Quantitative Proteomic Analyses of Tear Fluid in Graves' Orbitopathy

Thesis by Cecilie Aass Aasaaren

Hormone Laboratory

Oslo University Hospital

Department of Endocrinology, Morbid Obesity and Preventive Medicine

Oslo University Hospital

Faculty of Medicine

University of Oslo

Oslo, Norway

2017

© Cecilie Aass Aasaaren, 2017

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-114-5

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Table of contents

Acknowledgements……………………………………………………………………….....9

Abbreviations………………………………………………………………………………11

List of papers………………………………………………………………………………13

1. Introduction………………………………………………………………………………15

 1.1. Graves' disease…………………………………………………………………………15

 1.2. Graves' orbitopathy……………………………………………………………….....16

 1.3. Mass spectrometry based quantitative proteomics ……………………………..……..24

  1.3.1. Pre-fractionation.……………………………………………………………...26

  1.3.2. Electrospray ionization (ESI)……………………………………………..…….....27

  1.3.3. Mass analyzers……………………………………………………………….......28

  1.3.4. Mass spectrometry-based quantitative proteomics.......................................…....29

  1.3.5. Data analysis…………..………………………………………………………...32

 1.4. Proteomic analysis of tear fluid…………………………………………………...35

 1.5. Targeted quantitative proteomics analyses………..…………...................…………..37

2. Aims………………………………………………………………………………………40

 2.1. General aim………………………………………………………………………….40

 2.2. Specific aims……………………………………………………………………….....40

3. Subjects and Methods……………………………………………………………………41

 3.1. Patients and design…………………………………………………………………...41

  3.1.1. Patient recruitment…………………………………………………………….42

  3.1.2. Healthy volunteers…………………………………………………………….42

  3.1.3. Graves' patients with and without clinical signs of GO (Paper II/Aim 2)….....43

  3.1.4. Graves' patients with and without clinical signs of GO (Paper III/Aim 3)…...44

3.2. Methods……………………………………………………………………………..46

3.2.1. Materials………………………………………………………………………46

3.2.2. Tear collection…………………………………………………………………..46

3.2.3. Protein extraction for LC-MS/MS analysis……..……..……...………………47

3.2.4. Protein extraction for ELISA analyses………………………………………..47

3.2.5. Total protein concentration……………………………………………………49

3.2.6. ELISA measurements…………………………………………………………49

3.2.7. Digestion before LC-MS/MS analysis……...……………..…………………50

3.2.8. Dimethyl labeling……………………………………………………………..50

3.2.9. Offline fractionation on a porous graphite carbon column…………………...51

3.2.10. Nano LC-MS/MS...……….…………………………….……………………51

3.2.11. Data analysis…………………………………………………………………..52

3.2.12. Statistical analyses……………………………………………………………54

4. Main results-summary of papers…...……….………………………………………57

4.1. Paper I…………………………………………………………………………..57

4.2. Paper II…………………………………………………………………………58

4.3. Paper III………………………………………………………………………...59

5. Discussion……………...………...…………………………………………………61

5.1. Methodological considerations……………………..…………………………...61

5.1.1. Internal and external validity…………………………………………………61

5.1.2. Patient characteristics…………………………………………………………63

5.1.3. Sample size and power………………………………………………………..66

5.1.4. Sample preparation and optimization…………………………………………67

5.2. Quantitative comparison between GD patients with and without GO-Main findings..70

5.2.1. Dimethyl labeling based comparative proteomics of tear fluid……..……......70

5.2.2. ELISA measurements………..……………………………………..………73

6. Conclusions……………………………………………………………………………77

6.1. Future research…………………………………………………………………78

7. References…………………………………………………………………………...79

8. Papers and manuscripts……………………………………………………………...95

Acknowledgements

This graduate study has been carried out at the Hormone Laboratory at Aker, the Department of Endocrinology and the Department of Eye Diseases at Oslo University Hospital, in the period from June 2013 to June 2016. Financial support was kindly provided by the South-Eastern Norway Regional Health Authority (Helse Sør-Øst), and also from the Hormone Laboratory, Aker, Oslo University Hospital.

First and foremost I would like to thank my supervisors: Ingrid Norheim, Erik Fink Eriksen and Milaim Pepaj, for the outstanding guidance and continuous support throughout my PhD study. I would like to offer my special thanks to Milaim Pepaj for challenging me to learn more and helping me throughout this work. I would also like to express my gratitude to my co-authors Ellen Charlotte Børnick and Per Medbøe Thorsby for their contributions. Special thanks to the staff members, especially Tone Hagen, and the doctors at the Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital for their help with recruitment of control patients and sample collection.

My sincere thanks also go to the ophthalmologists, especially Ellen Charlotte Børnick, for guidance and recruitment of GO patients. A warm thanks to all of my colleagues for good cooperation, technical support and for making the working environment so great on the 4th floor. Special thanks to May Kristin Lyamouri Bredahl, Finn Erik Aas, Sandra Rinne Dahl and Jens Petter Berg for reviewing my work, and to Manuela Zucknick for statistical guidance.

Thanks to all of my friends for always being there for me. My sincere thanks to all my study patients for volunteering your time and bodies to my research.

Finally I wish to thank my family for always supporting me. Special thanks to my dear sister Kristina and my father for all the help. Last, but not least, a warm and special thanks goes to

Ola for being so understanding, patient and for making me laugh every day. You are the perfect husband.

Cecilie Aass Aasaaren

Oslo, June 2017

Abbreviations

1D: one dimensional

2D: two dimensional

$^{13}CD_2O$: carbon 13 deuterated formaldehyde

ANOVA: analysis of variance

APOD: apolipoprotein D

AUC: area under the curve

AZGP1: zinc-alpha 2 glycoprotein

BD_3CN: cyanoborodeuteride

BH_3CN: cyanoborohydride

CAS: clinical activity score

CD_2O: deuterated formaldehyde

CH_2O: formaldehyde

CT: computed tomography

Cys: cysteine

DIGE: difference gel electrophoresis

Da: Dalton

DMBT1: deleted in malignant brain tumors 1

DON: dysthyroid optic neuropathy

ELISA: enzyme-linked immunosorbent assay

ESI: electrospray ionization

EUGOGO: European group of Graves' orbitopathy

GAG: glycosaminoglycans

GD: Graves' disease

GE: Gel-electrophoresis

GO: Graves' orbitopathy

HILIC: hydrophilic interaction liquid chromatography

IEF: isoelectric focusing

IgA: secretory immunoglobulin A

IFN-ɣ: interferon-gamma

IL: interleukin

IL-1α: interleukin-1 alpha

IM: ion-mobility

iCAT: isotope-coded affinity tag

iTRAQ: isobaric tagging for relative and absolute quantification

IT: ion trap

IQR: interquartile range

LACRT: lacritin

LC: liquid chromatography

LYZ: lysozyme C

MALDI: matrix-assisted laser desorption

Met: methionine

MRM: multiple reaction monitoring

MS: mass spectrometry

MS/MS: tandem mass spectrometry

m/z: mass-to-charge ratio

NGF: nerve growth factor

NOSPECS: no physical signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle signs, corneal involvement, and sight loss

PAGE: polyacrylamide gel electrophoresis

PLOD2: procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2

PPS: PPS Silent Surfactant

Q: quadrupole

QoL: quality of life

RAI: radioactive iodine

RG SF: RapiGest SF

RP: reversed phase

Re: Remainder of the peptide

RENATES: regulated upon activation, normal T-cell expressed, and secreted

ROC: receiver operator characteristics

SCX: strong cation exchange

SDS: Sodium dodecyl sulfate

SILAC: isotopic labeling of proteins in eukaryotic cells

SLPI: antileukoproteinase

T_3: triiodothyronine

T_4: thyroxine

TMT: tandem mass tag

TNF-α: tumor necrosis factor-alpha

TOF: time-of-flight

TRAb: thyrotropin receptor autoantibodies

TSH: thyroid stimulating hormone/thyrotropin

TSHR: thyroid-stimulating hormone receptor

VISA: vision, inflammation, strabismus, and appearance

ZG16B: zymogen granule protein 16 homolog B

List of papers

Paper I

Aass C, Norheim I, Eriksen EF, Thorsby PM and Pepaj M. **Single unit filter-aided method for fast proteomic analysis of tear fluid**. Analytical Biochemistry 2015 April; 480:1-5.

Paper II

Aass C, Norheim I, Eriksen EF, Børnick EC, Thorsby PM and Pepaj M. **Comparative proteomic analysis of tear fluid in Graves' disease with and without orbitopathy**. Clinical Endocrinology 2016 May; 85:805-812.

Paper III

Aass C, Norheim I, Eriksen EF, Børnick EC, Thorsby PM and Pepaj M. **Establishment of a tear protein biomarker panel differentiating between Graves' disease with or without orbitopathy**. PLOS ONE 2017 April; 18:1-12.

1 Introduction

1.1 Graves' disease

Graves' disease (GD) is an autoimmune disease and the most common cause of hyperthyroidism with about 50-80 % of cases of hyperthyroidism worldwide [1-3]. GD is caused by thyrotropin receptor autoantibodies (TRAb), which bind to and activate the thyroid-stimulating hormone receptor (TSHR) in the thyroid gland [4]. The pathogenesis of GD is unclear, however, it is believed that T-cells recognize multiple epitopes of the thyrotropin receptor [5]. TRAb activation of the thyrotropin receptor causes an overproduction of thyroid hormones (thyroxine and triiodothyronine) which suppresses pituitary thyroid-stimulating hormone (TSH) secretion (Figure 1) [6]. Additionally, intra-thyroidal inflammatory cells produce and induce inflammatory cytokines which help maintain the intra-thyroidal autoimmune process [1].

Both genetic and environmental factors are important in the etiology of GD. The strongest evidence for genetic involvement in the development of GD comes from twin studies, which have reported a significantly higher rate of concordance for GD in monozygotic twins than in dizygotic twins [7]. This indicates that genes contribute significantly to GD, yet, no single gene is known to cause GD. The prevalence of hyperthyroidism among females is 2.5 % (4-10 times more frequent than in men), and 0.5 % for men [8-12]. The incidence of GD has been observed to be 21/100 000 in a clinical study in Sweden [13]. Peak ages have been observed between 30-50 years [13-15], and it may occur in all ethnic groups (Caucasians, Asians and Africans), with lowest prevalence reported in Africans [16]. The diagnosis of GD is based on clinical and biochemical manifestations of hyperthyroidism. Common symptoms are tachycardia, poor heat tolerance, diarrhea, enlargement of the thyroid, and weight loss. Current treatments for GD consist of antithyroid drugs, radioactive iodine (RAI) and

thyroidectomy. There are several clinical manifestations of GD, such as diffuse goiter, localized dermopathy, lymphoid hyperplasia, thyroid acropachy and Graves' orbitopathy [17]. These manifestations reside in localized regions, and are believed to reflect the underlying autoimmune processes and not necessarily a direct consequence of alterations in the thyroid function [17]. In more than 95 % of patients with GD, TRAb levels is usually found positive (>1.8 IU/L).

1.2 Graves' orbitopathy

Graves' orbitopathy (GO), also known as Graves' ophthalmopathy, thyroid-associated ophthalmopathy, Graves' eye disease or endocrine eye disease, is an autoimmune inflammatory disorder in the orbit and a frequent manifestation of GD [17-19]. Around 20 % of individuals with GD will develop some kind of ocular involvement [20]. However, only about 5 % of patients develop moderate to severe GO [20, 21]. Usually, GO and GD occur at the same time, but GO may occur months before or within 18 months after the diagnosis of GD [22, 23]. This suggests that GO and GD may share a common etiology. However, a minority of patients (less than 10 %) are euthyroid or hypothyroid at the time when the eye disease appear [24].

The pathogenesis (Figure 1) of GO is unclear, but believed to be caused and initiated by autoreactive T lymphocytes, binding to one or more TSHR antigens, shared by the thyroid and orbital muscle and connective tissue [6, 25]. The majority of these cells are T lymphocytes, macrophages, and B lymphocytes [26]. After reaching the orbit and recognizing the shared antigen, T lymphocytes trigger a cascade of events, including secretion of inflammatory cytokines [25]. Cytokines, such as interferon gamma (IFN-ɣ), tumor necrosis factor alpha (TNF-α) and interleukin 1 alpha (IL-1α), have all been demonstrated in the

retroocular tissues in patients with severe GO [27]. The cytokine IL-1α has also been found in the fatty connective tissues from mononuclear cell infiltrates, suggesting that IL-1α is produced by infiltrating cells and fibroblasts from within the orbit [27]. The cytokines also stimulate the secretion of hydrophilic glycosaminoglycans (GAG), which, when activated, direct the process of fibrosis. Hence, GAG may initiate some of the very early events that lead to an inflammatory response and tissue remodeling [28]. Orbital fibroblasts produce significantly more GAG than cells in other tissues, which may be due to an increased proportion of THY-1 + cells (which differentiate into myofibroblasts) [29]. THY-1 + cells comprise the majority of the fibroblast population within the fatty connective tissues of the posterior orbit [29]. The high amounts of GAG, particularly the hydrophilic hyaluronan, bind water and lead to volume increase in orbital tissues, which may cause many of the clinical manifestations in GO [3, 30]. The increase in ocular muscle volume has been suggested to be a result of tissue expansion from additional muscle cells and adipocytes differentiating from the respective fibroblast compartments [30]. However, the mechanism involved in fat enlargement in GO is not understood, even though an increase in fat cell number or size may contribute to fat expansion in GO as GAGs may accumulate in both muscle and connective tissues [31].

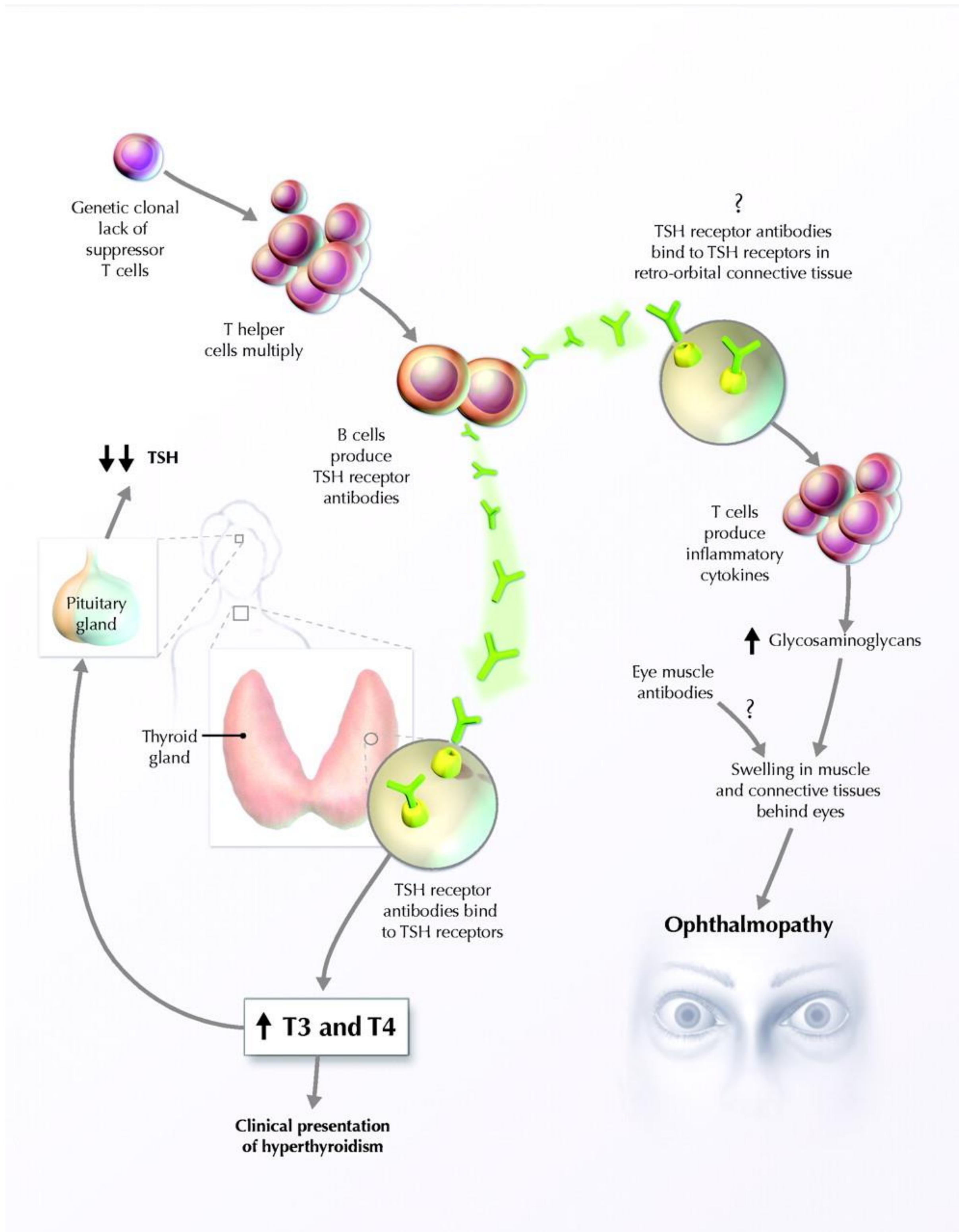


Figure 1. Possible pathogenesis of Graves' disease and Graves' orbitopathy. TRAb = thyrotropin receptor autoantibodies (TSH receptor antibodies), TSH = thyroid-stimulating hormone/thyrotropin, T_3 = triiodothyronine, T_4 = thyroxine. Adapted from Ginsberg et al. [6] with premission.

The inflammatory changes in the retrobulbar tissue result in symptoms and clinical characteristics in the patient, such as orbital pain, edema, excessive tearing, photophobia, pressure sensation behind the eyes, erythema of the perorbital tissues and conjunctivae, upper eyelid retraction, eye proptosis, diplopia and even sight loss due to compression on the optic nerve (dysthyroid optic neuropathy-DON) or breakdown of cornea leading to ulcers (Figure 2) [6, 32, 33]. Proptosis, also called exophthalmos, is a forward displacement of the globe due to enlargement of tissues in the bony orbit [30]. Swelling and thickening of the intra orbital eye muscles may result in diplopia. The stimulation of levator muscles may cause upper eyelid retraction, together with increased tear evaporation. Incomplete eyelid closure during sleep causes corneal dryness [30].

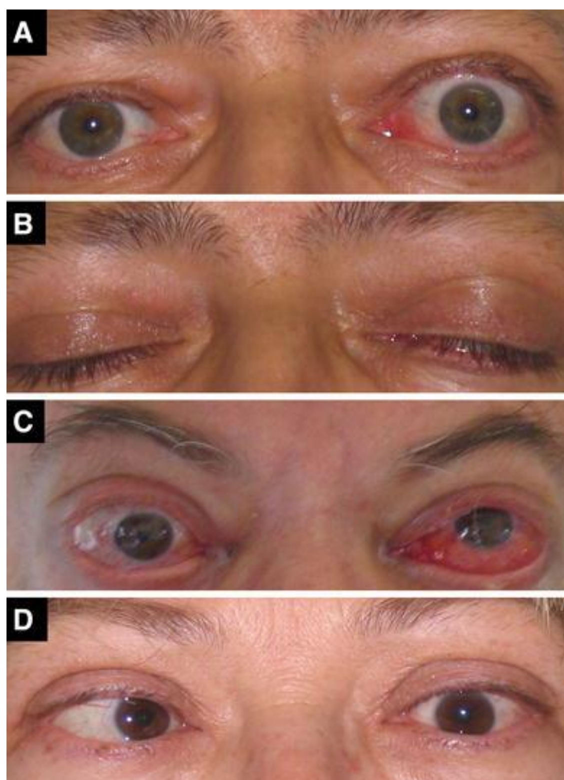


Figure 2. Clinical characteristics of Graves' orbitopathy. A) Mild GO with proptosis and lid retraction, B) Incomplete eye closure (left eye), C) Moderate-to-severe GO with inflammatory signs and reduced movement in upward gaze (right eye), D) Strabismus (right eye) in a patients with chronical eye disease. Adapted from Bartalena et al. [17] with premission.

Because autoantibodies directed against TSHR (TRAb) are detected in patients with GO, TRAb are currently considered as an important contributor to GO [3]. TSHR is regarded as one of the major antigens in the orbit and several authors have shown that serum levels of TRAb directly correlate with severity and activity of GO disease [34-40].

There are limited epidemiological data available on GO, and the prevalence of GO depends on the methodology used [20, 21, 41-43]. A review of clinical reports from 1960-1990 showed a general decline in the prevalence of GO [20]. Others have also shown a decline in incidence of GO, assumed to be a consequence of decrease in smoking prevalence [21, 41]. Bartley [42] found an overall age-adjusted incidence rate of 16 cases/100 000 population/year for women and 2.9 cases/100 000 population/year for men. GO may occur in patients at any age, however, the onset of the disease may differ between women and men [42]. The eye disease severity tends to be higher in older patients, and men seem to have a more severe ocular involvement than women [14]. This could be a result of the increased smoking habits among men.

Several classification systems have been developed to classify and describe GO in detail. The severity is categorized into mild, moderate and severe (and sight-threatening) GO. In 1977, a modified NO SPECS classification [44] was devised as a way to summarise the severity of GO. Today, summary scores such as NO SPECS are of little value in assessing outcomes [45] but remains a useful reminder of the features that should be assessed. The current evaluation systems for GO is the VISA classification (vision, inflammation, strabismus and appearance), and the European Group of Graves` Orbitopathy (EUGOGO) classification [46]. Both systems are grounded in the NO SPECS and CAS (Clinical Activity Score) classifications, and use indicators to assess signs of activity and the degree of severity to help the clinicians treating these patients. The EUGOGO recommendations are often used in Europe and the VISA recommendations are more commonly used in North America and Canada [46].

Although a general consensus for assessment of GO in clinical practice has been established, the clinical evaluation might be subjective and the disease often lead to therapeutic challenges [46-48]. The severity is based on the features of eyelid swelling, eyelid aperture, proptosis, eye motility, visual acuity and color vision [49]. Clinical activity score (CAS) for evaluation of the patient is based on pain, redness, swelling and impaired function of the eyes (Table 1) [48, 50]. CAS score of ≥ 3/7 at the first examination or ≥ 4/10 in successive examinations is considered clinically active GO [48].

Table 1. Clinical Activity Score (CAS)

For initial CAS, only score items 1/ 7	
1	Spontaneous orbital pain
2	Gaze evoked orbital pain
3	Eyelid swelling that is considered to be due to active (inflammatory phase) GO
4	Eyelid erythema
5	Conjunctival redness that is considered to be due to active (inflammatory) GO
6	Chemosis
7	Inflammation of caruncle OR plica
Patients assessed after follow/ up can be scored out of 10 by including items 8- 10	
8	Increase of ≥ 2 mm in proptosis
9	Decrease in uniocular excursion in any one direction of ≥ 8°
10	Decrease of acuity equivalent to 1 Snellen line

* Amended by EUGOGO after Mourits et al. [48, 50].

As GO is cosmetically disfiguring and functionally debilitating, the quality of life (QoL) in these patients is usually significantly reduced, and GO may last for several years after the diagnosis and treatment for GD [51, 52]. Especially patients with moderate to severe GO have reduced QoL and trouble with daily activities. Furthermore, the QoL in patients with mild GO is also reduced, as many patients experience changes in their appearance. A validated disease-specific QoL questionnaire for GO has been developed by EUGOGO [51].

Several studies have shown that cigarette smoking is strongly associated with GO [53-61], and approximately 40 % of patients with GO smoke [60]. The smoke extract is known to increase adipogenesis [61], and is suggested to directly damage the tear film lipid layer [62]. Furthermore, smoking is also associated with deteriorating outcome of GD [63]. Additionally, it has been shown that treatment with radioactive iodine treatments (RAI) is less favourable in smokers than non-smokers [63, 64].

Today, GO is still considered a mystery and its treatment is non-specific and challenging. The medical treatment of the patients with GD is important regarding risk for developing GO. Antithyroid drugs and thyroidectomy have been shown to give no risk or benefit with respect to GO [65, 66]. Medical treatment of patients with GO is usually given as block and-replace regiment, which combines a high dose of antithyroid drugs that completely blocks endogenous thyroid hormone production and a replacement dose of thyroxine, resulting in good control of thyroid hormone levels. Treatment with RAI carries a risk of about 15 % for developing or worsening GO [65, 66]. The risk is greatest if RAI is administrated during active GO, or to patients who smoke. The risk is also increased in patients with severe GD and high TRAb levels, or if the patient develops hypothyroidism with increasing TSH after treatment. However, RAI treatment in combination with glucocorticoids have shown to reduce this increased risk for GO, though RAI is usually not given to high-risk patients. Patients with severe GO are often remitted to endocrine surgery for total thyroidectomy.

Notably, for the treatment of active orbitopathy, intravenous pulsed prednisolone is favoured to oral prednisone. The response is better and the side effects reduced [67, 68]. A dose of 4.5-6 g methyl-prednisolone is the recommended total dose. In very severe disease, doses up to 8 g total methyl-prednisolone may be administrated. However, lower dose is preferred compared to a high-dose prednisolone, as the risk for liver dysfunction is dose-related [69]. At 10-12 g total methyl-prednisolone doses, death caused by liver complications has been observed [70]. Patients with DON (dysthyroid optic neuropathy) are best treated with intravenous prednisolone, with a higher dose initially. These patients may also need immediate decompression operation of the orbit to decrease the pressure on the optic nerves. Many patients with active GO and eye muscle involvement are in addition treated with local irradiation therapy of the orbit [71]. The radiation field is directed behind the eyeball, in thickened muscle and adipose tissue in the orbita.

The treatment of patients with GO depends primarily on the activity and the severity of the eye disease. Patients with signs of active GO should be referred to a multidisciplinary team with ophthalmologists and endocrinologists [72], taking care of both the GD and the inflammatory disease in the orbit. For patients with mild activity in the eye disease, a wait-and-see policy is often recommended. However, the activity in the eye disease may increase over weeks or months, and the patient should be followed up by CAS. The active phase of GO is considered up to 2 years. Once the GO is inactive (stable for 6 months) and GD is under good control, several rehabilitative surgery options for the eye disease is available. The full treatment consists of decompression surgery of the orbit, squint surgery, eyelid lengthening and blepharoplasty. Artificial teardrops, dark glasses and prisms are other helpful tools.

1.3 Mass spectrometry based quantitative proteomics

Proteomics is the large-scale study of proteins expressed in a given biological compartment at a given time, using advanced separation methods (and equipment) resolving the biological sample complexity [73-75]. Large-scale untargeted (discovery-driven) proteomic analysis is considered important in the clinical research to identify and quantify potential biomarkers for earlier detection of diseases. MS-based methodology is usually the methodology of choice to perform untargeted proteomics, which is often the first step in a biomarker discovery pipeline. Due to advantages of mass spectrometry (MS) and extensive technological and instrumental developments, such as the invention of fast, high-resolution mass analyzers, ionization techniques, and labeling approaches, MS based proteomics have increased the ability to analyze complex protein samples in-depth [76]. However, analyzing complex biological samples using MS-based proteomics may also provide several challenges. A major challenge is the identification of low-abundant proteins due to the dynamic range of measurements in mass spectrometers. This limitation can be circumvented by performing depletion of high-abundant proteins and/or performing protein fractionation. Other limitations with MS-based methodology are the cost and the requirement of trained personnel.

Proteomics can be classified into gel-based proteomics and gel-free proteomics [77]. Two-dimensional gel electrophoresis (2D-GE) is based on an initial separation of proteins by isoelectric focusing (IEF) in the first dimension followed by an orthogonal separation via sodium-dodecyl sulfate polyacrylamide gel electrophoresis in the second dimension (SDS-PAGE) [77, 78]. Due to its excellent resolving power, 2D-GE has been the method of choice for separation of complex protein mixtures followed by MS identification after in-gel proteolysis, using matrix-assisted laser desorption time of flight (MALDI-TOF) or electrospray ionization (ESI) tandem MS. However, the 2D-GE based methods suffer from several limitations. This approach is expensive, and involves labor-intensive and time-

consuming protocols. There are also challenges with limited dynamic range, low gel-to-gel reproducibility, automation and poor representation of low abundant proteins, highly acidic/basic proteins, and proteins with extreme size or hydrophobicity [79]. To overcome the low gel-to-gel reproducibility, 2D-fluorescent difference gel electrophoresis (2D-DIGE) [77] was introduced. This technique relies on pre-electrophoretic labeling of samples with one of three spectrally resolvable fluorescent dyes (Cy2, Cy3, and Cy5). The separation is performed on a single gel, and this makes spot matching and quantitation much simpler, more sensitive and accurate. The dyes are comparable in sensitivity to silver staining methods and are compatible with MS [77].

To overcome some of the above challenges and limitations with gel-based proteomics, several gel-free high throughput technologies for proteome analysis have been developed. Gel-free proteomic field can be divided into top-down and bottom-up approaches. The top-down approach identifies proteins by MS through the fragmentation of intact proteins within the mass spectrometer, followed by the measurement of these fragment ions [80]. In bottom-up proteomics (also known as shotgun proteomics), proteins are digested to peptides, which then are analyzed by using a combination of LC and MS [80, 81]. The LC-MS/MS approach is the preferred technique for "shotgun proteomics" and allows direct analysis of very complex peptide mixtures with high sensitivity and a broad dynamic range [80]. After determining peptide sequences using MS/MS, proteins can be identified using protein sequence databases. Shotgun proteomics also avoids the modest separation efficiency and poor mass spectral sensitivity associated with intact protein analysis [80]. Since peptides can be more easily separated by liquid chromatography than proteins, a peptide based proteomic analysis can be performed much faster and cheaper than a complete gel-based analysis. Most LC-based separation techniques are readily interfaced to MS and readily automated compared to 2D-GE [77]. In complex mixtures of peptides, the higher incidence of co-elution is the limiting factor

for the number of peptides that can be identified in an LC-MS/MS analysis. Another limitation is that abundant peptides may suppress the signal of the less abundant proteins, referred to as ion suppression. However, many of these problems may be solved by pre-fractionation of the samples (further described in Section 1.3.1).

1.3.1 Pre-fractionation

To achieve greater analytical depth for a proteome of interest, pre-fractionation of the sample prior to LC-MS/MS is important. Pre-fractionation can be performed either at protein or peptide level. There are several pre-fractionation methods based on chromatographic or electrophoretic principles.

Intact protein fractionation

For pre-fractionation of proteins, chromatographic principles such as ion exchange, reversed phase (RP), hydrophobic interaction affinity or size exclusion can be used [79, 82]. Gel electrophoresis can also be used where proteins are separated by charge and/or size [77]. In addition, liquid-based electrophoretic techniques can also exploit the same separation mechanism as IEF. In liquid-phase IEF, all proteins are separated based on their isoelectric point (pI). An advantage of liquid-phase IEF is the ability to fractionate a complex mixture of proteins according to their p*I* in a non-gel medium [79]. A disadvantage with IEF is that the ampholytes used to establish the pH gradient may interfere with ESI-MS [79]. Since SDS-PAGE is considered time-consuming, many groups use gel-free methods such as IEF and strong cation exchange chromatography (SCX).

Tryptic peptide fractionation

As for intact proteins, both chromatographic and electrophoretic approaches can be used for fractionation of tryptic digests [79, 82]. In IEF, tryptic peptide samples are applied to an immobilized pH gradient (IPG) strips containing a fixed pH gradient [77]. After IEF

fractionation, peptides can easily be recovered in solution for further analysis [77]. Liquid-phase IEF has high loading capacity and resolution power [77]. However, long running time of liquid-phase IEF is a limitation with this technique. Several chromatographic columns are available for fractionation of tryptic peptides, such as SCX, reversed phase (RP) or hydrophilic interaction liquid chromatography (HILIC) [79, 82]. Regardless of the method used for pre-fractionation, a significant increase in the number of identified proteins is achieved when such a step is included in the workflow prior to LC-MS analysis [83, 84].

1.3.2 Electrospray ionization (ESI)

Recent progress in the field of MS allows a fast and accurate analysis of complex protein mixtures and the detection of low abundant proteins. The principle of MS is based on ionization of the analytes in the ion source, followed by separation according to their mass-to-charge ratio (*m/z*) in the mass analyzer.

ESI is a widely used ion source in combination with LC. Together with the solvent, the sample components are introduced into the nebulizer capillary, which ends in a fine tip [85]. Analytes and matrix components are ionized when a high voltage is applied to the capillary tip (2-5 kV). Aided by a coaxial flow of nitrogen gas, a fine aerosol of charged droplets is formed. The droplets in the spray are reduced in size by evaporation and divided into even smaller droplets until the droplets are completely evaporated and gaseous charged molecules are obtained (Figure 3). ESI gives little (in-source) fragmentation because the energy is transferred gradually to the analyte by thermal energy at low temperatures. Therefore ESI is referred to as a soft ionization technique.

Noteworthy, non-volatile salts should be avoided in the mobile phase to prevent ion-suppression and clogging of the MS-inlet. The MS has a limited mass range, though an advantage of ESI is that the peptide analytes are mostly multiple charged [85].

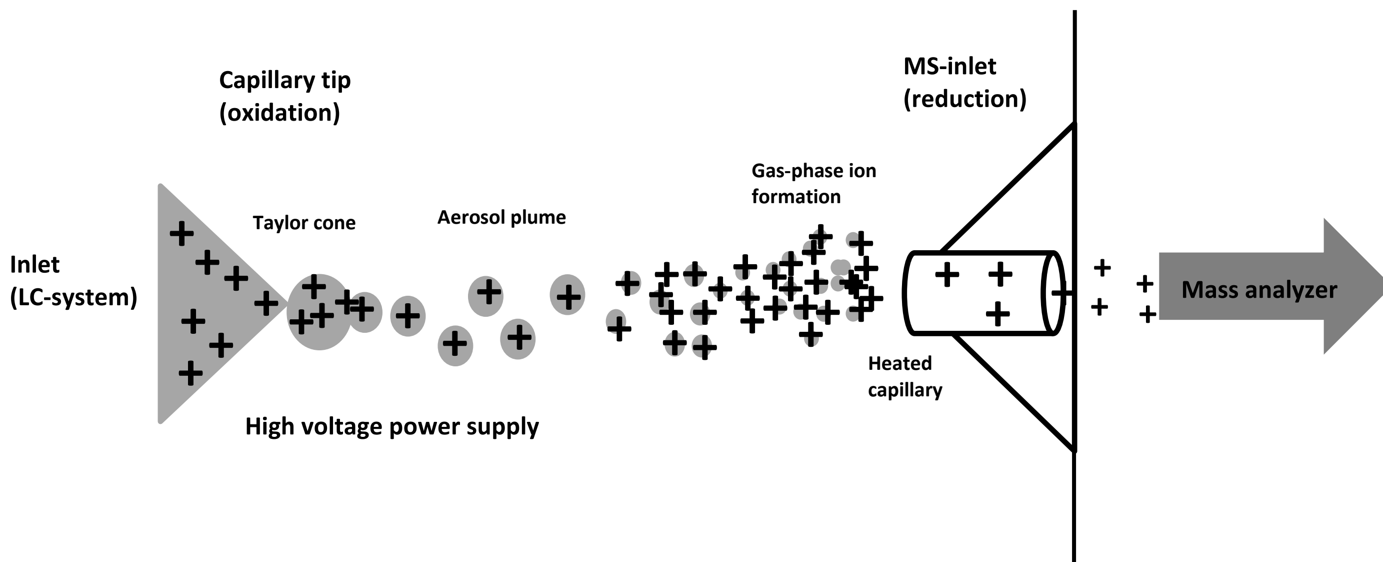


Figure 3. Schematic drawing of an ESI. When high voltage is applied, ions will move towards the opposite charged electrode. Repulsive forces between the ions will break the cone into small highly charged droplets, which will undergo fission and create even smaller highly charged droplets. At last, gas phase ions will be produced. (Aass C 2016).

1.3.3 Mass analyzers

There are several types of mass analyzers used in proteomics, including, LTQ Orbitrap XL, quadrupole-time-of-flight (Q-TOF) and ion-mobility MS (IM-MS). LTQ Orbitrap XL combines a linear ion trap (IT) MS and an Orbitrap mass analyzer. Ions generated by the ionization source (ESI) are collected in the linear ion trap followed by an axial ejection to the C-shaped storage trap [86]. Then ions are transferred from the C-trap and captured in the orbital trap (Figure 4) by rapidly increasing the electric field and reach the detector in small packages [86]. The frequency signal can be converted into a mass-to-charge (*m/z*) spectrum through a fast Fourier-transform algorithm, and finally converted into a mass spectrum.

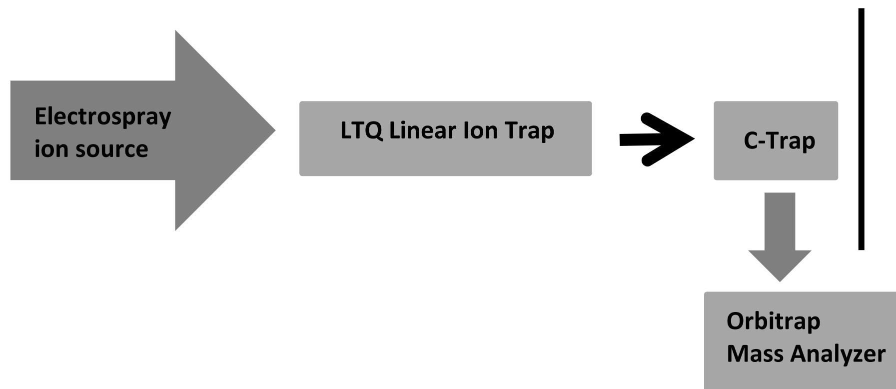


Figure 4. Schematic overview of an LTQ Orbitrap XL. Ions generated by the ionization ESI source are collected in the linear ion trap followed by ejection into the C-shaped storage trap, before the ions are trapped in an electrostatic field around a central electrode in the Orbitrap. (Aass C 2016).

1.3.4 Mass spectrometry-based quantitative proteomics

MS-based quantitative proteomics is a technique for determining the amount of proteins in a sample. MS-based quantitative proteomics are divided into label-based (incorporation of stable isotopes) and label-free quantitative proteomics [87]. In label-free approaches, the intensities of mass spectral peaks are typically compared directly between samples, based on their mass and retention times [87, 88]. This approach is well suited for multiple replicate analyses, and is simpler and faster to perform than label-based approaches. However, limitations regarding robustness and reproducibility may be an issue [88]. Especially, retention time during the chromatographic fractionation and separation, and ESI ion intensity (response) can vary significantly between replicates [87, 88]. Normalization and chromatographic alignment procedures to compensate for retention time shifts and to avoid mismatching peptide abundances across runs are applied (LC-MS/MS) [89]. The quality of the quantification is strongly dependent on the reproducibility of the LC-MS/MS data and on the bioinformatics tools for processing. Label-based approaches, using stable isotopes, are a good alternative to overcome some of the limitations associated with label-free approaches,

e.g. retention time shifts and ESI response. Based on stable isotopes, proteins may be chemically, metabolically, or enzymatically labeled with molecules that have a combination of light and heavy isotopes of ^{15}N, ^{13}C, ^{18}O, and ^{2}H in their composition [90, 91]. Examples of chemical isotopic labeling procedures are isotope coded affinity tag (ICAT), isotope-coded protein label (ICPL), and dimethyl labeling (will be discussed later in this chapter). Typical chemical isobaric labeling procedures are isobaric tags for relative and absolute quantification (iTRAQ) and tandem mass tags (TMT). Other common labeling procedures are enzymatic labeling with ^{18}O, and metabolic labeling with isotopic labeling of proteins in eukaryotic cells (SILAC). A major advantage with labeled strategies is that samples are mixed together after labeling and analyzed as one sample (in a single run). Consequently, precision is highly improved. Important, label-based approaches do not compensate for sample loss prior to the labeling, and in some label-based approaches retention shifts may occur due to incorporation of light and heavy isotopes [90].

Chemical stable isotopic dimethyl labeling, is a fast and cost-effective procedure that is applicable to any biological sample [92-94]. The labeling is usually performed at peptide level, after protein digestion, where all primary amines (the N terminus and the side chain of lysine residues) in the peptide mixture are converted to dimethylamines [95, 96]. The labels have near identical chemical properties; however, each contains a unique stable isotope composition resulting in mass differences of at least 4 Dalton (Da). Hence, a mixture of samples containing the different labels can be distinguished in a single MS run by comparing the peptide mass differences. Quantification is performed by comparing the signal intensity of the differentially labeled peptides. The procedure is based on the reaction of peptide primary amines with formaldehyde (CH_2O) to generate a Schiff base that is rapidly reduced by the addition of cyanoborohydride (BH_3CN) to the mixture (pH-range= 5-8.5) [97]. As shown in Figure 5 [92], the light label is generated by the combination of regular formaldehyde (CH_2O)

and cyanoborohydride (BH_3CN), providing a mass increase of 28 Da per primary amine on a peptide. By incorporation of deuterated formaldehyde (CD_2O), a new label is generated with a mass increase of 32 Da per primary amine. Through combining deuterated and ^{13}C-labeled formaldehyde ($^{13}CD_2O$) with cyanoborodeuteride (BD_3CN), a mass increase of 36 Da is obtained, resulting in the heavy label. To increase the accuracy of the labeling step, both a forward and reverse label-swap can be performed [98] by swapping the stable isotope labels in for instance, two samples (Figure 6).

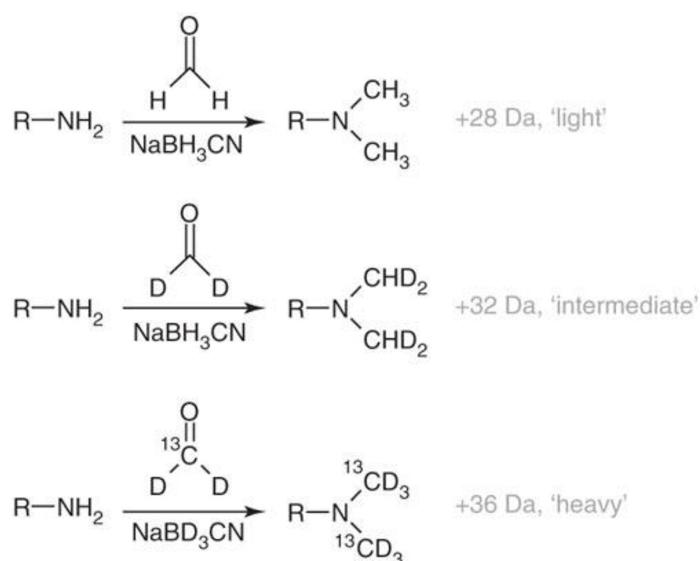


Figure 5. Dimethyl labeling with light, intermediate and heavy isotope dimethyl labeling. The primary amines reacts with formaldehyde, which generates a Schiff base that is reduced by the addition of cyanoborohydride. The different mass labels are generated by applying different combinations of formaldehyde and cyanoborohydride. Re = remainder of the peptide. Adapted from Boersema et al. [92] with premission.

Since dimethyl labeling is performed at peptide level and as one of the last steps in the proteomic analyses pipeline, variations introduced in the earlier steps of the sample preparation may be lost. Another limitation with this labeling procedure is small isotope effects in the LC separation, so called deuterium isotope effects [96, 99]. Although these

effects generally result in minimal LC retention shifts, they should be taken into consideration especially during sample fractionation in the first dimension. Bioinformatically, dimethyl labeling is a relatively straightforward modification for quantitative proteomics and can be performed by integrating extracted ion chromatograms of the differentially labeled peptides. This can be carried out manually, or by using both open-source software packages such as MSQuant, and/or commercially available software such as Proteome Discoverer.

1.3.5 Data analysis

Protein identification and quantification are performed through computerized search algorithms and extensive databases with biological information. By combining information regarding sample origin and treatment, computer programs are able to use the peptide mass and product ion mass data to identify the proteins. All algorithms calculate a matching sore that is used as a measure to say something about how close a given peptide sequence matches the masses [100, 101]. *The score is dependent on the number and relative intensity of sequence-specific ions present in the mass list and the algorithm parameters used for the particular calculation* [100]. The two most common proteomic search algorithms used for protein identification are SEQUEST and Mascot. Both apply similar approaches to assign peptides in a sequence database to measured MS/MS spectra, but they use fundamentally different principles in their mathematical operations [102, 103]. While SEQUEST is an algorithm that assigns scores to matches based on empirical and correlation measurements, Mascot is an algorithm which incorporates scores which accurately reflect true probabilities [104]. Probability-based scoring calculates the probability that the observed match between the experimental data set and each sequence database entry is a chance event. SEQUEST uses a descriptive model for peptide fragmentation and correlative matching to score the alignment

between observed and predicted spectra [105]. The algorithm identifies amino acid sequences in the database that match the measured mass of the peptide ion and then predicts the fragment ions expected for each sequence [105]. A score is calculated using a two-tiered scoring scheme to assess the quality of the match between the spectrum and amino acid sequence from a database [103]. Briefly described, the first score is the preliminary score (S_P), an empirically derived score, and the second score is a cross-correlation score (XCorr) of the experimental and theoretical spectra [103]. In addition to the preliminary- and the cross-correlation scores, ΔC_n is a quantity value describing the uniqueness of the match (match unique to a sequence, $\Delta C_n > 0.1$) by normalizing the difference of XCorr values between the best sequence and each of the other sequences [103]. While the ΔC_n is database dependent and reflects the quality of the match relative to near misses, XCorr is independent of database size and reflects the quality of the match between spectrum and sequence [103].

Due to the use of different score algorithms, inconsistencies between search results may appear. The algorithm's coverage or sensitivity and specificity are two aspects that may interfere with the results obtained. The sensitivity of a search algorithm demonstrates the ability to make correct identifications using any data, independent of the quality of the data, whereas specificity indicates whether the correct hit is significant relative to the other hits. For example, one search algorithm may be able to correctly identify peptides from poorer spectra than another search algorithm, which result in identification of a higher number of peptides and therefore higher sensitivity [106]. Even though a search algorithm is more sensitive than another algorithm, it may be less specific in discriminating between correct and incorrect peptide hits [106]. Often a good concordance between search algorithms is observed, showing peptides identified in both algorithms, when comparing different search algorithms on the same sample and LC-MS/MS instrument. However, different peptides may also be identified by only a single algorithm and therefore affect the results. With tryptic searches, one must

primarily rely on score filter criteria to distinguish correct from incorrect matches [102]. However, more stringent criteria may exclude correct proteins and thereby lower the sensitivity [102]. Kapp et al. [106] showed a clear distinction between correct and incorrect peptide hits for Mascot, and more overlap between correct and incorrect hits for SEQUEST, indicating a lower specificity for SEQUEST. The choice of search algorithm may also affect the results regarding low-abundance peptides, and identification of as many proteins as possible. In such cases, complementary analytical platforms and/or complementary search algorithms may be used. However, this may be insufficient if the platform reproducibility is low [102], and the two search algorithms are unable to identify the same peptides. Therefore, the choice of search algorithm depends on the MS/MS spectra collected and on the instrument used [102].

Detecting significant changes in protein abundance is a fundamental task in MS-based experiments when trying to compare samples from diseased and non-diseased subjects. Identifying significant changes between two groups is essential, however, at the risk of identifying false positives or false negatives. In high-throughput experiments with potentially many differentially expressed proteins, a false discovery rate (FDR) is used [101]. The FDR is designed to control the proportion of false positives among a set of proteins declared differentially expressed. To reduce the impact of possible false positive identifications, the number of peptides quantified and the FDR are important factors, and the chance of truly identifying differentially expressed proteins is higher. Identifying differentially expressed proteins is a task in proteomic studies commonly carried out by using statistics (target-decoy searching or empirical Bayes approaches) [99]. The identified peptide ratio of each differentially expressed protein is assigned a fold change (ratio) which correlates with the statistical significance of the change in protein expression [100]. Usually, cut-off values are set to avoid false positive results. However, if the cut-off values are strict, there is a possibility

that other identified proteins with fold changes not exceeding the cut-off values, are still significantly altered between samples (fold changes around 1, assumes experiments with little changes).

1.4 Proteomic analysis of tear fluid

Tear fluid is a thin layer of extracellular fluid covering and protecting the cornea and conjunctiva epithelia, and provides an optically smooth surface [107]. Normal tear fluid volume is between 5 μL and 10 μL [108], the secretion rate is about 1.2 μL/min [109] and the tear film thickness is between 3 μm and 40 μm [110]. The quality and production of tears are controlled and coordinated by the lacrimal glands, the ocular surface including both the cornea and conjunctiva, and the meibomian glands [107]. The tear film can be divided into three layers, including an inner mucin layer, a middle aqueous layer and an outer lipid layer [111]. In the aqueous middle layer, there is a large amount of proteins and the composition of proteins differs between open and closed eye, and between reflex and basal tears. Most proteins in tears originate from the lacrimal gland, but they may also come from epithelial cells that are shed or leaked from blood vessels during inflammation, wounding or irritation [112]. Tear proteins can be classified according to their abundance: major tear proteins at the high abundant region (concentrations of mg/mL to μg/mL), proteins secreted from ocular surface cells in the middle (μg/mL to ng/mL) and signaling molecules such as cytokines and growth factors at the low abundance region (ng/mL to pg/mL) [73]. The major tear proteins can be grouped into two categories, including proteins produced by the lacrimal gland, and serum proteins that leak from the conjunctival capillaries. The major tear proteins lysozyme (LYZ), lactotransferrin, secretory immunoglobulin A (IgA), lipocalin, albumin and lipophilin constitute about 80-90 % of the total amount of tear proteins [73]. These proteins are secreted by the lacrimal glands, meibomian glands, conjunctival goblet cells, or derive from leakage

from plasma [73]. Tear fluid is considered suitable for proteomic analysis because of its high concentration of proteins, ranging from 6 mg/mL to 11 mg/mL [113]. A wide range of techniques have been applied for the study of the tear proteome, including enzyme-linked immunosorbent assays (ELISA) [114], one- and two-dimensional electrophoresis [115, 116], and chromatographic techniques [117, 118]. Several studies have reported numerous proteins in tears [121, 122], and several studies have demonstrated that the tear fluid proteome is closely related to the ocular health. Moreover, it is believed that analysis of body fluids such as tear fluid, closer to the site of interest (instead of sera), may be useful and an accessible source for discovering biomarkers associated with ocular diseases, such as dry eye, keratoconus and GO. For instance, the tear fluid composition has shown to be altered in GO [121-127]. Early studies by Khalil et al. [122] found an increase in one or more of the five tear protein peaks using LC, and a raised tear IgA/LYZ ratio was seen in patients with GO compared to healthy controls, suggesting lacrimal gland involvement. Baker et al. [123] used 1D-gel electrophoresis to compare tear protein profiles between patients with GO and healthy subjects. Matrix-assisted laser desorption (MALDI) TOF-MS revealed that the levels of zinc-alpha 2 glycoprotein (AZGP1) and lactotransferrin were increased in tears in patients with GO. Okrojek et al. [124] reported down regulation of several proteins in the molecular weight range of 3000-20 000 Da and showed a clear discrimination between patients with GO and healthy controls. Additionally, Matheis et al. [125] found an altered regulation of proinflammatory and protective proteins in tears of patients with GO compared to healthy controls. The same group also identified a protein panel which significantly differentiated between GO patients, patients with dry eyes and controls [126]. They found an up-regulation of inflammatory proteins and a down-regulation of protective proteins in GO patients compared with controls. Moreover, Ujhelyi et al. [127] investigated the levels of 7 cytokines and Plasminogen Activator Inhibitor-1 in tear samples of GD patients with GO, in patients

with GD without GO, and in healthy controls. They observed a significant increase in the release of interleukin (IL), TNF-α, and RENATES (regulated upon activation, normal T-cell expressed, and secreted) in GO patients compared to controls. However, they did not find significant differences in tear fluid cytokine release between the GO group and the GD group.

1.5 Targeted quantitative proteomics analyses

Generally, there are two approaches for targeted quantification in proteomics: antibody-based immunoassays and MS-based assays. In a singleplex immunoassay, one analyte per assay is quantified, whereas in multiplex assays many candidates can be evaluated simultaneously. Enzyme-linked immunosorbent assay (ELISA), Western blot analysis and immunohistochemical analysis are the most widely used singleplex assays [130]. Assay formats exist to measure proteins in solutions (ELISA and immunospot assays), on the surface of cells, within cells (immunohistochemical and immunofluorescent microscopy) and in organs (*in vivo* imaging with labeled antibodies) [128]. Immunoassays rely on the ability of an antibody to bind to the specific structure of a molecule (epitope). ELISA is a simple and sensitive analytical technique for qualitative and/or quantitative determination of proteins, and has become one of the most widely used immunological assays today. ELISA involves the specific interaction between an antibody and an antigen, which allows the detection and quantitation of the antigen [129]. Quantification is obtained by a color change using enzyme-conjugated reagents and a chromogenic substrate (the color change is proportional to the number/amount of antibody-antigen complexes formed) [129]. Though well-characterized ELISA assays are suitable tools for verification analysis, the availability of high-quality ELISAs for biomarker candidates is limited, and the development of ELISA assays is expensive and time-consuming [130].

Antibody-based assays may also be used in the multiplexed format. These quantitative multiplex immunoassays couple the basic principle of antigen-antibody interactions to a wide variety of detection methods. In comparison to singleplex ELISA for a single analyte, multiplex assays offer the possibility to identify combinations of biomarkers with higher disease specificity than any single established biomarker alone. Multiplex immunoassays can have reasonable sensitivity, reproducibility, a wide dynamic range (pg/mL to ng/mL), high throughput and robustness. However, as with singleplex immunoassays, candidate biomarker verification is limited by the availability of good antibodies, cross-reactivity and the time and resources needed to develop multiplex assays [128, 130]. The analytical sensitivity for multiplexed protein measurement is much lower than singleplex ELISA assays, and limits the usefulness of the technology [130].

MS-based assays, such as selected reaction monitoring (SRM), also known as multiple reaction monitoring (MRM), is an alternative to ELISA for targeted protein quantification. MRM is highly selective (targeted), allowing researchers to fine tune an instrument to specifically look for peptides, or protein fragments, of interest [130-132]. Since potential disease-specific biomarker proteins in humans are often in the ng/mL to pg/mL concentration range, fractionation or enrichment strategies are necessary to increase the sensitivity, especially for low-abundance proteins to reduce sample complexity and ion suppression. MRM is only applicable to previously well-characterized peptides, and any sensitivity improvements achieved by introducing additional enrichment or fractionation steps may lead to sacrifice in analytical throughput. Further technology advances on both the MS interface and MS instrumentation are important for eventually achieving sufficient sensitivity without sacrificing analytical throughput [131]. One of the advantages of this approach is the multiplexing feature that allows high-throughput and rapid characterizations of multiple proteins simultaneously in a single LC-MRM-MS run [132,133]. A major advantage of MRM

based quantification over multiplex immunoassays is the relatively short time frame, low cost and ease in developing the multiplex MRM assay. The choice of assay to be used in targeted quantification of proteins relies basically on whether the kit and platform can provide reliable (precision), reproducible quantitative results.

2 Aims

2.1 General aim

The aim of this study was to compare tear protein levels between GD patients with and without GO using quantitative proteomics.

2.2 Specific aims

1. To develop a method for fast proteomic analysis of tear fluid and systematically investigate the most suitable conditions for protein extraction from Schirmer strips (Paper I).

2. To quantitatively compare pooled tear fluid from Graves' patients with moderate to severe GO and Graves' patients without GO using untargeted quantitative proteomics based on dimethyl labeling in combination with 2D LC-MS/MS (Paper II).

3. To evaluate tear levels of LYZ, LACRT and AZGP1 in individual Graves' patients with and without GO, using ELISA. And to investigate the diagnostic performance of these proteins either alone or combined in a panel (Paper III).

3 Subjects and Methods

3.1 Patients and design

This thesis is based on samples collected from patients at the department of ophthalmology at Oslo University Hospital-Ullevål and patients at the Department of Endocrinology, Morbid Obesity and Preventive Medicine at Oslo University Hospital-Aker during the years 2014-2015. The study was a cross-sectional study comparing GD patients with and without GO (Paper II and Paper III). An overview of the papers, groups and methods used are shown in Table 2. Since analysis of tear fluid from Schirmer strips are challenging due to sample loss during extraction, a study of tear samples from healthy laboratory volunteers was included prior to the studies to develop a single unit filter-aided method for fast proteomic analysis of tear fluid (Paper I).

Patients with GD were divided in three groups (the patients were only included in one group):

I. Moderate to severe GO-CAS 4 or more of 7
II. Mild GO-CAS 3/7
III. Patients without GO at end of medical treatment for GD

Table 2. Overview of the different papers, groups and methods used.

Paper	Patients evaluated	Methods used
II	Moderate to severe GO patients and patients without GO	LC-MS/MS in combination with dimethyl labeling on pooled tear fluid samples
III	Moderate to severe GO patients and patients without GO	ELISA analyses on individual tear fluid samples
III	Mild GO patients and patients without GO	ELISA analyses on individual tear fluid samples

3.1.1 Patient recruitment

All Graves' patients with and without GO were encouraged to participate in the study by their treating ophthalmologist at Ullevål and endocrinologist at Aker. Patients willing to hear more about the study were then given further description (written and oral) by the investigator (who was contacted by the treating doctors after oral consent was obtained by the patient). The patients then signed a consent form and filled out a questionnaire regarding their treatments performed, duration of their GD/GO, smoking habits and other diseases. After the written consent was obtained, tear fluid samples were collected. The same person performed the description of the study to each patient, the collection of consent forms and questionnaire, and collected tear fluid throughout the whole recruitment period.

3.1.2 Heathy volunteers

Three healthy volunteers with no history of GD or ocular involvement were included in the introduction phase of this thesis for evaluation of sample preparation considerations and method optimization (Paper I).

3.1.3 Graves' patients with and without clinical signs of GO (Paper II/Aim 2)

Design

In this cross-sectional study, pooled tear fluid from patients with moderate to severe GO and patients without GO were evaluated using quantitative proteomics.

Patients

Twenty one patients with moderate to severe GO (15 female, 6 male, median age 57 years, range 20-77 years; 9 smokers, 7 women and 2 men) were included (Table 3). Another group consisted of 21 Graves' patients without GO (17 female, 4 male, median age 44 years, range 26-69 years; 3 smokers, 3 women and 0 men, Table 3).

The inclusion criteria for all the patients with GO (Paper II and III) depended on CAS. All patients were evaluated by ophthalmologists and endocrinologists, and grouped into moderate to severe GO, mild GO, and patients without GO, based on activity and severity by CAS (eyelid swelling, eyelid aperture, eye proptosis, eye motility, visual acuity and color vision) [11], at the same day as Schirmer tear collection was done. All patients filled out the questionnaire about disease duration, clinical history, smoking habits, treatment with thyrostatic drugs, earlier treatments with RAI and operation of the thyroid. Dermatologic, clinical and serological data including TRAb, FT4, FT3 and TSH levels are summarized in Table 3. Exclusion criteria for patients with GD were other forms of hyperthyroidisms than GD, patients treated with steroids, long lasting or chronic GO and subjects unwilling or unable to give informed consent.

Inclusion criteria for patients without GO were Graves' patients in remission and earlier treated with thyrostatic medications, and no symptoms or signs of GO during their GD period. All should have elevated TRAb levels during the period of active Graves' disease. Exclusion criteria were hyperthyroidism of other reasons than GD and age below 18 years.

3.1.4 Graves' patients with and without clinical signs of GO (Paper III/Aim 3)

Design

This study was divided into two parts. The first part included patients with moderate to severe GO and patients without GO. The second part included patients with mild GO and patients without GO. Due to limited sample volume, the two studies included two separate groups of patients without GO.

Patients

The patients included in the moderate to severe GO group were the same as the patients in Section 3.1.3 (Section 3.1.3, Table 3, Paper II). Furthermore, tears from 18 patients with mild GO (15 female, 3 male, median age 51.5 years, range 28-63 years; 8 smokers, 8 women and 0 men) were examined (Table 4).

As described in Section 3.1.3, all patients with GO went through examination by an ophthalmologist and filled out a questionnaire regarding clinical history, smoking, symptoms and medications. Based on the ophthalmologists CAS examination on activity and severity, the patients were grouped into moderate to severe GO, mild GO and patients without GO. The inclusion and exclusion criteria for GO patients are described in Section 3.1.3. Disease duration, treatment, smoking habits, TRAb, FT4, FT3, and TSH levels are summarized in Table 3 for the patients with moderate to severe GO and patients without GO, and in Table 4 for the patients with mild GO. The patients without GO are described in Table 3 and 4.

Because of limited sample volume available, there were two separate groups of Graves' patients without GO. The first group consisted of 21 patients (17 female, 4 male, median age 44 years, range 26-69 years; 3 smokers, 3 women and 0 men, Table 3). This was the same group as described in Section 3.1.3. The second group consisted of 9 patients without GO (5 female, 4 male, median age 45 years, range 27-68 years; 2 smokers, 2 women and 0 men,

Table 4). The inclusion criteria for the groups without GO were the same as described in Section 3.1.3.

Table 3. Demographic, clinical and serological data of patients with moderate to severe GO and without GO (Paper II and III).

	Moderate to severe GO (n=21)	Without GO (n=21)	p-value
Females, n (%)	15 (71.4)	17 (81)	
Males, n (%)	6 (28.6)	4 (19)	
Age, y (median)	57 (20-77)	44 (26-69)	0.029
Smokers, n (%)	9 (42.8)	3 (14.3)	0.043
Duration GD, month (median)	14 (3-108)	21 (10-35)	0.018
TRAb, IU/L (reference range, <1.8)	19.9 (3.7-156.4)	<0.9 (<0.9-4.9)	<0.001
Radioiodine treatment, n (%)	5 (23.8)	0 (0)	0.019
Thyroid surgery, n (%)	1 (4.8)	1 (4.8)	1
TSH, mIU/L (median; reference range, 0.5-3.6)	0.05 (0.01-1.7)	0.88 (0.03-4.34)	<0.001
Serum FT4, pmol/L (reference range, 8-21)	17.3 (10.1-49.2)	13.5 (11.4-21.3)	0.004
Serum FT3, pmol/L (reference range, 3.6-8.3)	7.4 (4.4-20.0)	5.4 (4.5-8.6)	0.005

Table 4. Demographic, clinical and serological data of patients with mild GO and without GO (Paper III).

	Mild GO (n=18)	Without GO (n=9)	p-value
Females, n (%)	15 (83.3)	5 (55.6)	
Males, n (%)	3 (16.7)	4 (44.4)	
Age, y (median)	51.5 (28-63)	45 (27-68)	0.54
Smokers, n (%)	8 (44.4)	2 (22.2)	0.27
Duration GD, month (median)	8.5 (3-40)	18 (7-42)	0.047
TRAb, IU/L (reference range, <1.8)	5.4 (0.9-30.0)	<0.9 (<0.9-4.7)	0.003
Radioiodine treatment, n (%)	1 (5.5)	0 (0)	0.48
Thyroid surgery, n (%)	2 (11.1)	0 (0)	0.31
TSH, mIU/L (median; reference range, 0.5-4.4)	0.04 (0.03-1.7)	0.1 (0.03-2.33)	0.65
Serum FT4, pmol/L (reference range, 8-21)	21.3 (12.2-28.2)	18.6 (11.4-26.1)	0.37
Serum FT3, pmol/L (reference range, 3.6-8.3)	6.3 (4.6-11.7)	6.2 (4.8-8.6)	0.93

3.2 Methods

3.2.1 Materials

All chemicals and solvents were purchased from Sigma (St. Louis, MO, USA) unless stated otherwise. Milli Q water was obtained from a Milli-Q Ultrapure water purification system (Millipore, Bedford, MA, USA). Calcium chloride ($CaCl_2$), sodium chloride (NaCl) and ammonia solution (NH_3) were purchased from Merck (Darmstadt, Germany). PPS Silent Surfactant (PPS) and RapiGest SF (RG SF) acid-labile surfactants were obtained from Kem-En-Tec Nordic (Tåstrup, Denmark) and Waters (Milford, MA, USA), respectively. Invitrosol protein solubilizer was purchased from Invitrogen (Carlsbad, CA, USA). Schirmer strips used for tear collection were obtained from Haag-Streit UK (Harlow, Essex, UK). The proteases endoproteinase Lys-C/trypsin and PNGase F were obtained from Promega Corporation (Madison, WI, USA). Formic acid (FA) was purchased from Fisher Scientific (Geel, Belgum). Bio-Rad protein assay dye reagent concentrate and ɣ-microglobulin standard were obtained from Bio-Rad Laboratories (Hercules, CA, USA). Cut-off size (0.1 µM) ultrafree-MC (0.5 mL) VV centrifugal filter used for the extraction of proteins was from Millipore (Bedford, MA, USA). Lysozyme C (LYZ) (detection range; 0,156-2.5 ng/mL, inter-assay % CV 8.2 and intra-assay % CV 4.2) was purchased from abcam, Cambridge, UK. Zinc-alpha-2-glycoprotein 1 (AZGP1) (detection range; 4.7-300 ng/mL, inter-assay % CV < 12 and intra-assay % CV < 10) and lacritin (LACRT) (detection range; 1,56-100 ng/mL, inter-assay % CV < 12 and intra-assay % CV < 10) were obtained from Cloud-clone corp., Houston, TX, USA. The materials used in the different experiments are also described in Paper I-III.

3.2.2 Tear collection

Tear samples were collected using Schirmer's type I tear test (Haag-Streit UK Ltd, Harlow, Essex, UK) without local anesthesia. The Schirmer strips were inserted for five minutes in the

lower eyelid in the standard fashion in both eyes (open eye tears) or until the filter strips were full. The same investigator carried out the collection of tears from all the subjects included. After collection, the Schirmer strips were placed in a 2 ml microcentrifuge tube and immediately frozen at -80 °C until analysis (described in Paper II).

3.2.3 Protein extraction for LC-MS/MS analysis

Protein extraction from Schirmer strips for LC-MS/MS analysis is shown in Figure 6 and in Figure 1 Paper II. As described in Paper II Supplementary 1.1, twenty-one Schirmer strips, one from each of the subjects in each group were pooled (moderate to severe GO and without GO). Two strips were randomly collected from the pool and placed in a 0.1 µm centrifugal filter unit giving a total of 11 sample vials for each group. Tear proteins were then extracted from the Schirmer strips with 500 µL 100 mM triethylammonium bicarbonate (TEAB) and 50 mM NaCl, and mixed for 4 h at 25 °C at 300 rpm. The samples were then centrifuged at 7500 rpm for 5 min and the Schirmer strips were removed. The filtrate (protein solution) was evaporated and dissolved in 50 µL of denaturing buffer consisting of 8 M Urea in 50 mM TEAB (pH 8.5). Eleven samples were collected for each group and pooled.

3.2.4 Protein extraction for ELISA analyses

The protein extraction from Schirmer strips used in ELISA measurements is shown in Figure 6 and Figure 1 in Paper II. One Schirmer strip from each patient was collected and placed in a 0.1 µm centrifugal filter unit. Tear proteins were then extracted from the Schirmer strips with 500 µL 100 mM TEAB and 50 mM NaCl, and mixed for 4 h at 25 °C at 300 rpm. Next, the samples were centrifuged at 7500 rpm for 5 min and the Schirmer strips were removed.

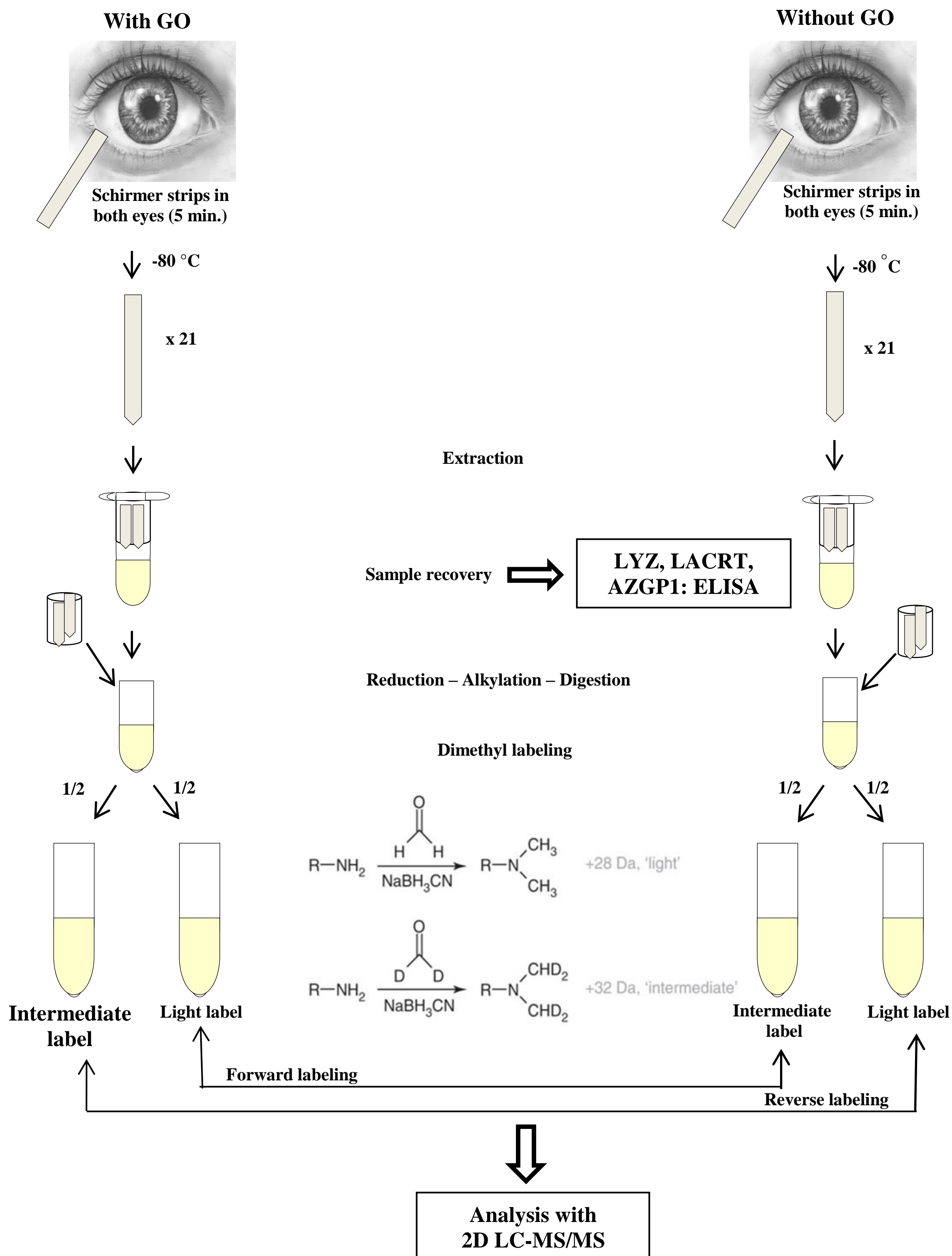


Figure 6. Schematic illustration of the single- unit filter- aided method for both sample handling and protein extraction from Schirmer strips with quantitative proteomics analysis using dimethyl labeling in combination with 2D LC-MS/MS or ELISA measurements. For untargeted 2D LC-MS/MS analysis: Twenty-one Schirmer strips, one from each of the subjects in each group, were pooled and extracted in a centrifugal filter unit. Two

independent experiments were performed by dimethyl labeling in forward and reverse directions prior to LC-MS/MS analysis to relatively quantify proteins altered between the patients with moderate to severe GO and without GO (Paper II). ELISA measurements were performed immediately after the extraction step (Paper II and III).

3.2.5 Total protein concentration

After protein extraction from Schirmer strips, total protein concentration measurements were performed by colorimetric protein assay using a Microplate absorbance reader (Tecan Austria GmbH, Grödig, Austria) and γ-microglobulin as standard.

3.2.6 ELISA measurements

The concentration of LYZ, AZGP1 and LACRT proteins was assessed with commercially available ELISA kits (on moderate to severe GO patients and without GO). LYZ levels were also assessed in the mild GO group and without GO. Protein concentrations obtained with ELISA assays were normalized with respect to total protein concentration and calculated and expressed as follows in µg/mg:

$$\text{Protein concentration } (\mu g/mg)_{\text{Normalized}} = \frac{\text{Protein concentration}}{\text{Total protein concentration}}$$

3.2.7 Digestion before LC-MS/MS analysis

As described in Supplementary 1.2 in Paper II, pooled tear protein samples from each group were reduced by adding 25 µL 100 mM dithiothreitol (DTT) solution to a final concentration of 10 mM and incubated for 1 h at 37 °C. Alkylations of free sulfhydryl groups were done by adding 22 µL 250 mM iodoacetamide (IAA) solution to a final concentration of 20 mM and incubated at 25 °C for 45 min in the dark. The sample was then diluted by adding 750 µL of 50 mM TEAB (pH 8.5) buffer to reduce the urea concentration to below 1 M, and PNGase F was added to a final concentration of 2 mM (adding 25 µL of a 500 mM PNGase F solution) before the sample was left overnight at 37 °C. Next, the reduced and alkylated proteins were digested with Lys-C/trypsin (1:20) (Pierce Thermo) for 6 h at 37 °C.

3.2.8 Dimethyl labeling

After digestion, dimethyl labeling was performed as described in detail in Supplementary 1.3 in Paper II. Digested samples from each group (Figure 6 and Figure 1 in Paper II) were dissolved in 250 µL of 100 mM TEAB buffer (pH 8.5) and divided in two equal volumes of 125 µL each, generating four samples. Then, 140 µL of 4 % (v/v) CH_2O (light label) and 140 µL of CD_2O (intermediate label) was added to one of the two sample vials, respectively, for each group (Figure 6). All four sample vials were then mixed briefly and centrifuged. Next, 140 µL of 0.6 M $NaBH_3CN$ was added and the samples were incubated in a fume hood for 1 h at 22 °C while mixing at 400 rpm. The labeling reaction was quenched by adding 560 µL of 1 % (v/v) ammonia solution and then mixed and centrifuged briefly. A further quenching and acidification was performed by adding 280 µL of 5 % formic acid and mixed. The labeled samples were then mixed in 1:1 ratio as follows; i) in the forward labeling experiment, the pooled GO sample (Graves' patients with moderate to severe GO) was light labeled while the without GO sample was intermediate labeled, and ii) in the reverse labeling experiment the

pooled GO sample was intermediate labeled (Graves' patients with moderate to severe GO) while the without GO sample was light labeled. Finally, the mixed samples were evaporated and dissolved in 0.1 % formic acid before 2D LC-MS/MS analysis. Before labeling of our samples, the labeling efficiency was tested several times on albumin, and showed an efficiency of 99 % for both the light and intermediate label.

3.2.9 Offline fractionation on a porous graphite carbon column

Mixed dimethyl labeled digests was fractionated on a porous graphitic carbon column (Hypercarb 100 x 4.6 mm Thermo) as described in Supplementary 1.4 in Paper II. The fractionation was performed at a flow rate of 0.8 mL/min using a Dionex Ultimate 3000 LC analytical system connected to an UV detector (Dionex, Thermo Scientific). A total of 25 fractions were collected manually every minute using mobile phases consisting of 0.1 % FA in water as solvent (A) and ACN with 0.1 % FA as solvent (B). The separation was run for 60 min using a multi-step gradient (0-80 % eluent B in 35 min followed by 80 % B in 5 min).

3.2.10 Nano LC-MS/MS

Described in Supplementary 1.5 Paper II, tryptic digest separation was performed on an PepMap RSLC Easy-spray C18 column (2 µm, 100 Å, 75 µm x 150 mm) using the EASY-nLC 1000 nano UHPLC system (Thermo fisher Scientific, Bremen, Germany) connected to a LTQ-Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific) equipped with a nano EASY-Spray source (Thermo Fisher Scientific). The analytical separation was run for 180 min using a multi-step gradient of 0.1 % FA in water as solvent (A) and 0.1 % FA in ACN as solvent (B) (0-25 % eluent B in 150 min and 25 %-60 % B in 20 min followed by 60 % B in 10 min) at a flow rate of 300 nL/min and a column temperature at 45 °C. The mass

spectrometer was operated in positive mode with a spray voltage set at 2.0 kV and the heated capillary temperature was kept at 200 °C. The LTQ-Orbitrap XL was operated in data-dependent mode in which one cycle of experiments consisted of one full-MS survey scan using the Orbitrap mass analyzer and subsequently five sequential MS/MS events of the most intense peaks using collision-induced dissociation (CID) in the LTQ. The MS survey scans were performed on the high resolution Orbitrap (R = 30 000) with an *m/z* range of 350-2000. Precursor ions with charge 1 or unassigned charge were rejected and the precursor ion isolation width was set to 3 *m/z* units.

3.2.11 Data analysis

Described in Supplementary 1.6 Paper II, Proteome Discoverer computational proteomics platform (Version 1.4, Thermo Scientific) with default settings was used to search raw files with MS/MS spectra created by Xcalibur software 2.07 (Thermo). Raw data files were searched against the UniProtKB *human* database (updated June 2015) using the SEQUEST HT search engine with the precursor and fragment mass tolerances set to 10 ppm and 0.6 Da, respectively. Only the peptides resulting from the tryptic cleavages were searched and two trypsin missed cleavage sites were allowed. As shown in Table 5, carbamidomethylation of cysteine/+ 57.021 Da was selected as a fixed modification. The variable modifications were as follows; + 15.995 Da for methionine oxidation, + 28.031 for dimethyl (K and N-term) light label and + 32.056 for dimethyl (K and N-term) intermediate label (Table 5). Peptide and protein false discovery rates (FDR) were set to 1 % using default filters. Dimethyl datasets were quantified using peak area with the *precursor ions quantifier* node integrated in the Proteome Discoverer with *RT tolerance of isotope pattern multiplets* set to 1 min. To correct for possible experimental bias, protein ratio distribution was normalized on protein median. The cut-off ratio for up- and down-regulated proteins was set at ≥ 2.0 and ≤ 0.5, respectively,

and only proteins quantified in both forward and reverse experiments were considered. Abundance ratios for proteins reported as differentially expressed in this study were confirmed by manual inspection of the MS spectra intensities of the labeled peptide pairs. Briefly, the criteria for passing the manual inspection were as follows: i) signal- to- noise ratios of both light and intermediate labeled peptide pairs ≥ 20, ii) light and intermediate labeled peptide ion spectra must show similar isotope patterns and expected mass shift between doublet clusters. For both forward and reverse experiments only unique peptides were considered for protein quantification.

Table 5. An overview of the fixed and variable modifications used.

Modification (Da)	Amino acid	Reagent	Product	Mass shift
Oxidation	Methionine (Met)		Methionine sulfoxide	+ 15. 995
Dimethyl labeling	K and N-termini	Formaldehyde and cyanoborohydride	Dimethylamine (light)	+ 28. 031
Dimethyl labeling	K and N-termini	Formaldehyde- d_2 cyanoborohydride	Dimethylamine (intermediate)	+ 32. 056
Carbamido-methylation	Cysteine (Cys)	Iodoacetamide	Carbamide cysteine	+ 57. 021

3.2.12 Statistical analyses

All statistical analyses were performed with IBM SPSS Statistics 22 (Armonk, NY, USA).

The study reported in **Paper I** aimed to identify as many proteins as possible in tears extracted from Schirmer strips by systematically investigate the influence of different extraction parameters. The number of proteins identified with different extraction solvents (expressed as the mean of three replicates) was compared using one-way of variance (ANOVA).

The study reported in **Paper II** aimed to quantitatively compare pooled tear fluid from Graves' patients with moderate to severe GO and Graves' patients without GO using untargeted quantitative proteomics based on dimethyl labeling in combination with 2D LC-MS/MS. Baseline demographics of continuous variables between groups were performed using Student's *t* test or Mann-Whitney U test, as appropriate. Due to normally distributed data, Pearson correlation was performed to determine the significance of the relationship between forward and reverse labeling MS experiments. A two-sided p-value < 0.05 was considered significant.

The study in **Paper III** aimed to evaluate tear levels of LYZ, LACRT and AZGP1 in individual Graves' patients with and without GO, using ELISA. And to investigate the diagnostic performance of these proteins either alone or combined in a panel. The data are presented as median unless otherwise specified due to the somewhat small sample size. We analyzed non-normally distributed data, using log-transformation or non-parametric methods, as appropriate. For comparison of continuous variables between groups, student's *t* tests or two-sided Mann-Whitney U tests were used. Spearman's correlation coefficients rank test was used for TRAb correlations due to the non-normally distributed data. Multiple logistic regression analysis was performed, with log-transformation of parameters when needed. This

was to ensure no violation of the assumptions of normality and to assess if there were any potential effects from independent variables such as age, smoking, gender and protein concentration on GO (dependent variable). A two-sided p-value < 0.05 was considered to be significant and uncorrected values are presented. Bonferroni-Holm corrections were also performed for the significance level for multiple testing, and showed that uncorrected p-values < 0.01 remained < 0.05 after correction (avoid type 1 error). However, the groups were not matched for age, gender and smoking so we cannot rule out the possibility that these factors contribute to the significance obtained in protein concentrations. The diagnostic value of LYZ, LACRT and AZGP1, individually and together, was assessed by ROC-curve analyses. However, this is a proof of concept study and further analysis on a larger cohort of patients is necessary and has been started.

In **Paper II and III** no power calculations were performed, but tear fluid from as many subjects as possible were collected during a year. The sample size was somewhat small with 21 patients in each group in the moderate to severe GO study (21 patients with moderate to severe GO and 21 patients without GO), and 18 and 9 patients in the mild GO study (18 patients with mild GO and 9 patients without GO). However, in these exploratory studies, the potential biomarkers of interest were analyzed with two different methods (LC-MS/MS and ELISA), which strengthens the interpretation of the results.

4 Main results-Summary of papers

4.1 Paper I

The aim of this study was to develop a single-unit filter-aided method for both sample handling and protein extraction, and to investigate the most suitable conditions for protein extraction. Tear fluid from healthy laboratory volunteers with no history of ocular disease were collected using Schirmer tear test. A total of 48 strips were collected and before extraction, the strips were mixed together and two strips were randomly collected and placed in each centrifugal filter unit. The whole sample preparation procedure, including protein extraction, reduction, alkylation, and digestion, was performed in a single-unit centrifugal filter before analyzed on a 1D LC-MS/MS system. Eight different extraction solvents were evaluated, and an increased number of identified proteins were seen with ammonium carbonate in combination with 50 mM NaCl. In general, there were a high number of proteins overlapping between the extractions solvents. The effect of extraction volume on the number of identified proteins was examined, showing that 450 µL resulted in the highest number of identified proteins compared with 200 µL and 1000 µL. Extraction time (4 h and overnight) and extraction temperature (25 °C, 37 °C and 50 °C) were also evaluated. The results showed an increased number of identified proteins with a 4 h extraction at 25 °C. Extracted proteins were also subjected to 2D LC-MS/MS analysis employing an SCX column in the first dimension and a RP column in the second dimension. 2D LC-MS/MS analysis resulted in 1526 protein identifications compared with 309 with 1D LC-MS/MS analysis. These findings suggest that a single-unit filter-aided extraction method in combination with 2D LC-MS/MS increased the number of identified proteins. This dataset of identified proteins represents a comprehensive catalogue of the human tear fluid proteome and may serve as a list for future biomarker research.

4.2 Paper II

This paper is based on the method development described in Paper I. The aim of this paper was to quantitatively compare pooled tear fluid from Graves' patients with moderate to severe GO and Graves' patients without GO using untargeted quantitative proteomics based on dimethyl labeling in combination with 2D LC-MS/MS. Quantitative comparison between the two groups was performed using dimethyl labeling (both forward and reverse labeling) in combination with 2D LC-MS/MS (Figure 6). 1212 proteins were identified at a 1 % FDR, and a total of 1136 proteins were quantified in both forward and reverse measurements. Using a cut-off value of ≥ 2.0 and ≤ 0.5 for up- and down-regulated proteins, 16 proteins showed significant altered regulation in tear samples from the moderate to severe GO group compared to the Graves' patients without GO. Caspase 14, dermcidin, antileukoproteinase (SLPI), procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (PLOD2), mesothelin, deleted in malignant brain tumors 1 (DMBT1), apolipoprotein D (APOD), zymogen granule protein 16 homolog B (ZG16B), lysozyme C (LYZ), glutathione peroxidase 3, extracellular glycoprotein lacritin (LACRT, Figure 7) and zinc-alpha-2-glycoprotein1 (AZGP1) were all up-regulated in both forward and reverse experiments in the sample from the GO group with fold-changes ranging from 2.1-5.6. Cystatin D, periplakin, mammaglobin A and prelamin A/C were down-regulated in the moderate to severe GO group compared with the Graves' patients without GO. We showed that a number of regulated proteins differed in tear fluid from patients with moderate to severe GO compared with Graves' patients without GO. These new observations may be useful to better understand the disease and may be used as biomarkers and indicators for disease activity and risk for developing GO.

4.3 Paper III

Based on the results from Paper II, this paper is a validation study describing the three proteins, LYZ, LACRT and AZGP1, in tear fluid from individual patients with moderate to severe GO, and patients without GO, using ELISA measurements. The mild GO group was assessed for only LYZ levels. The moderate to severe GO group had a higher median age ($p = 0.024$), increased number of smokers ($p = 0.043$) and increased TRAb levels than the patients without GO ($p < 0.001$). For the mild GO group there were no significant differences in smoking ($p = 0.27$) and age ($p = 0.50$) compared to the patients without GO. However, the TRAb levels were significantly higher in 16 patients in the mild GO group compared to the patients without GO ($p = 0.003$), but no significant difference in TRAb levels were observed between the two groups when comparing the highest TRAb levels between the mild GO group and patients without GO.

Significantly higher tear levels were found in the moderate to severe GO group compared with the patients without GO, in LYZ (median, 269 µg/mg ($n = 21$) vs. 84 µg/mg ($n = 21$), $p < 0.001$, Figure 8A), LACRT (median, 7 µg/mg ($n = 20$) vs. 0.9 µg/mg ($n = 20$), $p = 0.004$, Figure 8B) and AZGP1 (median, 42 µg/mg ($n = 20$) vs. 22 µ/mg ($n = 20$), $p = 0.001$, Figure 8C). Higher levels of LYZ, AZGP1 and LACRT were significantly associated with GO after adjustment for age, smoking and gender. Using ROC analysis, LYZ showed high accuracy with an AUC value of 0.91, while LACRT and AZGP1 showed AUC values of 0.77 and 0.80, respectively. When combined, LYZ, LACRT and AZGP1 tear fluid concentrations could detect patients with moderate to severe GO in GD with an AUC of 0.93. Tear levels of LYZ (median, 51 µg/mg vs. 14 µg/mg, $p = 0.003$), were significantly higher in patients with mild GO compared to patients without GO. LYZ were significantly associated with GO, and this association was still significant after adjustment for age, smoking and gender. ROC analysis showed an AUC of 0.86 for LYZ. ROC analysis of the highest TRAb levels revealed an AUC

of 0.73 for the moderate to severe GO group and an AUC of 0.46 for the mild GO group. These findings suggest that this novel three protein biomarker panel is able to differentiate between Graves' patients with and without GO and are comparable to serum TRAb.

5 Discussion

The objective of this thesis was to search for protein biomarkers for GO by comparing tear protein levels between patients with and without GO using quantitative proteomics.

In the following discussion part of this thesis, firstly, validity of the study is discussed. Furthermore, demographic, clinical and serological data of the patients included in this study will be discussed. Then sample preparation and 2D LC-MS/MS proteomic analyses using quantification with dimethyl labeling of tear fluid from patients with and without GO is discussed (pooled samples). Finally, confirmation with ELISA on selected proteins will be discussed (individual samples).

5.1 Methodological considerations

5.1.1 Internal and external validity

Validity is divided into internal validity and external validity (generalizability). Internal validity refers to the reliability of the study results, which is determined by the degree to which a study minimizes bias. To obtain a high internal validity, the study group should be as homogeneous as possible and is usually obtained by rigid inclusion and exclusion criteria, blinding and randomization/matching [134]. One of the most common types of bias in clinical research is those related to subject selection. Selection bias refers to systematic differences between baseline characteristics (age, gender, smoking, treatment) of the groups being compared. In our study, we enrolled every patient with the pre-set inclusion and exclusion criteria (see Section 3.3.3-3.3.4). Selection bias occurs when the criteria used to recruit patients into separate groups are inherently different, and this can be a problem with studies such as this since exposure and outcome (GD with and without GO) already have occurred at the time individuals were selected for study inclusion [134].

Due to our inclusion and exclusion criteria, there were differences between our groups' demographics (age and gender), disease specific characteristics (disease duration, TRAb levels and thyroid hormone levels), RAI treatment and smoking. Based on the differences between GD patients with and without GO, we cannot rule out the possibility that some of the variables in the inclusion/exclusion criteria may have affected our results regarding protein concentrations. We enrolled as many patients as possible during a year that met the inclusion and exclusion criteria set before the recruitment. Due to the somewhat loose inclusion and exclusion criteria set, matching of the groups and more rigid inclusion/exclusion criteria was not prioritized. However, this is necessary to reduce possible selection bias, and increase the internal validity of the study. Additionally, this study was an unblinded study where all parties were aware of the group assignment (with/without GO), which may have biased the study. However, in our opinion the outcome is already defined and would not reflect the results [134]. Another possible bias in this study is misplacement of recruited GO patients in the wrong GO group. For example a GO patient with seemingly mild GO will be placed in the mild GO group, whereas it actually may belong to the moderate to severe GO group.

External validity refers to the degree to which the findings are able to be generalized to other groups or populations outside the study group. Increased external validity can be obtained by widening the exclusion criteria so that the study population is very similar to the general population. However, wide inclusion criteria may compromise the internal validity of the study [134, 135]. In this study, all patients were recruited from one clinic Oslo University Hospital (OUS) which may impact the external validity of this study since we cannot rule out that clinicians at OUS perform different evaluations than clinicians in other hospitals. Achieving balance between internal and external validity is difficult, and an ideal study would randomize patients and blind those collecting and analyzing data, while keeping exclusion criteria to a minimum. As mentioned, we included all patients willing to participate in the

study using the pre-set inclusion and exclusion criteria to obtain as large groups as possible and make them as similar to the general patient population with GD and GO as possible. Since we did that, the internal validity may have been reduced but the external validity somewhat increased. However, it would have been an option to widen the inclusion and exclusion criteria in this study to reduce selection bias, even though this would have been on the expense of the external validity.

5.1.2 Patient characteristics

The choice of patients and differences in activity and severity

The most widely used comparison of patient groups in proteomic GO studies is the comparison between GO patients with different activity and severity and healthy subjects [124-126]. Since GD patients are the target population in which the biomarkers will be measured in, healthy controls are often inappropriate for defining disease-specific biomarkers in GO (Paper II). Therefore, patients with low activity GD without GO were chosen as the control group instead of healthy subjects. However, there may be a chance that a GD patient with no clinical apparent GO has some GO involvement and signs of orbital inflammation.

GO passes through several phases, where the first phase involves worsening of symptoms and signs. Active GO refers to the presence of inflammation, and is the phase when the patient is symptomatic and notice changes in the severity, which may result in double vision and/or proptosis [136]. In contrast, when there is no longer a change in severity of any feature (for up to 2 years), the disease has passed on to the chronic inactive phase. Considering these two stages of the disease manifestation, Planck et al. [137] found a clear difference in gene expression between the active and chronic phase of GO. While many genes involved in inflammation and immune response were up-regulated in active GO, up-

regulation of acute inflammatory genes was not observed in chronic GO. With this in mind, all GO patients in this study had a CAS score ≥ 3/7. The active GO patients were further divided into moderate to severe GO and mild GO to obtain as homogeneous groups as possible. The severity of GO was determined by an ophthalmologist in collaboration with an endocrinologist, and based on the EUGOGO classification (eyelid swelling, eyelid aperture, proptosis, eye mobility, visual acuity and color vision) [47-49]. Patients considered as having moderate and severe GO were evaluated for treatment with glucocorticoids, whereas patients with mild GO were not treated by glucocorticoids. However, the patients with mild GO were also strictly followed up by CAS assessments, and encouraged to stop smoking, use tear substitutes, and avoid bright light and dust.

Differences in TRAb levels

TRAb activity is currently considered as important contributors to GO, and enhanced expression of the autoantigen to TSHR within the orbit may play a role in the initiation or propagation of the autoimmune response in GO [137]. Several experimental and clinical studies supports the theory of TRAb involvement in GO [34-37, 40]. For instance, Eckstein et al. [34] observed a relationship between TRAb levels and GO severity and outcome. In our study, increased TRAb levels were observed in all patients with moderate to severe GO and in 16 of the 18 patients in the mild GO group, compared to patients without GO at the time of sample collection. In addition to the above observations on TRAb at the time of sample collection, we also evaluated the highest TRAb levels in patients during their disease course. Significantly higher TRAb levels were observed in all patients with moderate to severe GO compared to patients without GO, when comparing the highest TRAb levels in the two groups. However, there was no significant difference in the TRAb levels between mild GO patients compared to patients with no sign of GO, when comparing the highest TRAb levels in the two

groups. This indicates that TRAb levels are higher in patients with moderate to severe GO than in patients with mild GO and this is in line with findings observed by Eckstein et al [34].

Differences in smoking

The observed association between smoking and autoimmune diseases may indicate that there is a generalized stimulation of autoimmune processes in smokers [59]. Baker et al. [123] demonstrated that the tear composition of GO patients was altered, and that similar changes could be found in tears of smokers. The number of smokers was significantly higher among the patients with moderate to severe GO than patients without GO. However, the number of former and current smokers was not significantly higher in the mild GO group compared to patients without GO. As suggested in earlier research [123], the significantly higher number of smokers in the moderate to severe GO group indicates that smoking is a strong risk factor contributing to GO development and progression in patients with GD. This indicates that cigarette smoking has a strong and dose-responsive relationship with the course and severity of GO, and suggests that smokers generally suffer from more severe GO than non-smokers [60, 63]. Because smoking is a significant risk factor for the development of GO, and our groups are not matched for smoking, we cannot rule out the possibility that smoking may affect the levels of the dysregulated proteins in tear fluid observed between the groups. Additional studies with larger groups, matched for smoking, are necessary to determine the impact of smoking on tear protein composition. Since smoke extract is known to increase adipogenesis [60] and ex-smoking GO patients reduced progression in diplopia and proptosis, quitting smoking should be the first recommendation for patients with GO. Moreover, treatment with glucocorticoids has shown to have increased effect on non-smokers compared to smokers [64].

Differences in age and gender

In the two study groups, the age inclusion criteria were the same. Patients with moderate to severe GO group were significantly older than those with no signs of GO. However, patients in the mild GO group were not significantly older than those with no signs of GO. Usually, GD patients with GO are older than patients without GO (mean age: 46.4 years vs. 40.0 years) [14]. In this study the moderate to severe GO group was older than the mild GO group by 5.5 years, which supports the observation associated between severity of GO and advancing age [14].

As is the case in most studies on autoimmune disease [21], there were more women than men in the GO group compared to the patients without GO. There was no restriction on gender in the inclusion criteria in our study, and the increased number of women included in the GO groups may be reasonable due to a higher frequency of GD in women (with a prevalence in general women population of 2.5 %) [8-12]. Usually, higher incidence in women compared with men has been observed. However, a recent study observed no difference in gender distribution among patients without GO, mild GO and moderate to severe GO [20].

5.1.3 Sample size and power

This study is exploratory, and we did not perform power calculations [138]. No matching of the groups concerning smoking, age or gender was made. We intended to include 40 patients with GD and GO and 40 patients with GD without GO, but the recruitment of patients (especially GD patients without GO) proved more difficult than anticipated. Initially we included 39 GD patients with GO. We decided to divide the group into mild and moderate to severe GO patients which, resulted in two GO groups containing 21 and 18 patients, respectively. It was challenging to find subjects with GD without any clinical history of GO. Tear samples were collected from the two control groups, with 29 and 9 GD patients.

However, after a closer look, 8 patients of the totally 29 patients in the first control group had previously shown clinically apparent GO involvement and were excluded from the study.

5.1.4 Sample preparation and optimization

Due to high concentrations of proteins in tear fluid and the advantages of MS, there has been an increased attention to proteomic analysis of tear fluid. However, there are several analytical aspects to consider in proteomic analysis that may contribute to protein loss during processing, such as sample preparation and optimization.

The choice of collection procedure

In the present study, the well-established Schirmer tear test was used for tear collection (Figure 6 and Paper I). Schirmer tear test is rapid and easy to handle, does not require experienced personnel, is free of risk for injury, and it is considered more pleasant for the patient compared to capillary [139]. The procedure is routinely used in the ophthalmology clinic as a standard clinical test for dry eye disease, and has shown to be a good collection procedure, as shown in previous studies [139-141]. However, there are some challenges with this collection procedure; one is the contamination of proteins from epithelial cells on the ocular surface (conjunctiva and cornea) [107], and since an extraction step is required, a 100 % tear fluid recovery is not possible from the Schirmer strips. An alternative collection procedure of tear fluid is collection with glass capillary. With this method no proteins from the conjunctiva or cornea are collected, however, a degree of practice and experience is required, and subjects have shown to be more hesitant about the rigid capillary tube in this procedure compared with the flexible Schirmer tear test. Collection with glass capillary is more time-consuming and requires that the subjects have normal tear fluid secretion to obtain enough tear fluid. The opinion on which collection procedure is best suited differs [139-146],

and both the subject and the aim of measurement should be considered. Although sample handling and protein extraction from Schirmer tear test have been challenging, several studies on the human tear proteome have used this collection technique. Cutting and removal of the Schirmer strips after extraction, and several sample transfer-steps may lead to extensive sample loss prior to downstream analysis [120, 146-151].

The choice of extraction solvent

We examined different extraction solvents to find the solvent resulting in the highest number of identified proteins in tears. We observed an increased number of identified tear proteins with ammonium bicarbonate in combination with NaCl. This shows that using a non-optimal extraction solvent can result in serious sample loss already in the first step of the analysis. Ammonium carbonate as extraction solvent has previously been used for protein extraction from Schirmer strips by Green-Church et al. [146]. They used ammonium bicarbonate to extract proteins from Schirmer strips and then precipitated the extract with acetone [146]. Since NaCl is present in tears, we thought it may increase the solubility of tear proteins and promote the extraction from the strips [152]. Additionally, we evaluated three MS-compatible surfactants RapiGest SF (RG SF), Invitrosol, and PPS Silent Surfactant (PPS), to try to improve the protein extraction further, as suggested by Chen et al [153]. However, all three surfactants turned out to be less effective compared with ammonium bicarbonate containing NaCl. In general, the number of protein identities overlapping between the extraction solvents applied was high (Paper I), and typical tear proteins, such as LYZ, lipocalin 1, mammaglobin B and AZGP1, were identified in all extraction solvents applied. However, some differences were observed, indicating that some solvents extracted specific proteins. This shows that the choice of extraction solvent affects both the number of identified proteins as well as which proteins are extracted.

The choice of extraction volume, time and temperature

The effect of extraction volume on the number of identified proteins was examined. We observed solvent breakthrough (solvent not retained by the filter) with volumes larger than 600 µL, which shows that too high extraction volumes, may contribute to sample loss and reduced number of identified proteins. The highest number of identified proteins was observed using 450 µL, and was therefore used for further analyses.

It has been shown that 3-h extraction is sufficient to achieve time-independent protein extraction from Schirmer strips [148, 149, 154]. In our study, we experienced a slight increase in the number of identified proteins with overnight extraction compared with a 4-h extraction. Despite this slight increase in the number of identified proteins with overnight extraction, a 4-h extraction was selected because it is less time-consuming.

Usually, diffusion-based extraction improves when the temperature is increased [155]. We evaluated increased temperatures; however, the extraction efficiency did not improve with increasing temperatures in our study. In fact, due to the limitations related to the filter unit, showing a maximum temperature limit of 40 °C, lower temperatures turned out to be the most suitable for the number of identified proteins from Schirmer strips using this type of filter unit.

Two dimensional liquid chromatography mass spectrometry

As mentioned, to achieve greater analytical depth for a proteome of interest, pre-fractionation of the sample prior to LC-MS/MS is important and would preferably increase the number of proteins identified. After evaluation of the different extraction parameters on the number of identified proteins on a 1D LC-MS/MS system, extracted proteins were subjected to 2D LC-MS/MS analysis. As expected, we were able to identify 1217 more proteins with the optimized extraction parameters.

5.2 Quantitative proteomic comparison between GD patients with and without GO-Main findings

5.2.1 Dimethyl labeling based comparative proteomics of tear fluid

Quantification of protein expression in biological systems is an important part of proteomics research, and there are several techniques for quantifying the protein abundance between samples. In this study (Paper II), untargeted quantitative proteomics with 2D LC-MS/MS analysis was applied on pooled tear samples from GD patients with moderate to severe GO and without GO.

Dimethyl labeling based quantitative proteomics

Dimethyl labeling is a fast and cost-effective procedure that is applicable to any biological sample, and was therefore considered a good choice for quantification of proteins in this study. We labeled our samples with light and intermediate labels and chose to reverse label (forward and reverse label-swap can be performed by swapping the stable isotope labels, Figure 6, 7, and Paper II) the samples to increase the accuracy of the labeling step. Since we only needed two labels, we chose to label with light and intermediate labels due to recommendations from previous studies due to the interpretation of both deuterium and ^{13}C isotope in the heavy label [93]. Since this labeling procedure is performed on peptide level, variations introduced in the earlier steps of the sample preparation may have been lost. Additionally, the introduction of deuterium in the intermediate label may cause some isotopic effects [96, 99].

Up- and down-regulated proteins in patients with GO

LC-MS/MS analysis clearly showed differences in tear protein composition between the two groups investigated (Paper II). Several proteins were found to be significantly up-regulated or down-regulated in the GO group compared to the controls (Table 6). However, other proteins identified in this study can be related to GO even though the fold changes were outside the pre-set cut-off values (≥ 2.0 and ≤ 0.5 for up- and down-regulated proteins). The majority of the proteins identified and quantified in both sets were typical tear proteins, such as LYZ and LACRT. For instance, LACRT was 2.2-fold up-regulated (Table 6) in both forward and reverse experiments as shown in Figure 7. Many of the proteins we found dysregulated are involved in inflammation (antileukoproteinase (SLPI), LYZ, apolipoprotein D (APOD)) and immune response (deleted in malignant brain tumors 1 (DMBT1), LYZ and AZGP1) of the ocular surface (Table 6). For instance, LYZ has shown to play a role in the immune response [156], and increased tear levels of LYZ in patients with GO have been reported in previous studies. Khalil et al. found increased levels of LYZ in tears from patients with GO and suggested lacrimal gland involvement in patients with GO [122]. Matheis et al. showed elevated levels of LYZ in patients with GO compared with healthy controls, suggesting increased inflammatory process of the orbit and the lacrimal gland [125]. Although most of the proteins have been reported in tears earlier, proteins such as caspase 14, periplakin, mammaglobin A, and prelamin A/C, have to our knowledge not been reported in tears of GO patients before (Paper II). These new observations may be useful to better understand the disease.

Table 6. Altered proteins (expressed as fold changes) in patients with moderate to severe GO compared with patients without GO (Paper II).

Proteins	Mean fold- change [a]	Function in tears/ eye
Up- regulated		
Caspase 14	5.6	apoptosis
Dermcidin	3.4	defense
Antileukoproteinase (SLPI)	3.1	inflammation
Procollagen-lysine, 2- oxoglutarate 5- dioxygenase 2 (PLOD2)	2.9	unknown
Mesothelin	2.9	unknown
Deleted in malignant brain tumors 1 (DMBT1)	2.9	immune response
Apolipoprotein D (APOD)	2.8	maintenance
Zymogen granule protein 16 homolog B (ZG16B)	2.8	protection
Lysozyme C (LYZ)	2.7	immune response, inflammation
Glutathione peroxidase 3	2.6	unknown
Extracellular glycoprotein lacritin (LACRT)	2.2	wound healing
Zinc- alpha- 2- glycoprotein 1 (AZGP1)	2.1	immune response
Down- regulated		
Periplakin	0.5	unknown
Cystatin D	0.4	proteolysis
Mammaglobin A	0.4	unknown
Prelamin A/C	0.4	unknown

[a] Mean fold- change obtained for forward and reverse experiments.

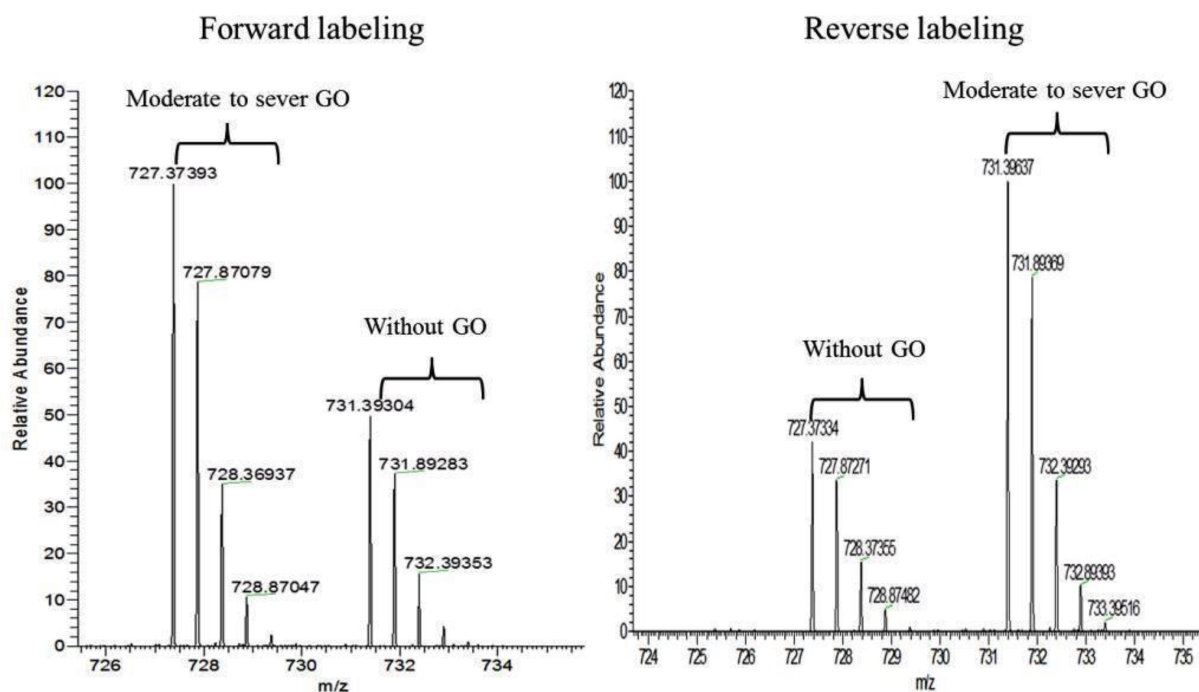


Figure 7. Mass spectra for light and intermediate labeled peptide (QFIENGSEFAQK) pairs from forward and reverse labeling experiments for LACRT. The peaks at *m/z* 727.37 and 731.39 correspond to light and intermediate labels for active GO, respectively (Supplementary Paper II).

5.2.2 ELISA measurements

In the third part of this thesis, we examined the protein levels of LYZ, LACRT and AZGP1 as potential protein biomarkers for GO on individual patient samples, using commercially available ELISA assays (Paper III). These proteins were chosen because of their up-regulation in GO in combination with their high abundance in tear fluid which is highly important when dealing with limited sample volume available. Other low-abundant proteins were also analyzed with ELISA (Table 6) but their detection proved to be difficult. As mentioned, the concentration of LYZ, AZGP1 and LACRT was assessed in tears from patients with moderate to severe GO and without GO. Additionally, only LYZ levels were assessed in the mild GO group and Graves' patients without GO due to limited sample volume available.

Tear levels of LYZ, LACRT and AZGP1 in patients with moderate to severe GO

The results from ELISA measurements in the moderate to severe GO group showed elevated levels of LYZ in the GO group (Figure 8A), which are in line with findings from former studies [125, 133]. We also observed elevated levels of LACRT in the GO group compared with patients without GO (Figure 8B). LACRT levels were only detected in 13 out of 20 tear samples in the GO group and 11 out of 20 tear samples in the patients without GO. This was probably a result of limited sample volume available. In this study we also found increased AZGP1 levels in the GO group compared to those without GO (Figure 8C), which is in line with previously reported up-regulation of AZGP1 in tears from patients with GO and healthy smokers [125]. Our logistic regression analysis showed significant association between GO and AZGP1, both with and without adjustment for smoking, age and gender. However, we did not match our groups for smoking, and can therefore not rule out the possibility that smoking may affect AZGP1 tear levels in this study, since smoking is a significant risk factor for the development of GO (Paper III).

Further analysis to study the possibility to differentiate between GD patients with and without GO was performed by measuring tear levels of LYZ, LACRT and AZGP1 using receiver operator characteristic (ROC) analysis. Even though each protein individually had reasonably high values for area under the curve (AUC), the combination of the three biomarkers combined in a panel performed better as a diagnostic test than a single protein alone. However, the diagnostic performance of our biomarker panel needs to be further validated prospectively in a larger population. ROC analysis was also used to evaluate if the highest TRAb levels could differentiate between GD patients with and without GO, and if this differentiation was better as a diagnostic test than the diagnostic performance of the three protein biomarker panel. The ROC analysis revealed an AUC of 0.73 for the moderate to severe GO patient group, which is lower than that obtained for the biomarker panel (AUC of 0.93). This

indicates that the three protein panel may differentiate GD patients with moderate to severe GO better than TRAb.

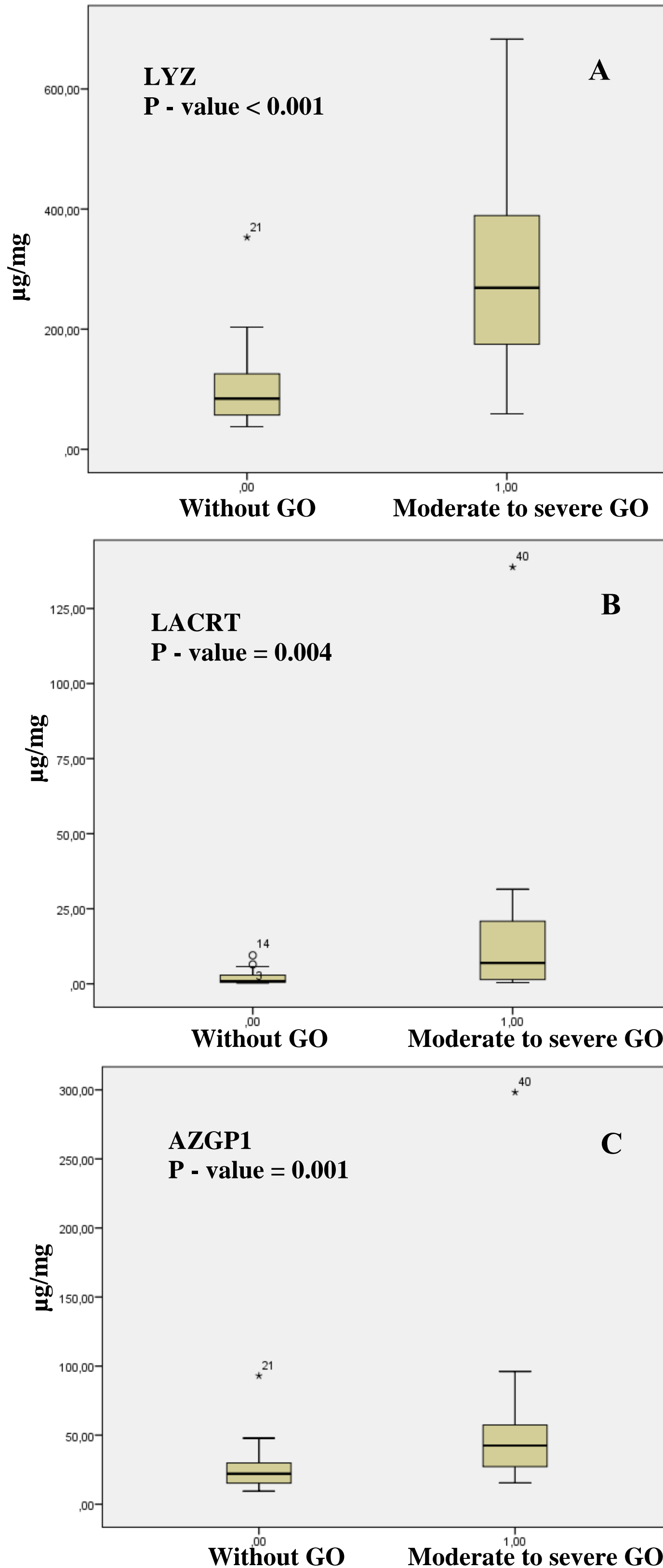


Figure 8. Levels of LYZ, LACRT and AZGP1 in tear fluid in moderate to severe GO patients vs patients without GO obtained with ELISA. Box and-whisker plot showing median and interquartile range (IQR). **(A)** LYZ (269 µg/ mg vs 84 µg/ mg), **(B)** LACRT (7 µg/ mg vs 0.9 µg/ mg) and **(C)** AZGP1 (42 µg/ mg vs 22 µ/ mg). Mann-Whitney U tests were performed and *p*- values are indicated on each graph (Paper III).

Tear levels of LYZ in patients with mild GO

Since LYZ showed the best performance in discriminating between GD patients with moderate to severe GO and those with no signs of GO, in combination with limited sample volume available, we wanted to examine LYZ tear levels in GD patients with mild GO. Significantly higher LYZ levels were observed in patients with mild GO compared to those without GO. Also for this group, the highest TRAb levels during disease were examined using ROC analysis to see if a better diagnostic performance was obtained with TRAb compared with LYZ. The results showed an AUC of 0.46, which is much lower than the AUC observed for LYZ (AUC of 0.86). These observations may represent a possible manifestation of subclinical stage of the disease. Usually, patients with mild GO are not recommended for glucocorticoid treatments and a wait- and- see policy is favored in hope of a spontaneous improvement [157]. However, a few patients progress into moderate to severe stages of GO and these patients are difficult to detect with clinical examination methods. We therefore believe that measurement of LYZ tear levels may help the clinician in deciding which patients need extra care concerning development of GO in the clinically mild GO group. Additionally, LYZ levels may help the clinician to decide which permanent treatment (thyroidectomy or RAI) that should be considered. Further studies to find clinically relevant cut-off values predicting patients at risk for developing GO are desirable (Paper III).

6 Conclusions

- To easy evaluate the tear fluid proteome in future studies, we have developed a robust single-unit filter-aided method for both sample handling and protein extraction from Schirmer strips.
- A number of novel regulated proteins in pooled tear fluid from patients with moderate to severe GO were identified using 2D LC-MS/MS in combination with dimethyl labeling.
- Tear levels of LYZ, LACRT and AZGP1 have been evaluated in patients with GO at an individual level using ELISA, and we have established a panel based on these 3 proteins that is able to differentiate between GD patients with moderate to severe GO from those without GO.
- High LYZ levels in patients with GD may suspect GO development and may indicate that additional follow-up and treatment of the patient is necessary.
- Up-regulation of lacrimal gland proteins such as LYZ and LACRT suggest involvement of the lacrimal gland in the pathogenesis of GO.
- Our groups were not matched for smoking, age and gender. Consequently, the possibility that smoking, age and gender may affect tear levels of proteins found regulated in the present study cannot be ruled out. Therefore, they need to be reproduced in a larger independent sample group matched for smoking, age and gender.

6.1 Future research

There is a need for further studies of the proteomic profile in tear fluid from GD patients with GO to verify the results of this thesis. Furthermore, the patient groups should be matched for smoking, age and gender. Other groups such as dry eye patients and patients with Sjøgren's syndrome should also be investigated. Finally, further evaluation on a larger prospective cohort of patients with power calculations, would be necessary to increase the diagnostic accuracy of the biomarker panel to evaluate the risk of developing GO among patients with GD.

7 References

[1] Weetman AP. Graves' disease. N Engl J Med 2000 October; 343:1236-1248.

[2] Cooper DS. Hyperthyroidism. Lancet 2003 August; 362:459-468.

[3] Brent GA. Graves' disease. N Engl J Med 2008 June; 358:2594-2605.

[4] Eckstein AK, Johnson KTM, Esser J, Ludgate M. Current insights into the pathogenesis of Graves' orbitopathy. Horm Metab Res 2009 May; 41:456-464.

[5] Martin A, Nakashima M, Zhou A, Aronson D, Werner AJ, Davis TF. Detection of major T cell epitopes on human thyroid stimulating hormone receptor by overriding immune heterogeneity in patients with Graves' disease. J Clin Endocrinol Metab 1997 June; 82:3361-3366.

[6] Ginsberg J. Diagnosis and management of Graves' disease. CMAJ 2003 March; 168:575-585.

[7] Brix TH, Kyvik KO, Christensen K, Hegedus L. Evidence for a major role of heredity in Graves' disease: A population-based study of two Danish twin cohorts. J Clin Endocrinol Metab 2000 October; 86:930-934.

[8] Tunbridge WM, Evered DC, Hall R, Appleton D, Brewis M, Clark F, Evans JG, Young E, Bird T, Smith PA. The spectrum of thyroid disease in a community: the Whickham survey. Clin Endocrinol (Oxf) 1977 August; 7:481-493.

[9] Vanderpump MPJ, Tunbridge WMG, French JM, Appleton D, Bates D, Clark F, Evans JG, Hasan DM, Rodgers H, Tunbridge F, Young ET. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham survey. Clin Endocrinol (Oxf) 1995 March; 43:55-68.

[10] Laurberg P, Pedersen KM, Hreidarsson A, Sigfusson N, Iversen E, Knudsen PR. Iodine intake and the patterns of thyroid disorders: a comparative epidemiological study of thyroid abnormalities in the elderly in Iceland and Jutland, Denmark. J Clin Endocrinol Metab 1997 December; 83:765-769.

[11] Bjoro T, Holmen J, Kruger O, Midthjell K, Hunstad K, Schreiner T, Sandnes L, Brochman H. Prevalence of thyroid disease, thyroid dysfunction and thyroid peroxidase

antibodies in a large, unselected population. The Healt Study of Nord-Trondelag (HUNT). Eur J Endocrinol 2000 June; 143:639-647.

[12] Hollowell JG, Staehling NW, Hannon WH, Gunter EW, Spencer CA, Braverman LE. Serum TSH, T4, and Thyroid Antibodies in the United States Population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). J Clin Endocrinol Metab 2002 January; 87:489-499.

[13] Abraham-Nordling M, Bystrom K, Torring O, Lantz M, Berg G, Calissendorff, Nystrøm HF, Jansson S, Jørneskog G, Karlsson FA, Nystrøm E, Ohrling H, Ørm T, Hallengren B, Wallin G. Incidence of hyperthyroidism in Sweden. Eur J Endocrinol 2011 September; 165:899-905.

[14] Perros P, Crombie AL, Matthews JNS, Kendall-Taylor P. Age and gender influence the severity of thyroid-associated ophthalmopathy: a study of 101 patients attending a combined thyroid-eye clinic. Clin Endocrinol 1993 January; 38:367-372.

[15] Abraham-Nordling M, Torring O, Lantz M, Hallengren B, Ohrling H, Lundell G. Incidence of hyperthyroidism in Stockholm, Sweden, 2003-2005. Eur J Endocrinol 2008 February; 158:823-827.

[16] Kalk WJ, Kalk J. Incidence and causes of hyperthyroidism in blacks. S Afr Med J 1989 June; 75:114-117.

[17] Bartalena L, Fatourechi V. Extrathyroidal manifestations of Graves' disease: a 2014 update. J Endocrinol Invest 2014 June; 37:691-700.

[18] Wiersinga WM, Bartalena L. Epidemiology and prevention of Graves' Ophthalmopathy. Thyroid 2002 June; 12:855-860.

[19] Bahn RS. Current Insights into the Pathogenesis of Graves' Ophthalmopathy. Hormone and Metabolic Research 2015 June; 47:773-778.

[20] Tanda ML, Piantanida E, Liparulo L, Veronesi G, Lai A, Sassi L, Pariani N, Gallo D, Azzolini C, Ferrario M, Bartalena L. Prevalence and Natural History of Graves' Orbitopathy

in a Large Series of Patients With Newly Diagnosed Graves' Hyperthyroidism Seen at a Single Center. J Clin Endocrinol Metabo 2013 April; 98:1443-1449.

[21] Laurberg P, Berman DC, Pedersen IB, Andersen S, Carle A. Incidence and clinical presentation of moderate to severe Graves' oebitopathy in a Danish population before and after iodine fortification of salt. J Clin Endocrinol Metab 2012 July; 97:2325-2332.

[22] Wiersinga WM, Smit T, Van der Gaag R, Koornneef L. Temporal relationship between onset of Graves‘ ophthalmopathy and onset of thyroidal Graves‘ disease. J Endocr Invest 1988 September; 11:615-619.

[23] Marcocci C, Bartalena L, Bogazzi F, Panicucci M, Pinchera A. Studies on the occurrence of ophthalmopathy in Graves' disease. Acta Endocrinol (Copenh) 1989 January; 120:473-478.

[24] Salvi M, Zhang ZG, Haegert D, Woo M, Liberman A, Cadarso L, Wall JR. Patients with endocrine ophthalmopathy not associated with overt thyroid disease have multiple thyroid immunological abnormalities. J Clin Endocrinol Metab 1990 January; 70:89-94.

[25] Bartalena L, Tanda ML. Graves ophthalmopathy. N Engl J Med 2009 March; 360:994-1001.

[26] Weetman AP, Cohen S, Gatter KC, Fellis P, Shine B. Immunohistochemical analysis of the retrobulbar tissues in Graves' ophthalmopathy. Clin exp. Immunol 1989 January; 75:222-227.

[27] Heufelder AE, Bahn RS. Detection and localization of cytokine immunoreactivity in retroocular connective tissue in Graves` ophthalmopathy. Eur J Clin Invest 1993 January; 23:10-17.

[28] Virakul S, Steensel L, Dalm VASH, Paridaens D, Van Hagen PM, Dik WA. Platelet-derived growth factor: A key factor in the pathogenesis of Graves ophthalmopathy and potential target for treatment. Eur Thyroid J 2014 August; 3:217-226.

[29] Khoo TK, Coenen MJ, Schiefer AR, Kumar S, Bahn RS. Evidence for enhanced Thy-1 (CD90) expression in orbital fibroblasts of patients with Graves`ophthalmopathy. Thyroid 2008 June; 18:1291-1296.

[30] Bahn RS. Pathogenesis of Graves' ophthalmopathy. N Engl J Med 1993 November; 329:1468-1475.

[31] Kemp EG, Rootman J. Lipid deposition within the extra-ocular muscles of a patient with dysthyroid ophthalmopathy. Orbit 1989 January; 8:45-48.

[32] Bahn RS. Mechanisms of disease Graves`ophthalmopathy. N Engl J Med 2010 February; 362:726-738.

[33] Teng CSS, Yeo PPB. Ophthalmic Graves' disease: Natural history and detailed thyroid function studies. Br Med J 1977 January; 1:273-275.

[34] Eckstein AK, Plicht M, Lax H, Nauhuser M, Mann K, Lederbogen S, Heckmann C, Esser J, Morgenthaler NG. Thyrotropin receptor autoantibodies are independent risk factors for Graves' ophthalmopathy and help to predict severity and outcome of the disease. J Clin Endocrinol Metabo 2006 July; 91:3464-3670.

[35] Gerding MN, van der Meer JW, Broenink M, Bakker O, Wiersinga WM, Prummel MF. Association of thyrotropin receptor antibodies with the clinical features of Graves' ophthalmopathy. Clin Endocrinol (Oxf) 2000 January; 52:267-271.

[36] Khoo DH, Ho SC, Seah LL, Fong KS, Tai ES, Chee SP, Eng PH, Aw SE, Fok AC. The combination of absent thyroid peroxidase antibodies and high thyroid-stimulating immunoglobulin levels in Graves' disease identifies a group at markedly increased risk of ophthalmopathy. Thyroid 1999 June; 9:1175-1180.

[37] Noh JY, Hamada N, Inoue Y, Abe Y, Ito K. Thyroid-stimulating antibody is related to Graves' ophthalmopathy, but thyrotropin-binding inhibitor immunoglobulin is related to hyperthyroidism in patients with Graves' disease. Thyroid 2000 June; 10:809-813.

[38] Wakelkamp IIMJ, Bakker O, Baldeschi L, Wiersinga WM, Prummel MF. TSH-R expression and cytokine profile in orbital tissue of active vs. inactive Graves ophthalmopathy patients. Clinical endocrinol 2003 January; 58:280-287.

[39] Bahn RS, Dutton CM, Natt N, Joba W, Spitzweng C, Heufelder AE. Thyrotropin Receptor Expression in Graves' Orbital Adipose/Connective Tissues: Potential Autoantigen in Graves' Ophthalmopathy. J Clin Endocrinol Metabo 1998 January; 83:998-1002.

[40] Eckstein AK, Plicht M, Lax H, Hirche H, Quadbeck B, Mann K, Steuhl KP, Esser J, Morgenthaler NG. Clinical results of anti-inflammatory therapy in Graves' ophthalmopathy and association with thyroidal autoantibodies. Clin Endocrinol (Oxf) 2004 September; 61:612-618.

[41] Lazarus JH. Epidemiology of Graves' ophthalmopathy (GO) and relationship with thyroid disease. Best Pract Res Clin Endocrinol Metab 2012 June; 26:273-279.

[42] Bartley GB. The epidemiological characteristics and clinical course of ophthalmology associated with autoimmune thyroid disease in Olmsted Country, Minnesota. Tr Am Ophtal Soc 1994 June; 92:477-588.

[43] Abraham-Nordling M, Torring O, Lantz M, Hallengren B, Ohrling Lundell G, H Calissendorff, Jørneskog G, Wallin G. Incidence of hyperthyroidism in Stockholm, Sweden, 2003-2005. Eur J Endocrinol 2008 February; 158:823-827.

[44] Werner SC. Modification of the classification of the eye changes in Graves' disease. Am J Ophthalmol 1977 June; 83:725-727.

[45] Dickinson AJ, Perros P. Controversis in the clinical evaluation of active thyroid-associated orbitopathy: use of a detailed protocol with comparative photographs for objective assessment. Clin Endocrinol 2001 June; 55:283-303.

[46] Barrio-Barrio J, Sabater AL, Bonet-Farriol E, Velazquez-Villoria A, Galofre JC. Graves' Ophthalmopathy: VISA versus EUGOGO classification, assessment, and management. J Ophthalmol 2015 July; 25:1-16.

[47] Bartalena L, Baldeschi L, Boboridis K, Eckstein A, Kahaly GJ, Marcocci C, Perros P, Salvi M, Wiersinga WM on behalf of the Eropean group on Graves' orbitopathy (EUGOGO). The 2016 European thyroid association/European group on Graves' orbitopathy guidelines for the management of Graves' orbitopathy. Eur Thyroid J 2016 March; 5:9-26.

[48] European Group of Graves' orbitopathy. EUGOGO Atlas. EUGOGO.

[49] Dickinson AJ. Clinical manifestations. A multidisciplinary approach-questions and answers. Second revised edition ed. Basel, Karger: 2010. p. 1-25.

[50] Mourits MP, Koornneef L, Wiersinga WM, Prummel MF, Berghout A, Van der Gaag R. Clinical criteria for the assessment of disease activity in Graves' ophthalmopathy: a novel approach. Br J Ophthalmol 1989 January; 73:639-644.

[51] Terwee CB, Gerding MN, Dekker FW, Prummel MF, Wiersinga WM. Development of a disease-specific quality of life questionnaire for patients with Graves' ophthalmopathy. Br J Ophthalmol 1998 January; 82:773-779.

[52] Terwee CB, Dekker FW, Prummel MF, Wiersinga WM. Graves' ophthalmopathy through the eyes of the patient: a state of the art on health-related quality of life assessment. Orbit 2001 June; 20:281-290.

[53] Hagg E, Asplund K. Is endocrine ophthalmopathy related to smoking? BMJ 1987 September; 295:634-635.

[54] Bartalena L, Martino E, Marcocci C, Bogazzi F, Panicucci M, Velluzzi F, Loviselli A, Pinchera A. More on smoking habits and Graves' ophthalmopathy. J Endocrinol Invest 1989 June; 12:733-737.

[55] Shine B, Fells P, Edwards OM, Weetman AP. Association between Graves' ophthalmopathy and smoking. Lancet 1990 May; 335:1261-1263.

[56] Prummel MF, Wiersinga WM. Smoking and risk of Graves' disease. JAMA 1993 January; 269:479-482.

[57] Tellez M, Cooper J, Edmonds C. Graves' ophthalmopathy in relation to cigarette smoking and ethnic origin. Clin Endocrinol (Oxf) 1992 January; 36:291-294.

[58] Wisna B, Mandahl A, Karlsson FA. Graves' disease, endocrine ophthalmopathy and smoking. Acta Endocrinol 1993 January; 128:156-160.

[59] Thornton J, Kelly SP, Harrison RA, Edwards R. Cigarette smoking and thyroid eye disease: a systematic review. Eye 2007 January; 21:1135-1145.

[60] Pfeilschifter J, Ziegler R. Smoking and endocrine ophthalmopathy: Impact of smoking severity and current vs lifetime cigarette consumption. Clin Endocrinol (Oxf) 1996 July; 45:477-481.

[61] Cawood TJ, Moriarty P, O`Farrelly C, O'Shea D. Smoking and thyroid-associated ophthalmopathy: a novel explanation of the biological linl. J Clin Endocrinol Metab 2006 October; 92:59-64.

[62] Altinors DD, Akca S, Akova YA, Bilzikci B, Goto E, Dogru M, Tsubota K. Smoking associated with damage to the lipid layer of the ocular surface. Am J Ophthalmol 2005 December; 141:2016-2021.

[63] Bartalena L, Marcocci C, Tanda ML, Manetti L, Dell Unto E, Bartolomei MP, Nardi M, Martino E, Pinchera A. Cigarette smoking and treatment outcomes in Graves' ophthalmopathy. Ann Intern Med 1998 October; 129:632-635.

[64] Eckstein A, Quadbeck G, Mueller G, Rettenmeier AW, Hoermann R, Mann K, Steuhl P, Esser J. Impact of smoking on response to treatment of thyroid associated ophthalmopathy. Br J Ophthalmol 2003 January; 87:773-776.

[65] Tallstedt L, Lundell G, Tørring O, Wallin G, Ljunggren J, Blomgren H, Taube A and the thyroid group. Occurrence of ophthalmopathy after treatment for Graves' hyperthyroidism. N Engl J Med 1992 June; 326:1733-1738.

[66] Bartalena L, Marcocci C, Bogazzi F, Manetti L, Tanda ML, Dell Unto E, Bruno-Bossi G, Nardi M, Bartolomei MP, Lepri A, Rossi G, Martino E, Pinchera A. Relation between therapy for hyperthyroidism and the course of Graves' ophthalmopathy. N Engl J Med 1998 January; 338:73-78.

[67] Marcocci C, Bartalena L, Tanda ML, Manetti L, Dell Unto E, Rocchi R, Barbesino G, Mazzi B, Bartolomei MP, Lepri P, Cartei F, Nardi M, Pinchera A. Comparison of the effectiveness and tolerability of intravenous or oral glucocorticoids associated with orbital radiotherapy in the management of Graves' ophthalmopathy: results of a prospective, single-blind, randomized study. J Clin Endocrinol Metab 2001 August; 86:3562-3567.

[68] Kahaly GJ, Pitz S, Hommel G, Dittmar M. Randomized, single-blind trial of intravenous versus oral steroid monotherapy in Graves' orbitopathy. J Clin Endocrinol Metab 2005 July; 90:5234-5240.

[69] Le Moli R, Baldeschi L, Saeed P, Regensburg N, Mourits MP, Wiersinga WM. Determinants of liver damage associated with intravenous methylprednisolone pulse therapy in Graves' ophthalmopathy. Thyroid 2007 April; 17:357-362.

[70] Marino M, Morabito E, Brunetto MR, Bartalena L, Pinchera A, Marocci C. Acute and severe liver damage associated with intravenous glucocorticoid pulse therapy in patients with Graves' ophthalmopathy. Thyroid 2004 May; 14:403-406.

[71] Smith TJ, Kahaly GJ, Ezra DG, Fleming JC, Dailey RA, Tang RA, Harris GJ, Antonelli A, Salvi M, Goldberg RA, Gigantelli JM, Couch SM, Shriver EM, Hayek BR, Hink EM, Woodward RM, Gabriel K, Magni G, Douglas RS. Teprotumumab for Thyroid-Associated Ophthalmopathy. N Engl J Med 2017 May; 376:1748-1761.

[72] Wiersinga WM. The philosophy of Graves' ophthalmopathy. Orbit 2005 June; 24:165-171.

[73] Versura P, Campos EC. Update on human tear proteome. European ophthalmic review 2013 June; 7:36-41.

[74] Pandey A, Mann M. Proteomics to study genes and genomes. Nature 2000 June; 405:837-846.

[75] Aebersold R, Mann M. Mass spectrometry-based proteomics. NATURE 2003 March; 422:198-207.

[76] Gillet LC, Leitner A, Aebersold R. Mass spectrometry applied to bottom-up proteomics:Entering the high throughput era for hypothesis testing. Annu. Rev. Anal. Chem.. 2016 March; 9:449–72.

[77] Abdallah C, Dumas-Gaudot E, Renaut J, Sergeant K. Gel-based and Gel-free quantitative proteomics approaches at a glance. International J Plant Genomics 2012 October; 2012:1-17.

[78] Beranova-Giorgianni S. Proteome analysis by twodimensional gel electrophoresis and mass spectrometry: strengths and limitations. Trends in Analytical Chemistry 2003 May; 22:273-281.

[79] Issaq HJ, Conrads TP, Janini GM, Veenstra TD. Methods for fractionation, separation and profiling of proteins and peptides. Electrophoresis 2002 September; 23:3048-3061.

[80] Karpievitch YV, Polpitiya AD, Anderson GA, Smith RD, Dabney AR. Liquid chromatography mass spectrometry-based proteomics: biological and technological aspects. Ann Appl Stat 2011 May; 4: 1797–1823.

[81] Duncan MW, Aebersold R, Caprioli RM. The pros and cons of peptide-centric proteomics. Nature biotechnology 2010 July; 28:659-664.

[82] Mostovenko E, Hassan C, Rattke J, Deelder AM, Van Veelen PA, Palmblad M. Comparison of peptide and protein fractionation methods in proteomics. EuPa Open Proteomics I 2013 September; 1:13-37.

[83] Gilar M, Daly AE, Kele M, Neue UD, Gebler JC. Implications of column peak capacity on the separation of complex peptide mixtures in single-and-two-dimensional high-performance liquid chromatography. J chromatogr A 2004 October; 1061:183-192.

[84] Issaq HJ, Chan KC, Janini GM, Conrads TP, Veenstra TD. Multidimensional separation of peptides for effective proteomic analysis. J of Chrom B 2005 January; 817:35-47.

[85] Ishihama Y. Proteomic LC-MS systems using nanoscale liquid chromatography with tandem mass spectrometry. J Chromatogr A 2005 January; 1067:73-83.

[86] Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG. The Orbitrap: a new mass spectrometer. J Mass Spectrom 2005 March; 40:430-443.

[87] Wasinger VC, Zeng M, Yau Y. Current status and advances in quantitative proteomic mass spectrometry. International J Prot 2013 January; 2013:1-12.

[88] Kito K, Ito T. Mass spectrometry-based approaches toward absolute quantitative proteomics. Current Genomics 2008 April; 9:263-274.

[89] Callister SJ, Barry RC, Adkins JN, Johnson ET, Qian WJ, Webb-Robertson BJM, Smith RD, Lipton MS. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. J Prot Res 2006 January; 5:277-286.

[90] Nogueira FCS, Domont GB. Survey of Shotgun Proteomics. Shotgun proteomics: Methods and protocols. Ed Martins-de-Souza M: 2014. p. 3-23.

[91] Heck AJR, Krijgsveld. Mass spectrometry-based quantitative proteomics. Expert REV Proteomics 1 2004 March; 3:317-326.

[92] Boersema PJ, Aye TT, van Veen TAB, Heck AJR, Mohammed S. Triplex protein quantification based on stable isotope labeling by peptide demethylation applied to cell and tissue lysates. Proteomics 2008 August; 8:4624-4632.

[93] Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJR. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nature protocols 2009 March; 4:484-494.

[94] Kovanich D, Cappadona S, Raijmakers R, Mohammed S, Scholten A, Heck AJR. Applications of stable isotope dimethyl labeling in quantitative proteomics. Anal Bioanal Chem 2012 May; 404:991-1009.

[95] Raijmakers R, Berkers CR, de Jong A, Ovaa H, Heck AJR, Mohammed S. Automated online sequential isotope labeling for protein quantification applied to proteasome tissue-specific diversity. Molecular and cellular proteomics 2008 June; 7.9:1755-1762.

[96] Boutilier JM, Warden H, Doucette AA, Wentzell PD. Chromatographic behavior of peptides following demethylation with H_2/D_2-formaldehyde: Implications for comparative proteomics. J chromatogr B 2012 September; 908:59-66.

[97] Hsu JL, Huang SY, Chow NH, Chen S. Stable-Isotope Dimethyl Labeling for Quantitative Proteomics. Analytical Chemistry 2003 December; 75:6843-6852.

[98] Lau H, Suh HW, Golkowski M, Ong S. Comparing SILAC- and stable isotope dimethyl-labeling approaches for quantitative proteomics. J Proteome Res 2014 July; 13:4164-4174.

[99] Zhang R, Sioma CS, Thompson RA, Xiong L, Regnier FE. Controlling deuterium isotope effects in comparative proteomics. Anal Chem 2002 August; 74:3662-3669.

[100] Fenyo D, Beavis RCA. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. Anal Chem 2003 February; 75:768/774.

[101] Nesvizhskii AI, Vitek O, Aebersold R. Analysis and verification of proteomic data generated by tandem mass spectrometry. Nature Methods 2007 October; 4:787-797.

[102] Elias JE, Haas W, Faherty BK, Gygi SP. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nat Med 2005 September; 2:667-675.

[103] Sadygov RG, Cociorva D, Yates JR. Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. Nat Met 2004 November; 1:195-202.

[104] Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 1999 August; 20:3551-3567.

[105] Eng JK, McCormack AL, Yates JR. An approach to correlate tándem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 1994 June; 5:976-989.

[106] Kapp EA, Schütz F, Connolly LM , Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JN, Om enn GS, Simpson RJ. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: Sensitivity and specificity analysis. Proteomics 2005 May; 5: 3475–3490.

[107] Zhou L, Beuerman RW. Tear analysis in ocular surface diseases. Progress in retinal and eye research 2012 June; 31:527-550.

[108] Tiffany JM. The normal tear film. Dev Ophthalmol 2008 May; 41:1-20.

[109] Mishima S, Gasset A, Klyce SD Jr, Baum JL. Determination of tear volume and tear flow. Invest ophthalmol 1966 June; 5:264-276.

[110] King-Smith PE, Fink BA, Hill RM, Koelling KW, Tiffany JM. The thickness of the tear film. Curr Eye Res 2004 March; 29:357–368.

[111] Li N, Wang N, Zheng J, Liu XM, Lever OW, Erickson PM, Li L. Characterization of human tear proteome using multiple proteomic analysis techniques. J Proteome Res 2005 June; 4:2052-2061.

[112] Rolando M, Zierhut M. The ocular surface and tear film and their dysfunction in dry eye disease. Surv ophthalmol 2001 March; 45:S203-210.

[113] Ng V, Cho P. The relationship between total tear protein concentrations determined by different methods and standards. Graefes Arch clin exp ophthalmol 2000 February; 238:571-576.

[114] Kijlstra A, Jeurissen SH, Koning KM. Lactotransferrin levels in normal human tears. Br J Ophthalmol 1983 June; 67:199-202.

[115] Molloy MP, Bolis S, Herbert BR, Ou K, Tyler MI, van Dyk DD, Willcox MD, Gooley AA, Williams KL, Morris CA, Walsh BJ. Establishment of the human reflex tear two-dimensional polyacrylamide gel electrophoresis reference map: new proteins of potential diagnostic value. Electrophoresis 1997 June; 18*:*2811–2815.

[116] Reitz C, Breipohl W, Augustin A, Bours J. Analysis of tear proteins by one- and two-dimensional thin-layer iosoelectric focusing, sodium dodecyl sulfate electrophoresis and lectin blotting. Detection of a new component: cystatin C. Graefe's Arch Clin Exp Ophthalmol 1998 March; *236:*894–899.

[117] Baier G, Wollensak G, Mur E, Redl B, Stoffler G. Analysis of human tear proteins by different high-performance liquid chromatographic techniques. J. Chromatogr 1990 January; 2:319-328.

[118] Boonstra A, Kijlstra A. Separation of human tear proteins by high performance liquid chromatography. Current eye research 1984 October; 3:1461-1469.

[119] De Souza GA, de Godoy LMF, Mann M. Identification of 491 proteins in the tear fluid proteome reveals a large number of protease inhibitors. Genome biology 2006 August; 7:R72.1-R72.11.

[120] Zhou L, Zhao SZ, Koh SK, Chen L, Vaz C, Tanavde V, Li XR, Beuerman RW. In-depth analysis of the human tear proteome. Journal of proteomics 2012 May; 75:3877-3885.

[121] Li-hong, J, Rui-il, W. Analysis of Graves ophthalmopathy patients tear protein spectrum. Chinese medical journal 2013 June; 126:4493-4498.

[122] Khalil HA, De Keizer RJE, Bodelier VMQ, Kijlstra A. Secretory IgA and lysozyme in tears of patients with Graves ophthalmopathy. Documenta ophthalmologica 1989 July; 72:329-334.

[123] Baker GRC, Morton M, Rajapaska RS, Bullock M, Gullu S, Mazzi B, Ludgate M. Altered tear composition in smokers and patients with Graves ophthalmopathy. Arch ophthalmol 2006 October; 124:1451-1456.

[124] Okrojek R, Grus FH, Matheis N, Kahaly GJ. Proteomics in autoimmune thyroid eye disease. Hormone and metabolic research 2009 February; 41:465-470.

[125] Matheis N, Okrojek R, Grus FH, Kahaly GJ. Proteomics of tear fluid in thyroid-associated orbitopathy. Thyroid 2012 October; 22:1039-1045.

[126] Matheis N, Grus FH, Breitenfeld M, Knych I, Funke S, Pitz S, Ponto KA, Pfeiffer N, Kahaly GJ. Proteomics differentiate between thyroid-associated orbitopathy and dry eye syndrome. Investigative ophthalmology and visual science 2015 March; 56:2649-2656.

[127] Ujhelyi B, Gogolak P, Erdei A, Nagy V, Balazs E, Rajnavolgyi E, Berta A, Nagy EV. Graves orbitopathy results in profound changes in tear composition: A study of plasminogen activator inhibitor-1 and seven cytokines. Thyroid 2012 April; 22:407- 413.

[128] Kingsmore FS. Multiplexed protein measurement: technologies and applications of protein and antibody arrays. Nature Reviews Drug Discovery 2006 March; 1:1-11.

[129] Paulie S, Perlmann H. Enzyme-linked immunosorbent assay. eLS 2016 January; 1-5.

[130] Parker CE, Borchers CH. Mass spectrometry based biomarker discovery, verification, and validation- Quality assurance and control of protein biomarker assays. Molecular oncology 2014 March; 8:840-858.

[131] Gillette MA, Carr SA. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat methods 2013 January; 1:28-34.

[132] Picotti P, Aebersold R. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. Nature methods 2012 May; 9:555-566.

[133] Tong L, Zhou XY, Jylha A, Aapola U, Liu DN, Koh SK, Tian D, Quah J, Uusitalo H, Beuerman RW, Zhou L. Quantification of 47 human tear proteins using high resolution multiple reaction monitoring (HR-MRM) based- mass spectrometry. Journal of Proteome 2014 December; 115:36-48.

[134] Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. Plast reconstr surg 2010 August; 2:619-625.

[135] Bornhöft G, Maxion-Bergemann S, Wolf U, Kienle GS, Michalsen A, Vollmar HC, Gilbertson S, Matthiessen PF. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. BMC Medical Research Methodology 2006 December; 6:1-12.

[136] Pearce S, Kendall-Taylor. Natural history. A multidisciplinary approach-questions and answers. Second revised edition ed. Basel, Karger: 2010. p. 77-87.

[137] Planck T, Parikh H, Brorson H, Mårtensson T, Åsman P, Gropp L, Hallengren B, Lantz M. Gene expression in Graves ophthalmopathy and arm lymphedema: Similarities and differences. Thyroid 2011 June; 21:663-674.

[138] Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Statistics Downloaded 2017 May; 453-458.

[139] Posa A, Brauer L, Schicht M, Garreis F, Beileike S, Paulsen F. Schirmer strip vs. capillary tube method: non-invasive methods of obtaining proteins from tear fluid. Annals Anatomy 2013 January; 196:137-142.

[140] Stutshell RN, Feldman JJ, Farris RL, Mandel ID. The effect of collection technique on tear composition. Invest ophthalmol vis sci 1984 March; 25:374-377.

[141] Choy CKM, Cho P, Chung WY, Benzie IFF. Water-soluble antioxidants in human tears: effect of the collection method. Invest ophthalmol vis sci 2001 December; 42:3130-3134.

[142] Li K, Chen Z, Duan F, Liang J, Wu K. Quantification of tear proteins by SDS-PAGE with an internal standard protein: A new method with special reference to small volumes of tears. Graefes Arch Clin Exp Ophthalmol 2010 Feberuary; 248:853-862.

[143] Lam SM, Tong L, Duan X, Petznick A, Wenk MR, Shui G. Extensive characterization of human tear fluid collected using different techniques unravels the presence of novel lipid amphiphiles. J Lipid Research 2013 November; 55:289-298.

[144] Small D, Hevy J, Tang-Liu D. Comparison of tear sampling techniques for pharmacokinetics analysis: ofloxacin concentrations in rabbit tears after sampling with

Schirmer tear strips, capillary tubes, or surgical sponges. J Ocul Pharmacol ther 2000 May; 16:539-446.

[145] Dumortier G, Chaumeil JC (Review). Lachrymal determinations: methods and updates on biopharmaceutical and clinical applications. Ophthalmic res 2004 February; 36:183-194.

[146] Green-Church KB, Nichols KK, Kleinholz NM, Zhang L, Nichols JJ. Investigation of the human tear film proteome using multiple proteomic approaches. Mol Vis 2008 March; 14:456–470.

[147] Zhou L, Beuerman R, Chan C, Zhao S, Li X, Yang H, Tong L, Liu S, Stern ME, Tan D. Identification of tear fluid biomarkers in dry eye syndrome using iTRAQ quantitative proteomics. J Proteome Res 2009 August; 11:4889–4905.

[148] Chong RS, Jiang YZ, Boey PY, Yu SJ, Htoon HM, Aung T, Khaw PT, Wong TT. Tear cytokine profile in mediated glaucoma patients. Ophthalmology 2010 December; 117:2353–2358.

[149] Wong TT, Zhou L, Li J, Tong L, Zhao SZ, Li XR, Yu SJ, Kob SK, Beuerman RW. Proteomic profiling of inflammatory signaling molecules in tears of patients on chronic glaucoma medication Invest. Ophthalmol Vis Sci 2011 September; 52:7385–7391.

[150] VanDerMeid KR, Su SP, Krenzer KL, Ward KW, Zhang J. A method to extract cytokines and matrix metalloproteinases from Schirmer strips and analyze using Luminex. Mol Vis 2011 April; 17:1056–1063.

[151] Jumblatt MM, McKenzie RW, Jumblatt JE. MUC5AC mucin is a component of the human precorneal tear film. Invest Ophthalmol Vis Sci 1999 January; 40:43–49.

[152] Tiffany JM. Tears in health and disease. Eye 2003 February; 17:923-926.

[153] Chen EI, Cociorva D, Norris JL, Yates III JR. Optimization of mass spectrometry compatible surfactants for shotgun proteomics. J Proteome Res 2007 July; 6:2529–2538.

[154] Denisin AK, Karns K, Herr AE. Post-collection processing of Schirmer strip-collected human tear fluid impacts protein content. Analyst 2012 September; 137:5088–5096.

[155] Chan S, Choo W. Effect of extraction conditions on the yield and chemical properties of pectin from cocoa husks. Food Chem 2013 July; 141:3752–3758.

[156] Torsteindottir I, Hakansson L, Hallgren R, Gudbjørnsson B, Arvidson NG, Venge P. Serum lysozyme: a potential marker of monocyte/macrophage activity in rheumatoid arthritis. Reumatology 1999 July; 38:1249-1254.

[157] Wiersinga WM. Management of Graves` ophthalmopathy. Nature clinical practice 2007 May; 3:396-404.

8 Papers and manuscripts

III

PLOS ONE

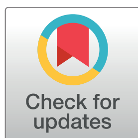
RESEARCH ARTICLE

Establishment of a tear protein biomarker panel differentiating between Graves' disease with or without orbitopathy

Cecilie Aass[1,2]*, Ingrid Norheim[3], Erik Fink Eriksen[2,3], Ellen Charlotte Børnick[4], Per Medbøe Thorsby[1], Milaim Pepaj[1]

1 Hormone Laboratory, Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway, **2** Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway, **3** Deptartment of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway, **4** University Department of Ophthalmology, Oslo University Hospital, Oslo, Norway

* cecilaas@medisin.uio.no



OPEN ACCESS

Citation: Aass C, Norheim I, Eriksen EF, Børnick EC, Thorsby PM, Pepaj M (2017) Establishment of a tear protein biomarker panel differentiating between Graves' disease with or without orbitopathy. PLoS ONE 12(4): e0175274. https://doi.org/10.1371/journal.pone.0175274

Editor: Paula Soares, Universidade do Porto Faculdade de Medicina, PORTUGAL

Received: July 8, 2016

Accepted: March 23, 2017

Published: April 18, 2017

Copyright: © 2017 Aass et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by a Grant from Helse SørØst 2013052. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Background

Graves' orbitopathy (GO) is an autoimmune inflammatory ocular complication and one of the most frequent manifestations of Graves' disease (GD). Clinical judgment of GO is subjective sometimes leading to clinical and therapeutic challenges. Better tools to diagnose this severe complication are warranted.

Patients and methods

The aim of the present study was to evaluate tear levels of LYZ, LACRT and AZGP1 in GD patients with or without GO, as possible biomarkers for GO. Tear samples were collected from GD patients with moderate-to-severe GO (n = 21) and no clinical signs of GO (n = 21). Additionally, 18 GD patients with mild GO and 9 patients without GO were included in a further part of the study.

Results

Tear levels of LYZ ($p < 0.001$), LACRT ($p = 0.004$) and AZGP1 ($p = 0.001$) were significantly elevated in GD patients with moderate-to-severe GO compared to GD patients without GO. The discriminatory power of the three biomarkers, combined in a panel was confirmed by ROC plot analysis, with an AUC value of 0.93 (sensitivity of 95%; specificity of 80%). Since LYZ showed the best performance in discriminating between GD patients with (moderate-to-severe) and without GO (in combination with limited sample volume available), LYZ levels were also measured in tears from GD patients with mild GO and without GO. Significantly higher levels of LYZ were measured in GD patients with mild GO compared to those without GO ($p = 0.003$).

Conclusions

We have established a novel three-protein biomarker panel that is able to discriminate between GD patients with and without GO, which might aid in diagnostic evaluation of GO as well as an indicator for disease activity.

Introduction

Graves' orbitopathy (GO) is an inflammatory autoimmune disorder of the orbit and one of the most frequent manifestations of Graves' disease (GD) [1, 2]. GD is a common disorder that occurs more frequently in women than in men, with a peak age between 30 and 50 years [3–5]. GO is clinically apparent in approximately 20% of patients with GD [4, 6], and is characterized by inflammatory changes and enlargement of extraocular muscles and orbital tissues. These changes, combined with the local production of cytokines and other mediators of inflammation, result in upper eyelid retraction, edema, proptosis, diplopia, erythema of the perorbital tissues and conjunctivae, and even sight loss due to compressive optic neuropathy or breakdown of cornea [7, 8]. Generally, increased levels of thyrotropin receptor autoantibodies (TRAb) are considered as an important contributor to GO [9]. The symptoms of GO are primarily due to inflammation, and the severity of GO is categorized into mild, moderate, and severe (including sight-threatening) GO [10]. Although a general consensus for assessment of GO in routine clinical practice is established, according to the 2016 recommendations of the European Group on Graves' Orbitopathy [10], the clinical judgment is subjective sometimes leading to clinical and therapeutic challenges. Therefore, from a clinical perspective, it is important to find disease-associated markers that identify GO patients.

In an earlier comprehensive quantitative proteomics study [11], we demonstrated increased levels of lysozyme C (LYZ), lacritin (LACRT) and zinc-alpha-2 glycoprotein 1 (AZGP1) in pooled tear fluid sample from GD patients with moderate-to-severe GO compared with GD patients without clinical signs of GO. The aim of the present study was therefore to assess individual levels of LYZ, LACRT and AZGP1 in tears from GD patients with moderate-to-severe GO and without GO using ELISA, and to investigate the diagnostic performance of these proteins on GO, either alone or combined in a panel.

Materials and methods

Subjects

As described in detail previously [11], tear samples from 21 patients with GD and moderate-to-severe GO (15 female, 6 male, median age 57 years, range 20–77 years; 9 smokers) were included in this study (moderate-to-severe GO, Table 1). Another group consisting of 21 patients with outburned GD and no sign of ocular symptoms during the disease (17 female, 4 male, median age 44 years, range 26–69 years; 3 smokers) served as the control group (Table 1) [11]. Patients in this group were examined by an endocrinologist (by the same endocrinologist as in the GO group), had normal TRAb values (except for two patients) and no history of GO [11].

Another study group consisting of 18 patients with GD and mild GO (15 female, 3 male, median age 51.5 years, range 28–63 years; 8 smokers) was also included in this study (mild GO, Table 2). Due to limited sample volume available, a new control group consisting of 9 patients with outburned GD and no sign of ocular symptoms during the disease (5 female, 4

Table 1. Demographic, clinical and serological data of GD patients with moderate-to-severe GO and without GO.

	Moderate-to-severe GO (n = 21)	Without GO (n = 21)	
Females, n (%)	15 (71.4)	17 (81)	
Males, n (%)	6 (28.6)	4 (19)	
Age, y (median)	57 (20–77)	44 (26–69)	
Smokers, n (%)	9 (42.8)	3 (14.3)	
Duration GD, month (median)	14 (3–108)	21 (10–35)	0.018
TRAb, IU/L (reference range, <1.8)	19.9 (3.7–156.4)	<0.9 (<0.9–4.9)	<0.001
Radioiodine treatment, n (%)	5 (23.8)	0 (0)	0.019
Thyroid surgery, n (%)	1 (4.8)	1 (4.8)	1
TSH, mIU/L (median; reference range, 0.5–3.6)	0.05 (0.01–1.7)	0.88 (0.03–4.34)	<0.001
Serum FT4, pmol/L (reference range, 8–21)	17.3 (10.1–49.2)	13.5 (11.4–21.3)	0.004
Serum FT3, pmol/L (reference range, 3.6–8.3)	7.4 (4.4–20.0)	5.4 (4.5–8.6)	0.005

https://doi.org/10.1371/journal.pone.0175274.t001

male, median age 45 years, range 27–68 years; 2 smokers) served as the control group in this further part of the study (mild GO, Table 2). Patients in this group were examined by an endocrinologist (by the same endocrinologist as in the GO groups), had normal TRAb values (except for two patients) and no history of GO.

All GO subjects went through a thorough examination by a trained ophthalmologist and an endocrinologist, and filled out a questionnaire regarding clinical history (other autoimmune,—and ocular diseases), smoking habits, symptoms and medications. However, ocular surface diseases, such as dry, which may alter the tear composition, cannot be completely excluded since GD and GO patients may have dry eyes. None of the patients with GO had started on medical treatment with steroids (which may affect the ocular disease). A clinical score (mild, moderate and severe) was given by the ophthalmologist to all the patients with GO, based on their overall severity (eyelid swelling, eyelid aperture, proptosis, eye motility, visual acuity and color vision) [12] and clinical activity, at the time of sample collection. All GO patients were in the active phase of the disease course, with a CAS score $\geq$ 3/10 (10 when severity is included). Written informed consent was obtained from all the subjects and all protocols were approved by Regional Committee for Medical and Health Research Ethics (REK south-east, 2013/839) in accordance with the ethical standards laid down in the Declaration of Helsinki [11, 13].

Table 2. Demographic, clinical and serological data of GD patients with mild GO and without GO.

	Mild GO (n = 18)	Without GO (n = 9)	
Females, n (%)	15 (83.3)	5 (55.6)	
Males, n (%)	3 (16.7)	4 (44.4)	
Age, y (median)	51.5 (28–63)	45 (27–68)	
Smokers, n (%)	8 (44.4)	2 (22.2)	
Duration GD, month (median)	8.5 (3–40)	18 (7–42)	0.047
TRAb, IU/L (reference range, <1.8)	5.4 (0.9–30.0)	<0.9 (<0.9–4.7)	0.003
Radioiodine treatment, n (%)	1 (5.5)	0 (0)	0.48
Thyroid surgery, n (%)	2 (11.1)	0 (0)	0.31
TSH, mIU/L (median; reference range, 0.5–3.6)	0.04 (0.03–1.7)	0.1 (0.03–2.33)	0.65
Serum FT4, pmol/L (reference range, 8–21)	21.3 (12.2–28.2)	18.6 (11.4–26.1)	0.37
Serum FT3, pmol/L (reference range, 3.6–8.3)	6.3 (4.6–11.7)	6.2 (4.8–8.6)	0.93

https://doi.org/10.1371/journal.pone.0175274.t002

Materials, tear collection and protein extraction

Materials, tear collection and protein extraction are described in detail previously [11] with some alterations. See details in S1 File.

ELISA measurements

One Schirmer strip from each individual was extracted as described in S1 File and the concentration of the three proteins was assessed with commercially available ELISA kits; AZGP1 (detection range; 4.7–300 ng/ml), LACRT (detection range; 1.56–100 ng/ml) and LYZ (detection range; 0.156–25 ng/ml). Additionally LYZ levels were also assessed in both mild GO group and controls. Protein concentrations obtained with ELISA assays were normalized with respect to total protein concentration and expressed as protein/total protein concentration in μg/mg.

Statistical analysis

ELISA results were exported to SPSS statistical analysis program (IBM SPSS statistics 22, Armonk, NY, USA). Data are presented as median and interquartile range (IQR). We analyzed non-normally distributed data, log-transformed, or used non-parametric methods, as appropriate. For comparison of continuous variables between groups, student's *t* tests or two-sided Mann-Whitney U tests were used. Spearman correlation coefficients rank test were used for TRAb correlations. A two-sided p-value < 0.05 was considered to be significant and uncorrected values are presented. Bonferroni-Holm correction was also performed. The diagnostic value for LYZ, LACRT and AZGP1, individually and together, was assessed by ROC-curve analyses. Multiple logistic regression analysis were performed with log-transformation of parameters when needed, to assess if there was any potential effects from independent variables such as age, smoking, gender and protein concentration on GO (dependent variable).

Results

Moderate-to-severe GO

Demographic, serological and clinical data. Twenty one GD patients with moderate-to-severe GO and 21 GD patients without signs of GO were included in the study, as shown in Table 1. The number of smokers was significantly higher in the moderate-to-severe GO group (p = 0.043) than those without GO, and the moderate-to-severe GO group were significantly older than those with no signs of GO (p = 0.024). The mean values for duration of GD, TRAb, TSH, serum T3 and serum T4 were significantly different between the groups (see Table 1 for p-values). Among the patients with moderate-to-severe GO, five had been treated with radioiodine and one had gone through thyroid surgery. On the other hand, none of the GD patients without GO were treated with radioiodine, while one had gone through thyroid surgery (Table 1).

Tear levels of LYZ, LACRT and AZGP1. We determined tear levels of LYZ, LACRT and AZGP1 with commercially available ELISA assays. The levels of LYZ ((median, 268.9 μg/mg (n = 21) vs. 84.3 μg/mg (n = 21), p $<$ 0.001)), LACRT ((median, 6.9 μg/mg (n = 20) vs. 0.9 μg/mg (n = 20), p = 0.004)) and AZGP1 ((median, 42.5 μg/mg (n = 20) vs. 22.1 μ/mg (n = 20), p = 0.001)) were significantly higher in GD patients with moderate-to-severe GO compared to GD patients without GO, as shown in Fig 1A–1C. These results validate our previous findings, that levels of these three proteins are elevated in GD patients with moderate-to-severe GO [11]. Importantly, LACRT levels were detected in only 13 of 20 tear samples from the GO group and 11 of 20 tear samples from the control group, probably due to limited sample

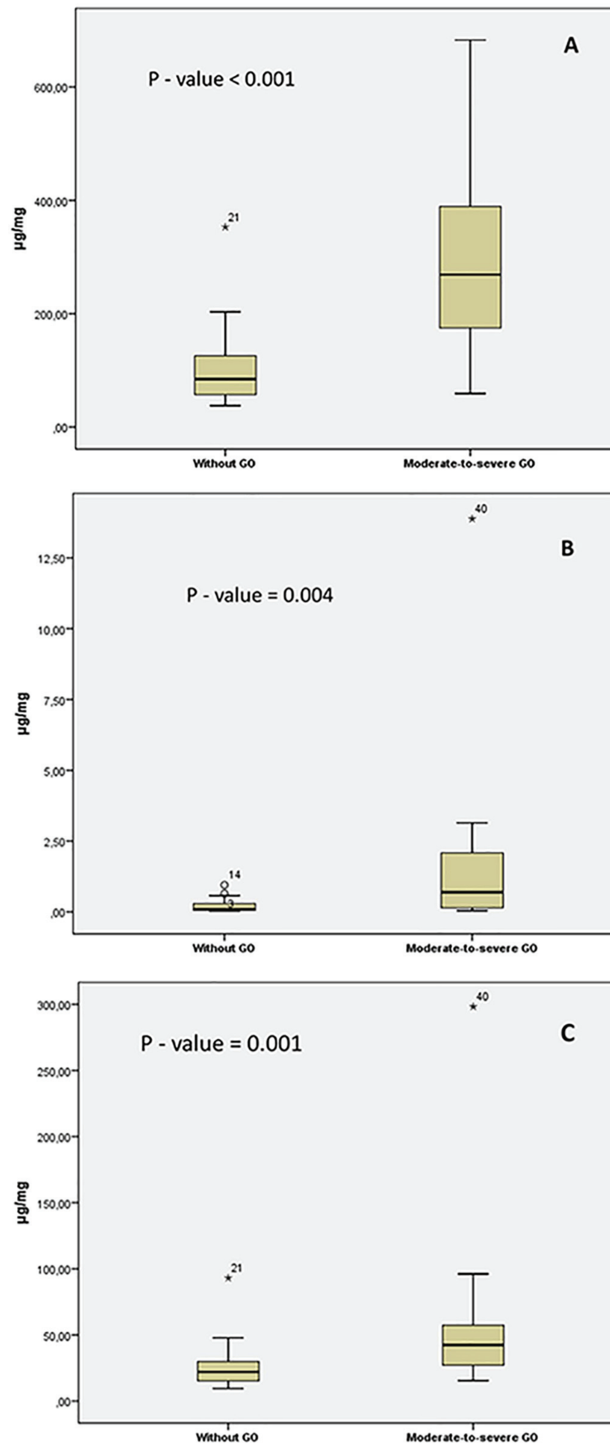


Fig 1. Tear levels of LYZ, LACRT and AZGP1 in moderate-to-severe GO patients vs GD patients without GO obtained with ELISA. Box and-whisker plot showing median and interquartile range (IQR). (A) The median for LYZ (268.9 µg/mg vs 84.3 µg/mg), (B) LACRT (6.9 µg/mg vs 0.9 µg/mg) and (C) AZGP1 (42.5 µg/mg vs 22.1 µ/mg). Mann-Whitney U-tests were performed and *p*-values are indicated on each graph.

https://doi.org/10.1371/journal.pone.0175274.g001

Table 3. Logistic regression analysis for moderate-to-severe GO patients in relation to LYZ, LACRT and AZGP1 both with and without adjustment for smoking, gender and age using log-transformed data.

	No adjusted effect (OR)	95% CI	p-value	Adjusted effect (OR)	95% CI	p-value
LYZ	8.90	2.58–30.70	0.001	10.53	2.38, 46.58	0.002
AZGP1	4.81	1.58–14.69	0.006	4.99	1.35, 18.57	0.016
LACRT	1.70	1.16–2.50	0.006	1.59	1.03, 2.44	0.035

https://doi.org/10.1371/journal.pone.0175274.t003

volume available. However, tear levels for the samples not detected with ELISA were set to the lowest obtained concentration (0.2 μg/mg) for ROC analysis.

Logistic regression analysis showed that elevated levels of LYZ ($p = 0.001$), LACRT ($p = 0.006$) and AZGP1 ($p = 0.006$) were significantly associated with GO (Table 3). Moreover, higher levels of LYZ, AZGP1 and LACRT were significantly associated with GO after adjustment for age, smoking and gender ((p = 0.002, OR = 10.53, 95% CI = (2.38, 46.58)), ((p = 0.016, OR = 4.99, 95% CI = (1.35, 18.57)) and ((p = 0.035, OR = 1.59, 95% CI = (1.03, 2.44)), respectively in a multivariable logistic regression analysis (Table 3). Increased tear levels of LYZ, AZGP1 and LACRT correlated with increased TRAb values in patients with moderate-to-severe GO (LYZ, $p < 0.001$, LACRT, $p = 0.002$, AZGP1, $p < 0.001$).

The diagnostic performance of LYZ, LACRT and AZGP1 was established by calculating area under the curves (AUC) representing receiver operator characteristic (ROC) plots. We first assessed their individual performance in discriminating between GD patients with and without GO. LYZ showed high accuracy with an AUC value of 0.91 (sensitivity of 95% and specificity of 71%), while LACRT and AZGP1 showed AUC values of 0,77 (sensitivity of 95% and specificity of 70%) and 0,80 (sensitivity of 95% and specificity of 75%), respectively. A better performance than any of the individual biomarkers was achieved when the biomarkers where combined into a panel, with an AUC value of 0.93 (sensitivity of 95%, specificity of 80%, p-value < 0.001, 95% CI = 0.83, 1.0), as shown in Fig 2.

Mild GO

Demographic, serological and clinical data. In this further study, eighteen GD patients with mild GO and 9 GD patients without signs of GO (Table 2) were included. Notably, none of these individuals overlapped with the other study group. Due to limited sample volume available, a new group served as a control group in this further part of the study. Patients in the mild GO group were not significantly older than those with no signs of GO ($p = 0.50$), and the number of former and current smokers was not significantly higher in the mild GO group ($p = 0.27$). The mean values for duration of GD and TRAb were significantly different between the groups (see Table 2 for p-values). Moreover, among the patients with mild GO, one had been treated with radioiodine and two had gone through thyroid surgery. None of the GD patients without GO were treated with radioiodine or had gone through thyroid surgery (Table 2), and none of the patients with mild GO had started treatment with steroids at the time of sample collection.

Tear levels of LYZ. Since LYZ showed the best performance in discriminating between GD patients with moderate-to-severe GO and those with no signs of GO, we next tested whether GD patients with mild GO also showed a pronounced increase in LYZ levels compared to GD patients without clinical signs of GO. ELISA analysis showed that levels of LYZ (median, 51.4 μg/mg vs. 13.9 μg/mg, $p = 0.003$) were significantly higher in the mild GO group (Fig 3). Additionally, a ROC curve was generated, showing an AUC of 0.86 (sensitivity of

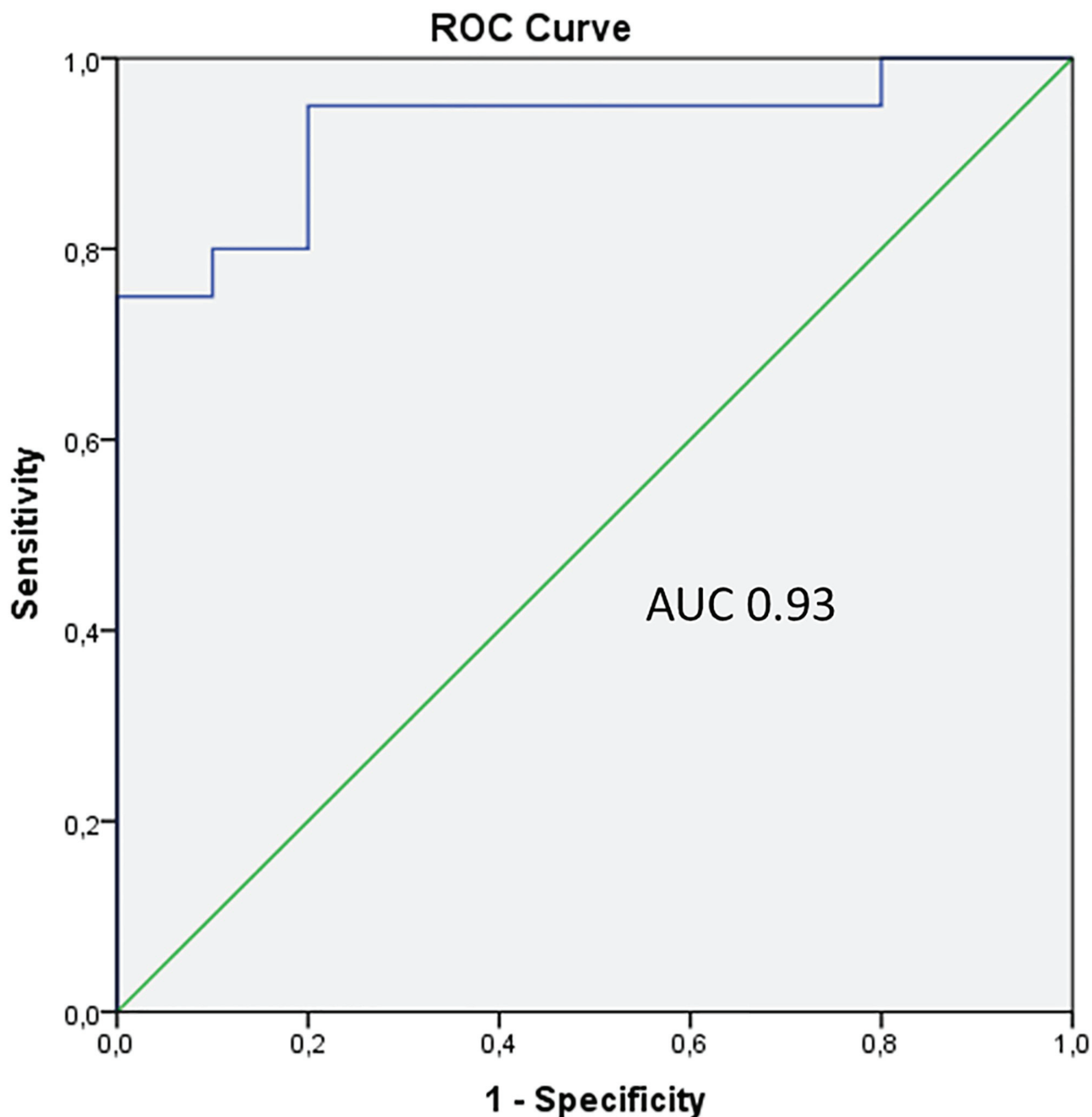


Fig 2. Diagnostic performance of the 3-protein biomarker panel in discriminating GD patients with moderate-to-severe GO from GD patients without GO. ROC curve analysis showing LYZ, LACRT and AZGP1 combined in a panel, with an AUC of 0.93. The curve is created by plotting the true positive rate (sensitivity) against the false positive rate (1 –specificity).

https://doi.org/10.1371/journal.pone.0175274.g002

100% and specificity of 78%, S1 Fig). Notably, increased tear levels of LYZ did not correlate with TRAb levels (p = 0.57) in the mild GO group.

Using logistic regression analysis, elevated levels of LYZ were significantly associated (p = 0.010) with GO, and this association was still significant after adjustment for age, smoking and gender ((p = 0.032, OR = 7.58, 95%—CI = (1.19, 48.40)).

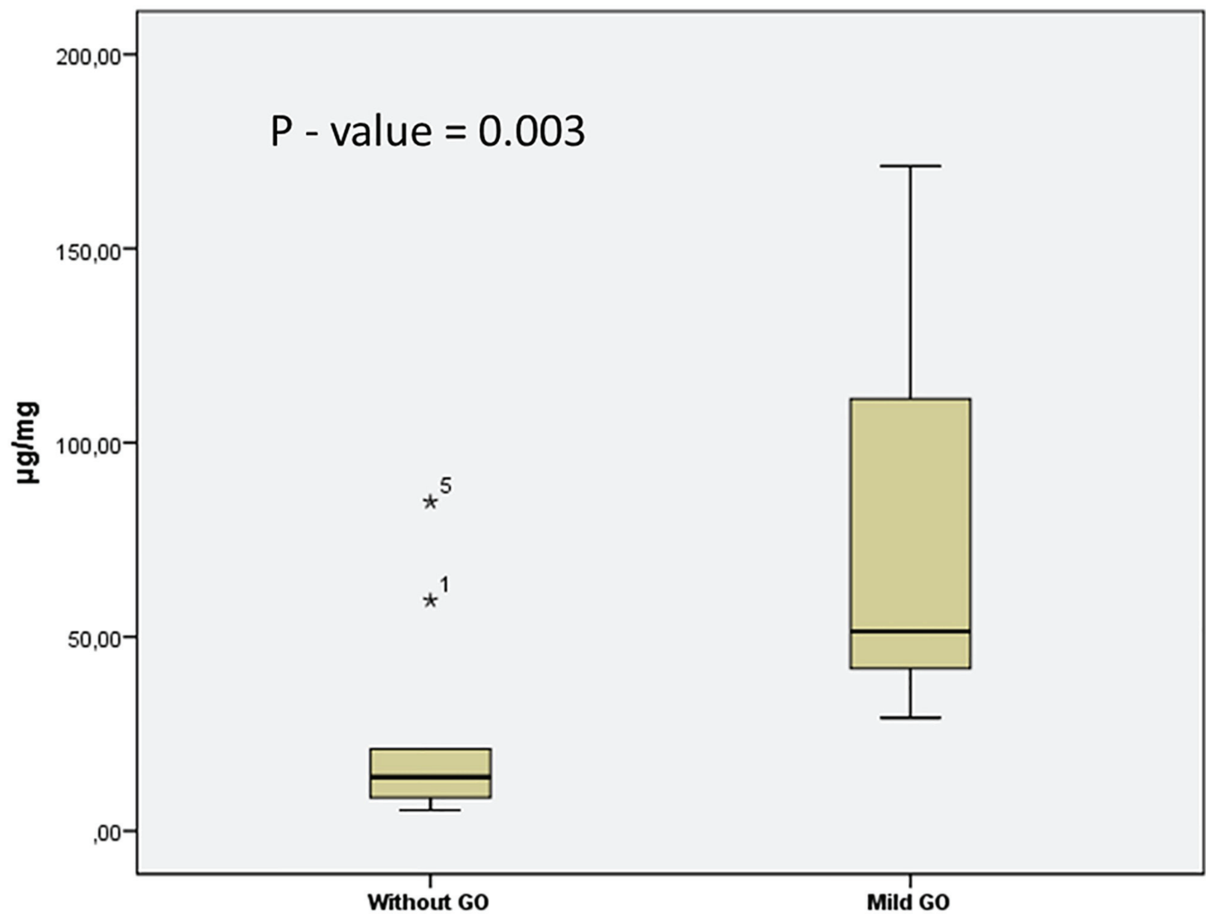


Fig 3. Tear levels of LYZ in mild GO patients vs GD patients without GO obtained with ELISA. Box and-whisker plot showing median and interquartile range (IQR). The median for LYZ (51.4 µg/mg vs. 13.9 µg/mg). Mann-Whitney U-test was performed and the *p*-value is indicated on the graph.

https://doi.org/10.1371/journal.pone.0175274.g003

Discussion

In our previous proteomic study [11], we showed that tear levels of LYZ, AZGP1 and LACRT were elevated in pooled samples of GD patients with moderate-to-severe GO compared to those without GO. The main reason why these three proteins were chosen was their high abundance in tear fluid, which is highly important when dealing with limited sample volume available, in combination with ELISA. The objective of this study was therefore to examine the protein levels of LYZ, LACRT and AZGP1 as potential protein biomarkers for GO and to validate our previous findings [11] using individual protein profiling.

Moderate-to-severe GO

Out of the 21 patients with moderate-to-severe GO, smoking were more prevalent in the GO group compared to the control group (9 vs. 3), showing that about 40% of the GO patients smoke, which has also been suggested in previous studies [14]. Smoking is a well-established risk factor for developing GO, and the odds increase progressively as the severity of eye disease increases. It is believed that GD is caused by TRAb, which activate TSHR, and enhanced

expression of the autoantigen TSHR within the orbit may play a role in the initiation or propagation of the autoimmune response in GO [15]. In this study, increased TRAb levels were observed in all patients with moderate-to-severe GO, and median TRAb levels were significantly higher in these patients in comparison to patients with no signs of GO. Although measurement of serum TRAb levels in patients with GO is recommended, assessment of activity and severity of GO are measured by the CAS and the features of eyelid swelling, eyelid aperture, proptosis, eye motility, visual acuity and color vision [12]. Thus, the clinical judgment in estimating the activity and severity of GO is highly subjective leading to major clinical and therapeutic challenges.

Tear levels of LYZ, a major lacrimal gland protein, were elevated in the GO group (Fig 1A). The present results with increased tear levels of LYZ in patients with GO corroborates with results observed in previous studies [11, 16, 17]. In logistic regression analysis, a significant association between GO and LYZ was obtained, both with and without adjustment for age, gender and smoking. Additionally, a significant correlation was observed between LYZ levels and TRAb values.

Moreover, tear levels of LACRT, a major lacrimal gland protein, were also elevated in the GO group (Fig 1B). Notably, LACRT levels were detected in only 13 of 20 tear samples from the GO group and 11 of 20 tear samples from the control group with ELISA assay, probably due to limited sample volume available. Further analyses of LACRT are necessary to support these current findings. Though well-characterized ELISA assays are suitable tools for verification analysis, sometimes sample volume needed is an issue, especially with tear fluid as only a few microliters are obtained from an individual. A highly viable alternative to ELISA is the use of targeted MS methods using multiple reaction monitoring (MRM) approaches. This methodology is highly suited for candidate verification as it is possible to quantify many proteins with high sensitivity in a single LC-MRM-MS run [18].

Similarly, AZGP1, a multidisciplinary tear protein [19], was found increased in the GO group (Fig 1C). These results are in line with previously reported up-regulation of AZGP1 in tears from patients with GO and healthy smokers [20]. In logistic regression analysis, a significant association between GO and AZGP1 was obtained, both with and without adjustment for age, gender and smoking. Since our groups are not matched for smoking, we cannot not rule out the possibility that smoking may affect the AZGP1 tear levels since smoking is a significant risk factor for the development of GO [11]. Although we did not find any correlation between AZGP1 levels and smoking in the GO group, additional studies with larger groups, matched for smoking, are necessary to determine the impact of smoking on AZGP1 tear level.

As mentioned earlier, clinically, it is very valuable for the ophthalmologist to have disease-associated markers to aid in diagnosis of GO. The results from this study have shown that it is possible to differentiate between GD patients with and without GO by measuring tear levels of LYZ, LACRT and AZGP1. As it is well-known that ocular diseases may alter the tear composition, it should be mentioned that the secretion of LYZ, LACRT and AZGP1 have been detected in tears of healthy subjects [21, 22]. Additionally, LYZ has also shown to be down-regulated in patients with dry-eye syndrome [23]. Although the diagnostic performance (ROC curve analysis) of the three proteins separately was satisfactory (reflected by high values for AUC), in a complex disease such as GO, it is difficult to discriminate patients using individual biomarkers. Hence, a combination of biomarkers rather than a single one likely performs better as a diagnostic test. When combined, LYZ, LACRT and AZGP1 form a powerful tear fluid panel able to discriminate between GD patients with moderate-to-severe GO and those without signs of GO with 93% accuracy. The diagnostic performance of our biomarker panel now needs to be further validated in a prospective longitudinal study with a larger population. Tear collection from such a cohort are underway in our lab.

Mild GO

Since LYZ showed the best performance in discriminating between GD patients with moderate-to-severe GO and those with no signs of GO, and in combination with limited sample volume available, we next examined the ability of LYZ to distinguish between GD patients with mild GO from GD patients without clinical signs of GO. The levels of LYZ were significantly higher in GD patients with mild GO compared to GD patients without GO (Fig 3), representing a possible manifestation of subclinical stage of the disease. Moreover, LYZ was significantly associated with GO, and this association was still significant after adjustment for age, smoking and gender. With an AUC of 0.86, LYZ also show potential for differentiating GD patients with mild GO from those with no signs of GO. The difference between mean age and the number of smokers was not significantly different between the patients included in the mild GO group and patients with no signs of GO. However, increased TRAb levels were observed in 16 patients in the mild GO group, and median TRAb levels were significantly higher in this group compared to patients with no signs of GO.

Today, patients with mild GO are usually not recommended for anti-inflammatory treatment with steroids, however, many patients suffer from changes in their appearance and can have troublesome diplopia [24]. Generally, a wait-and-see policy is favored, in hope of a spontaneous improvement [24]. However, some of these patients progress to more moderate to severe stage of GO disease, representing a subgroup which cannot be detected with clinical examination methods. We believe that tear fluid measurement of LYZ may help the clinician in deciding which of the patients that needs extra care concerning development of GO in the clinically mild GO group.

Concluding remarks

The aim of the present study was to assess levels of LYZ, LACRT and AZGP1 in tear fluid from GD patients with moderate-to-severe and without GO using ELISA, and to investigate the diagnostic performance of these proteins either alone or combined in a panel. We have established a novel three protein biomarker panel that is able to differentiate between GD patients with GO from those without GO. This diagnostic marker panel could enable non-invasive diagnosis and be of tremendous benefit for patients who are being evaluated for this debilitating disease. Additionally, LYZ measured in tears from GD patients with mild GO were significantly higher compared to GD patients without GO.

Since dry eye is common among patients with GD and GO, and the groups are not matched for age, gender and smoking we cannot rule out at this point, that the altered tear composition are only a result of GO. Furthermore, further evaluation of these biomarkers on a larger cohort of GD patients, and patients with other forms of hyperthyroidism, would be necessary to increase the diagnostic accuracy of the biomarker panel.

Supporting information

S1 File. Materials, tear collection, and protein extraction.
(DOCX)

S1 Fig. Diagnostic performance of LYZ in discriminating GD patients with mild GO from GD patients without GO.
(TIFF)

Acknowledgments

Thank you to the staff members, especially Tone Hagen, and doctors at the Department of Endocrinology, Oslo University Hospital for their help with recruitment of control patients and sample collection. Additionally, thank you to the ophthalmologists at Oslo University Hospital, for guidance and recruitment of active GO patients. We also thank the staff members at the Norwegian Hormone Laboratory, Oslo University Hospital, for their technical support and Manuela Zucknick for statistical guidance.

Author Contributions

Formal analysis: CA.

Investigation: MP CA.

Methodology: CA MP.

Project administration: MP.

Software: CA.

Supervision: MP PMT ECB IN EFE.

Writing – original draft: CA.

Writing – review & editing: PMT MP EFE ECB IN.

References

1. Bartalena L, Fatourechi V. Extrathyroidal manifestations of Graves' disease: a 2014 update. J Endocrinol Invest. 2014; 37: 691–700. https://doi.org/10.1007/s40618-014-0097-2 PMID: 24913238
2. Wiersinga WM, Bartalena L. Epidemiology and prevention of Graves' Ophthalmopathy. Thyroid. 2002; 12: 855–860. https://doi.org/10.1089/105072502761016476 PMID: 12487767
3. Abraham-Nordling M, Bystrom K, Torring O, Lantz M, Berg G, Calissendorff J, et al. Incidence of hyperthyroidism in Sweden. Eur J Endocrinol. 2011; 165: 899–905. https://doi.org/10.1530/EJE-11-0548 PMID: 21908653
4. Perros P, Crombie AL, Matthews JNS, Kendall-Taylor P. Age and gender influence the severity of thyroid-associated ophthalmopathy: a study of 101 patients attending a combined thyroid-eye clinic. Clin Endocrinol. 1993; 38: 367–372.
5. Abraham-Nordling M, Torring O, Lantz M, Hallengren B, Ohrling H, Lundell G. Incidence of hyperthyroidism in Stockholm, Sweden, 2003–2005. Eur J Endocrinol. 2008; 158: 823–827. https://doi.org/10.1530/EJE-07-0877 PMID: 18505903
6. Tanda ML, Piantanida E, Liparulo L, Veronesi G, Lai A, Sassi L, et al. Prevalence and natural history of Graves' orbitopathy in a large series of patients with newly diagnosed Graves' hyperthyroidism seen at a single center. J Clin Endocrinol Meatb. 2013; 98: 1443–1449.
7. Ginsberg J. Diagnosis and management of Graves' disease. CMAJ. 2003; 168: 575–585. PMID: 12615754
8. Bahn RS. Mechanisms of disease Graves'ophthalmopathy. N Engl J Med. 2010; 362:726–738.
9. Eckstein AK, Johnson KTM, Thanos M. Esser J, Ludgate M. Current insights into the pathogenesis of graves orbitopathy. Horm metab res.2009; 41: 456–464. PMID: 19530272
10. Bartalena L, Baldeschi L, Boboridis K, Eckstein A, Kahaly GJ, Marcocci C, et al. The 2016 European thyroid association/European group on Graves' orbitopathy guidelines for the management of Graves' orbitopathy. Eur Thyroid J. 2016; 5: 9–26. https://doi.org/10.1159/000443828 PMID: 27099835
11. Aass C, Norheim I, Eriksen EF, Børnick EC, Thorsby PM, Pepaj M. Comparative proteomic analysis of tear fluid in Graves' disease with and without orbitopathy. Clinical Endocrinology. 2016; 0: 1–8.
12. Dickinson AJ. Clinical manifestations. A multidisciplinary approach-questions and answers. In: Basel Karger,editors. 2010. pp.1–25.
13. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bulletin of the World Health Organization. 2001; 79: 373–374. PMID: 11357217

14. Pfeilschifter J, Ziegler R. Smoking and endocrine ophthalmopathy: Impact of smoking severity and current vs lifetime cigarette consumption. Clin Endocrinol (Oxf). 1996; 45: 477–481.
15. Heufelder AE, Dutton CM, Sarkar G, Donovan KA, Bahn RS. Detection of TSH receptor RNA in cultured fibroblasts from patients with Graves' ophthalmopathy and perorbital dermopathy. Thyroid. 2009; 3: 297–300.
16. Matheis N, Okrojek R, Grus FH, Kahaly GJ. Proteomics of tear fluid in thyroid-associated orbitopathy. Thyroid. 2012; 22: 1039–1045. https://doi.org/10.1089/thy.2012.0119 PMID: 22873942
17. Khalil HA, De Keizer RJE, Bodelier VMQ, Kijlstra A. Secretory IgA and lysozyme in tears of patients with Graves ophthalmopathy. Documenta ophthalmologica. 1989; 72: 329–334. PMID: 2625093
18. Tong L, Zhou XY, Jylha A, Aapola U, Liu DN, Koh SK, et al. Quantification of 47 human tear proteins using high resolution multiple reaction monitoring (HR-MRM) based- mass spectrometry. Journal of Proteome. 2014; 115: 36–48.
19. Hassan I, Waheed A, Yadav S, Singh TP, Ahmad F. Zinc α2-glycoprotein: a multidisciplinary protein. Mol cancer res. 2008; 6: 892–906. https://doi.org/10.1158/1541-7786.MCR-07-2195 PMID: 18567794
20. Baker GRC, Morton M, Rajapaska RS, Bullock M, Gullu S, Mazzi B, et al. Altered tear composition in smokers and patients with Graves ophthalmopathy. Arch ophthalmol, 2006; 124: 1451–1456. https://doi.org/10.1001/archopht.124.10.1451 PMID: 17030713
21. De Souza GA, de Godoy LMF, Mann M. Identification of 491 proteins in the tear fluid proteome reveals a large number of protease inhibitors. Genome biology. 2006; 7: R72.1–R72.11.
22. Zhou L, Zhao SZ, Koh SK, Chen L, Vaz C, Tanavde V, et al. In-depth analysis of the human tear proteome. Journal of proteomics. 2012; 75: 3877–3885. https://doi.org/10.1016/j.jprot.2012.04.053 PMID: 22634083
23. Zhou L, Beuerman RW, Chan CM, Zhao SZ, Li XR, Yang H, et al. Identification of Tear Fluid Biomarkers in Dry Eye Syndrome Using iTRAQ Quantitative Proteomics. J Proteome Res. 2009; 8: 4889–4905. https://doi.org/10.1021/pr900686s PMID: 19705875
24. Wiersinga WM. Management of Graves' ophthalmopathy. Nature clinical practice. 2007; 3: 396–404. https://doi.org/10.1038/ncpendmet0497 PMID: 17452966

Errata

Name of candidate: Cecilie Aass Aasaaren

Dissertation title: Quantitative proteomic analyses of tear fluid in Graves' orbitopathy

Abbreviations for type corrections:

Cor – correction

Celtf – change of text format

Space – space removed

Ins – inserting missing words

Side	Linje	Orginal tekst	Type rettelse	Korrigert tekst
6	11	….LC- MS/MS…	Space	….LC-MS/MS…
13	2		Ins	**Paper I**
13	10	…Establishment of a tear fluid protein biomarker…	Cor	…Establishment of a tear protein biomarker…
16	6	….<1.8 IU/L…	Cor	…>1.8IU/L…
16	8	…thyroid-associated …	Space	..thyroid-associated..
22	20	…The risk is greatest if RAI is administrated during active GO, to patients who smoke..	Cor	…The risk is greatest if RAI is administrated during active GO, or to patients who smoke..
24	2	…large- scale…	Space	…large-scale…
24	19	...sodium- dodecyl…	Space	…sodium-dodecyl…
24	20	…2D- GE…	Space	…2D-GE…
26	11	…pre-fractionation…	Space	…pre-fractionation …
26	16	…. liquid- phase…	Space	… liquid-phase…
30	24	…(pH- range= 5-8.5)..	Space	....(pH-range= 5-8.5)..
33	6	…cross-correlation…	Space	…cross-correlation…
36	14	..1D- gel…	Space	…1D-gel…
38	12	…MS- based….	Space	…MS-based…
41	6	…cross- sectional…	Space	…cross-sectional..
43	10	…The inclusion criteria for all the patients with GD…	Cor	…The inclusion criteria for all the patients with GO…
46	7	…..acid- labile…	Space	…acid-labile…

50	2	….paper II…	Celtf	…Paper II…
51	8	….supplementary 1.4…	Celtf	…Supplementary 1.4…
51	14	**3.2.10 Nano LC- MS/ MS**	Space	**3.2.10 Nano LC-MS/MS**
54	6	…one- way…	Space	…one-way…
54	20	….log- transformation…	Space	…log- transformation..
55	3	…two- sided…	Space	…two-sided…
57	9	…sigle- unit…	Space	…single-unit…
68	9	…non- optimal…	Space	…non-optimal..
68	24	…as well as which proteins which are extracted.	Cor	…as well as which proteins are extracted.
69	7	..time- independent…	Space	…time- independent..
75	10	A) The for LYZ…	Cor	A) LYZ….
89	25	[111]…Journal of proteome research…	Cor	[111]…J Proteome Res…
92	5	[135]…*BMC Medical Research Methodology*…	Celtf	[135]…BMC Medical Research Methodology…