphate buffer saline (PBS) pH 7.0 – 7.4.

Table 7-2: Characteristics of target proteins (4th FAS-1 domain of TGFBIp)

Protein	Average mass (Da)	Molar absorbance at 280 nm (cm ⁻¹)	Isoelectric point	Hydrophobicity index
WT	14112	0.18	6.32	0.10
R555W	14142	0.58	5.75	0.13
H572R	14131	0.18	7.62	0.09

The Maybridge library

The used Maybridge library (Ro3-Rule of three) included 2500 fragments with molecular weights (MW) ranged from 82-296 Daltons (Da) and an average mass of 180 Da. ClogP values were from -2.5 to 3.0. Each fragment contained 0-4 rotatable bonds, 0-3 hydrogen bond donors or acceptors. Polar surface area of each molecule was from 0-70.04 Å². Fragments were dissolved in dimethyl sulfoxide (DMSO) as individual compounds at original concentration of 100 mM. The properties of the compound from the library are listed in the below **Table 7-3**.

Table 7-3: Properties of the compounds included in the R03 Maybridge library

	Rule of Three Criteria	Maybridge Ro3 Fragments Range of Values	Maybridge Ro3 Fragments Average Value
Molecular Weight	≤ 300	93.1 – 298.4	178.3
cLogP	≤ 3.0	-2.5 – 3.0	1.5
H-Bond acceptors	≤ 3	0 - 3	2.4
H-Bond Donors	≤ 3	0 - 2	0.9
Rotatable bonds (Flexibility Index)	≤ 3	0 - 3	0.9
Polar Surface Area	≤ 60 $\text{\AA}^2$	4.9-69.1 $\text{\AA}^2$ (99% < 60 $\text{\AA}^2$)	37.0 $\text{\AA}^2$

Samples in the primary screening were the entire Maybridge fragment library that was distributed into 50 mixtures, each comprised of 50 fragments with non-overlapped Molecular weight. The mixtures were diluted 100 times with water to give a final concentration of 20 μM for each fragment in 1 % DMSO. In the confirmation single solution study, the samples were prepared by diluting original fragments 1000 times with water to a final concentration of 0.1 mM in 0.1 % DMSO.

LC-MS water and extra pure ammonium acetate were purchased from Merck Millipore (Billerica, MA, USA). Other chemicals were of analytical grade or higher and purchased from Sigma-Aldrich (St. Louis, MO, USA), if not otherwise stated.

Immobilization of target on silica particles and packing affinity column

The proteins were immobilized on spherical porous diol silica particles (Kromasil, from EKA, Bohus, Sweden, at 5 μm in diameter, 300 $\text{\AA}$ pore size, $\sim 100 \text{ m}^2$ surface area per gram) by reductive amination reaction (Schiff base method). First, the diol silica ($\sim 7 \text{ mg}$ for each protein, see [Table 7-4](#) for details) was sonicated in 0.3 mL isopropanol for 2 min, and then washed by centrifugation ($2000 \times g$ in 2 min) with $4 \times 0.3 \text{ mL}$ water followed by another 2 min sonication. Periodic acid was added in the slurry to a concentration of 100 mg/mL to oxidize diols on silica into aldehyde group. The slurry was rotated at room temperature (22°C) for 3 h before being washed with $4 \times 0.3 \text{ mL}$ water then 0.3 mL of 100 mM sodium phosphate buffer, pH 7.5 for silica to be coupled with WT or R555W, pH 7.0 for H572R (immobilization buffer). Resulting

aldehyde silica batches were let reacted with proteins in PBS containing 10 mg/mL NaCNBH₃ (final concentration). Reaction mixtures were rotated at 4° C for 92 h (WT and H572R) or 96 h (R555W) before they were stopped by centrifugation (2000 × g in 2 min) and removal of supernatants, and then washed with 6 × 0.3 mL immobilization buffer. Supernatants and washings were pooled, which were later measured absorbance at 280 nm (NanoDrop 1000, Thermo Fisher Scientific, Wilmington, DE, USA) to calculate the amount of immobilized proteins.

Another batch of diol silica was treated the same way except no protein was added into the immobilization reaction mixture. This batch served as blank silica (produced blank column 2015-04-02).

Table 7-4: Details of immobilizations of target proteins and obtained columns

Silica batch	WT	R555W	H572R	Blank
Amount of starting diol silica (mg)	7.3	7.1	7.2	7.6
pH of immobilization buffer	7.5	7.5	7.0	7.5
Amount protein added (mg)	1.14	0.89	0.48	0
Volume of reaction slurry (μL)	102	299	139	200
Incubation time (h)	92	96	92	66

Each resulting silica batch was packed into a stainless steel capillary (50 × 0.5 mm) by an air-driven liquid pump (Haskel MS-110, Burbank, CA, USA) at 340 bar for 2 h. Packing mobile phase was 100 mM sodium phosphate buffer pH 7.5. Obtained columns were WT, R555W, H572R and blank columns, packed with corresponding silica batches. Immobilization and packing were carried out shortly before the primary screening in order to keep the proteins fresh. If not in use, the affinity columns were filled with PBS pH 7.4 and stored at 4° C.

7.2.4 Screening of Maybridge library using WAC

Primary screening of mixtures

Screening by WAC was performed on a 1200 Agilent HPLC system that coupled to a 6530 Accurate-Mass quadrupole time-of-flight mass spectrometer (MS),

which employed an Agilent Jet Stream Thermal Focusing electrospray ionization source (Agilent Technologies, Santa Clara, CA, USA) (LC-Q-tof). Mobile phase flow rate was 20 $\mu\text{L}/\text{min}$ and injection volume was 0.4 μL . Mobile phase was 10 mM ammonium acetate, pH 6.6. Room temperature was controlled at 18-19° C during the analyses. MS experiments were performed in single MS, positive or negative ionization mode. Fragmentor voltage was set at 100 V; capillary voltage (V_{cap}) at 1800 V, drying nitrogen gas was at 12 L/min at 250° C and nebulizer pressure at 40 psig. Sheath gas flow was at 12 L/min; 300° C. Skimmer voltage was at 65 V. A reference mixture solution containing ions m/z 121.050873 and 922.00979 in positive mode was used to correct the masses in real time. Chromatograms were acquired in the mass range m/z 70 – 930.

The LC-Qtof was used for the screening experiment and MassHunter B02.01 was used to acquire data; data analyses were performed using MassHunter B06.00 (both versions were from Agilent Technologies). Fragment chromatograms were retrieved by extracted ion chromatogram (EIC) function with their accurate masses, using symmetric expansion at ± 0.5 m/z . Extracted chromatogram was smoothed and integrated to obtained time at peak apex, which was considered as retention time. All injections were duplicated and analysis time was 45 min for each injection.

First, the full Maybridge library was screened by injecting the above 50 mixtures onto the target columns. Primary active retention times (t_{active}) of fragments from this experiments were used to calculate primary dissociation constant (K_d), and subsequently primary ligand efficiency (LE) based on equation (1) and (2), respectively.

$$K_d = B_{\text{tot}} / (t_{\text{active}} \times F) \quad (1)$$

$$LE = -(RT \ln K_d) / NHA \quad (2)$$

In which B_{tot} is the total number of active sites in R555W column, assuming to be 50 % of immobilized protein amount; t_{active} is the net retention time (t'_R) of a fragment on active column minus net retention time of that fragment on the blank column. Net retention time is retention time minus the void time for DMSO (ion m/z 79.0212 in positive mode). Negative t_{active} from the subtraction were normalized to zero. F is mobile phase flow rate. LE is ligand efficiency in kcal mol^{-1} per atom, R is the universal gas constant ($R = 1.99 \times 10^{-3} \text{ kcal K}^{-1} \text{ mol}^{-1}$), T is the absolute temperature in Kelvin (K), NHA is the number of heavy atoms in the fragment. In the primary hit selection, only fragments with highest LE for each target (55 fragments with $LE \geq 0.38$

kcalmol⁻¹ per heavy atom for WT, 53 fragments with $LE \geq 0.43$ kcalmol⁻¹ per heavy atom for H572R, and 50 fragments with $LE \geq 0.46$ kcalmol⁻¹ per heavy atom for R555W) were chosen to enter single solution confirmation step in consideration of our WAC and NMR hit confirmation capacity.

Primary hit confirmation as singletons

The primary hits were then run as singles to confirm detection. The LC-Q-tof conditions were kept the same as above, except the new room temperature was 25-28° C. Acquisition time and polarity mode to analyze each fragment were set according to information from the mixture screening. The standard deviation (SD) values of retention times of all detected fragments including DMSO on duplicate injections in mixtures were calculated; the average SD on each column were used to set a limit of affinity detection on that column. Fragments were confirmed as hits when their $LE \geq 0.30$ kcalmol⁻¹ per heavy atom and t_{active} were higher than ten times of the average SD.

7.2.5 Sample preparation for NMR and NMR experimentation

The top 10 compounds that gave a good LE value for the R555W mutant protein was used for secondary validation by NMR assays. Uniformly ¹⁵N-labeled proteins (R555W) were used at a concentration of 0.1 to 0.2 mM proteins in PBS buffer. The NMR samples consisted of 90 % H₂O, 10 % D₂O, and 5 % DSS as internal standard and transferred to 5 mm NMR tubes (Norell). All NMR spectra were acquired using 600 or 800 MHz Bruker Avance II spectrometers. Chemical shifts were referenced to DSS directly for ¹H and indirectly for ¹³C and ¹⁵N spins. The NMR data were processed using TopSpin 2.0 (www.bruker-biospin.com) and analyzed using CARA (cara.nmr.ch).

¹⁵N-HSQC experiments giving a correlation between amide nitrogen and proton pairs were acquired as fingerprint spectra for the protein. The mutant R555W protein was then titrated with various concentrations of DMSO). ¹⁵N-HSQC experiments were then recorded again at different fragment concentrations (1:1, 1:5 and 1:10) for the top 10 fragments identified by WAC experiments. The interactions of the proteins to the ligand were measured by peak shifts or changes in peak intensity when the ¹⁵N-HSQC spectrum of the free protein was overlaid with spectra of DMSO and various fragment concentrations. The NMR assignments for the 4th-FAS1 domain of R555W mutant

protein (2LTC) spectra were obtained from the BMRB website and the HSQC spectrum was calibrated accordingly. The peak shifts were analyzed and the amino acid residues that caused the shifts were mapped on to the structure of the protein.

The ^{15}N -HSQC spectrum of 1:10 protein to drug concentration was used for calculating peak shifts. Chemical shift perturbations along the proton and nitrogen dimensions for each amino acid of mutant R555W protein alone were subtracted from the chemical shift perturbations of protein with the compound. The chemical shift values obtained in terms of ppm values were converted to Hertz by multiplying the perturbations with magnetic strength of the NMR machine used. Root Mean Square Deviation (RMSD) values were also calculated based on the below mentioned formula and used for heat map generation of chemical shift changes noted in each amino acid of the TGFBIp-R555W mutant protein.

$\text{RMSD} = \text{Square root } [(\text{difference in chemical shifts along proton dimension in Hertz})^2 + (\text{difference in chemical shifts along nitrogen dimension in Hertz})^2]$

7.2.6 Limited proteolysis by Trypsin

The purified 4th FAS-1 domain proteins (50 μg) were subjected to proteolysis by trypsin for 1 hour. The proteins were pre-treated with the 3 lead compounds from NMR analysis (MO07617, RJF00203, BTB05094) for about 30 minutes at room temperature. 250 μg of each of the protein was subjected to trypsin digest (1:500) protein-trypsin ratio at 37° C. 50 μg of protein sample was withdrawn after 15 min, 30 min, 45 min and 1 hour and the proteolytic reaction was stopped by addition of formic acid. The samples were air-dried and stored at -20 ° C before MALDI analysis. The samples were suspended in 0.1% TFA /50% acetonitrile solution and spotted on to MALDI matrices and peptide fragments generated by trypsin digest were further analyzed by MALDI-TOF Mass spectrometry.

7.2.7 Peptide Mass fingerprinting (PMF) of tryptic peptides using MALDI-TOF/TOF-MS Analysis

The peptide sample was mixed with an equal amount of matrix solution containing 10 mg/ mL R-cyano-4-hydroxycinnamic acid in 0.1% TFA/50% ACN and spotted onto a 384-well stainless steel MALDI target plate (Applied Biosystems, Foster City, CA). An ABI 4800 Proteomics Analyzer MALDI TOF/TOF mass spectrometer (Applied Biosystems) was used to analyze the spotted samples. Both MS and MS/MS spectra were recorded in the combination mode and submitted for database searching. Typically, 1800 shots were accumulated for each MS spectrum while 4500 shots were accumulated for each MS/MS spectrum. For protein ID, combined MS and MS/MS data were submitted via GPS Explorer (version 3.6, Applied Biosystems) to Mascot server (version 2.0, Matrix Science) for database searching. Swissprot database (version 54.4, including 287 050 entries, and 105 598 259 residues) was utilized for the search. Cysteine carbamidomethylation and methionine oxidation were selected as fixed and variable modifications, respectively, and two missing cleavages were allowed. Precursor ion tolerances were set to < 0.05 (Meng, Zhang et al. 2008).

7.3 Results

7.3.1 Expression and Purification of 4th FAS1_WT and mutant TGFBIp

In order to proceed with all biophysical experiments, NMR and drug screening assays, good quality of protein was essential. The 4th FAS1 domain of WT and the mutants R555W and H572R were purified by large-scale protein expression. Large-scale cultures were grown with and without the isotopes, and the first step of purification was done using the affinity chromatography column making use of the His tag. The proteins eluted from the affinity chromatography column were then purified further using a size exclusion chromatography column (**Figures 7-1 and 7-2**). The pure proteins were collected and tested for correct mass and protein sequence was verified.

We used nano-LC–MS/MS to confirm the substitution of amino acids in the mutant proteins. MALDI mass spectrometry was used to confirm the identity of the purified mutants. The mass of the WT protein showed a single intact species at a mass/charge ratio of $16,666 \pm 10$ Da. The mass of the R555W protein showed a single intact species at a mass/charge ratio of $16,696 \pm 10$ Da, and this is in agreement with the predicted molecular mass of the mutant R555W, and the mass of the H572R protein showed a single intact species at a mass/charge ratio of $16,686 \pm 10$ Da, and this is in agreement with the predicted molecular mass of the mutant H572R (**Figure 7-3**). The pure protein was analyzed using SDS-PAGE that resulted in pure protein as indicated in **Figure 7-4**. The secondary structure of the protein was also confirmed by CD spectropolarimetry. The WT and the mutants showed a mixture of alpha and beta sheet structure. All the quality control assays like SDS-PAGE, mass spectrometry and CD experiment were done to have a pure protein for other biophysical and NMR assays. Proteins were quantified using a calorimetric assay based on the modified Bradford reagent. Proteins were quantified and made into 100µl aliquots and frozen until further use.

Figure 7-1: Size exclusion Chromatography (SEC) elution profile for TGFB1p-WT protein.

Figure 7-2: Size exclusion Chromatography (SEC) elution profile for TGFBIp-R555W and H572R protein.

Figure 7-3: Mass spectrometry analysis of 4th FAS-1 domain of TGFB1p-WT, R555W and H572R proteins. The single peak shows purity of protein and experimental mass of the pure protein.

Figure 7-4: A. SDS-PAGE analysis showing a single band for each of the 4th FAS-1 domain of TGFBIp-WT, R555W and H572R proteins around 15 KDa.

B. The circular Dichroism (CD) spectra of TGFBIp-WT, R555W and H572R. The dips around 222 nm, 218 nm and another dip around 208 nm shows a mixture of alpha helix and beta- sheet secondary structure of protein.

7.3.2 F-Pocket analysis of 4th FAS1 domain of TGFBIp-WT and mutant R555W

The deposited structures available from PDB of the 4th FAS1 domain of TGFBIp WT (2LTB) and mutant R555W (2LTC) were prepared as inputs for the F-pocket analysis. These were the recently published solution NMR structures, and the pymol analysis of overlaying the WT and R555W did not show a major change in the overall secondary structure of the protein except local changes around the mutations (Underhaug, Koldso et al. 2013). The predictions from the F-pocket analysis identified 3 potential binding pockets for WT protein and 4 potential binding pockets for R555W protein (**Figure 7-5**). Pocket 1 of the WT protein had a good druggability score of 0.934 (between scales of 0 to 1), indicating that it is very likely for a compound/fragment to bind to the pocket. It had a pocket score of 30.751. The other 2 pockets predicted for the WT did not have a good druggability score. For the 4th FAS1 domain of the R555W mutant, the F-pocket analysis came up with 4 potential binding pockets with a druggability score of 0.907 for pocket 1. It also had a pocket score of 33.981. Pocket 1 of both the WT and mutant R555W were used as inputs for the In-silico docking studies. Both the WT and the R555W mutant shared a similar pocket that gave the highest druggability score (**Table 7-5**). The pocket may be due to the cavity in the protein formed by the 3 alpha helices. The In-silico docking studies using the tool auto dock-Vina listed all the 2500 compounds from the 2500 compound Maybridge library based on the binding affinity to this pocket.

Table 9-5: Output from the F-Pocket analysis for 4th FAS-1 domain of TGFBIp-WT and mutant R555W

4 th FAS1 domain WT		4 th FAS1 domain R555W	
Pocket 1	Pocket score = 30.751 Druggability score = 0.934	Pocket 1	Pocket score = 33.981 Druggability score = 0.907
Pocket 2	Pocket score = 24.9220 Druggability score = 0.0782	Pocket 2	Pocket score = 24.6410 Druggability score = 0.0495
Pocket 3	Pocket score = 18.0877 Druggability score = 0.0505	Pocket 3	Pocket score = 23.9584 Druggability score = 0.0124
		Pocket 4	Pocket score = 23.2564 Druggability score = 0.6165

Figure 7-5: F-pocket analysis of 4thFAS-1 domain of TGFB1p-WT and R555W mutant protein. The analysis shows the presence of 3 pockets in TGFB1p-WT protein and 4 pockets in R555W mutant protein.

7.3.3 Screening of compounds using Weak Affinity Chromatography

Immobilization yield

The results of immobilization and column packing are described in [Table 7-6](#). In general, immobilization yield was satisfying and protein density could reach a high level. All obtained active columns have the capacity to detect affinities in mM range. During the packing of column R555W, there was a leak that resulted in 10 mm of the column not being filled. The accidental void was then filled with the starting diol silica and the amount of protein in the column was recalculated accordingly.

Table 7-6: Immobilization results and affinity column characteristics

Silica batch	WT	R555W	H572R	Blank
Immobilization yield (%)	88	100	71	Not applicable
Density of immobilized protein (mg/g silica); (mM)	138; 5.69	125; 5.16	47; 1.92	0
Amount of protein in column* (nmol)	55.8	40.5	18.8	0

*Assuming 5.71 mg silica stayed in the column. For R555W, calculation was based on 4.57 mg of silica.

Primary screening of mixtures

Among 2500 fragments of the Maybridge library that were screened on the 4th FAS-1 domain of TGFBIp-WT, R555W, H572R and blank columns, more than 88% were detected in each column ([Table 7-7](#)). The remaining fragments were not detectable, either due to their poor ionization in the MS, excessive forming of adduct ions that resulted in new m/z values and therefore escaped detection, or because they bound so tightly to the column that their retention time exceeded the monitoring window of 90 min (duplicate injections of 45 min each). Some fragments were detected only once in the duplicate due to a weak signal, which caused missing SD values. The interference of contaminating ions, especially from DMSO, was not considered in the counting. Note that contaminating ions from DMSO only interfere with the detection of non-retained fragments that are co-eluted with DMSO. In affinity screening by WAC, we often focus more on the retained fragments.

Table 7-7: SDs of retention times of fragments by duplicate mixture injections and detection of mixture screening on affinity columns

Column	WT	R555W	H572R	Blank
Average SD (including DMSO) (min)	0.0530 (n= 2299)	0.0584 (n = 2337)	0.0425 (n= 2243)	0.0402 (n=2393)
Highest value of SD (min)	1.415	1.596	2.351	1.477
Number of detected fragments	2252 (90 %)	2298 (92 %)	2199 (88 %)	2345 (94 %)
K _d range detected (μM)	19 - 2629	13 – 1732	9 – 1101	Not applicable

In this report, a fragment is considered as a binder when its t_{active} is equal to or higher than ten times of average SD of retention times of fragments on the column (Table 7-7). Results showed that the obtained affinity columns could reliably detect binding in the range of μM to mM. The detection capacity of the columns covered more than two orders of magnitude (Table 7-7).

Besides K_d, LE is an important parameter to consider when selecting hits. It is actually the average binding energy per heavy atom in a fragment. One small molecule that binds weakly to the protein target may have better LE in comparison with a larger molecule with higher affinity; therefore, they can be more valuable as hits. Considering the importance of LE, fragments with the highest LE value in primary mixture screening were picked out for the single solution confirmation study. Table 7-8 briefly describes the characteristics of primary hits on each target. A combination of taking into account both the LE and binding affinities, a list of about 134 fragments were selected. Since the initial screening was done in the form of 50 compounds in one mixture, the compound that gave a good affinity was also tested in the form of single compounds to confirm binding of the compounds to the proteins.

Table 7-8: Hits from the primary screening in mixtures

Target	WT	R555W	H572R
Number of selected hits	55	50	53
LE range (kcalmol ⁻¹ per heavy atom)	0.38 – 0.70	0.46 – 0.67	0.43 – 0.68

Hits from all three targets were combined together with 34 fragments that were suspected as tight binders to R555W (detected on blank and other columns but not on R555W) to form a primary mixture screening hit list. The resulting list included 134 fragments that proceeded to the single solution confirmation step.

Hit confirmation as single solutions

Analyzing samples as single solutions simplifies the readout in MS and removes the possible interaction of compounds in the mixture. This step is necessary to eliminate errors that may occur while reading MS signal from the mixture screening. Among 134 fragments that were analyzed as singletons, 110 fragments were detected on a blank column, 105 fragments on R555W and H572R, and 104 fragments on WT. The discrepancy just indicates some fragments bound strongly to the target protein so that they were not eluted, or the signals were weak and unstable in the Mass Spectrometer. A few fragments, which had been suspected as tight binders, showed up when their analysis times were extended. However, there were still some fragments that did not appear after 300 min on protein columns even though they showed good signals with the blank column. These fragments were actually good hits and, therefore, included in the final list. **Table 7-9** shows the results of the hit confirmation. More details on the structure of the compound, its retention time on the column and interaction with each of the 4th FAS-1 domain of TGFBIp proteins, the affinity of binding and Ligand efficiency of compound hits (61 compounds) can be found in **Appendix 1**.

Table 7-9: Characteristic of final hits after confirmation by analyzing single solution

Target	WT	R555W	H572R
Number of selected hits	45	59	48
Detected t_{active} range (min)	0.541 – 87.899	0.588 – 203.533	0.522 – 284.497
LE range (kcalmol ⁻¹ per heavy atom)	0.30 – > 0.61	0.30 – 0.70	0.30 – 0.78
K _d range (μM)	< 5 – 2579	< 3.4 – 1722	< 1.6 – 900

The combination of hits on all three targets resulted in a list of 61 fragments, which can proceed to the confirmation experiments with secondary

screening techniques such as NMR. Among the hits, probably compounds that bind to both mutated proteins with higher LE than LE of binding to the WT are the most interesting, considering the nature of the disease. It is observed that there was a discrepancy in the binding behaviour of fragments to the proteins, even though the proteins differed from each other in just one amino acid (point mutation). **Figure 7-6** depicts LE of the 61 final fragment hits that show good LE and binding affinity towards the 4th FAS-1 domain of WT and both the mutant proteins.

It has to be noted that only ~ 50 fragments that had highest LE in the primary mixture screening were selected for the confirmation. The constraint in number was due to confirmation capacity of WAC and NMR of our labs. The primary hit list could actually be expanded to $LE \geq 0.30 \text{ kcal mol}^{-1}$ per heavy atom, which resulted in 652 primary hits for R555W, 259 for H572R, and 240 for WT. The detailed list of all the 61 final leads is in **Appendix 1**. The initial screening by WAC showed 10 compounds that had a high-affinity binding towards both the mutants as in **Figure 7-7**. Out of the 10 compounds, 5 compounds had a common structural motif that could aid in the interaction of protein and compounds.

Figure 7-6: Ligand Efficiency (LE) values of interactions between the final hits and targets R555W, H572R and WT. An interaction should have $LE \geq 0.30 \text{ kcalmol}^{-1}$ per heavy atom and $t_{\text{active}} \geq$ ten times average SD retention times to be considered as hits.

Figure 7-7: The list of top 10 compounds that showed higher affinities towards both the 4th FAS-1 domain of TGFBIp-R555W and H572R mutant protein. The red boxed group is the common structural motif identified in the compounds.

7.3.4 Selection of compounds for validation screening by NMR

A total of about more than 50 compounds were identified by initial Weak Affinity chromatography screening for each of the WT and the 2 mutant TGFBIp. Due to the limited capacity to handle a large number of compounds screening by ^{15}N -HSQC-based NMR analysis, the top 10 compounds that showed good binding to R555W mutant protein and compounds that had good ligand efficiency towards R555W mutant were chosen for secondary screening by NMR. The top 10 compounds as listed in [Table 7-10](#) with their binding efficiencies were chosen for the secondary NMR-based screening. The elution profile based on the retention time of these 10 compounds with R555W, H572R mutant proteins and WT-TGFBIp are shown in [Figure 7-8](#).

There were no good ^{15}N -HSQC spectra available for the WT-TGFBIp and H572R mutant due to the poor stability of proteins and the greater aggregation propensity of the proteins. The proteins aggregated so quickly that by the time good ^{15}N -HSQC spectra could be recorded there was not enough soluble protein for the spectra to be resolved well without peak broadening [Figure 7-9](#). Due to the non-availability of good ^{15}N -HSQC spectra for the WT and H572R protein to observe chemical shift perturbations and to study the possible drug binding mechanism, the H572R mutant and WT-TGFBIp were not available for the secondary NMR analysis. For ^{15}N -HSQC-based secondary validation, R555W mutant protein was used. Although the 3-D structure of the 4th FAS-1 domain of WT-TGFBIp is available from PDB (protein code 2LTB) determined by solution NMR, the published constructs and our designed constructs were different (Underhaug, Koldso et al. 2013).

We also worked on the same TGFBIp amino acid sequence as published by Underhaug et al. and designed similar constructs to be cloned into a good expression vector system, but we did not succeed with protein expression and protein precipitation and aggregation issues. To overcome this problem, new constructs were designed with the introduction of GB1 protein solubility enhancement tags (SET) as published by Zhou et al. (Zhou and Wagner 2009). The introduction of such solubility enhancement tags were found to improve the solubility and stability of the proteins and the samples can be used for structural studies.

Table 7-10: List of compounds used for secondary NMR screening

No	CODE	Structure	K _d on R555W (μM)	K _d on H572R (μM)	K _d on WT (μM)	LE on R555W (kcalmol ⁻¹ per heavy atom)	LE on H572R (kcalmol ⁻¹ per heavy atom)	LE on WT (kcalmol ⁻¹ per heavy atom)
1	RJF00203		< 3	< 2	< 5	> 0.63	> 0.66	> 0.61
2	CC03813		5	2	17	0.61	0.66	0.54
3	CC07101		382	No binding	No binding	0.39	No binding	No binding
4	BTB07813		184	No binding	No binding	0.32	No binding	No binding
5	MO07617		164	No binding	No binding	0.43	No binding	No binding
6	BTB05094		< 3	< 2	< 5	> 0.50	> 0.53	> 0.49
7	BTB10358		7	4	66	0.51	0.53	0.41
8	SEW04861		< 3	< 2	< 5	> 0.54	> 0.57	> 0.52
9	RDR00814		28	36	55	0.48	0.47	0.45
10	S02118		29	44	82	0.41	0.40	0.37

Figure 7-8: Weak Affinity Chromatography- Retention time profiles of compounds and DMSO with the column immobilised with proteins of the 4th FAS-1 domain of TGFBIP H572R, WT and R555W

Weak Affinity Chromatography- Retention time profiles of compounds and DMSO with the column immobilised with proteins of the 4th FAS-1 domain of TGFB1p H572R, WT and R555W

Weak Affinity Chromatography- Retention time profiles of compounds and DMSO with the column immobilised with proteins of the 4th FAS-1 domain of TGFB1p H572R, WT and R555W

Figure 7-9: ^{15}N -HSQC spectra of 4th FAS1 domain of TGFBIp-WT and the mutant H572R protein recorded using a 700 MHz magnet and the samples in PBS buffer pH= 7.0.

Screening of Compounds by ^{15}N -HSQC NMR assays

Chemical shift mapping or chemical shift perturbation may be a very useful tool when an interaction between the protein and ligand has to be monitored or studied. The screening is usually done with a ^{15}N -labelled protein and an unlabelled chemical compound or ligand. With a good assigned ^{15}N -HSQC spectra of the target protein, it is possible to study the interaction of protein and ligand by titrating the protein with different concentrations of the ligand and recording ^{15}N -HSQC spectra at each step. An overlay of the ^{15}N -HSQC spectra of free protein over different concentrations of the ligand gives an idea of the binding and interaction. The analysis of chemical shifts of individual amino acids and following the movement of amino acid peaks upon each titration gives the binding efficiency of the compound to protein. The amino acids with the maximum change are most likely to be the binding or interaction site of the protein-ligand complex.

The top 10 compounds were used to screen against the TGFBIp-R555W mutant. The protein compound interaction of the top compounds identified by the WAC screening was verified by ^{15}N -HSQC-based NMR assays. The assays predict the exact location where the compounds could bind and the possible amino acids involved in the compound binding based on the chemical shift perturbations. Since the compounds were dissolved in DMSO, titration of the protein to various concentrations of DMSO was done, and no big changes in chemical shifts were observed for protein-DMSO interaction. Based on the WAC screening, the compounds are relatively weak binders or have a transient interaction with the protein; hence the exchange observed is a fast exchange of protein-ligand bound and unbound states. The fast exchanges between the protein and compound at the binding regions give rise to chemical shift perturbations, where an individual amino acid peak shows gradual movement upon titration with various concentrations of the ligand until all binding sites of the protein reach saturation with an excess amount of ligand.

Out of the 10 compounds analysed by NMR, compounds MO07617, RJF00203 and BTB05094 showed good chemical shift perturbations around the binding pockets as predicted by f-pocket assays. **Figures 7-10 to 7-16** show the chemical shifts for each of the amino acids along the proton dimension and also along the nitrogen dimension. The amino acid changes are also marked in the 3D structure of the mutant R555W protein as in **figures 7-19 to 7-12**.

For compound MO07617, the 1:10 protein to compound ^{15}N -HSQC chemical shift perturbations of the individual amino acids shows good binding of the compound/ligand to the protein as shown in **Figure 7-10**. Any change in the chemical shifts above +10 hertz or -10 hertz was identified as good binding, and the amino acids may be involved in an interaction of the compound/ligand with the protein, as the local interaction of the compound with the amino acids causes huge changes in the chemical shifts. The chemical shift changes along the proton dimension may be due to the interaction through a hydrogen bond. The changes may also be caused by the ring-current effect of the aromatic side chain of compounds. The chemical shifts also coincide with the predicted binding pocket of the protein. Amino acids 545-550 of the R555W mutant protein are near the binding pocket, and the chemical shift changes may be due to the compound/ligand binding to the pocket.

An observed significant change in chemical shift perturbations across both the proton and nitrogen dimensions are marked in the 3D-structure of the protein **Figure 7-10**. This local change that takes place also confirms the prediction by f-pocket analysis on the presence of a potential druggable pocket. There are other amino acids that also show some chemical shift perturbations, but with the maximum chemical shifts observed near amino acids 545 to 550 of the protein, it is likely that the interactions take place near the binding pocket of the protein.

For compound RJF00203 and BTB05094 (**Figure 7-11 and 7-12**), there are good chemical shift perturbations observed around the binding pocket near amino acids 545 to 550. The magnitude of the chemical shifts for the top three compounds near the binding pocket is quite comparable. Since a similar effect is found in 3 compounds, it is possible that these compounds interact strongly with the mutant protein. The other seven compounds (CC03813, CC07101, BTB07813, BTB10358, SEW04861, RDR00814 and S02118) showed some chemical shift perturbations that were not as intense as the first 3 compounds and did not show preferential binding near the binding pockets.

The RMSD values for each of the amino acids on how the chemical shift perturbations were extended along the proton and nitrogen dimension were calculated for all the 10 compounds as shown in **Figures 7-17 and 7-18**. Since the protein had 132 amino acids, the RMSD graphs were split into 2 parts with the first 60 amino acids in the first and the rest of the amino acids in the second part. The colour index shows the maximum chemical shift perturbations pertaining to individual amino acids. A colour

index with red colour shows no changes or very little changes in that particular amino acid when the compound/ligand is added. The blue to magenta colour index shows greater chemical shift perturbations for that particular amino acid residue. The RMSD values also conclude that the compounds MO07617, RJF00203 and BTB05094 showed good binding near the predicted binding pocket regions.

The other seven compounds identified by WAC did not show specific binding of the ligands to one of the predicted binding pockets. Although these compounds showed a greater retention time profile in the initial WAC screening, the principle of WAC screening is to identify binding or interaction between the protein and ligand and not specificity of binding or interaction. In WAC screening, many small interactions or changes in the solvent-amino acid environment throughout the surface of the protein may add up to a delay in the retention time profiles. This might have contributed to binding affinity calculations and hence identified as hits, as WAC is only able to identify compound binding but ^{15}N -HSQC can give us more information on the binding site, mode of binding and affinity of binding. This strategy might also be useful to eliminate false positives picked up by WAC.

The initial screening by WAC showed 10 compounds that have a higher binding affinity towards both the mutants (**Figure 7-8 and Table 7-10**), and a common chemical motif was identified that could possibly interact with the proteins. The 5 compounds that showed a common motif were analysed using ^{15}N -HSQC assays, but only 3 compounds with that motif showed good binding to the R555W mutant protein. Out of the 3 compounds, 2 compounds (RJF00203 and BTB05094) showed very good chemical shift perturbations along the binding pocket region. The other compounds with the similar motif (SEW04861, RDR00814 and S02118) did not have significant chemical shift perturbations along the binding pocket region. The compounds that gave a good chemical shift response are the ones also having a highly electronegative fluorine molecule in their chemical structure. The fluorine molecules might contribute to the strong binding of the ligand/compound to the mutant R555W protein

Figure 7-10: (A) Chemical shift perturbations of individual amino acid of 4th FAS-1 domain of TGFB1p with **MO07617** compound. (B) Chemical shift perturbations that showed a difference of above ± 10 Hertz along the proton dimension marked in cyan on the available 3D structure from PDB (2LTC) of TGFB1p-R555W mutant protein. The blue spheres are the predicted binding pockets by f-pocket analysis. (C) Chemical shift perturbations that showed a difference of above ± 10 Hertz along the nitrogen dimension marked in magenta on the available 3D structure from PDB (2LTC) of TGFB1p-R555W mutant protein.

Figure 7-11: (A) Chemical shift perturbations of individual amino acid of 4th FAS-1 domain of TGFB1p with RJF00203 compound. (B) Chemical shift perturbations that showed a difference of above ± 10 Hertz along the proton dimension marked in cyan on the available 3D structure from PDB (2LTC) of TGFB1p-R555W mutant protein. The blue spheres are the predicted binding pockets by f-pocket analysis. (C) Chemical shift perturbations that showed a difference of above ± 10 Hertz along the nitrogen dimension marked in magenta on the available 3D structure from PDB (2LTC) of TGFB1p-R555W mutant protein.

Figure 7-12: (A) Chemical shift perturbations of individual amino acid of 4th FAS-1 domain of TGFB1p with BTB05094 compound. (B) Chemical shift perturbations that showed a difference of above ± 10 Hertz along the proton dimension marked in cyan on the available 3D structure from PDB (2LTC) of TGFB1p-R555W mutant protein. The blue spheres are the predicted binding pockets by f-pocket analysis. (C) Chemical shift perturbations that showed a difference of above ± 10 Hertz along the nitrogen dimension marked in magenta on the available 3D structure from PDB (2LTC) of TGFB1p-R555W mutant protein.

Figure 7-13: Chemical shift perturbations of individual amino acid of 4th FAS-I domain of TGFB1p with compounds CC03813 and CC07101.

Figure 7-15: Chemical shift perturbations of individual amino acid of 4th FAS-1 domain of TGFBIp with compounds SEW04816 and RDR00814.

Figure 7-16: Chemical shift perturbations of individual amino acid of 4th FAS-1 domain of TGFB1p with compounds BTB07813 and BTB10358.

Figure 7-17: RMSD values for the first half of the 132 amino acid of TGFBIp-R555W mutant protein to top 10 compound binding validated by ^{15}N -HSQC based NMR assays.

Figure 7-18: RMSD values for the second half of the 132 amino acid of TGFBIp-R555W mutant protein to top 10 compound binding validated by ¹⁵N-HSQC based NMR assays.

7.3.5 Limited proteolysis of 4th FAS-1 domain of TGFBIp-WT, R555W and H572R proteins

The mutant protein experiences a certain level of local changes in the secondary structure due to the altered amino acid residue. This may expose certain hydrophobic residues that are otherwise buried inside the protein fold. The limited proteolysis of the 4th FAS-1 domain of TGFBIp WT and the mutants R555W and H572R may provide information on the secondary structure of the protein and the exposure of amino acid residues susceptible to proteolysis. The addition of the lead compounds from the drug screening was used to test if the addition of compounds could prevent the formation of amyloidogenic peptides in the mutant protein when the proteins are subjected to limited proteolysis by trypsin. The tryptic peptides generated from the mutant proteins may be compared to the WT protein and if the addition of compounds produced tryptic peptides similar to the WT-protein.

The addition of chemical compounds to the mutant proteins and subjecting them to limited proteolysis may be used as a therapeutic option if the compound could stabilise the secondary structure of the mutant protein and prevent the formation of amyloidogenic peptide seeds that could possibly delay the aggregation process. The 4th FAS-1 domain of TGFBIp-WT and the mutants were subjected to limited proteolysis by using trypsin enzyme with and without the lead compounds from drug screening and the peptide patterns generated by trypsin enzyme digest were analysed at different time points ([Table 7-11](#) and [Table 7-12](#)). For the R555W mutant protein, when the protein was subjected to proteolysis by trypsin, around 15 min of proteolysis produced 5 peptides that could be detected by MALDI-Mass Spectrometry. After 30 min, 45 min and 1 hour, while the number of peptides remains the same, there is a production of the highly amyloidogenic peptide sequence (E⁶¹¹PVAEPDIMATNGVVHVITNVL⁶³²) after 30 minutes of the enzyme digest. The observed peptide sequence is known to be highly amyloidogenic as predicted by various algorithms, and our group (Anandalakshmi, Murugan et al. 2017) has studied the aggregation properties of this peptide sequence in much detail. The peptide sequence of 11 reported clinical mutations responsible for the amyloid phenotype is located around this region (amino acid 611 to 632) of the protein.

When the R555W mutant protein was treated with the lead compounds (MO07617, RJF00203 and RDR00814) for 30 min at room temperature prior to the trypsin digest, the highly amyloidogenic peptide sequence was not produced. This phenomenon was observed for all the 3 lead compounds. The compounds might have interacted with the exposed hydrophobic amino acids and form chemical interaction involving the residues and protecting the amino acid from trypsin enzymatic digestion. This process of chemical stabilisation of the mutant protein with the lead compounds prevented the formation of amyloidogenic peptide seeds.

Similarly, the H572 mutant protein was subjected to limited proteolysis by trypsin, and the peptide fragments were analysed by MALDI mass spectrometry. The mutant H572R protein in the absence of any compounds produced 4 different peptides 15 minutes after enzyme digestion. The longer amyloidogenic peptide (N⁶⁰³NVVSVNKEPVAEPDIMATNGVVHVITNVL⁶³²) was observed. The same peptide fragments were observed after 30 min, 45 min and 1 hour, and they included the longer amyloidogenic peptide. This peptide is an extended version of the amyloidogenic peptide observed in the R555W mutant (E⁶¹¹PVAEPDIMATNGVVHVITNVL⁶³²). When the H572R mutant protein is treated with compound MO07617, there is still a generation of the amyloidogenic peptide, whereas when the protein is treated with the other 2 compounds RJF00203 and RDR00814, the long amyloidogenic fragment is not generated 45 minutes after the enzyme digest. The compound MO07617 does not bind to the H572R protein according to the initial WAC screening, whereas the other 2 compounds bind with a higher affinity to the mutant protein.

The ability of the compounds to prevent the formation of highly amyloidogenic peptide seeds can be explored as a possible treatment option. As previously reported by our group in the mass spectrometry analysis of the protein composition of the amyloid corneal deposits from dystrophic patients, we identified the serine protease HTRA1 to be highly enriched in the amyloid deposits. These compounds could be used as a potential modulator of HTRA1 protease activity when they bind to mutant proteins and do not expose amino acid residues susceptible to enzyme proteolysis activity.

Table 7-11: Peptide fragments generated by limited proteolysis of R555W with trypsin and analysis by MALDI mass fingerprinting

	15 min	30 min	45 min	1 hour
R555W No compound	SLQGDKLEVSLK EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR EPVAEPDIMATNGVVHVITNVL NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR EPVAEPDIMATNGVVHVITNVL NNVSVNKEPVAEPDIMATNGVVHVITNVL	ELANILK EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR EPVAEPDIMATNGVVHVITNVL
R555W with MO07617	ALPPREWSR EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL
R555W with RJF00203	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL
R555W with RDR00814	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	SLQGDKLEVSLK EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	ELANILK SLQGDKLEVSLK EGVYTVFAPTNEAFR YHIGDEILVSGGIGALVR FSMLVAAIQSAGLTETLNR

Table 7-12: Peptide fragments generated by limited proteolysis of H572R with trypsin and analysis by MALDI mass fingerprinting

	15 min	30 min	45 min	1 hour
H572R No compound	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL
H572R with MO07617	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL
H572R with RJF00203	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR
H572R with RDR00814	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR NNVSVNKEPVAEPDIMATNGVVHVITNVL	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR	IGDEILVSGGIGALVR EGVYTVFAPTNEAFR FSMLVAAIQSAGLTETLNR

7.4 Discussion

Any new remedy to treat *TGFBI*-associated corneal dystrophy is the goal of every corneal researcher, as corneal transplant is the option primarily sought for to ameliorate the disease condition. Most of the treatment options are invasive and require the time and effort of a trained surgeon to perform the procedures leading to increased patient care costs. Even after the corneal transplants, there is a problem of recurrence of the corneal protein deposits (Ellies, Renard et al. 2002) based on the type of mutation the patient has inherited. A new method of treating the disease should be simple to handle, or possibly a topical application of a drug administered by the patients themselves should be sought after.

In order to develop a treatment methodology, the disease mechanism and pathology of the disease should be studied well. Since only limited information is available on the stromal-Bowman membrane *TGFBI*- associated corneal dystrophies, the aim of the thesis was to understand the cornea-specific mutant protein aggregation more. When the project started, the only 3D structure of TGFBIp available on PDB was the low resolution of the 4th FAS-1 domain of WT-TGFBIp (2VXP and 1X3B). Later, 3D structures of the 4th Fas-1 domain of TGFBIp -WT and R555W mutant (2LTB and 2LTC) were available on PDB and published (Underhaug, Koldso et al. 2013). The publication reported the presence of a hydrophobic cavity near the helices $\alpha 3$, $\alpha 3'$, and $\alpha 4$ forming a “helix-triangle module” in the protein. The publication also suggested that the stability of the mutant R555W protein increased when the side chain of tryptophan fully occupied the cavity, whereas in the WT protein the arginine residue was away from the cavity (Underhaug, Koldso et al. 2013). The finding made it interesting to work on the drug discovery approach by filling the available cavity of the protein since many of the mutant proteins of TGFBIp are known to have local changes in the secondary structure near the mutation site and unstable upon substitution of a different amino acid.

The first aim of the study was to identify any possible cavity with a good druggability score. The published structure of TGFBIp WT and R555W were used as input for the pocket prediction by f-pocket analysis. The study identified 3 pockets for WT and 4 pockets for R555W with one pocket each having a good druggability score near the helix-triangle region. The presence of a good binding pocket made it possible

for a compound that could fill the cavity of the protein to thermodynamically stabilise the protein or possibly prevent/delay protein aggregation. The Ro3 (rule of three) library with 2500 compounds was used for screening with the 4th FAS-1 domain of TGFBIp WT and mutants R555W and H572R.

The transient interaction of the 2500 compounds (50 mixtures with 50 compounds in each mixture with differences in the masses) was screened using Weak affinity chromatography against WT and the 2 mutants. The initial screening identified about 134 lead compounds from the mixtures, and these compounds were also verified as singletons to confirm protein binding. The singleton assays confirmed about 61 lead compounds. The binding of top 10 compounds with good LE and affinity towards R555W mutant was also verified by ¹⁵N-HSQC NMR assays. A good HSQC spectrum was not available for ¹⁵N-HSQC NMR assays for WT and H572R although the 3D-NMR structure for the WT protein was published by Underhaug et al. We also used similar constructs with different expression systems as reported in the publication but experienced problems with protein purification. The protein was not very stable in PBS solution in which all the drug binding ¹⁵N-HSQC NMR assays were tested.

The secondary NMR assays using chemical shift perturbations were analysed for each of the compounds to R555W mutant titration at different compound concentrations. The perturbations for 1:10 protein to compound ratio were used for RMSD calculations. The secondary NMR assays were useful in eliminating false positives produced by WAC screening and identified 3 lead compounds that bound with greater affinity near the binding pockets predicted by f-pocket analysis. The proteins were tested for melting temperature (T_m) with and without the presence of compounds, but due to the nature of the proteins and the compounds both Circular dichroism assays and sypro-orange dye assays were not able to determine the T_m of the proteins.

The 3 lead compounds were added to the protein, a limited proteolysis of the protein using trypsin enzyme was done for a maximum of one hour, and the peptide fragments were analysed by MALDI-mass spectrometry. The addition of lead compounds to both the mutants prevented the formation of amyloidogenic peptide sequences, which can trigger the whole aggregation cascade. For a protein to act as a substrate for any proteolytic enzyme, the peptide stretch of the substrate should be available to fit into the active site of the enzyme. In a well-folded protein, the stereochemistry of the protein cannot be modified to fit the protease activity due to its rigid structure; hence most of the well-folded proteins are resistant to proteolysis

activity (Fontana, de Laureto et al. 2004). Proteolysis does not occur at random places on the protein surface but in very specific loops that offer conformational flexibility for the binding of the protease to the protein. Hence these experiments are used to probe the local secondary structure of the protein.

The addition of compounds to the mutant proteins stabilises the local secondary structure on the surface of the protein. The compounds may have made new chemical bonds or interactions with amino acids along the binding pocket (which is the target in this study), thereby not providing access to the protease to act on the protein. This may prevent the formation of amyloid peptides with high aggregation propensity that could, in turn, act as amyloid fibril seeds to accelerate amyloid fibril formation or protein aggregation.

These 3 lead compounds could be tested against the proteolytic processing of HTRA1 since HTRA1 is known to process the mutant protein (Karring, Runager et al. 2012, Karring, Poulsen et al. 2013, Poulsen, Runager et al. 2014, Courtney, Toftgaard Poulsen et al. 2015, Anandalakshmi, Murugan et al. 2017). The exploration of lead compounds to prevent the formation of amyloidogenic peptides or help in the protein turnover and processing of the mutant protein would be interesting. Such a compound could be used in treating the protein aggregation related to TGFBI-corneal dystrophies. The secondary screening of the 4th FAS-1 domain of mutant H572R and WT with the lead compounds could also provide more clues on compound binding to the mutant whose 3D structure is not available as of now to do a detailed drug binding study. The new constructs, if stable, can also be used for the 3D structure determination of mutant H572R with and without the compound bound forms. Since the drug discovery process is long and time-consuming, the aim of the thesis is to identify lead compounds that could be taken to the next stages of the drug discovery process.

Chapter 6

8 Future Work and conclusion

Corneal Dystrophy is a unique protein aggregation disorder and much was not known about the disease pathology and the causes for protein aggregation. Here in this report we tried to explore more on the disease mechanism by analyzing the composition of patient amyloid corneal deposits, characterize highly amyloidogenic TGFBIp peptides that could serve as a model to test inhibitory compounds and a drug discovery approach to prevent protein aggregation. There are always possibilities to explore each of these aims in much larger details. Thus a list of possible future work options is discussed for each of the aims reported in this thesis.

8.1 Analysis of protein composition of the amyloid corneal deposits from corneal dystrophy patients

The composition of amyloid protein deposits from corneal dystrophy patients with H626R and R124C mutations were analysed using LC-MS/MS based mass spectrometry. The study reports the presence of several amyloidogenic proteins, non-fibrillar amyloidogenic proteins and TGFBIp as the major constituent of the corneal amyloid deposits. We also identified the possible involvement of serine protease HTRA1 in the processing of mutant TGFBIp. From the mass spectrometry based analysis of the trypsin digestion of TGFBIp, many peptide sequences were found to be enriched in amyloid deposits from patient samples compared to the healthy controls. Three peptide sequences (G⁵¹⁵DNRFSMLVAAIQSAGLTETLN⁵³², Y⁵⁷¹HIGDEILVSGGIGALVR⁵⁸⁸ and E⁶¹¹PVAEPDIMATNGVVHVITNVLQPPANRPQER⁶⁴²) were of higher abundance and could be mapped to the 4th FAS-1 domain of the protein. This region is highly aggregation prone and contains more than 80% of the reported mutations.

The project could be expanded by studying the composition of corneal protein deposits from other mutations not reported thus far. The mutations may include amyloidogenic, non amyloidogenic and mixed phenotypes. The characterization of other tryptic peptides that were found to be enriched in patient deposits is another possible option to know more about the amyloidogenic peptide fragments, their aggregation propensity and how these peptide fragments can act as potential amyloid

fibril seeds and aid in the aggregation of the mutant proteins. The involvement of HTRA1 processing of the amyloid mutant protein could also be studied in much detail as no such data on HTRA1 processing on the over-expressed TGFBIp is available. This data could be useful to understand the proteolytic processing of the mutant proteins by serine protease HTRA1.

8.2 Bio-physical and structural characterization of peptide fragments derived from 4th FAS-1 domain of TGFBIp

It has been hypothesized that the shorter peptide fragments have greater aggregation potential than the larger proteins. In this report, mass spectrometry analysis of the composition of corneal deposits from dystrophic patients with H626R and R124C mutations was analysed. Tryptic digestion of the amyloid corneal deposits showed enrichment of certain peptide fragments in dystrophic patients compared to the normal controls. Detailed biophysical characterization of one of the peptide sequence (TGFBIp- E⁶¹¹PVAEPDIMATNGVVHVITNVLQ⁶³²) that has a greater aggregation potential predicted by several prediction software was done and reported (Anandalakshmi, Murugan et al. 2017). Bio physical characterization of the peptide by circular dichroism, ThT assays and TEM assays were done. Reports on the Thermal stability of 2 clinically relevant peptides (TGFBIp⁶¹¹⁻⁶³³ G623R and TGFBIp⁶¹¹⁻⁶³³ N622K) mutations were assessed by CD and TEM. We have reported the first solid state-NMR spectra of the amyloid fibrils derived from one of the most stable mutant peptide (G623R) and the secondary structure of the amyloid fibrils were also reported. Cytotoxicity assays of these amyloid fibrils on the Human corneal fibroblast cells were also studied to understand the toxicity of mutant peptide oligomers vs amyloid fibrils.

Since uniform Isotopic labelling of all amino acids in the peptide sequence were very expensive when synthetically produced, we labelled only a few amino acids and hence were not able to report the complete 3D structure of the amyloid fibrils from the stable G623R mutant. We have designed an over expression of the peptide sequence model using pET-28 expression system. The protein expression and purification protocols have been optimized (Figure 8-1 to 8-4). Uniform ¹⁵N and ¹³C isotope labelled protein expression and purification protocol optimization is underway. This expression system can produce uniformly labelled G623R mutant peptide and this can

be used to monitor amyloid fibrillation process. The amyloid fibrils derived from this peptide can be used for structure determination by solid state NMR.

The expression system can also be extended to other clinically relevant mutant peptides and the amyloid fibrillation process could be monitored and the structure of amyloid fibrils can be determined by solid state NMR. This recombinant short peptides can also be used as a model system to screen for any drug candidates that could delay the aggregation process or disaggregate the amyloid fibrils.

pET-28a (+) Vector

Aminoacid Sequence of TGFBIp:

MGSSHHHHSSGLVPRGSHMENLYFQGDKLEVSLKNNVSVNKEP
VAEPDIMATNRVHVITNVLQ

Figure 8-1: Vector map of pET 28a (+) vector and the sequence of TGFBI used in the study

Figure 8-2: Expression optimization of target protein with Histidine tag in E-coli BL21 (DE3).
Expression optimization of target protein with tag in E-coli BL21 (DE3).

Figure 8-3: SDS-PAGE analysis of 2ug purified protein with Histidine tag

M- Prestained protein marker

1- The target protein.

Figure 8-4: Western blot analysis of the purified protein using anti-His tag antibody. Goat anti-rabbit IgG-HRP was used as secondary antibody.

8.3 Screening of compounds to prevent protein aggregation by modulating protein stability and prevent formation of amyloidogenic peptide fragments

In the drug discovery approach, we identified a possible drug binding pocket with a good druggability score for both the 4th FAS-1 domain WT and R555W mutant protein as the 3D structure was available only for the two proteins. The primary drug discovery approach was to screen 2500 compounds from the Maybridge library using Weak affinity chromatography (WAC) method and to identify lead compounds that show good binding to WT and the 2 mutant proteins (R555W and H572R). We identified about 61 lead compounds to be verified by secondary ¹⁵N-HSQC NMR based assays. TGFBIp protein is a difficult protein to over-express and purify from E.coli cells. The protein has high aggregation potential and is highly instable in solution, hence we could not have good HSQC spectra for WT and H572R proteins.

The top compounds that showed good binding to R555W mutant with good LE were chosen for the secondary analysis by NMR. Three lead compounds which shows good binding near the binding pocket region as predicted by f-pocket analysis was identified and among the 10 compounds, false positives predicted by WAC were also eliminated. Limited proteolysis assays to understand the mechanism of protein unfolding and proteolysis using trypsin enzyme were also studied. The three top compounds were used in limited proteolysis assay by trypsin and the tryptic peptides generated by the enzyme digest were analysed by MALDI-peptide fingerprinting. The addition of compounds prevented the formation of highly amyloidogenic peptide fragments.

The secondary drug validation by ¹⁵N-HSQC NMR based assays requires stable proteins in solution throughout the experimentation process. We have designed new construction for both 4th FAS-1 domain of TGFBIp-WT and R555W mutant with solubility enhancement tags as recommended by Wagner et al. The optimization on the production and purification of proteins for the new H572R and WT protein is underway. These new constructs will be used to re-screen and validate the compound binding by ¹⁵N-HSQC NMR based assays. Limited proteolysis of mutant proteins with lead compounds showed the possible prevention of protein aggregation by restricting the generation of highly amyloidogenic peptide seeds. The peptide seeds could possibly

contribute to rapid aggregation as shorter peptides have greater aggregation potential. The lead compounds is being tested on the limited proteolysis assay using serine protease HTRA1. The mass spectrometry analysis showed the possible involvement of HTRA1 in processing the mutant protein. In such a case if the lead compound could inhibit the activity of HTRA1 and amyloidogenic peptide generation, that compound could be a therapeutic lead compound. Such a compound could be extended to increase binding affinity and activity and taken to the next stages of drug discovery program.

Measuring the Melting temperature (T_m) of protein in the presence and absence of compounds were measured by circular dichroism and fluorimetric sypro orange assays. But due to the nature of the protein, the T_m of the WT and mutant protein could not be determined in the first place (results not included) without the drug. Although the circular dichroism assays were able to determine the T_m of the proteins alone, with the addition of compounds, the spectrum became noisy due to the absorption induced by the compounds in the near and far UV-circular dichroism measurements. Other techniques are being explored to check if the addition of compounds affected the T_m of the mutant proteins.

Chapter 7

Publications related to the project

1. Lakshminarayanan, R., S. S. Chaurasia, **V. Anandalakshmi**, S. M. Chai, E. Murugan, E. N. Vithana, R. W. Beuerman and J. S. Mehta (2014). "Clinical and genetic aspects of the TGFBI-associated corneal dystrophies." *Ocul Surf* **12**(4): 234-251.
2. Lakshminarayanan, R., S. S. Chaurasia, E. Murugan, **A. Venkatraman**, S. M. Chai, E. N. Vithana, R. W. Beuerman and J. S. Mehta (2015). "Biochemical properties and aggregation propensity of transforming growth factor-induced protein (TGFBIp) and the amyloid forming mutants". *Ocul Surf* **13**(1): 9-25.
3. Murugan, E., **A. Venkatraman**, Z. Lei, V. Mouvet, R. Rui Yi Lim, N. Muruganatham, E. Goh, G. Swee Lim Peh, R. W. Beuerman, S. S. Chaurasia, L. Rajamani and J. S. Mehta (2016). "pH Induced Conformational Transitions in the Transforming Growth Factor beta-Induced Protein (TGFbetaIp) Associated Corneal Dystrophy Mutants". *Sci Rep* **6**: 23836.
4. **Anandalakshmi, V.**, E. Murugan, E. G. T. Leng, L. W. Ting, S. S. Chaurasia, T. Yamazaki, T. Nagashima, B. L. George, G. S. L. Peh, K. Pervushin, R. Lakshminarayanan and J. S. Mehta (2017). "Effect of position-specific single-point mutations and biophysical characterization of amyloidogenic peptide fragments identified from lattice corneal dystrophy patients". *Biochem J* **474**(10): 1705-1725.
5. **Venkatraman Anandalakshmi** ; Dutta, Bamaprasad; Murugan, Elavazhagn; Piliang, Hao ; Lakshminarayanan, Rajamani; Sook Yee , Anita; Pervushin, Konstantin; Sze, Siu Kwan; mehta, Jodhbir. "Proteomic analysis of amyloid corneal aggregates from TGFBI-H626R Lattice corneal dystrophy patient implicates serine-protease HTRA1 in mutation-specific pathogenesis of TGFBIp". *J Proteome Res.* 2017 Jul 20. doi: 10.1021/acs.jproteome.7b00188. [Epub ahead of print]
6. **Venkatraman Anandalakshmi**, MinH-Dao Duang-Thi, Bamaprasad Dutta, Andreas Larsson, Sze Siu Kwan, Sten Ohlson, Konstantin Pervushin and J.S. Mehta. "Prevention of aggregation of mutant TGFBIp with compounds that modulate protein stability and proteolytic pathway". (Manuscript under preparation)

Posters and abstracts

1. Biomedical Engineering Society (Singapore) 9th Scientific Meeting (BES9SM), 16 May 2015
Topic: TGFBIp associated Corneal Dystrophy- exploring structure-based drug development strategies for disease prevention and treatment
Authors: **Venkatraman Anandalakshmi**, Konstantin Pervushin and Jodhbir S. Mehta

2. 9th Alpine Congress on solid state NMR (Chamonix Mont-Blanc, France), 13 - 17 September 2015
Topic: TGFBIp associated Corneal Dystrophy- structure determination of amyloid fibrils by Solid State NMR
Authors: **Venkatraman Anandalakshmi**, Konstantin Pervushin and Jodhbir S. Mehta
3. SingHealth Duke-NUS Scientific Congress 2016 (Singapore) 23-24 September 2016
Topic: Investigation of the in-vitro aggregation process of TGFBIp derived peptides and intermediate aggregates related to Corneal Dystrophy
Authors: **Venkatraman Anandalakshmi**, Elavazhagan Murugan, Lakshminarayanan R, Konstantin Pervushin, Jodhbir S. Mehta
4. Science Improving Lives Symposium 2017 (Singapore), 27th February 2017
Topic: Analysis of Composition and proteolytic processing of amyloid corneal aggregates associated with mutations in the TGFBI gene.
Authors: **Venkatraman Anandalakshmi**, Lakshminarayanan R, Konstantin Pervushin, Siu Kwan Sze, Jodhbir S. Mehta

Awards

1. Singapore Eye Research Institute- graduate Scholarship (2013-2017)
2. Outstanding Graduate Student (2014)
3. Research Talent Development fund (TDF) to fund conference travel to 9th Alpine Congress on solid state NMR (Chamonix Mont-Blanc, France), 13 - 17 September 2015
4. Second runner-up NTU-3 Minute Thesis Competition, Singapore (2017)
5. Runner-up in the National level 3 Minute Thesis Competition, Singapore (2017)

Pilot Research funding with author as co-PI

1. SNEC Health Research Endowment Fund (R1071/86/2013), 2014, “Analysis of Composition and processing of corneal aggregates associated with mutations in the TGFBI gene”. (20,000 SGD)
2. SNEC Health Research Endowment Fund (R1076/91/2013), 2014, “Structure and Biophysics driven Ligand screening for potential compounds that modulate TGFBIp aggregates”. (20,000 SGD)
3. SNEC Health Research Endowment Fund (R1209/15/2015), 2015 “TGFBIp associated Corneal Dystrophy-structure determination of amyloid fibrils by Solid State NMR” (20,000 SGD)
4. SNEC Health Research Endowment Fund (R1265/71/2015), 2015, “Analysis of protein, lipid and metabolite composition of corneal aggregates associated with mutations in the TGFBI gene by Matrix-Assisted Laser Desorption Ionization (MALDI) Imaging Mass Spectrometry (IMS)”. (20,000 SGD)

5. SNEC Health Research Endowment Fund (R1376/62/2016), 2016, “Characterization of peptides enriched in the corneal deposits of TGFBI associated corneal dystrophy patients”. (20,000 SGD)
6. SNEC Health Research Endowment Fund (R1377/63/2016), 2016, “Metagenomics survey of microbial population present in the corneal epithelium of TGFBI associated corneal dystrophy patients”. (20,000 SGD)
7. SNEC Health Research Endowment Fund (R1429/12/2017), 2017, 2017, “Identification of lead compounds that increase stability of TGFBIp and inhibit HTRA1 enzymatic activity in Lattice Corneal Dystrophy patients”. (20,000 SGD)

Appendix

Appendix I - WAC on TGFBIp- Primary screening results
CONFIRMED HITS FROM WAC

No	CODE	Structure	t _R on blank (min)	t _{active} on R555W (min)	t _{active} on H572R (min)	t _{active} on WT (min)	K _d on R555W (μM)	K _d on H572R (μM)	K _d on WT (μM)	LE on R555W (kcalmol ⁻¹ per heavy atom)	LE on H572R (kcalmol ⁻¹ per heavy atom)	LE on WT (kcalmol ⁻¹ per heavy atom)
1	RJF00203		2.885	> 300	> 300	> 300	< 3	< 2	< 5	> 0.63	> 0.66	> 0.61
2	CC09513		0.424	3.74	4.562	1.478	271	103	944	0.70	0.78	0.59
3	SB01744		49.136	1.98	34.551	37.043	511	14	38	0.41	0.61	0.55
4	SEW04211		4.665	57.872	>300	87.899	17	< 2	16	0.54	> 0.66	0.55

5	CC03813		3.27	203.533	284.497	81.782	5	2	17	0.61	0.66	0.54
6	DP01481	Br H	0.612	0.614	0.630	0.75	1649	746	1860	0.54	0.61	0.53
7	SEW04861		1.993	> 300	> 300	> 300	< 3	< 2	< 5	> 0.54	> 0.57	> 0.52
8	CC03613		0.607	2.779	5.344	1.403	364	88	994	0.59	0.69	0.51
9	SB01578	Cl H Cl H	0.845	1.839	2.914	1.309	551	161	1066	0.56	0.65	0.51
10	BTB05094		3.788	> 300	> 300	> 300	< 3	< 2	< 5	> 0.50	> 0.53	> 0.49

11	SB02117		0.331	34.39	8.389	2.183	29	56	639	0.69	0.65	0.49
12	AC39765		0.507	Un detectable	3.471	4.8	Not available	135	291	Not available	0.53	0.48
13	CC13813		3.905	3.166	4.552	4.255	320	103	328	0.48	0.55	0.48
14	AC10948		0.423	4.042	1.163	1.819	250	404	767	0.55	0.52	0.47
15	TL00533		0.401	1.326	0.646	0.783	764	728	1782	0.53	0.54	0.47
16	CC29914		0.233	0.588	0.357	0.346	1722	No binding	No binding	0.54	No binding	No binding

17	CC23613		0.58	0.744	0.695	0.741	1361	676	1883	0.49	0.54	0.47
18	CD08326		0.027	1.44	1.053	1.355	703	446	1030	0.48	0.51	0.45
19	CC55513		0.337	1.06	0.637	0.595	955	738	2345	0.52	0.54	0.45
20	RDR00814		5.507	36.278	13.067	25.34	28	36	55	0.48	0.47	0.45
21	SB00549		1.25	2.364	0.994	1.206	428	473	1157	0.51	0.51	0.45
22	SPB02126		0.164	0.782	0.332	0.541	1295	No binding	2579	0.49	No binding	0.44
23	SEW04056		1.155	24.094	24.960	4.102	42	19	340	0.55	0.59	0.43

24	AC39241		0.14	1.428	0.661	0.957	709	711	1458	0.48	0.48	0.43
----	---------	---	------	-------	-------	-------	-----	-----	------	------	------	------

25	AC42628		0.59	1.732	0.826	0.945	585	569	1476	0.49	0.49	0.43
26	MO01163		1.524	10.621	4.789	6.961	95	98	200	0.46	0.46	0.42
27	AC39711		0.008	3.964	1.047	0.396	255	449	No binding	0.62	0.57	No binding
28	CD09752		1.65	3.655	1.623	2.958	277	290	472	0.44	0.44	0.41
29	TL00368		0.573	0.296	0.651	0.709	No binding	722	1968	No binding	0.48	0.41
30	CC24814		0.62	2.537	0.708	1.401	399	664	996	0.47	0.44	0.41

31	BTB10358		1.237	151.762	120.923	21.209	7	4	66	0.51	0.53	0.41
32	SEW02215		3.109	12.306	5.689	10.004	82	83	139	0.43	0.43	0.41
33	AC44411		0.795	1.014	0.741	1.072	999	634	1301	0.41	0.44	0.40
34	SB01940		0.837	13.664	1.644	3.493	74	286	399	0.47	0.40	0.39
35	AC12636		0.759	8.587	1.322	1.795	118	356	777	0.49	0.43	0.39
36	CC46114		0.804	4.529	2.042	3.073	224	230	454	0.42	0.42	0.38
37	CC48414		1.467	7.737	1.420	3.039	131	331	459	0.44	0.40	0.38

38	RDP00249		0.752	3.151	0.818	1.6	321	575	872	0.44	0.40	0.38
39	NH00697		4.267	33.399	5.771	9.772	30	81	143	0.44	0.40	0.38
40	S02118		3.139	34.503	10.575	17.1	29	44	82	0.41	0.40	0.37
41	RJC01075		0.277	0.609	0.217	0.38	1663	No binding	No binding	0.42	No binding	No binding
42	SEW05368		1.827	27.944	2.515	4.141	36	187	337	0.47	0.39	0.37
43	AC33596		0.23	6.727	2.270	0.606	151	207	2302	0.52	0.50	0.36

44	CC75946		1.152	24.339	27.003	6.731	42	17	207	0.43	0.47	0.36
45	BTB09946		3.894	107.102	10.115	18.763	9	46	74	0.43	0.37	0.35
46	CC29946		0.722	0.782	0.230	0.231	1295	No binding	No binding	0.44	No binding	No binding
47	HAN00443		3.431	30.432	5.456	9.842	33	86	142	0.38	0.35	0.33
48	TL00552		0.285	0.621	0.171	0.349	1630	No binding	No binding	0.38	No binding	No binding
49	MAY00295		0.09	0.702	0.301	0.341	1442	No binding	No binding	0.39	No binding	No binding
50	MO07019		0.426	1.41	0.522	0.838	718	900	1665	0.36	0.35	0.32

51	AC40357		0.936	7.859	1.217	2.427	129	386	575	0.38	0.33	0.32
52	AC43278		0.092	5.238	1.407	0.593	193	334	2352	0.42	0.40	0.30
53	CC35523		2.343	55.909	1.666	3.868	18	282	361	0.41	0.30	0.29
54	S06035		0.248	1.457	0.105	0.298	695	No binding	No binding	0.39	No binding	No binding
55	BTB03828		0.935	5.149	1.148	2.201	197	409	634	0.32	0.29	0.27
56	CD11814		0.35	1.536	0.125	0.319	659	No binding	No binding	0.36	No binding	No binding

57	MO07617		0.498	6.165	0.137	0.299	164	No binding	No binding	0.43	No binding	No binding
58	JFD01174		0.212	0.67	0.248	0.378	1511	No binding	No binding	0.30	No binding	No binding
59	SPB03698		2.076	44.126	0.749	2.253	23	628	619	0.35	0.24	0.24
60	BTB07813		0.561	5.488	0.214	0.497	184	No binding	No binding	0.32	No binding	No binding
61	CC07101		0	2.65	0.002	0.017	382	No binding	No binding	0.39	No binding	No binding

Reference

1. Afshari, N. A., R. P. Bahadur, D. E. Eifrig Jr, I. B. Thøgersen, J. J. Enghild and G. K. Klintworth (2008). "Atypical asymmetric lattice corneal dystrophy associated with a novel homozygous mutation (Val624Met) in the TGFBI gene." Molecular Vision **14**: 495-499.
2. Afshari, N. A., J. E. Mullally, M. A. Afshari, R. F. Steinert, A. P. Adamis, D. T. Azar, J. H. Talamo, C. H. Dohlman and T. P. Dryja (2001). "Survey of patients with granular, lattice, avellino, and Reis-Bücklers corneal dystrophies for mutations in the BIGH3 and gelsolin genes." Archives of Ophthalmology **119**(1): 16-22.
3. Akiya, S. and S. I. Brown (1970). "Granular dystrophy of the cornea. Characteristic electron microscopic lesion." Arch Ophthalmol **84**(2): 179-192.
4. Aldave, A. J. (2007). "Elucidating the Molecular Genetic Basis of the Corneal Dystrophies." Archives of Ophthalmology **125**(2): 177.
5. Aldave, A. J. (2008). "A Novel Variant of Combined Granular-Lattice Corneal Dystrophy Associated With the Met619Lys Mutation in the TGFBI Gene." Archives of Ophthalmology **126**(3): 371.
6. Aldave, A. J., S. A. Rayner, J. A. King, J. A. Affeldt and V. S. Yellore (2005). "A unique corneal dystrophy of Bowman's layer and stroma associated with the Gly623Asp mutation in the transforming growth factor β -induced (TGFBI) gene." Ophthalmology **112**(6): 1017-1022.
7. Aldave, A. J., B. Sonmez, S. L. Forstot, S. A. Rayner, V. S. Yellore and B. J. Glasgow (2007). "A clinical and histopathologic examination of accelerated TGFBIp deposition after LASIK in combined granular-lattice corneal dystrophy." Am J Ophthalmol **143**(3): 416-419.
8. Anandalakshmi, V., E. Murugan, E. G. T. Leng, L. W. Ting, S. S. Chaurasia, T. Yamazaki, T. Nagashima, B. L. George, G. S. L. Peh, K. Pervushin, R. Lakshminarayanan and J. S. Mehta (2017). "Effect of position-specific single-point mutations and biophysical characterization of amyloidogenic peptide fragments identified from lattice corneal dystrophy patients." Biochem J **474**(10): 1705-1725.
9. Andersen, R. B., H. Karring, T. Møer-Pedersen, Z. Valnickova, I. B. Thøgersen, C. J. Hedegaard, T. Kristensen, G. K. Klintworth and J. J. Enghild (2004). "Purification and structural characterization of transforming growth factor beta induced protein (TGFBIp) from porcine and human corneas." Biochemistry **43**(51): 16374-16384.
10. Anderson, N. J., R. F. Beran and T. L. Schneider (2002). "Epi-LASEK for the correction of myopia and myopic astigmatism." J Cataract Refract Surg **28**(8): 1343- 1347.
11. Andreasen, M., S. B. Nielsen, K. Runager, G. Christiansen, N. C. Nielsen, J. J. Enghild and D. E. Otzen (2012). "Polymorphic fibrillation of the destabilized fourth fasciclin-1 domain mutant A546T of the Transforming growth factor-beta-induced protein (TGFBIp) occurs through multiple pathways with different oligomeric intermediates." J Biol Chem **287**(41): 34730-34742.
12. Auw-Haedrich, C., H. Agostini, I. Clausen, T. Reinhard, P. Eberwein, D. F. Schorderet and C. Gruenauer-Kloevekorn (2009). "A Corneal Dystrophy Associated with Transforming Growth Factor β -Induced Gly623Asp Mutation. An Amyloidogenic Phenotype." Ophthalmology **116**(1): 46-51.
13. Awwad, S. T., M. A. Di Pascuale, R. N. Hogan, S. L. Forstot, J. P. McCulley and H. D. Cavanagh (2008). "Avellino Corneal Dystrophy Worsening after Laser In

- Situ Keratomileusis: Further Clinicopathologic Observations and Proposed Pathogenesis." American Journal of Ophthalmology **145**(4): 656-661.
14. Ayres, B. D. and C. J. Rapuano (2006). "Excimer Laser Phototherapeutic Keratectomy." The Ocular Surface **4**(4): 196-206.
 15. Baldwin, A. J., T. P. Knowles, G. G. Tartaglia, A. W. Fitzpatrick, G. L. Devlin, S. L. Shammas, C. A. Waudby, M. F. Mossuto, S. Meehan, S. L. Gras, J. Christodoulou, S. J. Anthony-Cahill, P. D. Barker, M. Vendruscolo and C. M. Dobson (2011). "Metastability of native proteins and the phenomenon of amyloid formation." J Am Chem Soc **133**(36): 14160-14163.
 16. Basaiawmoit, R. V., C. L. Oliveira, K. Runager, C. S. Sørensen, M. A. Behrens, B. H. Jonsson, T. Kristensen, G. K. Klintworth, J. J. Enghild, J. S. Pedersen and D. E. Otzen (2011). "SAXS models of TGFBIp reveal a trimeric structure and show that the overall shape is not affected by the Arg124His mutation." J Mol Biol **408**(3): 503-513.
 17. Bemporad, F., G. Calloni, S. Campioni, G. Plakoutsi, N. Taddei and F. Chiti (2006). "Sequence and structural determinants of amyloid fibril formation." Accounts of Chemical Research **39**(9): 620-627.
 18. Bergström, J., C. Murphy, M. Eulitz, D. T. Weiss, G. T. Westermark, A. Solomon and P. Westermark (2001). "Codeposition of apolipoprotein A-IV and transthyretin in senile systemic (ATTR) amyloidosis." Biochem Biophys Res Commun **285**(4): 903-908.
 19. Bergström, J., C. L. Murphy, D. T. Weiss, A. Solomon, K. Sletten, U. Hellman and P. Westermark (2004). "Two different types of amyloid deposits--apolipoprotein A-IV and transthyretin--in a patient with systemic amyloidosis." Lab Invest **84**(8): 981-988.
 20. Bettelheim, F. A. and B. Plessy (1975). "The hydration of proteoglycans of bovine cornea." Biochimica et Biophysica Acta (BBA) - General Subjects **381**(1): 203-214.
 21. Billings, P. C. (2002). "The Transforming Growth Factor-beta -inducible Matrix Protein beta ig-h3 Interacts with Fibronectin." Journal of Biological Chemistry **277**(31): 28003-28009.
 22. Billings, P. C., D. J. Herrick, U. Kucich, B. N. Engelsberg, W. R. Abrams, E. J. Macarak, J. Rosenbloom and P. S. Howard (2000). "Extracellular matrix and nuclear localization of β ig-h3 in human bladder smooth muscle and fibroblast cells." Journal of Cellular Biochemistry **79**(2): 261-273.
 23. Biswas, S. (2001). "Missense mutations in COL8A2, the gene encoding the alpha2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy." Human Molecular Genetics **10**(21): 2415-2423.
 24. Bitan, G., M. D. Kirkitadze, A. Lomakin, S. S. Vollers, G. B. Benedek and D. B. Teplow (2002). "Amyloid -protein (A) assembly: A 40 and A 42 oligomerize through distinct pathways." Proceedings of the National Academy of Sciences **100**(1): 330-335.
 25. Botto, M., P. N. Hawkins, M. C. Bickerstaff, J. Herbert, A. E. Bygrave, A. McBride, W. L. Hutchinson, G. A. Tennent, M. J. Walport and M. B. Pepys (1997). "Amyloid deposition is delayed in mice with targeted deletion of the serum amyloid P component gene." Nat Med **3**(8): 855-859.
 26. Boutboul, S., G. C. M. Black, J. E. Moore, J. Sinton, M. Menasche, F. L. Munier, L. Laroche, M. Abitbol and D. F. Schorderet (2006). "A subset of patients with epithelial basement membrane corneal dystrophy have mutations in TGFBI/BIGH3." Human Mutation **27**(6): 553-557.

27. Breydo, L., J. W. Wu and V. N. Uversky (2012). " α -Synuclein misfolding and Parkinson's disease." Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease **1822**(2): 261-285.
28. Callebaut, I., V. Mignotte, M. Souchet and J. P. Mornon (2003). "EMI domains are widespread and reveal the probable orthologs of the *Caenorhabditis elegans* CED-1 protein." Biochem Biophys Res Commun **300**(3): 619-623.
29. Cao, W., H. Ge, X. Cui, L. Zhang, J. Bai, S. Fu and P. Liu (2009). "Reduced penetrance in familial Avellino corneal dystrophy associated with TGFBI mutations." Molecular Vision **15**: 70-75.
30. Carrwik, C. and U. Stenevi (2009). "Lattice corneal dystrophy, gelsolin type (MeretojaâTMs syndrome)." Acta Ophthalmologica **87**(8): 813-819.
31. Chakravarthi, S. V. V. K., C. Kannabiran, M. S. Sridhar and G. K. Vemuganti (2005). "TGFBI gene mutations causing lattice and granular corneal dystrophies in Indian patients." Investigative Ophthalmology and Visual Science **46**(1): 121-125.
32. Chang, L., W. Zhiquan, D. Shijing, Z. Chen, L. Qingfeng, L. Li and S. Xuguang (2009). "Arg124Cys mutation of the TGFBI gene in 2 Chinese families with Thiel-Behnke corneal dystrophy." Archives of Ophthalmology **127**(5): 641-644.
33. Chau, H. M. (2003). "H626R and R124C mutations of the TGFBI (BIGH3) gene caused lattice corneal dystrophy in Vietnamese people." British Journal of Ophthalmology **87**(6): 686-689.
34. Chiti, F. and C. M. Dobson (2006). "Protein Misfolding, Functional Amyloid, and Human Disease." Annual Review of Biochemistry **75**(1): 333-366.
35. Chiti, F., M. Stefani, N. Taddei, G. Ramponi and C. M. Dobson (2003). "Rationalization of the effects of mutations on peptide and protein aggregation rates." Nature **424**(6950): 805-808.
36. Chiti, F., N. Taddei, F. Baroni, C. Capanni, M. Stefani, G. Ramponi and C. M. Dobson (2002). "Kinetic partitioning of protein folding and aggregation." Nat Struct Biol **9**(2): 137-143.
37. Choi, S., Y. M. Yoo, B. Y. Kim, T. I. Kim, H. J. Cho, S. Y. Ahn, H. K. Lee, H. S. Cho and E. K. Kim (2010). "Involvement of TGF- β receptor- and integrin-mediated signaling pathways in the pathogenesis of granular corneal dystrophy II." Investigative Ophthalmology and Visual Science **51**(4): 1832-1847.
38. Choi, S. I., T. I. Kim, K. S. Kim, B. Y. Kim, S. Y. Ahn, H. J. Cho, H. K. Lee, H. S. Cho and E. K. Kim (2009). "Decreased catalase expression and increased susceptibility to oxidative stress in primary cultured corneal fibroblasts from patients with granular corneal dystrophy type II." Am J Pathol **175**(1): 248-261.
39. Clausen, T., C. Southan and M. Ehrmann (2002). "The HtrA family of proteases: implications for protein composition and cell fate." Mol Cell **10**(3): 443-455.
40. Clout, N. J., D. Tisi and E. Hohenester (2003). "Novel fold revealed by the structure of a FAS1 domain pair from the insect cell adhesion molecule fasciclin I." Structure **11**: 197-203.
41. Corden, L. D., O. Swensson, B. Swensson, R. Rochels, B. Wannke, H. J. Thiel and W. H. McLean (2000). "A novel keratin 12 mutation in a German kindred with Meesmann's corneal dystrophy." Br J Ophthalmol **84**(5): 527-530.
42. Coria, F., F. Prelli, E. M. Castaño, M. Larrondo-Lillo, J. Fernandez-Gonzalez, S. G. van Duinen, G. T. A. M. Bots, W. Luyendijk, M. L. Shelanski and B. Frangione (1988). " β -Protein deposition: a pathogenetic link between Alzheimer's disease and cerebral amyloid angiopathies." Brain Research **463**(1): 187-191.
43. Courtney, D. G., E. Toftgaard Poulsen, S. Kennedy, J. E. Moore, S. D. Atkinson, E. Maurizi, M. A. Nesbit, C. B. T. Moore and J. J. Enghild (2015). "Protein

- Composition of TGFBI-R124C- and TGFBI-R555W- Associated Aggregates Suggests Multiple Mechanisms Leading to Lattice and Granular Corneal Dystrophy." Investigative Ophthalmology & Visual Science **56**(8): 4653.
44. Coutu, D. L., J. H. Wu, A. Monette, G. E. Rivard, M. D. Blostein and J. Galipeau (2008). "Periostin, a member of a novel family of vitamin K-dependent proteins, is expressed by mesenchymal stromal cells." J Biol Chem **283**(26): 17991-18001.
 45. Das, S., A. Langenbucher and B. Seitz (2005). "Delayed Healing of Corneal Epithelium After Phototherapeutic Keratectomy for Lattice Dystrophy." Cornea **24**(3): 283-287.
 46. David G. Courtney, E. T. P., Susan Kennedy, Johnny E. Moore, Sarah D. Atkinson, Eleonora Maurizi, M. Andrew Nesbit, C. B. Tara Moore, and a. J. J. Enghild (2015). "Protein composition of TGFBI-R124C and R555W associated aggregates suggest multiple mechanism." IOVS **56**: 4653-4661.
 47. DeMattos, R. B., A. O'Dell M, M. Parsadanian, J. W. Taylor, J. A. Harmony, K. R. Bales, S. M. Paul, B. J. Aronow and D. M. Holtzman (2002). "Clusterin promotes amyloid plaque formation and is critical for neuritic toxicity in a mouse model of Alzheimer's disease." Proc Natl Acad Sci U S A **99**(16): 10843-10848.
 48. Diaper, C. J., D. F. Schorderet, P. Chaubert and F. L. Munier (2005). "Clinical and immunopathological corneal phenotype in homozygotes for the BIGH3 R124H mutation." Eye (Lond) **19**(1): 92-96.
 49. Dighiero, P., F. Niel, P. Ellies, F. D'Hermies, M. Savoldelli, G. Renard, M. Delpech and S. Valleix (2001). "Histologic phenotype-genotype correlation of corneal dystrophies associated with eight distinct mutations in the TGFBI gene11The authors have no commercial interest in the development or marketing of any of the products referred to in this article." Ophthalmology **108**(4): 818-823.
 50. Dikiy, I. and D. Eliezer (2012). "Folding and misfolding of alpha-synuclein on membranes." Biochimica et Biophysica Acta (BBA) - Biomembranes **1818**(4): 1013-1018.
 51. Dinh, R., C. J. Rapuano, E. J. Cohen and P. R. Laibson (1999). "Recurrence of corneal dystrophy after excimer laser phototherapeutic keratectomy1☆." Ophthalmology **106**(8): 1490-1497.
 52. Dobson, C. M. (2003). "Protein folding and misfolding." Nature **426**(6968): 884-890.
 53. Dobson, F. C. a. C. M. (2006). "Protein Misfolding, Functional Amyloid, and Human Disease." The Annual Review of Biochemistry.
 54. Dogru, M., C. Katakami, M. Miyashita, E. Hida, M. Uenishi, K. Tetsumoto, S. Kanno, T. Nishida and A. Yamanaka (2000). "Ocular surface changes after excimer laser phototherapeutic keratectomy11The authors have no proprietary interest in any of the products mentioned in this paper." Ophthalmology **107**(6): 1144-1152.
 55. Doliana, R. (2000). "EMI, a novel cysteine-rich domain of EMILINs and other extracellular proteins, interacts with the gC1q domains and participates in multimerization." FEBS Letters **484**: 164-168.
 56. Dyrlund, T. F., E. T. Poulsen, C. Scavenius, C. L. Nikolajsen, I. B. Thogersen, H. Vorum and J. J. Enghild (2012). "Human cornea proteome: identification and quantitation of the proteins of the three main layers including epithelium, stroma, and endothelium." J Proteome Res **11**(8): 4231-4239.
 57. Ehrmann, M. and T. Clausen (2004). "Proteolysis as a Regulatory Mechanism." Annual Review of Genetics **38**(1): 709-724.

58. El-Ashry, M. F., M. M. Abd El-Aziz, L. A. Ficker, A. J. Hardcastle, S. S. Bhattacharya and N. D. Ebenezer (2004). "BIGH3 mutation in a Bangladeshi family with a variable phenotype of LCDI." *Eye* **18**(7): 723-728.
59. El Kochairi, I., I. Letovanec, S. Uffer, F. L. Munier, P. Chaubert and D. F. Schorderet (2006). "Systemic investigation of keratoepithelin deposits in TGFBI/BIGH3-related corneal dystrophy." *Molecular Vision* **12**: 461-466.
60. Elavazhagan, M., R. Lakshminarayanan, L. Zhou, L. W. Ting, L. Tong, R. W. Beuerman, S. S. Chaurasia and J. S. Mehta (2012). "Expression, purification and characterization of fourth FAS1 domain of TGF β 1p-associated corneal dystrophic mutants." *Protein Expression and Purification* **84**(1): 108-115.
61. Ellies, P., R. A. Bejjani, J. L. Bourges, P. Y. Boelle, G. Renard and P. Dighiero (2003). "Phototherapeutic keratectomy for BIGH3-linked corneal dystrophy recurring after penetrating keratoplasty." *Ophthalmology* **110**(6): 1119-1125.
62. Ellies, P., G. Renard, S. Valleix, P. Yves Boelle and P. Dighiero (2002). "Clinical outcome of eight BIGH3-linked corneal dystrophies." *Ophthalmology* **109**(4): 793-797.
63. Fändrich, M., M. A. Fletcher and C. M. Dobson (2001). "Amyloid fibrils from muscle myoglobin." *Nature* **410**(6825): 165-166.
64. Feizi, S., M. Pakravan, A. R. Baradaran-Rafiee and S. Yazdani (2007). "Granular corneal dystrophy manifesting after radial keratotomy." *Cornea* **26**(10): 1267-1269.
65. Fink, A. L. (1998). "Protein aggregation: folding aggregates, inclusion bodies and amyloid." *Fold Des* **3**(1): R9-23.
66. Finke, J. M., M. Roy, B. H. Zimm and P. A. Jennings (2000). "Aggregation events occur prior to stable intermediate formation during refolding of interleukin 1 β ." *Biochemistry* **39**(3): 575-583.
67. Folberg, R., E. Alfonso, J. O. Croxatto, N. G. Driezen, N. Panjwani, P. R. Laibson, S. A. Boruchoff, J. Baum, E. S. Malbran, R. Fernandez-Mejide, J. A. Morrison, V. B. Bernardino, V. V. Arbizio and D. M. Albert (1988). "Clinically Atypical Granular Corneal Dystrophy with Pathologic Features of Lattice-like Amyloid Deposits." *Ophthalmology* **95**(1): 46-51.
68. Fontana, A., P. P. de Laureto, B. Spolaore, E. Frare, P. Picotti and M. Zambonin (2004). "Probing protein structure by limited proteolysis." *Acta Biochim Pol* **51**(2): 299-321.
69. Francois, J. and J. Feher (1972). "Light microscopical and polarisation optical study of the lattice dystrophy of the cornea." *Ophthalmologica* **164**(1): 1-18.
70. Frantz, S. (2005). "Drug discovery: Playing dirty." *Nature* **437**(7061): 942-943.
71. Freund, D. E., R. A. McCally RI Fau - Farrell, S. M. Farrell Ra Fau - Cristol, N. L. Cristol Sm Fau - L'Hernault, H. F. L'Hernault NI Fau - Edelhauser and H. F. Edelhauser "Ultrastructure in anterior and posterior stroma of perfused human and rabbit corneas. Relation to transparency." (0146-0404 (Print)).
72. Fujiki, K., Y. Hotta, K. Nakayasu and A. Kanai (1998). "Homozygotic patient with BIGH3 gene mutation in Corneal Dystrophy." *Cornea*.
73. Fujiki, K., Y. Hotta, K. Nakayasu, T. Yamaguchi, T. Kato, Y. Uesugi, N. T. Ha, S. Endo, N. Ishida, W. N. Lu and A. Kanai (2000). "Six different mutations of TGFBI (β ig-h3, keratoepithelin) gene found in Japanese corneal dystrophies." *Cornea* **19**(6): 842-845.
74. Fujiki, K., K. Nakayasu and A. Kanai (2001). "Corneal dystrophies in japan." *Journal of Human Genetics* **46**(8): 431-435.
75. Fukuoka, H., S. Kawasaki, K. Yamasaki, A. Matsuda, A. Fukumoto, A. Murakami and S. Kinoshita (2010). "Lattice corneal dystrophy type iv (P.Leu527Arg) Is

- caused by a founder mutation of the TGFBI gene in a single Japanese ancestor." Investigative Ophthalmology and Visual Science **51**(9): 4523-4530.
76. Gallo, G., T. Wisniewski, N. H. Choi-Miura, J. Ghiso and B. Frangione (1994). "Potential role of apolipoprotein-E in fibrillogenesis." Am J Pathol **145**(3): 526-530.
 77. Gerling, U. I. M., E. Brandenburg, H. v. Berlepsch, K. Pagel and B. Koksche (2011). "Structure Analysis of an Amyloid-Forming Model Peptide by a Systematic Glycine and Proline Scan." Biomacromolecules **12**(8): 2988-2996.
 78. Gibson, M. A., G. Hatzinikolas, J. S. Kumaratilake, L. B. Sandberg, J. K. Nicholl, G. R. Sutherland and E. G. Cleary (1996). "Further Characterization of Proteins Associated with Elastic Fiber Microfibrils Including the Molecular Cloning of MAGP-2 (MP25)." Journal of Biological Chemistry **271**(2): 1096-1103.
 79. Gibson, M. A., J. S. Kumaratilake and E. G. Cleary (1997). "Immunohistochemical and ultrastructural localization of MP78/70 (beta ig-h3) in extracellular matrix of developing and mature bovine tissues." Journal of Histochemistry & Cytochemistry **45**(12): 1683-1696.
 80. Gilbert, R. E., J. L. Wilkinson-Berka, D. W. Johnson, A. Cox, T. Soulis, L. L. Wu, D. J. Kelly, G. Jerums, C. A. Pollock and M. E. Cooper (1998). "Renal expression of transforming growth factor-beta inducible gene-h3 (beta ig-h3) in normal and diabetic rats." Kidney Int **54**(4): 1052-1062.
 81. Gratchev, A., P. Guillot, N. Hakiy, O. Politz, C. E. Orfanos, K. Schledzewski and S. Goerdts (2001). "Alternatively Activated Macrophages Differentially Express Fibronectin and Its Splice Variants and the Extracellular Matrix Protein betaIG-H3." Scandinavian Journal of Immunology **53**(4): 386-392.
 82. Grau, S., A. Baldi, R. Bussani, X. Tian, R. Stefanescu, M. Przybylski, P. Richards, S. A. Jones, V. Shridhar, T. Clausen and M. Ehrmann (2005). "Implications of the serine protease HtrA1 in amyloid precursor protein processing." Proc Natl Acad Sci U S A **102**(17): 6021-6026.
 83. Grau, S., A. Baldi, R. Bussani, X. Tian, R. Stefanescu, M. Przybylski, P. Richards, S. A. Jones, V. Shridhar, T. Clausen and M. Ehrmann (2005). "Implications of the serine protease HtrA1 in amyloid precursor protein processing." Proceedings of the National Academy of Sciences **102**(17): 6021-6026.
 84. Grothe, H. L., M. R. Little, A. S. Cho, A. J. Huang and C. Yuan (2009). "Denaturation and solvent effect on the conformation and fibril formation of TGFBIp." Molecular vision **15**: 2617-2626.
 85. Gruenauer-Kloevekorn, C., S. Braeutigam, U. G. Froster and G. I. W. Duncker (2009). "Surgical outcome after phototherapeutic keratectomy in patients with TGFBI-linked corneal dystrophies in relation to molecular genetic findings." Graefe's Archive for Clinical and Experimental Ophthalmology **247**(1): 93-99.
 86. Gruenauer-Kloevekorn, C., I. Clausen, E. Weidle, M. Wolter-Roessler, F. Tost, H. E. Völcker, D. P. Schulze, W. Heinritz, T. Reinhard, U. Froster, G. Duncker, D. Schorderet and C. Auw-Haendrich (2009). "TGFBI (BIGH3) gene mutations in German families: Two novel mutations associated with unique clinical and histopathological findings." British Journal of Ophthalmology **93**(7): 932-937.
 87. Guo, Z. and D. Eisenberg (2007). "The Mechanism of the Amyloidogenic Conversion of T7 Endonuclease I." Journal of Biological Chemistry **282**(20): 14968-14974.
 88. Ha, B. J., T.-i. Kim, S.-i. Choi, R. D. Stulting, D. H. Lee, H.-S. Cho and E. K. Kim (2010). "Mitomycin C Does Not Inhibit Exacerbation of Granular Corneal

- Dystrophy Type II Induced by Refractive Surface Ablation." *Cornea* **29**(5): 490-496.
89. Ha, N. T., L. X. Cung, H. M. Chau, T. K. Thanh, K. Fujiki, A. Murakami and A. Kanai (2003). "A novel mutation of the TGFB1 gene found in a Vietnamese family with atypical Granular Corneal Dystrophy." *Japanese Journal of Ophthalmology* **47**(3): 246-248.
 90. Ha, S. W., H. J. Kim, J. S. Bae, G. H. Jeong, S. C. Chung, J. G. Kim, S. H. Park, Y. L. Kim, S. Kam, I. S. Kim and B. W. Kim (2004). "Elevation of urinary betaig-h3, transforming growth factor-beta-induced protein in patients with type 2 diabetes and nephropathy." *Diabetes Res Clin Pract* **65**(2): 167-173.
 91. Han, K. E., S. Choi, W. S. Chung, S. H. Jung, N. Katsanis, T. Kim and E. K. Kim (2012). "Extremely varied phenotypes in granular corneal dystrophy type 2 heterozygotes." *Molecular Vision* **18**: 1755-1762.
 92. Han, K. E., S. I. Choi, T. I. Kim, Y. S. Maeng, R. D. Stulting, Y. W. Ji and E. K. Kim (2016). "Pathogenesis and treatments of TGFB1 corneal dystrophies." *Prog Retin Eye Res* **50**: 67-88.
 93. Han, K. E., W. S. Chung, T. Kim, K. S. Kim, T. I. Kim and E. K. Kim (2013). "Changes of clinical manifestation of granular corneal deposits because of recurrent corneal erosion in granular corneal dystrophy types 1 and 2." *Cornea* **32**(5): e113-120.
 94. Han, Y. P., A. J. Sim, S. C. Vora and A. J. Huang (2012). "A unique TGFB1 protein in granular corneal dystrophy types 1 and 2." *Curr Eye Res* **37**(11): 990-996.
 95. Han, Y. P., A. J. Sim, S. C. Vora and A. J. W. Huang (2011). "Unique TGFB1 protein in lattice corneal dystrophy." *Investigative Ophthalmology and Visual Science* **52**(11): 8401-8406.
 96. Hanssen, E., B. Reinboth and M. A. Gibson (2003). "Covalent and non-covalent interactions of betaig-h3 with collagen VI. Beta ig-h3 is covalently attached to the amino-terminal region of collagen VI in tissue microfibrils." *J Biol Chem* **278**(27): 24334-24341.
 97. Hashimoto, K., M. Noshiro, S. Ohno, T. Kawamoto, H. Satakeda, Y. Akagawa, K. Nakashima, A. Okimura, H. Ishida, T. Okamoto, H. Pan, M. Shen, W. Yan and Y. Kato (1997). "Characterization of a cartilage-derived 66-kDa protein (RGD-CAP/big-h3) that binds to collagen1The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number D55717.1." *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1355**(3): 303-314.
 98. Hedegaard, C. J., I. B. Thøgersen, J. J. Enghild, G. K. Klintworth and T. Möller-Pedersen (2003). "Transforming growth factor beta induced protein accumulation in granular corneal dystrophy type III (Reis-Bücklers dystrophy). Identification by mass spectrometry in 15 year old two-dimensional protein gels." *Molecular Vision* **9**: 355-359.
 99. Hirano, K., Y. Hotta, K. Fujiki and A. Kanai (2000). "Corneal amyloidosis caused by Leu518Pro mutation of betaig-h3 gene." *Br J Ophthalmol* **84**(6): 583-585.
 100. Hirate, Y., H. Okamoto and K. Yamasu (2003). "Structure of the zebrafish fasciclin I-related extracellular matrix protein (betaig-h3) and its characteristic expression during embryogenesis." *Gene Expr Patterns* **3**(3): 331-336.
 101. Hoersch, S. and M. A. Andrade-Navarro (2010). "Periostin shows increased evolutionary plasticity in its alternatively spliced region." *BMC Evol Biol* **10**: 30.

102. Holland, E. J., S. M. Daya, E. M. Stone, R. Folberg, A. A. Dobler, J. D. Cameron and D. J. Doughman (1992). "Avellino corneal dystrophy. Clinical manifestations and natural history." Ophthalmology **99**(10): 1564-1568.
103. Holtzman, D. M. (2004). "In vivo effects of ApoE and clusterin on amyloid-beta metabolism and neuropathology." J Mol Neurosci **23**(3): 247-254.
104. Hopfer, U., N. Fukai, H. Hopfer, G. Wolf, N. Joyce, E. Li and B. R. Olsen (2005). "Targeted disruption of Col8a1 and Col8a2 genes in mice leads to anterior segment abnormalities in the eye." Faseb j **19**(10): 1232-1244.
105. Hopkins, A., J. Mason and J. Overington (2006). "Can we rationally design promiscuous drugs?" Current Opinion in Structural Biology **16**(1): 127-136.
106. Hwang, J. Y., J. S. Shim, M. Y. Song, S. V. Yim, S. E. Lee and K. S. Park (2016). "Proteomic analysis reveals that the protective effects of ginsenoside Rb1 are associated with the actin cytoskeleton in beta-amyloid-treated neuronal cells." J Ginseng Res **40**(3): 278-284.
107. Ing, J. J., H. H. Ing, L. R. Nelson, D. O. Hodge and W. M. Bourne (1998). "Ten-year postoperative results of penetrating keratoplasty." Ophthalmology **105**(10): 1855-1865.
108. Inoue, T., H. Watanabe, S. Yamamoto, N. Maeda, Y. Inoue, Y. Shimomura and Y. Tano (2002). "Recurrence of Corneal Dystrophy Resulting from an R124H Big-h3 Mutation After Phototherapeutic Keratectomy." Cornea **21**(6): 570-573.
109. Ishihama, Y., Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber and M. Mann (2005). "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 **4**(9): 1265-1272.
110. Jakus, M. A. (1956). "Studies on the cornea. II. The fine structure of Descemet's membrane." J Biophys Biochem Cytol **2**(4 Suppl): 243-252.
111. Jarrett, J. T., E. P. Berger and P. T. Lansbury, Jr. (1993). "The C-terminus of the beta protein is critical in amyloidogenesis." Ann N Y Acad Sci **695**: 144-148.
112. Jarrett, J. T. and P. T. Lansbury, Jr. (1993). "Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?" Cell **73**(6): 1055-1058.
113. Jue, B. and D. M. Maurice (1986). "The mechanical properties of the rabbit and human cornea." Journal of Biomechanics **19**(10): 847-853.
114. Jun, R. (2004). "Avellino corneal dystrophy after LASIK." Ophthalmology **111**(3): 463-468.
115. Karring, H., E. T. Poulsen, K. Runager, I. B. Thøgersen, G. K. Klintworth, P. Højrup and J. J. Enghild (2013). "Serine protease HtrA1 accumulates in corneal transforming growth factor beta induced protein (TGFB1p) amyloid deposits." Mol Vis **19**: 861-876.
116. Karring, H., E. T. Poulsen, K. Runager, I. B. Thøgersen, G. K. Klintworth, P. Højrup and J. J. Enghild (2013). "Serine protease HtrA1 accumulates in corneal transforming growth factor beta induced protein (TGFB1p) amyloid deposits." Molecular Vision **19**: 861-876.
117. Karring, H., K. Runager, I. B. Thøgersen, G. K. Klintworth, P. Højrup and J. J. Enghild (2012). "Composition and proteolytic processing of corneal deposits associated with mutations in the TGFB1 gene." Exp Eye Res **96**(1): 163-170.
118. Karring, H., K. Runager, I. B. Thøgersen, G. K. Klintworth, P. Højrup and J. J. Enghild (2012). "Composition and proteolytic processing of corneal deposits associated with mutations in the TGFB1 gene." Experimental Eye Research **96**(1): 163-170.

119. Karring, H., K. Runager, Z. Valnickova, I. B. Thøgersen, T. Møller-Pedersen, G. K. Klintworth and J. J. Enghild (2010). "Differential expression and processing of transforming growth factor beta induced protein (TGFBIp) in the normal human cornea during postnatal development and aging." Exp Eye Res **90**(1): 57-62.
120. Karring, H., I. B. Thøgersen, G. K. Klintworth, J. J. Enghild and T. Møller-Pedersen (2004). "Proteomic analysis of the soluble fraction from human corneal fibroblasts with reference to ocular transparency." Mol Cell Proteomics **3**(7): 660-674.
121. Kato, T., J. H. Chang and D. T. Azar (2003). "Expression of type XVIII collagen during healing of corneal incisions and keratectomy wounds." Invest Ophthalmol Vis Sci **44**(1): 78-85.
122. Kato, Y., H. Yagi, Y. Kaji, T. Oshika and Y. Goto (2013). "Benzalkonium chloride accelerates the formation of the amyloid fibrils of corneal dystrophy-associated peptides." J Biol Chem **288**(35): 25109-25118.
123. Kawamoto, T., M. Noshiro, M. Shen, K. Nakamasu, K. Hashimoto, Y. Kawashima-Ohya, O. Gotoh and Y. Kato (1998). "Structural and phylogenetic analyses of RGD-CAP/ β ig-h3, a fasciclin-like adhesion protein expressed in chick chondrocytes." Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression **1395**(3): 288-292.
124. Kim, B. Y., J. A. Olzmann, S. I. Choi, S. Y. Ahn, T. I. Kim, H. S. Cho, H. Suh and E. K. Kim (2009). "Corneal dystrophy-associated R124H mutation disrupts TGFBI interaction with Periostin and causes mislocalization to the lysosome." J Biol Chem **284**(29): 19580-19591.
125. Kim, J. E., H. W. Jeong, J. O. Nam, B. H. Lee, J. Y. Choi, R. W. Park, J. Y. Park and I. S. Kim (2002). "Identification of motifs in the fasciclin domains of the transforming growth factor-beta-induced matrix protein betaig-h3 that interact with the α v β 5 integrin." J Biol Chem **277**(48): 46159-46165.
126. Kim, J. E., S. J. Kim, B. H. Lee, R. W. Park, K. S. Kim and I. S. Kim (2000). "Identification of motifs for cell adhesion within the repeated domains of transforming growth factor-beta-induced gene, betaig-h3." J Biol Chem **275**(40): 30907-30915.
127. Kim, J. E., R. W. Park, J. Y. Choi, Y. C. Bae, K. S. Kim, C. K. Joo and I. S. Kim (2002). "Molecular properties of wild-type and mutant betaIG-H3 proteins." Invest Ophthalmol Vis Sci **43**(3): 656-661.
128. Kim, J. W., H. M. Kim and J. S. Song (2008). "Phenotypic non-penetrance in granular corneal dystrophy type II." Graefes Arch Clin Exp Ophthalmol **246**(11): 1629-1631.
129. Kim, J. Y., S. W. Lim, C. Li, J. S. Kim, K. O. Ahn, H. J. Yang, B. S. Choi, Y. S. Kim, J. Kim, B. K. Bang and C. W. Yang (2005). "Effect of FTY720 on chronic cyclosporine nephropathy in rats." Transplantation **80**(9): 1323-1330.
130. Kim, T.-i., M. I. Roh, H. E. Grossniklaus, S. J. Kang, S. M. Hamilton, D. F. Schorderet, W. B. Lee and E. K. Kim (2008). "Deposits of Transforming Growth Factor- β -Induced Protein in Granular Corneal Dystrophy Type II After LASIK." Cornea **27**(1): 28-32.
131. Kim, T. I., S. I. Choi, H. K. Lee, Y. J. Cho and E. K. Kim (2008). "Mitomycin C induces apoptosis in cultured corneal fibroblasts derived from type II granular corneal dystrophy corneas." Molecular Vision **14**: 1222-1228.
132. Kitamura, M. and T. S. Suto (1997). "TGF-beta and glomerulonephritis: anti-inflammatory versus pro-sclerotic actions." Nephrol Dial Transplant **12**(4): 669-679.

133. Klintworth, G. K. (1967). "Primary Familial Corneal Amyloidosis (Lattice Corneal Dystrophy)." Annals of Internal Medicine **66**(6): 1288.
134. Klintworth, G. K. (1999). "Advances in the molecular genetics of corneal dystrophies." American Journal of Ophthalmology **128**(6): 747-754.
135. Klintworth, G. K. (2009). "Corneal dystrophies." Orphanet J Rare Dis **4**: 7.
136. Klintworth, G. K., W. Bao and N. A. Afshari (2004). "Two mutations in the TGFBI (BIGH3) gene associated with lattice corneal dystrophy in an extensively studied family." Investigative Ophthalmology and Visual Science **45**(5): 1382-1388.
137. Klintworth, G. K., Z. Valnickova and J. J. Enghild (1998). "Accumulation of beta ig-h3 gene product in corneas with granular dystrophy." Am J Pathol **152**(3): 743-748.
138. Knupp, C., C. Pinali, P. N. Lewis, G. J. Parfitt, R. D. Young, K. M. Meek and A. J. Quantock (2009). The Architecture of the Cornea and Structural Basis of Its Transparency. Advances in Protein Chemistry and Structural Biology, Elsevier BV: 25-49.
139. Koldso, H., O. J. Andersen, C. L. Nikolajsen, C. Scavenius, C. S. Sørensen, J. Underhaug, K. Runager, N. C. Nielsen, J. J. Enghild and B. Schjøtt (2015). "Early Events in the Amyloid Formation of the A546T Mutant of Transforming Growth Factor beta-Induced Protein in Corneal Dystrophies Compared to the Nonfibrillating R555W and R555Q Mutants." Biochemistry **54**(36): 5546-5556.
140. Konishi, M., M. Yamada, Y. Nakamura and Y. Mashima (1999). "Varied appearance of cornea of patients with corneal dystrophy associated with R124H mutation in the BIGH3 gene." Cornea **18**(4): 424-429.
141. Konishi, M., M. Yamada, Y. Nakamura and Y. Mashima (2000). "Immunohistology of kerato-epithelin in corneal stromal dystrophies associated with R124 mutations of the BIGH3 gene." Current Eye Research **21**(5): 891-896.
142. Koo, E. H., P. T. Lansbury, Jr. and J. W. Kelly (1999). "Amyloid diseases: abnormal protein aggregation in neurodegeneration." Proc Natl Acad Sci U S A **96**(18): 9989-9990.
143. Korvatska, E. (2000). "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." Journal of Biological Chemistry **275**(15): 11465-11469.
144. Krojer, T., M. Garrido-Franco, R. Huber, M. Ehrmann and T. Clausen (2002). Crystal Structure of DegP (HtrA), Protein Data Bank, Rutgers University.
145. Kuchle, M., W. R. Green, H. E. Volcker and J. Barraquer (1995). "Reevaluation of corneal dystrophies of Bowman's layer and the anterior stroma (Reis-Bucklers and Thiel-Behnke types): a light and electron microscopic study of eight corneas and a review of the literature." Cornea **14**(4): 333-354.
146. Lakshminarayanan, R., S. S. Chaurasia, V. Anandalakshmi, S.-M. Chai, E. Murugan, E. N. Vithana, R. W. Beuerman and J. S. Mehta (2014). "Clinical and Genetic Aspects of the TGFBI-associated Corneal Dystrophies." The Ocular Surface **12**(4): 234-251.
147. Lakshminarayanan, R., S. S. Chaurasia, V. Anandalakshmi, S. M. Chai, E. Murugan, E. N. Vithana, R. W. Beuerman and J. S. Mehta (2014). "Clinical and genetic aspects of the TGFBI-associated corneal dystrophies." Ocul Surf **12**(4): 234-251.
148. Lakshminarayanan, R., S. S. Chaurasia, E. Murugan, A. Venkatraman, S. M. Chai, E. N. Vithana, R. W. Beuerman and J. S. Mehta (2015). "Biochemical

- properties and aggregation propensity of transforming growth factor-induced protein (TGFBIp) and the amyloid forming mutants." *Ocul Surf* **13**(1): 9-25.
149. Lakshminarayanan, R., E. N. Vithana, S. M. Chai, S. S. Chaurasia, P. Saraswathi, A. Venkatraman, C. Rojare, D. Venkataraman, D. Tan, T. Aung, R. W. Beuerman and J. S. Mehta (2011). "A novel mutation in transforming growth factor-beta induced protein (TGFβIp) reveals secondary structure perturbation in lattice corneal dystrophy." *British Journal of Ophthalmology* **95**(10): 1457-1462.
 150. Land, M. F. and R. D. Fernald (1992). "The Evolution of Eyes." *Annual Review of Neuroscience* **15**(1): 1-29.
 151. Le Guilloux, V., P. Schmidtke and P. Tuffery (2009). "Fpocket: an open source platform for ligand pocket detection." *BMC Bioinformatics* **10**: 168.
 152. Lee, H. J., C. Choi and S. J. Lee (2002). "Membrane-bound alpha-synuclein has a high aggregation propensity and the ability to seed the aggregation of the cytosolic form." *J Biol Chem* **277**(1): 671-678.
 153. Lee, J. H., S.-H. Chung, R. D. Stulting, W. C. Kim, H. K. Lee and E. K. Kim (2006). "Effects of Corneal Neovascularization on the Manifestations of Avellino Corneal Dystrophy (Granular Corneal Dystrophy Type II)." *Cornea* **25**(8): 914-918.
 154. Lee, J. H., S. M. Cristol, W. C. Kim, E. S. Chung, H. Tchah, M. S. Kim, C. M. Nam, H. S. Cho and E. K. Kim (2010). "Prevalence of granular corneal dystrophy type 2 (Avellino corneal dystrophy) in the Korean population." *Ophthalmic Epidemiol* **17**(3): 160-165.
 155. Lee, J. H., R. D. Stulting, D. H. Lee, C. S. Lee, W. C. Kim and E. K. Kim (2008). "Exacerbation of granular corneal dystrophy type II (Avellino corneal dystrophy) after LASEK." *J Refract Surg* **24**(1): 39-45.
 156. Lee, S. H., J. S. Bae, S. H. Park, B. H. Lee, R. W. Park, J. Y. Choi, J. Y. Park, S. W. Ha, Y. L. Kim, T. H. Kwon and I. S. Kim (2003). "Expression of TGF-beta-induced matrix protein betaig-h3 is up-regulated in the diabetic rat kidney and human proximal tubular epithelial cells treated with high glucose." *Kidney Int* **64**(3): 1012-1021.
 157. Lee, W. B., K. S. Himmel, S. M. Hamilton, X. C. Zhao, R. W. Yee, S. J. Kang and H. E. Grossniklaus (2007). "Excimer laser exacerbation of Avellino corneal dystrophy." *Journal of Cataract & Refractive Surgery* **33**(1): 133-138.
 158. Leung, E. W., L. Rife, R. E. Smith and E. P. Kay (2000). "Extracellular matrix components in retrocorneal fibrous membrane in comparison to corneal endothelium and Descemet's membrane." *Mol Vis* **6**: 15-23.
 159. Li, C., S. W. Lim, B. S. Choi, S. H. Lee, J. H. Cha, I. S. Kim, J. Kim and C. W. Yang (2005). "Inhibitory Effect of Pravastatin on Transforming Growth Factor β1-Inducible Gene h3 Expression in a Rat Model of Chronic Cyclosporine Nephropathy." *American Journal of Nephrology* **25**(6): 611-620.
 160. Li, D., Y. Qi, L. Wang, H. Lin, N. Zhou and L. Zhao (2008). "An atypical phenotype of Reis-Bücklers corneal dystrophy caused by the G623D mutation in TGFBI." *Molecular Vision* **14**: 1298-1302.
 161. Lindsley, A., W. Li, J. Wang, N. Maeda, R. Rogers and S. J. Conway (2005). "Comparison of the four mouse fasciclin-containing genes expression patterns during valvuloseptal morphogenesis." *Gene Expr Patterns* **5**(5): 593-600.
 162. Lisch, W. and B. Seitz (2008). "[New international classification of corneal dystrophies and clinical landmarks]." *Klin Monbl Augenheilkd* **225**(7): 616-622.
 163. Lobley, A., L. Whitmore and B. A. Wallace (2002). "DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra." *Bioinformatics* **18**(1): 211-212.

164. Lu, J., J. Qian, K. I. Izvolsky and W. V. Cardoso (2004). "Global analysis of genes differentially expressed in branching and non-branching regions of the mouse embryonic lung." Dev Biol **273**(2): 418-435.
165. Lu, L., W. W. Reinach Ps Fau - Kao and W. W. Kao (2001). "Corneal epithelial wound healing." (1535-3702 (Print)).
166. Ma, K., G. Liu, Y. Yang, M. Yu, R. Sui, W. Yu, X. Chen, Y. Deng, N. Yan, G. Cao and X. Liu (2010). "TGFB1 gene mutation analysis in a Chinese pedigree of reis-bücklers corneal dystrophy." Molecular Vision **16**: 556-561.
167. MacRaid, C. A., C. R. Stewart, Y. F. Mok, M. J. Gunzburg, M. A. Perugini, L. J. Lawrence, V. Tirtaatmadja, J. J. Cooper-White and G. J. Howlett (2004). "Non-fibrillar Components of Amyloid Deposits Mediate the Self-association and Tangling of Amyloid Fibrils." Journal of Biological Chemistry **279**(20): 21038-21045.
168. Majmudar, P. A., S. L. Forstot, R. F. Dennis, V. S. Nirankari, R. E. Damiano, R. Brenart and R. J. Epstein (2000). "Topical Mitomycin-C for subepithelial fibrosis after refractive corneal surgery." Ophthalmology **107**(1): 89-94.
169. Martine Frising, M., * Gabriele Wildhardt, MD, P Lars Frisch, MD,* and Susanne Pitz, MD* (2006). "Recurrent Granular Dystrophy of the Cornea: An Unusual Case." Cornea.
171. Martinez Molina, D., R. Jafari, M. Ignatushchenko, T. Seki, E. A. Larsson, C. Dan, L. Sreekumar, Y. Cao and P. Nordlund (2013). "Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay." Science **341**(6141): 84-87.
172. Mashima, Y., M. Konishi, Y. Nakamura, Y. Imamura, M. Yamada, T. Ogata, J. Kudoh and N. Shimizu (1998). "Severe form of juvenile corneal stromal dystrophy with homozygous R124H mutation in the keratoepithelin gene in five Japanese patients." British Journal of Ophthalmology **82**(11): 1280-1284.
173. Mashima, Y., S. Yamamoto, Y. Inoue, M. Yamada, M. Konishi, H. Watanabe, N. Maeda, Y. Shimomura and S. Kinoshita (2000). "Association of autosomal dominantly inherited corneal dystrophies with BIGH3 gene mutations in Japan." American Journal of Ophthalmology **130**(4): 516-517.
174. Matsumura, S., S. Uemura and H. Mihara (2004). "Fabrication of Nanofibers with Uniform Morphology by Self-Assembly of Designed Peptides." Chemistry - A European Journal **10**(11): 2789-2794.
175. Maurice, D. M. (1972). "The location of the fluid pump in the cornea." J Physiol **221**(1): 43-54.
176. Meng, W., H. Zhang, T. Guo, C. Pandey, Y. Zhu, O. L. Kon and S. K. Sze (2008). "One-step procedure for peptide extraction from in-gel digestion sample for mass spectrometric analysis." Anal Chem **80**(24): 9797-9805.
177. Michelacci, Y. M. (2003). "Collagens and proteoglycans of the corneal extracellular matrix." Brazilian Journal of Medical and Biological Research **36**(8): 1037-1046.
178. Mishima, S. (1968). "Corneal thickness." (0039-6257 (Print)).
179. Monsellier, E. and F. Chiti (2007). "Prevention of amyloid-like aggregation as a driving force of protein evolution." EMBO Rep **8**(8): 737-742.
180. Monsellier, E., M. Ramazzotti, P. P. de Laureto, G. G. Tartaglia, N. Taddei, A. Fontana, M. Vendruscolo and F. Chiti (2007). "The distribution of residues in a polypeptide sequence is a determinant of aggregation optimized by evolution." Biophys J **93**(12): 4382-4391.

181. Moon, J. W., S. W. Kim, T.-i. Kim, S. M. Cristol, E. S. Chung and E. K. Kim (2007). "Homozygous Granular Corneal Dystrophy Type II (Avellino Corneal Dystrophy)." *Cornea* **26**(9): 1095-1100.
182. Morel, B., L. Varela and F. Conejero-Lara (2010). "The Thermodynamic Stability of Amyloid Fibrils Studied by Differential Scanning Calorimetry." *The Journal of Physical Chemistry B* **114**(11): 4010-4019.
183. Morishige, N., T. I. Chikama, Y. Ishimura, T. Nishida, M. Takahashi and Y. Mashima (2004). "Unusual phenotype of an individual with the R124C mutation in the TGFBI gene." *Archives of Ophthalmology* **122**(8): 1224-1227.
184. Morphy, R. and Z. Rankovic (2007). "Fragments, network biology and designing multiple ligands." *Drug Discovery Today* **12**(3-4): 156-160.
185. Moss, J., I. Shore and D. Woodrow (1998). "An ultrastructural study of the colocalization of biglycan and decorin with AA amyloid fibrils in human renal glomeruli." *Amyloid* **5**(1): 43-48.
186. Muller, L. J., C. F. Marfurt, F. Kruse and T. M. Tervo (2003). "Corneal nerves: structure, contents and function." *Exp Eye Res* **76**(5): 521-542.
187. Muller, L. J., G. F. Vrensen, L. Pels, B. N. Cardozo and B. Willekens (1997). "Architecture of human corneal nerves." *Invest Ophthalmol Vis Sci* **38**(5): 985-994.
188. Munier, F. L., B. E. Frueh, P. Othenin-Girard, S. Uffer, P. Cousin, M. X. Wang, E. Héon, G. C. M. Black, M. A. Blasi, E. Balestrazzi, B. Lorenz, R. Escoto, R. Barraquer, M. Hoeltzenbein, B. Gloor, M. Fossarello, A. D. Singh, Y. Arsenijevic, L. Zografos and D. F. Schorderet (2002). "BIGH3 mutation spectrum in corneal dystrophies." *Investigative Ophthalmology and Visual Science* **43**(4): 949-954.
189. Munier, F. L., E. Korvatska, A. Djemai, D. LePasier, L. Zografos, G. Prescia and D. F. Schorderet (1997). "Kerato-epithelin mutations in four 5q31-linked corneal dystrophies." *American Journal of Ophthalmology* **124**(2): 272.
190. Munishkina, L. A., A. L. Fink and V. N. Uversky (2004). "Conformational prerequisites for formation of amyloid fibrils from histones." *J Mol Biol* **342**(4): 1305-1324.
191. Murphy, C., J. Alvarado, R. Juster and M. Maglio (1984). "Prenatal and postnatal cellularity of the human corneal endothelium. A quantitative histologic study." *Invest Ophthalmol Vis Sci* **25**(3): 312-322.
192. Murugan, E., A. Venkatraman, Z. Lei, V. Mouvet, R. Rui Yi Lim, N. Muruganantham, E. Goh, G. Swee Lim Peh, R. W. Beuerman, S. S. Chaurasia, L. Rajamani and J. S. Mehta (2016). "pH Induced Conformational Transitions in the Transforming Growth Factor beta-Induced Protein (TGFbetaIp) Associated Corneal Dystrophy Mutants." *Sci Rep* **6**: 23836.
193. Nam, E. J., K. H. Sa, D. W. You, J. H. Cho, J. S. Seo, S. W. Han, J. Y. Park, S. I. Kim, H. S. Kyung, I. S. Kim and Y. M. Kang (2006). "Up-regulated transforming growth factor β -inducible gene h3 in rheumatoid arthritis mediates adhesion and migration of synoviocytes through $\alpha\beta 3$ integrin: Regulation by cytokines." *Arthritis & Rheumatism* **54**(9): 2734-2744.
194. Nam, J. O., J. E. Kim, H. W. Jeong, S. J. Lee, B. H. Lee, J. Y. Choi, R. W. Park, J. Y. Park and I. S. Kim (2003). "Identification of the $\alpha\beta 3$ integrin-interacting motif of betaig-h3 and its anti-angiogenic effect." *J Biol Chem* **278**(28): 25902-25909.
195. Namba, Y., M. Tomonaga, H. Kawasaki, E. Otomo and K. Ikeda (1991). "Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease." *Brain Research* **541**(1): 163-166.

196. Newington, J. T., T. Rappon, S. Albers, D. Y. Wong, R. J. Rylett and R. C. Cumming (2012). "Overexpression of pyruvate dehydrogenase kinase 1 and lactate dehydrogenase A in nerve cells confers resistance to amyloid beta and other toxins by decreasing mitochondrial respiration and reactive oxygen species production." J Biol Chem **287**(44): 37245-37258.
197. Niel-Buttschi, F., B. Kantelip, J. Iwaszkiewicz, V. Zoete, M. Boimard, M. Delpech, J. L. Bourges, G. Renard, F. D'Hermies, P. J. Pisella, C. Hamel, B. Delbosc and S. Valleix (2011). "Genotype-phenotype correlations of TGFBI p.Leu509Pro, p.Leu509Arg, p.Val613Gly, and the allelic association of p.Met502Val-p.Arg555Gln mutations." Molecular Vision **17**: 1192-11202.
198. Noble, B. A., A. Agrawal, C. Collins, M. Saldana, P. R. Brogden and B. Zuberbuhler (2007). "Deep Anterior Lamellar Keratoplasty (DALK)." Cornea **26**(1): 59-64.
199. Norris, R. A., C. B. Kern, A. Wessels, E. E. Wirrig, R. R. Markwald and C. H. Mjaatvedt (2005). "Detection of betaig-h3, a TGFbeta induced gene, during cardiac development and its complementary pattern with periostin." Anat Embryol (Berl) **210**(1): 13-23.
200. O'Brien, E. R., K. L. Bennett, M. R. Garvin, T. W. Zderic, T. Hinohara, J. B. Simpson, T. Kimura, M. Nobuyoshi, H. Mizgala, A. Purchio and S. M. Schwartz (1996). "ig-h3, a Transforming Growth Factor Inducible Gene, Is Overexpressed in Atherosclerotic and Restenotic Human Vascular Lesions." Arteriosclerosis, Thrombosis, and Vascular Biology **16**(4): 576-584.
201. Ohlson, S. (2008). "Designing transient binding drugs: a new concept for drug discovery." Drug Discov Today **13**(9-10): 433-439.
202. Ohlson, S., S. Shoravi, T. Fex and R. Isaksson (2006). "Screening for transient biological interactions as applied to albumin ligands: a new concept for drug discovery." Anal Biochem **359**(1): 120-123.
203. Ohno, S., T. Doi, K. Fujimoto, C. Ijuin, N. Tanaka, K. Tanimoto, K. Honda, M. Nakahara, Y. Kato and K. Tanne (2002). "RGD-CAP (β ig-h3) Exerts a Negative Regulatory Function on Mineralization in the Human Periodontal Ligament." Journal of Dental Research **81**(12): 822-825.
204. Ohno, S., M. Noshiro, S. Makihira, T. Kawamoto, M. Shen, W. Yan, Y. Kawashima-Ohya, K. Fujimoto, K. Tanne and Y. Kato (1999). "RGD-CAP (β ig-h3) enhances the spreading of chondrocytes and fibroblasts via integrin α 1 β 1." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1451**(1): 196-205.
205. Okada, M., S. Yamamoto, H. Watanabe, Y. Inoue, M. Tsujikawa, N. Maeda, Y. Shimomura, K. Nishida, S. Kinoshita and Y. Tano (1998). "Granular corneal dystrophy with homozygous mutations in the kerato-epithelin gene." American Journal of Ophthalmology **126**(2): 169-176.
206. Paliwal, P., J. Gupta, R. Tandon, A. Sharma and R. B. Vajpayee (2009). "Clinical and genetic profile of Avellino corneal dystrophy in 2 families from north India." Archives of Ophthalmology **127**(10): 1373-1376.
207. Park, K.-A., C.-S. Ki, E.-S. Chung and T.-Y. Chung (2007). "Deep Anterior Lamellar Keratoplasty in Korean Patients With Avellino Dystrophy." Cornea **26**(9): 1132-1135.
208. Park, S. J., S. Park, H. C. Ahn, I. S. Kim and B. J. Lee (2004). "Conformational resemblance between the structures of integrin-activating pentapeptides derived from betaig-h3 and RGD peptide analogues in a membrane environment." Peptides **25**(2): 199-205.

209. Park, S. W., J. S. Bae, K. S. Kim, S. H. Park, B. H. Lee, J. Y. Choi, J. Y. Park, S. W. Ha, Y. L. Kim, T. H. Kwon, I. S. Kim and R. W. Park (2004). "Beta ig-h3 promotes renal proximal tubular epithelial cell adhesion, migration and proliferation through the interaction with alpha3beta1 integrin." Exp Mol Med **36**(3): 211-219.
210. Park, S. Y., M. Y. Jung and I. S. Kim (2009). "Stabilin-2 mediates homophilic cell-cell interactions via its FAS1 domains." FEBS Lett **583**(8): 1375-1380.
211. Patel, D. A., S. H. Chang, G. J. Harocopos, S. C. Vora, D. H. Thang and A. J. W. Huang (2010). "Granular and lattice deposits in corneal dystrophy caused by R124C mutation of TGFB1p." Cornea **29**(11): 1215-1222.
212. Patro, S. Y. and T. M. Przybycien (1994). "Simulations of kinetically irreversible protein aggregate structure." Biophys J **66**(5): 1274-1289.
213. Penzes, P. and J. E. Vanleeuwen (2011). "Impaired regulation of synaptic actin cytoskeleton in Alzheimer's disease." Brain Res Rev **67**(1-2): 184-192.
214. Perrin, R. J. (2001). "Exposure to Long Chain Polyunsaturated Fatty Acids Triggers Rapid Multimerization of Synucleins." Journal of Biological Chemistry **276**(45): 41958-41962.
215. Plagg, B., D. Ehrlich, K. M. Kniewallner, J. Marksteiner and C. Humpel (2015). "Increased Acetylation of Histone H4 at Lysine 12 (H4K12) in Monocytes of Transgenic Alzheimer's Mice and in Human Patients." Curr Alzheimer Res **12**(8): 752-760.
216. Poepsel, S., A. Sprengel, B. Sacca, F. Kaschani, M. Kaiser, C. Gatsogiannis, S. Raunser, T. Clausen and M. Ehrmann (2015). "Determinants of amyloid fibril degradation by the PDZ protease HTRA1." Nat Chem Biol **11**(11): 862-869.
217. Poulsen, E. T., N. S. Nielsen, M. M. Jensen, E. Nielsen, J. Hjortdal, E. K. Kim and J. J. Enghild (2016). "LASIK surgery of granular corneal dystrophy type 2 patients leads to accumulation and differential proteolytic processing of transforming growth factor beta-induced protein (TGFB1p)." Proteomics **16**(3): 539-543.
218. Poulsen, E. T., K. Runager, M. W. Risor, T. F. Dyrlund, C. Scavenius, H. Karring, J. Praetorius, H. Vorum, D. E. Otzen, G. K. Klintworth and J. J. Enghild (2014). "Comparison of two phenotypically distinct lattice corneal dystrophies caused by mutations in the transforming growth factor beta induced (TGFB1) gene." Proteomics Clin Appl **8**(3-4): 168-177.
219. Poulsen, E. T., K. Runager, M. W. Risor, T. F. Dyrlund, C. Scavenius, H. Karring, J. Praetorius, H. Vorum, D. E. Otzen, G. K. Klintworth and J. J. Enghild (2014). "Comparison of two phenotypically distinct lattice corneal dystrophies caused by mutations in the transforming growth factor beta induced (TGFB1) gene." Proteomics Clin Appl **8**(3-4): 168-177.
220. Ramshini, H., C. Parrini, A. Relini, M. Zampagni, B. Mannini, A. Pesce, A. A. Saboury, M. Nemat-Gorgani and F. Chiti (2011). "Large proteins have a great tendency to aggregate but a low propensity to form amyloid fibrils." PLoS One **6**(1): e16075.
221. Rapuano, C. J. (2001). "Excimer laser phototherapeutic keratectomy." Current Opinion in Ophthalmology **12**(4): 288-293.
222. Reinboth, B., J. Thomas, E. Hanssen and M. A. Gibson (2006). "Beta ig-h3 interacts directly with biglycan and decorin, promotes collagen VI aggregation, and participates in ternary complexing with these macromolecules." J Biol Chem **281**(12): 7816-7824.

223. Ridgway, A. E. A., S. Akhtar, F. L. Munier, D. F. Schorderet, H. Stewart, R. Perveen, R. E. Bonshek, M. T. P. Odenthal, M. Dixon, R. Barraquer, R. Escoto and G. C. M. Black (2000). "Ultrastructural and molecular analysis of Bowman's layer corneal dystrophies: An epithelial origin?" Investigative Ophthalmology and Visual Science **41**(11): 3286-3292.
224. Roca, M., R. L. Althaus and M. P. Molina (2013). "Thermodynamic analysis of the thermal stability of sulphonamides in milk using liquid chromatography tandem mass spectrometry detection." Food Chem **136**(2): 376-383.
225. Roque, A., R. Sortino, S. Ventura, I. Ponte and P. Suau (2015). "Histone H1 Favors Folding and Parallel Fibrillar Aggregation of the 1-42 Amyloid-beta Peptide." Langmuir **31**(24): 6782-6790.
226. Rozsa, A. J. and R. W. Beuerman (1982). "Density and organization of free nerve endings in the corneal epithelium of the rabbit." Pain **14**(2): 105-120.
227. Runager, K., R. V. Basaiawmoit, T. Deva, M. Andreasen, Z. Valnickova, C. S. Sørensen, H. Karring, I. B. Thøgersen, G. Christiansen, J. Underhaug, T. Kristensen, N. C. Nielsen, G. K. Klintworth, D. E. Otzen and J. J. Enghild (2011). "Human phenotypically distinct TGFBI corneal dystrophies are linked to the stability of the fourth FAS1 domain of TGFBIp." J Biol Chem **286**(7): 4951-4958.
228. Runager, K., R. García-Castellanos, Z. Valnickova, T. Kristensen, N. C. Nielsen, G. K. Klintworth, F. X. Gomis-Rüth and J. J. Enghild (2009). "Purification, crystallization and preliminary X-ray diffraction of wild-type and mutant recombinant human transforming growth factor β -induced protein (TGFBIp)." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **65**(3): 299-303.
229. Sakimoto, T., H. Kanno, J. Shoji, Y. Kashima, S. Nakagawa, S. Miwa and M. Sawa (2003). "A novel nonsense mutation with a compound heterozygous mutation in TGFBI gene in lattice corneal dystrophy type I." Japanese Journal of Ophthalmology **47**(1): 13-17.
230. Sawada, H., H. Konomi and K. Hirosawa (1990). "Characterization of the collagen in the hexagonal lattice of Descemet's membrane: its relation to type VIII collagen." J Cell Biol **110**(1): 219-227.
231. Schittny, J. C. (1988). "High resolution immunoelectron microscopic localization of functional domains of laminin, nidogen, and heparan sulfate proteoglycan in epithelial basement membrane of mouse cornea reveals different topological orientations." The Journal of Cell Biology **107**(4): 1599-1610.
232. Schittny, J. C., H. Kresse and P. H. Burri (1995). "Immunostaining of a heterodimeric dermatan sulphate proteoglycan is correlated with smooth muscles and some basement membranes." Histochemistry and Cell Biology **103**(4): 271-279.
233. Schmitt-Bernard, C. F., A. Chavanieu, G. Herrada, G. Subra, B. Arnaud, J. G. Demaille, B. Calas and Á. Argilés (2002). "BIGH3 (TGFBI) Arg124 mutations influence the amyloid conversion of related peptides in vitro. Implications in the BIGH3-linked corneal dystrophies." European Journal of Biochemistry **269**(21): 5149-5156.
234. Schmitt-Bernard, C. F., C. Guittard, B. Arnaud, J. Demaille, A. Argiles, M. Claustres and S. Tuffery-Giraud (2000). "BIGH3 exon 14 mutations lead to intermediate type I/IIIA of lattice corneal dystrophies." Investigative Ophthalmology and Visual Science **41**(6): 1302-1308.
235. Schmitt-Bernard, C. F., C. Schneider and A. Argiles (2002). "Clinical, histopathologic, and ultrastructural characteristics of BIGH3(TGFBI) amyloid

- corneal dystrophies are supportive of the existence of a new type of LCD: the LCDi." *Cornea* **21**(5): 463-468.
236. Schorderet, D. F., M. Menasche, S. Morand, S. Bonnel, V. Büchillier, D. Marchant, K. Auderset, C. Bonny, M. Abitbol and F. L. Munier (2000). "Genomic characterization and embryonic expression of the mouse Bigh3 (Tgfb1) gene." *Biochemical and Biophysical Research Communications* **274**(2): 267-274.
 237. Seeliger, J., K. Estel, N. Erwin and R. Winter (2013). "Cosolvent effects on the fibrillation reaction of human IAPP." *Physical Chemistry Chemical Physics* **15**(23): 8902.
 238. Shah, R. A. and S. E. Wilson (2010). "Use of mitomycin-C for phototherapeutic keratectomy and photorefractive keratectomy surgery." *Current Opinion in Ophthalmology*: 1.
 239. Shao, G., J. Berenguer, A. C. Borczuk, C. A. Powell, T. K. Hei and Y. Zhao (2006). "Epigenetic inactivation of Betaig-h3 gene in human cancer cells." *Cancer Res* **66**(9): 4566-4573.
 240. Sher, N. A. (1991). "Clinical Use of the 193-nm Excimer Laser in the Treatment of Corneal Scars." *Archives of Ophthalmology* **109**(4): 491.
 241. Sherrard, E. S. (1976). "The corneal endothelium in vivo: its response to mild trauma." *Exp Eye Res* **22**(4): 347-357.
 242. Shihab, F. S., T. F. Andoh, A. M. Tanner, N. A. Noble, W. A. Border, N. Franceschini and W. M. Bennett (1996). "Role of transforming growth factor-beta 1 in experimental chronic cyclosporine nephropathy." *Kidney Int* **49**(4): 1141-1151.
 243. Shinoda, K., M. Tomita and Y. Ishihama (2010). "emPAI Calc--for the estimation of protein abundance from large-scale identification data by liquid chromatography-tandem mass spectrometry." *Bioinformatics* **26**(4): 576-577.
 244. Simon, G. M., M. J. Niphakis and B. F. Cravatt (2013). "Determining target engagement in living systems." *Nat Chem Biol* **9**(4): 200-205.
 245. Skonier, J., M. Neubauer, L. Madisen, K. Bennett, G. D. Plowman and A. F. Purchio (1992). "cDNA Cloning and Sequence Analysis of β ig-h3, a Novel Gene Induced in a Human Adenocarcinoma Cell Line after Treatment with Transforming Growth Factor- β ." *DNA and Cell Biology* **11**(7): 511-522.
 246. Solari, H. P., M. P. Ventura, A. B. A. Perez, J. M. F. Sallum, M. N. Burnier and R. Belfort (2007). "TGFB1 gene mutations in Brazilian patients with corneal dystrophy." *Eye* **21**(5): 587-590.
 247. Song, J., A. J. Tan H Fau - Perry, T. Perry Aj Fau - Akutsu, G. I. Akutsu T Fau - Webb, J. C. Webb Gi Fau - Whisstock, R. N. Whisstock Jc Fau - Pike and R. N. Pike "PROSPER: an integrated feature-based tool for predicting protease substrate cleavage sites." (1932-6203 (Electronic)).
 248. Sørensen, C. S., K. Runager, C. Scavenius, M. M. Jensen, N. S. Nielsen, G. Christiansen, S. V. Petersen, H. Karring, K. W. Sanggaard and J. J. Enghild (2015). "Fibril Core of Transforming Growth Factor Beta-Induced Protein (TGFB1p) Facilitates Aggregation of Corneal TGFB1p." *Biochemistry* **54**(19): 2943-2956.
 249. Sornson, E. T. (1965). "GRANULAR DYSTROPHY OF THE CORNEA: AN ELECTRON MICROSCOPIC STUDY." *Am J Ophthalmol* **59**: 1001-1007.
 250. Sorour, H. M., S. B. Yee, N. J. Peterson, F. T. Li, M. S. Macsai, X. C. Zhao and R. W. Yee (2005). "Recurrence of Chromosome 10 Thiel-Behnke Corneal Dystrophy (CDB2) After Excimer Laser Phototherapeutic Keratectomy or Penetrating Keratoplasty." *Cornea* **24**(1): 45-50.
 251. Speed, M. A., J. King and D. I. Wang (1997). "Polymerization mechanism of polypeptide chain aggregation." *Biotechnol Bioeng* **54**(4): 333-343.

252. Spiess, C., A. Beil and M. Ehrmann (1999). "A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein." Cell **97**(3): 339-347.
253. Srivastava, O. P. and K. Srivastava (1999). "cAMP-dependent phosphorylation of betaig-h3 protein in human corneal endothelial cells." Curr Eye Res **19**(4): 348-357.
254. Stefani, M. (2004). "Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world." Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease **1739**(1): 5-25.
255. Stenvang, M., G. Christiansen and D. E. Otzen (2016). "Epigallocatechin Gallate Remodels Fibrils of Lattice Corneal Dystrophy Protein, Facilitating Proteolytic Degradation and Preventing Formation of Membrane-Permeabilizing Species." Biochemistry **55**(16): 2344-2357.
256. Stewart, H., G. C. M. Black, D. Donnai, R. E. Bonshek, J. McCarthy, S. Morgan, M. J. Dixon and A. A. E. Ridgway (1999). "A mutation within exon 14 of the TGFBI (BIGH3) gene on chromosome 5q31 causes an asymmetric, late-onset form of lattice corneal dystrophy." Ophthalmology **106**(5): 964-970.
257. Stix, B., M. Leber, P. Bingemer, C. Gross, J. Ru'schoff, M. Fa'ndrich, D. F. Schorderet, C. K. Vorwerk, M. Zacharias, A. Roessner and C. Ro'cken (2005). "Hereditary Lattice Corneal Dystrophy Is Associated with Corneal Amyloid Deposits Enclosing C-Terminal Fragments of Keratoepithelin." Investigative Ophthalmology & Visual Science **46**(4): 1133.
258. Stokes, M. B., S. Holler, Y. Cui, K. L. Hudkins, F. Eitner, A. Fogo and C. E. Alpers (2000). "Expression of decorin, biglycan, and collagen type I in human renal fibrosing disease." Kidney International **57**(2): 487-498.
259. Stone, E. M., W. D. Mathers, G. O. D. Rosenwasser, E. J. Holland, R. Folberg, J. H. Krachmer, B. E. Nichols, P. D. Gorevic, C. M. Taylor, L. M. Streb, J. A. Fishbaugh, T. E. Daley, B. M. Sucheski and V. C. Sheffield (1994). "Three autosomal dominant corneal dystrophies map to chromosome 5q." Nature Genetics **6**(1): 47-51.
260. Surguchev, A. and A. Surguchov "Conformational diseases: looking into the eyes." (1873-2747 (Electronic)).
261. Surguchev, A. and A. Surguchov (2010). "Conformational diseases: Looking into the eyes." Brain Research Bulletin **81**(1): 12-24.
262. TakÁCs, L., P. Boross, J. TÖZsÉR, L. MÓDis Jr, G. TÓTh and A. Berta (1998). "Transforming Growth Factor-β Induced Protein, βIG-H3, is Present in Degraded Form and Altered Localization in Lattice Corneal Dystrophy Type I." Experimental Eye Research **66**(6): 739-745.
263. Tan, D. T. H. and J. S. Mehta (2007). "Future Directions in Lamellar Corneal Transplantation." Cornea **26**: S21-S28.
264. Tennent, G. A. (1999). [2] Isolation and characterization of amyloid fibrils from tissue. Methods in Enzymology, Elsevier BV: 26-47.
265. Tennent, G. A., L. B. Lovat and M. B. Pepys (1995). "Serum amyloid P component prevents proteolysis of the amyloid fibrils of Alzheimer disease and systemic amyloidosis." Proc Natl Acad Sci U S A **92**(10): 4299-4303.
266. Terry, M. A. and P. J. Ousley (2001). "Deep Lamellar Endothelial Keratoplasty in the First United States Patients." Cornea **20**(3): 239-243.
267. Thapa, N., B.-H. Lee and I.-S. Kim (2007). "TGFB1p/βig-h3 protein: A versatile matrix molecule induced by TGF-β." The International Journal of Biochemistry & Cell Biology **39**(12): 2183-2194.

268. Thapa, N., B. H. Lee and I. S. Kim (2007). "TGFB1p/βig-h3 protein: A versatile matrix molecule induced by TGF-β." International Journal of Biochemistry and Cell Biology **39**(12): 2183-2194.
269. Tian, X., K. Fujiki, W. Wang, A. Murakami, P. Xie, A. Kanai and Z. Liu (2005). "Novel mutation (V505D) of the TGFB1 gene found in a Chinese family with lattice corneal dystrophy, type I." Japanese Journal of Ophthalmology **49**(2): 81-88.
270. Underhaug, J., H. Koldso, K. Runager, J. T. Nielsen, C. S. Sorensen, T. Kristensen, D. E. Otzen, H. Karring, A. Malmendal, B. Schiott, J. J. Enghild and N. C. Nielsen (2013). "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 **1834**(12): 2812-2822.
271. Underhaug, J., H. Koldso, K. Runager, J. T. Nielsen, C. S. Sørensen, T. Kristensen, D. E. Otzen, H. Karring, A. Malmendal, B. Schiøtt, J. J. Enghild and N. C. Nielsen (2013). "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.
272. Urbanc, B., M. Betnel, L. Cruz, G. Bitan and D. B. Teplow (2010). "Elucidation of amyloid beta-protein oligomerization mechanisms: discrete molecular dynamics study." J Am Chem Soc **132**(12): 4266-4280.
273. Uversky, V. N. and A. L. Fink (2004). "Conformational constraints for amyloid fibrillation: the importance of being unfolded." Biochim Biophys Acta **1698**(2): 131-153.
274. Uversky, V. N., A. S. Karnoup, R. Khurana, D. J. Segel, S. Doniach and A. L. Fink (1999). "Association of partially-folded intermediates of staphylococcal nuclease induces structure and stability." Protein Sci **8**(1): 161-173.
275. Ventura, S., J. Zurdo, S. Narayanan, M. Parreno, R. Mangués, B. Reif, F. Chiti, E. Giannoni, C. M. Dobson, F. X. Aviles and L. Serrano (2004). "Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3) case." Proc Natl Acad Sci U S A **101**(19): 7258-7263.
276. Walsh, N. P., B. M. Alba, B. Bose, C. A. Gross and R. T. Sauer (2003). "OMP Peptide Signals Initiate the Envelope-Stress Response by Activating DegS Protease via Relief of Inhibition Mediated by Its PDZ Domain." Cell **113**(1): 61-71.
277. Wang, W. (2005). "Protein aggregation and its inhibition in biopharmaceutics." Int J Pharm **289**(1-2): 1-30.
278. Wang, X., T. Takata, X. Bai, F. Ou, K. Yokono and T. Sakurai (2012). "Pyruvate prevents the inhibition of the long-term potentiation induced by amyloid-beta through protein phosphatase 2A inactivation." J Alzheimers Dis **30**(3): 665-673.
279. Warren, J. F., R. L. Abbott, M. K. Yoon, J. B. Crawford, W. H. Spencer and T. P. Margolis (2003). "A new mutation (Leu569Arg) within exon 13 of the TGFB1 (BIGH3) gene causes lattice corneal dystrophy type I." American Journal of Ophthalmology **136**(5): 872-878.
280. Watanabe, H., Y. Hashida, K. Tsujikawa, M. Tsujikawa, N. Maeda, Y. Inoue, S. Yamamoto and Y. Tano (2001). "Two patterns of opacity in corneal dystrophy caused by the homozygous BIG-H3 R124H mutation." Am J Ophthalmol **132**(2): 211-216.
281. Weiss, J. S. (2010). "Corneal dystrophies: Molecular genetics to therapeutic intervention-fifth ARVO/Pfizer ophthalmics research institute conference." Investigative Ophthalmology and Visual Science **51**(11): 5391-5402.

282. Weiss, J. S., A. J. Moller Hu Fau - Aldave, B. Aldave Aj Fau - Seitz, C. Seitz B Fau - Bredrup, T. Bredrup C Fau - Kivela, F. L. Kivela T Fau - Munier, C. J. Munier Fl Fau - Rapuano, K. K. Rapuano Cj Fau - Nischal, E. K. Nischal Kk Fau - Kim, J. Kim Ek Fau - Sutphin, M. Sutphin J Fau - Busin, A. Busin M Fau - Labbe, K. R. Labbe A Fau - Kenyon, S. Kenyon Kr Fau - Kinoshita, W. Kinoshita S Fau - Lisch and W. Lisch (2015). "IC3D classification of corneal dystrophies--edition 2." (1536-4798 (Electronic)).
283. Weiss, J. S., H. U. Moller, W. Lisch, S. Kinoshita, A. J. Aldave, M. W. Belin, T. Kivela, M. Busin, F. L. Munier, B. Seitz, J. Sutphin, C. Bredrup, M. J. Mannis, C. J. Rapuano, G. Van Rij, E. K. Kim and G. K. Klintworth (2008). "The IC3D classification of the corneal dystrophies." Cornea **27 Suppl 2**: S1-83.
284. Wheeldon, C. E., B. H. de Karolyi, D. V. Patel, T. Sherwin, C. N. McGhee and A. L. Vincent (2008). "A novel phenotype-genotype relationship with a TGFBI exon 14 mutation in a pedigree with a unique corneal dystrophy of Bowman's layer." Molecular vision **14**: 1503-1512.
285. Wiberg, C. (2001). "Biglycan and Decorin Bind Close to the N-terminal Region of the Collagen VI Triple Helix." Journal of Biological Chemistry **276**(22): 18947-18952.
286. Wilken, C., K. Kitzing, R. Kurzbauer, M. Ehrmann and T. Clausen (2004). "Crystal Structure of the DegS Stress Sensor." Cell **117**(4): 483-494.
287. Yam, G. H., K. Wang, V. Jhanji, K. W. Choy, L. Baum and C. P. Pang (2012). "In vitro amyloid aggregate forming ability of TGFBI mutants that cause corneal dystrophies." Invest Ophthalmol Vis Sci **53**(9): 5890-5898.
288. Yamada, N., K. Kawamoto, N. Morishige, T. I. Chikama, T. Nishida, M. Nishioka, N. Okayama and Y. Hinoda (2009). "Double mutation (R124H, N544S) of TGFBI in two sisters with combined expression of Avellino and lattice corneal dystrophies." Molecular Vision **15**: 974-979.
289. Yang, J., X. Han, D. Huang, L. Yu, Y. Zhu, Y. Tong, B. Zhu, C. Li, M. Weng and X. Ma (2010). "Analysis of TGFBI gene mutations in Chinese patients with corneal dystrophies and review of the literature." Molecular Vision **16**: 1186-1193.
290. Yang, Q. N., Y. W. Zhao, L. H. Guo, N. H. Yan, X. Y. Liu and S. P. Cai (2011). "Arg124Cys mutation of the TGFBI gene in a Chinese pedigree of Reis-Bucklers corneal dystrophy." Int J Ophthalmol **4**(3): 235-238.
291. Yerbury, J. J., S. Poon, S. Meehan, B. Thompson, J. R. Kumita, C. M. Dobson and M. R. Wilson (2007). "The extracellular chaperone clusterin influences amyloid formation and toxicity by interacting with prefibrillar structures." The FASEB Journal **21**(10): 2312-2322.
292. Yoneyama, M., T. Tomizawa, N. Tochio, S. Koshiba, M. Inoue, T. Kigawa and S. Yokoyama (2005). Solution structure of the FAS1 domain of human transforming growth factor-beta induced protein IG-H3, Protein Data Bank, Rutgers University.
293. Yoo, J. H. and H. S. Cho (2009). The fourth FAS1 domain structure of human Bigh3, Protein Data Bank, Rutgers University.
294. Yoshida, S., A. Yoshida, S. Nakao, A. Emori, T. Nakamura, K. Fujisawa, Y. Kumano and T. Ishibashi (2004). "Lattice corneal dystrophy type I without typical lattice lines: role of mutational analysis." Am J Ophthalmol **137**(3): 586-588.
295. Yu, P., Y. Gu, Y. Yang, X. Yan, L. Chen, Z. Ge, M. Qi, J. Si and L. Guo (2006). "A clinical and molecular-genetic analysis of Chinese patients with lattice corneal dystrophy and novel Thr538Pro mutation in the TGFBI (BIGH3) gene." Journal of Genetics **85**(1): 73-76.

296. Yuan, C., H. L. Berscheid and A. J. Huang (2007). "Identification of an amyloidogenic region on keratoepithelin via synthetic peptides." FEBS Lett **581**(2): 241-247.
297. Zanetti, M., P. Braghetta, P. Sabatelli, I. Mura, R. Doliana, A. Colombatti, D. Volpin, P. Bonaldo and G. M. Bressan (2004). "EMILIN-1 deficiency induces elastogenesis and vascular cell defects." Mol Cell Biol **24**(2): 638-650.
298. Zenteno, J. C., V. Correa-Gomez, C. Santacruz-Valdez, R. Suarez-Sanchez and C. Villanueva-Mendoza (2009). "Clinical and genetic features of TGFBI-linked corneal dystrophies in Mexican population: Description of novel mutations and novel genotype-phenotype correlations." Experimental Eye Research **89**(2): 172-177.
299. Zenteno, J. C., A. Ramirez-Miranda, C. Santacruz-Valdes and R. Suarez-Sanchez (2006). "Expanding the mutational spectrum in TGFBI-linked corneal dystrophies: Identification of a novel and unusual mutation(Val113Ile) in a family with granular dystrophy." Molecular Vision **12**: 331-335.
300. Zerovnik, E. (2002). "Amyloid-fibril formation. Proposed mechanisms and relevance to conformational disease." Eur J Biochem **269**(14): 3362-3371.
301. Zhang, C., G. Zeng, H. Lin, D. Li, L. Zhao, N. Zhou and Y. Qi (2009). "A novel mutation I522N within the TGFBI gene caused lattice corneal dystrophy I." Molecular vision **15**: 2498-2502.
302. Zhao, Y., M. El-Gabry and T. K. Hei (2006). "Loss of Betaig-h3 protein is frequent in primary lung carcinoma and related to tumorigenic phenotype in lung cancer cells." Mol Carcinog **45**(2): 84-92.
303. Zhao, Y. L., C. Q. Piao and T. K. Hei (2002). "Overexpression of Betaig-h3 gene downregulates integrin alpha5beta1 and suppresses tumorigenicity in radiation-induced tumorigenic human bronchial epithelial cells." Br J Cancer **86**(12): 1923-1928.
304. Zhong, X., S. Chen, W. Huang, J. Yang, X. Chen, Y. Zhou, Q. Zhou and Y. Wang (2010). "Novel and known mutations of TGFBI, their genotype-phenotype correlation and structural modeling in 3 Chinese families with lattice corneal dystrophy." Molecular Vision **16**: 224-230.
305. Zhou, P. and G. Wagner (2009). "Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies." Journal of Biomolecular NMR **46**(1): 23.
306. Zinchenko, R. A., G. I. Elchinova, S. D. Nurbaev and E. K. Ginter (2001). Russian Journal of Genetics **37**(3): 290-301.
307. Zinn, K., L. McAllister and C. S. Goodman (1988). "Sequence analysis and neuronal expression of fasciclin I in grasshopper and Drosophila." Cell **53**(4): 577-587.