Proteomic Variability of the Eye Lens and its Effect on Cataract Onset and Development

vorgelegt von
Diplom-Biologe
Wolfgang Hoehenwarter
aus Berlin

von der Fakultät III Prozesswissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. Nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. R. Lauster
Berichter: Prof. Dipl.-Ing. Dr. U. Stahl
Berichter: Dr. P. Jungblut

Tag der wissenschaftlichen Aussprache: 31. Oktober 2006

Berlin 2007
D 83

Acknowledgments

I would like to thank Prof. Dr. sc. techn. L.-G. Fleischer and Prof. Dipl.-Ing. Dr. Ulf Stahl of the TU Berlin for accepting the proposition for attainment of a doctorate and the doctoral thesis.

I would like to thank Prof. Dr. Thomas F. Meyer, Prof. Dr. Karl Sperling and Prof. Dr. Peter M. Kloetzel for allowing me to perform the described work in their respective departments and institutes, the Department of Molecular Biology of the Max Planck Institute for Infection Biology, the Institute of Human Genetics of the Charite´ and the Institute of Biochemistry of the Charite´.

I would like to thank Dr. Peter R. Jungblut and Prof. Dr. Joachim Klose for initiating the work and for supervising the work in progress.

My special thanks go to Dr. Peter R. Jungblut for discussion and support, and for reading and correcting the manuscript.

I would also like to thank Prof. Dr. Nalin M. Kumar for supplying the lenses of the *Mus musculus* wild type and transgenic animals.

Thanks are also due to all of my colleagues and co-workers of the Core Facilities at the Max Planck Institute for Infection Biology, the Institute of Human Genetics of the Charite´ and the Institute of Biochemistry of the Charite´, especially Dr. Sevil Aksu, Renate Ackermann, Dr. Katharina Janek, Dr. Alexander Krah, Dr. Klaus-Peter Pleissner, Monika Schmid, Dr. Frank Schmidt, Dr. Robert Stein and Dr. Ursula Zimny-Arndt.

Thanks also to Robert Ahrends, Karina Barthel, Prof. Dr. Burkardt Dahlmann, Christian Falke, Anja Freiwald, Hajar Habibi, Christina Hempel, Dr. Robert Hurwitz, Alexander Kloß, Sabine Lange, Dr. Lei Mao, Dr. Jens Mattow, Dr. Limei Meng Okkels, Dr. Ljiljana Otasevic, Franziska Schiele, Frank Siejak, Kathrin Textoris-Taube, Dr. Bernd Thiede,

Dr. Maik Wacker, Ina Wagner, Dr. Wolfgang Wehrl, the team from Agilent Technologies and everyone else I worked with in the past five years and omitted because of forgetfulness.

Summary

In an attempt to determine the variability in the proteins of the eye lens of the mouse, a discovery-type approach with proteomics methods was used. The lens is composed almost entirely of proteins in various stages of solution. Around 90% of lens proteins belong to the crystallin protein superfamily, which is loosely classified into Alpha, Beta and Gamma crystallins. There are sixteen primary crystallin gene translation products in the eye lens of the mouse, so the unmodified proteome is easily managed.

The variability in two species, *Mus musculus* and *Mus spretus*, was examined using liquid chromatography (LC) and mass spectrometry (MS). Two types of isotope coded protein labeling reagents, the cleavable ICAT reagent and the iTRAQ reagent were used for the detection of variability and polymorphisms. The iTRAQ label was used in conjunction with nano-LC, automated MALDI plate spotting and MALDI/TOF/TOF-MS. The iTRAQ methodology proved superior to the ICAT based approach employing nano-LC and on-line ESI-MS. The investigation with the iTRAQ reagent identified and quantified 11 of the 16 crystallins ubiquitous to mammals and many of the major non-crystallin proteins in the lens fiber cells. Four polymorphisms in the crystallins (Alpha B, Gamma B, Gamma F and Gamma S crystallin) were also detected in the two mouse species, using the iTRAQ methodology.

The variability in two strains, C57BL/6J and 129/SvJ of the *Mus musculus* species, and its effects on the onset and severity of cataract of the lens nucleus (because of loss of the Alpha 3 connexin protein due to knockout of the Alpha 3 connexin gene *Gja3)* was analyzed using two-dimensional electrophoresis (2-DE), 2-DE image analysis software and MALDI-MS. This produced the largest and most comprehensive investigation of mouse lens proteins to date, comprising more than 100 identified proteins each in the two strains and in the mutants of both strains. The variability in the lens proteins was completely established. Seven factors (CP49, HSP27/25, Chaperonin subunit 6A, DnaK-type molecular chaperone hsc70 homolog, Endoplasmatic reticulum protein 29, Syntaxin binding protein and Annexin A1) which specifically determine the onset and severity of nuclear cataract in the two strains are identified.

The investigation shows that the healthy and pathological development in the lens

- cannot be attributed to certain proteins alone, but that
- similar but distinct networks of interactions, specific to the two mouse strains, are responsible for these phenomena in each of them.
- The view has to be extended to the protein species level.

A basis for the detailed understanding of the processes in healthy lens and cataract development is established.

Table of Contents

1. Introduction

Proteomics became a scientific term in 1995 when Wasinger coined the term proteome and defined it as a samples protein complement (Wasinger, 1995). Genes and the genome can be seen as a blueprint underlying and to a significant extent defining the life of an organism. It can be seen as an analogy that proteins put the genetic design into practice making them the major class of physiological effectors. It is proximate that the comprehensive analysis of proteins and their interactions is necessary to understand living organisms.

The lens of the eye is spherical and consists of many elongated fiber cells aligned with the general direction of light entry and a single layer of metabolically active cuboidal epithelial cells at the organs anterior perimeter. The oldest cells are located in the lens core and are the primary fiber cells of the lens nucleus. The younger, more peripheral secondary fiber cells of the lens cortex are derived from the epithelial cells located at the lens equator which undergo terminal differentiation into fiber cells and are subsequently layered as the lens grows. The fiber cells are devoid of most organelles and cellular structures and are filled with a mass of proteins in solution, the crystallins (Mörner, 1894). This peculiar arrangement achieves lens transparency (Delaye and Tardieu, 1983).

1.1. An Introduction to Proteomics

Proteomics as a science arbitrarily began in 1975 when O´Farrell developed the two-dimensional gel electrophoresis (2-DE) technique combining isoelectric focusing (IEF) and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for

protein separation (O´Farrell, 1975). This has become the hallmark proteomics technique. The polypeptides migrate to positions termed protein spots on the 2-DE gel as dictated by the relation of their chemical parameters to the applied 2-DE parameters.

High quality 2-DE has a remarkable resolution capacity, being able to separate up to 10000 sample constituents distinguished by a single amino acid or post translational modification (PTM) (Klose and Kobalz, 1995). This is also offered by other techniques but the visualization of the separated polypeptides in a 2-DE protein spot pattern by staining them with a dye is a premier advantage of 2-DE. Two very popular dyes are silver staining (Jungblut and Seifert, 1990) and Coomassie Brilliant Blue (CBB) G250 staining (Doherty et al., 1998). The 2-DE gels can be scanned and their protein spot patterns introduced into 2-DE image analysis softwares. These softwares facilitate protein spot matching and quantification of protein abundance from staining intensities. Usually CBB G250 stained proteins are excised from 2-DE gels because this dye is more compatible with mass spectrometry than silver staining. The proteins are then digested in the excised gel pieces with enzymes or reagents (Otto et al., 1996) and the resulting peptides analyzed by mass spectrometry. Alternatively, the proteins can be extracted from the gel pieces by electroelution (Ueda et al., 2002) or passive diffusion (Castellanos-Serra et al., 1996, Ueda et al., 2002) and then analyzed whole. An experienced user can readily pinpoint the differences between proteomes and truly characterize and understand them in detail expediting the detection of physiologically or clinically relevant molecules such as variants or disease markers (Klose et al., 2002; Krah et al., 2004). Note that at this analytical level modified and unmodified polypeptides should be termed protein species and not protein as this describes the translation products at the covalent molecular level (Jungblut and Thiede, 1997) which is entirely unambiguous.

The other premier technique employed for protein and peptide separation prior to identification is liquid chromatography. Liquid chromatography is one of the oldest and best techniques for molecular separation in biochemical analysis earning Archer Martin and Richard Synge the Nobel Prize for Chemistry in 1952. It is based on the individual behavior of soluble analytes in a mobile phase passed over and interacting differentially with a solid phase or matrix. Various combinations of liquid and solid phases featuring distinct molecular interaction and separation properties have been developed and tried over the years, leading to several popular liquid chromatography approaches: affinity chromatography, ion exchange chromatography, reverse phase chromatography and size exclusion chromatography to name a few.

With liquid chromatography procedures proteins or peptides are often eluted on-line or semi on-line into a mass spectrometer (Yates 3rd, 1998). The primary advantage offered by this technique is the relative speed at which complex protein mixtures can be assayed. However, often higher degrees of separation and consequently more than two chromatographic steps are required. This became known as multidimensional protein identification technology or MudPIT (Washburn et al., 2001). High end multidimensional chromatography often includes up to 15 or more chromatography steps for polypeptide separation (Link et al., 1999; Washburn et al., 2001; Wolters et al., 2001). Experiments of this type with peptides generate large amounts of data making comprehensive evaluation equally laborious as the complete analysis of the protein spots in 2-DE protein spot patterns (Hoehenwarter et al., 2006a; Swanson and Washburn., 2005). Also, modifications can not be clearly assigned to protein species making their characterization difficult.

Following separation proteins or peptides have to be identified. Originally, purified or separated proteins were mostly analyzed by Edman degradation, a procedure which realizes protein primary structure beginning at a polypeptides N-terminus (Edman, 1949).

Introduction

Primary structure information is the most basic and meaningful form of polypeptide characterization and presents the least degree of ambiguities.

A major breakthrough was reached in the late nineteen eighties, when mass spectrometry became available for polypeptide analysis. Originally developed late in the last century, mass spectrometry is a technique that measures an ionized molecules mass and charge. A mass spectrometer is composed of an ion source, an analyzer and an ion detector. Ionized molecules from the ion source are manipulated by electric fields where their recorded behavior in the analyzer prior to detection by the ion detector allows calculation of their physical parameters. So called soft ionization techniques convert polypeptides to gas phase ions without damaging the molecules. The principal soft ionization techniques employed in proteomics are matrix assisted laser desorption ionization (MALDI) (Karas and Hillenkamp, 1988; Tanaka et al., 1988) and electrospray ionization (ESI) (Fenn et al., 1989).

MALDI utilizes solid phase ionization effected by short laser pulses. Analytes are embedded in a crystal matrix which protects them from direct exposure to the laser. The matrix is broken apart by the laser and absorbs its energy. It is partially transferred onto the embedded analytes which in turn are ionized and desorbed into the gas phase through the broken matrix by an electric field.

ESI utilizes Coulomb fission or ion evaporation. Analytes in a liquid phase sprayed from a narrow bore needle tip are subjected to high temperature in the presence of a strong electric field. The resulting liquid phase evaporation and increased charge density results in droplets dividing into individual charged ions and/or the charge repulsion overcoming the liquid surface tension leading to the release of the charged ions from the liquid into the gas phase.

Some popular analytical components in a mass spectrometer are time-of-flight (TOF) (Olthoff et al., 1988), ion trap (Louris et al., 1987; Jonscher and Yates 3rd, 1997)

and Fourrier transform ion cyclotron resonance (FTICR) (Comisarow and Marshall, 1974) instruments. In TOF instruments the ionized molecules are accelerated by a strong electric field and propelled down a flight path in vacuum conditions. An ions flight duration is measured upon reaching the mass spectrometers detector at the end of the flight path allowing the calculation of the molecules mass and charge. Modern TOF instruments feature a mass accuracy of better than 30 ppm.

An ion trap is composed of three hyperbolic electrodes, a ring and two endcaps. The ions enter the trap sequentially through an electrostatic ion gate and are trapped by an oscillating potential applied to the ring electrode. They develop a stable secular frequency according to their mass and charge and subject to the size of the trap and the frequency and amplitude of the oscillating potential. By ramping the frequency of a supplementary potential applied to the endcap electrodes and matching (resonance conditions) and exceeding the ions stable secular frequency, the ions develop an unstable trajectory and are differentially ejected from the trap to a detector through holes in the endcap. Ion trap analyzers have a somewhat poorer mass accuracy than modern TOF instruments of around 0.4 Da.

In FTICR instruments ions from an external ion trap are ejected into the ICR ion trap centered in a superconducting magnet under vacuum conditions, where they are pulsed into a high radius cyclotron orbit. The ion cyclotron resonances refer to the number of orbits per second and are recorded as a function of the oscillating charge induced in a detector plate by the orbiting ions as they approach and recede from the plate. Fourrier transformation then yields the cyclotron frequencies and subsequently the mass and charge of all analyzed ions. FTICR instruments feature an unprecedented mass accuracy of less than 1 ppm.

Today's generation of mass spectrometers deliver not only molecular masses but also compositional and conformational information by fragmenting the analyzed ions and

applying a second round of mass spectrometry to the fragment ions. This is termed tandem mass spectrometry or MS/MS (Senn et al., 1966; Biemann et al., 1966).

TOF/TOF mass spectrometers feature two flight paths separated by a collision cell (Medzihradszky et al., 2000). Ions formed with a high internal energy dissociate after desorption. This is called metastable fragmentation. Ions with a lower internal energy can be fragmented by collision with a neutral gas introduced into the collision cell. This is termed collision induced dissociation (CID). Fragment ion flight durations are recorded in the second flight path of the TOF/TOF instrument and their mass and charge calculated accordingly.

Ion traps are routinely filled with helium to dampen the ions kinetic energies and to contract their trajectories towards the center of the trap. Application of a low amplitude supplementary potential to the endcap electrodes increases the ions kinetic energy leading to destabilization of their trajectory insufficient for excitation however conductive to frequent collision with the helium ions in the trap. Accordingly the ions are fragmented and the fragments can then be further excited and ejected from the trap to the detector.

FTICR mass spectrometers utilize electron capture dissociation (ECD) for ion fragmentation. Recent evidence attests that ECD is a nonergodic high energy process resulting in unimolecular dissociation under conformational control. Excited ions thus dissociate at defined sites before the excitation energy is randomized (Breuker et al., 2004). The fragments ion cyclotron resonances are then recorded and Fourrier transformed yielding their mass and charge.

Contrary to chemical sequencing, the data acquired by mass spectrometry must be converted into meaningful information leading to the identification of the protein species. A polypeptides mass and charge allow the calculation of its primary structure, however considering twenty proteinogenic amino acids and modifications the possibilities for large peptides and proteins are considerable. High mass accuracy is essential to

minimize ambiguities. Modern FTICR mass spectrometers featuring a mass accuracy of less than 1 ppm are capable of unambiguously defining the amino acid content of analyzed polypeptides but cannot shed light on the amino acid sequence.

In theory, the sequenced genomes of organisms contain the primary structures of all of an organisms primary translation products. This is an enormous asset for protein identification. Polypeptide masses recorded in mass spectra are used to search databases of conceptually translated genomics data according to primary structure segment masses, in many cases resulting in matches of the mass spectrometric data to amino acid sequences producing peptide sequence suggestions and conclusive protein identification. This is termed Peptide Mass Fingerprinting or Peptide Mass Mapping (PMF or PMM) (Thiede et al., 2005). Modifications can be assessed from mass shifts using the identification as a reference. A number of software suites are available commercially or free of charge on the world wide web that expedite this process. Nevertheless MS/MS spectra are invaluable because they contain amino acid sequence information and if available should be utilized for comprehensive results.

An MS/MS spectrum is composed of the signals of a peptides fragment ions resulting from covalent bond dissociation. Metastable dissociation of polypeptides in mass spectrometers occurs mainly at the peptide bond yielding peptide fragment ions termed b and y fragments. If a peptide ions charge is retained at the N-terminus following dissociation the resulting fragment ion is termed b ion while if the charge is retained at the C-terminus the fragment ion is termed y ion. Mass spectrometry can only detect ions so the fragments that do not retain a charge after dissociation are not detected and termed neutral loss. The peptide b and y ion series are the most important MS/MS ion series in proteomics mass spectrometry as polypeptide amino acid sequences can be reconstructed with them. They allow the additive calculation of mass differences between peaks of fragment ion signals and the correlation of the mass differences to amino acid

residues. A combination of mass, charge and MS/MS data or exclusively MS/MS data if sufficiently available can be used to determine protein primary structure including modifications and modification sites. CID also produces predominantly b and y ions as well as a number of low mass fragment ions such as immonium and side chain fragment ions of the a, c, x and z ion series which allow discrimnation of isobaric amino acid residues. ECD provides the most extensive characterization of protein primary structure. An overview of common MS/MS fragmentation patterns is given (Figure 1).

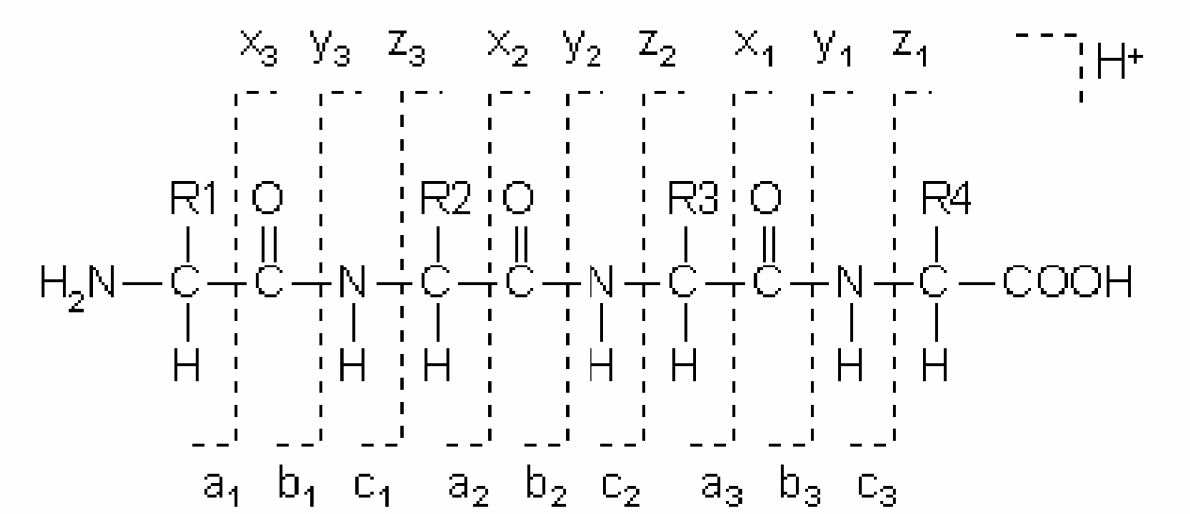


Figure 1. Overview of the most common MS/MS fragmentation patterns in polypeptide mass spectrometry.

Peptide based proteomics procedures are called bottom up procedures. Bottom up procedures using 2-DE for protein species separation can identify protein species by providing 100% sequence information and elucidating the exact site of possible multiple post translational modifications. They are however very labor intensive and almost entirely off-line. Automatic spot excision and in gel digestion machinery is available but needs to be improved.

In contrast to bottom up protein analysis top down proteomics is based on the analysis of undigested proteins. Modified proteins are easily detected by mass shifts, however exact modification sites remain elusive without sequence information. Most recent advances featuring FTICR instruments and ECD are beginning to bring top down

protein MS/MS into focus, however this is still visionary (Sze et al., 2002). Nevertheless routine full protein MS/MS would be a major milestone combining the speed of liquid chromatography on-line mass spectrometry with the *de facto* realization of the protein species level.

Reagents for directed in vitro chemical modification of proteins have been developed with the reduction of sample complexity, relative and absolute quantification of sample constituents and improvement of peptide MS/MS fragmentation in mind. These strategies are colloquially termed protein “tagging” or “labeling”. Most labeling reagents combine a reactive group targeting specific chemical groups, an affinity group such as a biotin moiety for selective enrichment or purification and a reporter group featuring different isotope compositions among the different forms of the same label for quantification in mass spectrometry. Generally, samples labeled with the different forms of the same reagent distinct in mass are combined. Their components are then separated and concomitantly analyzed with mass spectrometry. The comparison of the ion signals of the differentially labeled components with the same primary structure from each sample allows relative quantification of proteins. By incorporating and labeling a known amount of an internal standard absolute quantification can be achieved. Labeling chemistry can be introduced into a proteomics experiment at any stage of protein separation and is compatible with 2-DE and/or liquid chromatography and can be applied to proteins and/or peptides depending on the specific reagent.

Many labeling reagents specifically target cysteine amino acid residues. An improved version of the original cysteine specific labeling reagent, the isotope coded affinity tag ICAT (Gygi et al., 1999), the cleavable ICAT reagent, was made commercially available by Applied Biosystems (Foster City, CA, USA). The light and heavy forms of this reagent are carbon isotope coded and have a mass difference of 9 Da. The two forms of the reagent are chromatographically inert so the differentially labeled peptides

with the same primary structure are not separated in chromatography. The reagents biotin moiety is cleaved under acidic conditions making it smaller and eliminating reagent fragments in MS/MS.

In contrast to targeting specific amino acid residue side chains, reagents for unspecific labeling of peptides N- or C-termini are also available. This is termed global internal standard technology or GIST (Chakraborty and Regnier, 2002) and allows a broader realization of a samples constituents. Several reagents have been developed for N-terminal labeling of peptide amino groups including the iTRAQ label (Ross et al., 2004) which was also made commercially available by Applied Biosystems.

The iTRAQ strategy allows the discrimination of up to four samples. The label features an amino group specific reactive group, an isotope coded reporter group specific to each of one of four reagent forms and a neutral loss or balance group making the four reagent forms isobaric. The four reagent forms are carbon isotope coded, so differentially labeled peptides with the same primary structure are not chromatographically separated and acquired simultaneously in one MS spectrum. Furthermore, they are concomitantly fragmented and also combined in one MS/MS spectrum. The isotope coded reporter groups specific for each reagent form dissociate in MS/MS and allow the discrimination and quantification of the differentially labeled same peptides by their fragment ion signals in the low mass range of the MS/MS spectrum. The balance group also dissociates and is lost as it does not carry a charge. Finally, as the reporter and balance groups have dissociated, the peptide fragments with the same primary structure labeled with the remainder of the label are rendered isobaric, so the fragment ion signals from the same peptides are cumulative. This greatly increases spectral quality and peptide identification.

The methods for polypeptide separation and analysis described above are combined to build a proteomics strategy. In many cases these and other methods are

interchangeable and deliver similar but specialized results. Their optimal combination for comprehensive protein species separation and identification over a broad range and with a high throughput is the essence of proteomics and requires careful planning and experience. An overview of these proteomics techniques with the exception of FTICR mass spectrometry is given (Figure 2).

1.2. The Lens of the Eye: Development and Transparency, Proteins and Morphology

Proteomics research on the eye lens has a long history. Already in 1975 Kibbelaar and Bloemendal applied whole size exclusion chromatography water soluble α, β_H, β_L and γ and water insoluble fractions as well as the urea solubilized pellet from calf lens to urea PAGE combined with SDS-PAGE producing a first rudimentary protein spot pattern of the major protein components (Kibbelaar and Bloemendal, 1975). The lens is an avascular cellular conglomerate whose function is to guarantee proper light refraction and visual acuity (Bloemendal, 1977). Following the development of the lens nucleus from the posterior epithelial cells of the lens vesicle, cells cease mitosis with the exception of the cuboidal cells of the anterior epithelium. These cells are eventually passed from the central region of the anterior epithelium through the germinative region to the lens equator, where they undergo differentiation into fiber cells. The cuboidal cells elongate and synthesize large amounts of proteins of the crystallin protein superfamily (Mörner, 1894) and at the terminal stage of differentiation loose their organelles in a process reminiscent of apoptosis (Papaconstantinou, 1967; Bassnett, 2002). These so called secondary fiber cells are layered onto the lens nucleus and successively onto

each other forming the lens cortex. Lenticular development is essentially complete a few days after birth but continues at a slowed rate throughout life.

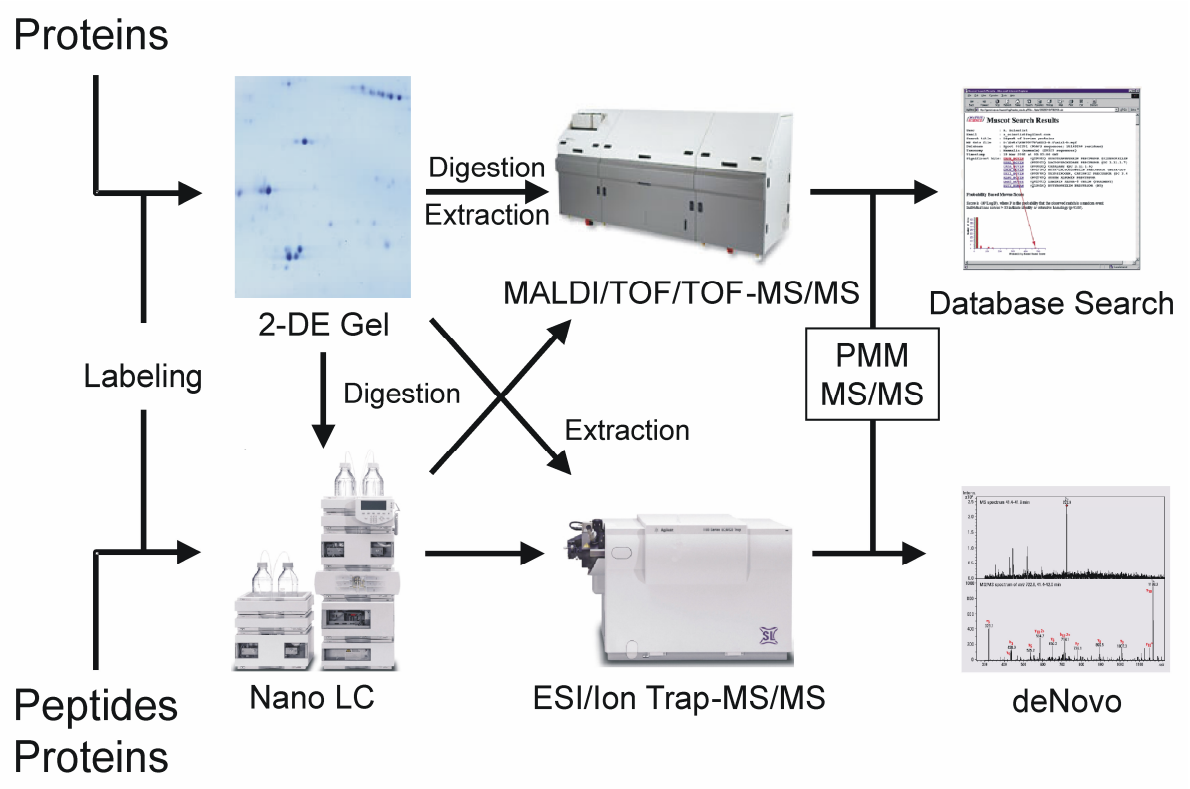


Figure 2. Building a proteomics strategy. The major techniques for protein and peptide separation, analysis and identification.

The lack of organelles, high protein concentrations and the short range order of the soluble proteins are necessary for lens transparency (Delaye and Tardieu, 1983). Naturally, aspects of lens biology vary greatly between species, reflecting optical specialization and environmental demands. The human and fish lens represent two extremes, the former, like diurnal animals having a softer more uniform mould, while the latter is clearly divided into a hard nucleus and a much softer cortical region (Koretz and Handelmann, 1988; Fernald and Wright, 1983). The rodent lens is very hard with evidence from Raman spectroscopy suggesting the proteins in its core are denatured (Yu et al., 1985). Concurrently, refractive indices also vary greatly between species (Kröger et al., 1994).

Thirty to thirty five percent of bovine lens wet weight is protein, the other 65% is virtually water. The crystallin protein superfamily comprises 90% of the lens protein mass (Bloemendal, 1977). The 16 crystallins ubiquitous to mammals undergo extensive post translational modification so that a vast number of protein species are in constant flux in various degrees of dynamic interaction both physically and physiologically (Bhat, 2004), maintaining the soluble basis for lenticular transparency. Serious perturbations of this balance leads to improper protein association, aggregation and insolubilization. Incident light is diffracted by the protein aggregates, resulting in lens opacification and a more or less grievous loss of visual acuity clinically termed cataract (Lambert, 1994).

The crystallins are divided into three major classes, Alpha (α), Beta (ß) and Gamma (γ) crystallins (Mörner, 1894). Evolutionary constraints have imparted a high degree of stability on this protein superfamily, as lenticular cells devoid of cellular machinery are unable to turnover proteins. Native Alpha crystallin is an oligomer with a Mw in excess of 300 kDa, consisting of monomeric subunits in dynamic exchange (van der Ouderaa et al., 1973; Puri et al., 1983; van den Oetelaar et al., 1990; Gesierich and Pfeil, 1996; Bova et al., 1997). The two major Alpha crystallin monomers are the products of two genes, Alpha A and Alpha B and share this terminology. They are 60 % identical in primary structure (Schoenmakers et al., 1969; Bloemendal and de Jong, 1991). The third Alpha crystallin monomer is an alternative splicing product of the Alpha A crystallin gene and is termed Alpha A insert crystallin (Hendriks et al., 1990). The small heat shock protein HSP27/25 can also interact with Alpha crystallins as a monomeric subunit in functional higher order assembly (Leroux et al., 1997a, 1997b).

The Alpha A chain has a Mw around 20 kDa, Alpha B around 22 kDa and Alpha A insert around 24 kDa. The monomers have a global, hydrophobic N-terminal domain characterized by α-helical structure function regions (Smith et al., 1996; Pasta et al., 2003), a ß-sheet structured C-terminal region containing the so called „Alpha crystallin

domain“ (Caspers et al., 1995) and a C-terminal extension. The Alpha crystallin domain is implicit to all small heat shock proteins and it was shown that Alpha B crystallin is indeed a member of this family (Klemenz et al., 1991) and that the Alpha A crystallin chain can function as a molecular chaperone (Horwitz, 1992).

Both native and recombinant Alpha crystallin has not been successfully crystallized due to its polydisperse nature, however models of the monomers tertiary structures (Farnsworth et al., 1997; Hoehenwarter et al., 2006a) and the functional oligomers quaternary structure have been proposed (Tardieu et al., 1986; Groenen et al., 1994; Haley et al., 1998; Horwitz et al., 1998; Abgar et al., 2000). In a typical mammalian lens Alpha crystallin constitutes about 35% of proteins and is synthesized throughout lens development and is evenly distributed throughout the organ (Aarts et al., 1989) with trace amounts being detected in other tissues (Iwaki et al., 1990; Srinivasan et al., 1992). Apparently Alpha crystallin is not only a major structural component of transparent lens architecture but its properties as a molecular chaperone are an important factor in maintaining lens homeostasis.

The Beta and Gamma crystallins are organized into two structural domains with a short connecting peptide. The domains are made up of two consecutive greek key motifs each consisting of four ß-strands which intercalate and arrange the domain in anti-parallel ß-sheet conformation (Driessen et al., 1981; Wistow et al., 1981; Wistow and Piatigorsky, 1988; Slingsby et al., 1996; Carver, 1999). Native Beta crystallins assemble in homo- or heterodimers which can dissociate into monomers or associate in higher Mw oligomers probably up to octamers while Gamma crystallins generally do not associate (Berbers et al., 1984; Bateman et al., 2001; Hejtmancik et al., 1997, 2004; Slingsby and Bateman, 1990; Zarina et al., 1994). These similarities in protein secondary and tertiary structure suggest a common ancestral gene encoding a single motif and a series of ancient gene duplication and recruitment events resulting in the genetic diversification of

Beta and Gamma crystallins in today's vertebrate species (Wistow and Piatigorsky, 1988; Lubsen et al., 1988; Norledge et al., 1996; Clout et al., 2000).

There are six mammalian Beta crystallin genes and similar genes detected in amphibian and fish expressing three basic (B1 – B3) and four acidic (A1 - A4) Beta crystallin proteins. Beta A1 and A3 crystallin are encoded by the same gene from distinct translation initiation points. Beta A1 crystallin lacks the 17 N-terminal amino acid residues of the full length translation product, Beta A3 crystallin. The Beta crystallin monomers have Mw ranging from 22 kDa (Beta A4) to 28 kDa (Beta B1) (Berbers et al., 1984). They are characterized by N- and C-terminal extensions reaching beyond the molecular domains. It is an accepted hypothesis that these extensions and the connecting peptide are instrumental in dimer- and oligomerization of Beta crystallins (Carver et al., 1993; Cooper et al., 1993a, 1993b) initiated by 3D domain swapping (Bennett et al., 1994, 1995) with the extensions promoting higher order association beyond the dimer (Lampi et al., 2001; Bateman et al., 2001, 2003). The Beta B1 and B3 crystallin genes are expressed early in lenticular development, consequently their proteins are found primarily in the lens nucleus. Beta B2 is often seen as the main Beta crystallin and the acidic Beta crystallins are found throughout the lens in the nucleus and the cortex (Aarts et al., 1989; Chambers and Russell, 1991; Lampi et al., 1998; Ueda et al., 2002).

The six Gamma crystallin proteins common to mammals, Gamma crystallins A through F, are linked to a gene cluster of highly homologous genes (Shinohara et al., 1982; Breitman et al., 1984; Lok et al., 1984; Goring et al., 1992; Graw et al., 1993). They are among the earliest gene products in lens development and like the Beta crystallins are fiber cell specific. It has been suggested that the mammalian Gamma crystallins constitute the tightest molecular arrangement among crystallins and that they exclude water to the highest degree being constituents of the harder lens nucleus

especially in rodents (Slingsby, 1985). Gamma S crystallin, the only Gamma crystallin detected in all vertebrate species and present in numerous tissues outside the lens is particularly suited to water rich tissue regions and can thus be regarded as a separate branch in vertebrate Gamma crystallin gene evolution (van Rens et al., 1989).

The crystallins are extensively modified in the course of healthy lens development. The concept is emerging that the crystallins in high concentration are more than physical structural lens components, but that their protein species possess additional manifold catalytical or regulatory functions also essential for lens transparency (Bhat, 2004). Thus the dynamics of this protein superfamily and indeed of the lens itself are only beginning to be understood.

Another central concern in guaranteeing light transmission without diffraction is proper lens architecture (Kuszak et al., 1994). Fiber cell shape is maintained by an extensive cytoskeleton whose protein components are synthesized and finalized prior to terminal differentiation (Kuwabara, 1975; FitzGerald, 1988). Actin, synthesized in the epithelial cells is polymerized from globular Actin (Ramaekers et al., 1981). The intermediate filament protein Vimentin also synthesized throughout the epithelium and the elongating fiber cells is eliminated during terminal differentiation and not present in the mature fiber cells (Bradley et al., 1979; Ellis et al., 1984; Sandilands et al., 1995). The major component of the mature cytoskeleton, the lens unique beaded filament is established last (Maisel and Perry, 1972). The beaded filament is made up of two proteins, Phakinin and Filensin also termed CP49 and CP95, and has a 5 – 7 nm backbone with 20 nm beads evenly spaced throughout. It associates with Alpha crystallin and is itself membrane associated (Ireland and Maisel, 1984, 1989; Carter, 1995).

Membrane proteins such as gap junctions also contribute to cellular stability and arrangement. The gap junction network is instrumental in circulating ionic currents thought to be followed by fluid flow throughout the lens and may direct the overall pattern

of current flow in the lens. This guarantees a homeostatic intracellular environment where active transport from peripheral epithelial cells maintains the milieu in the differentiated interior primary and secondary fiber cells. (Robinson and Patterson, 1982-1983; Mathias et al., 1997). Gap junctions consisting of two connexon hemi channels each made up of six connexin proteins are a large part of this network (Goodenough, 1992; Kumar and Gilula, 1996; Rae et al., 1996). The Alpha 3 and Alpha 8 connexin proteins form homo- and heterotypic as well as heteromeric gap junctions (Kumar and Gilula, 1996) in the differentiated fiber cells of the lens nucleus and cortex. The Alpha 1 connexin has been detected in the lens epithelial cells. A typical 2-DE protein spot pattern of the urea soluble fraction of the lens proteins of the ten day old *Mus musculus* C57BL/6J strain produced with a gel size of 23 x 30 x 0.25 cm and silver staining is shown (Figure 3).

1.3. Cataract Development and Modifying Factors

Lenticular opacification is clinically termed cataract. It is a protein conformational disease attributed to unfolding, improper folding, insolubilization and aggregation of lens proteins. The protein aggregates scatter incident light thus impairing vision. Cataract development is imminently connected to the calcium ion levels in the fiber cells of the lens. (Harding, 1972; Carrell and Lomas, 1997; Shearer et al., 1999; Harding, 2002). Cataract is a common age related disease in humans linked to a number of risk factors and genetic predisposition (Graw, 2003, 2004) and leaves 4 million people newly blind each year in India alone (Minassian and Mehra, 1990). Congenital hereditary cataracts appearing early in life are comparatively rare and occur with an incidence of 30 cases per 100000 live births in developed countries (Graw, 2004). Numerous dominant and recessive gene mutations involved in cataract onset have been categorized and mapped

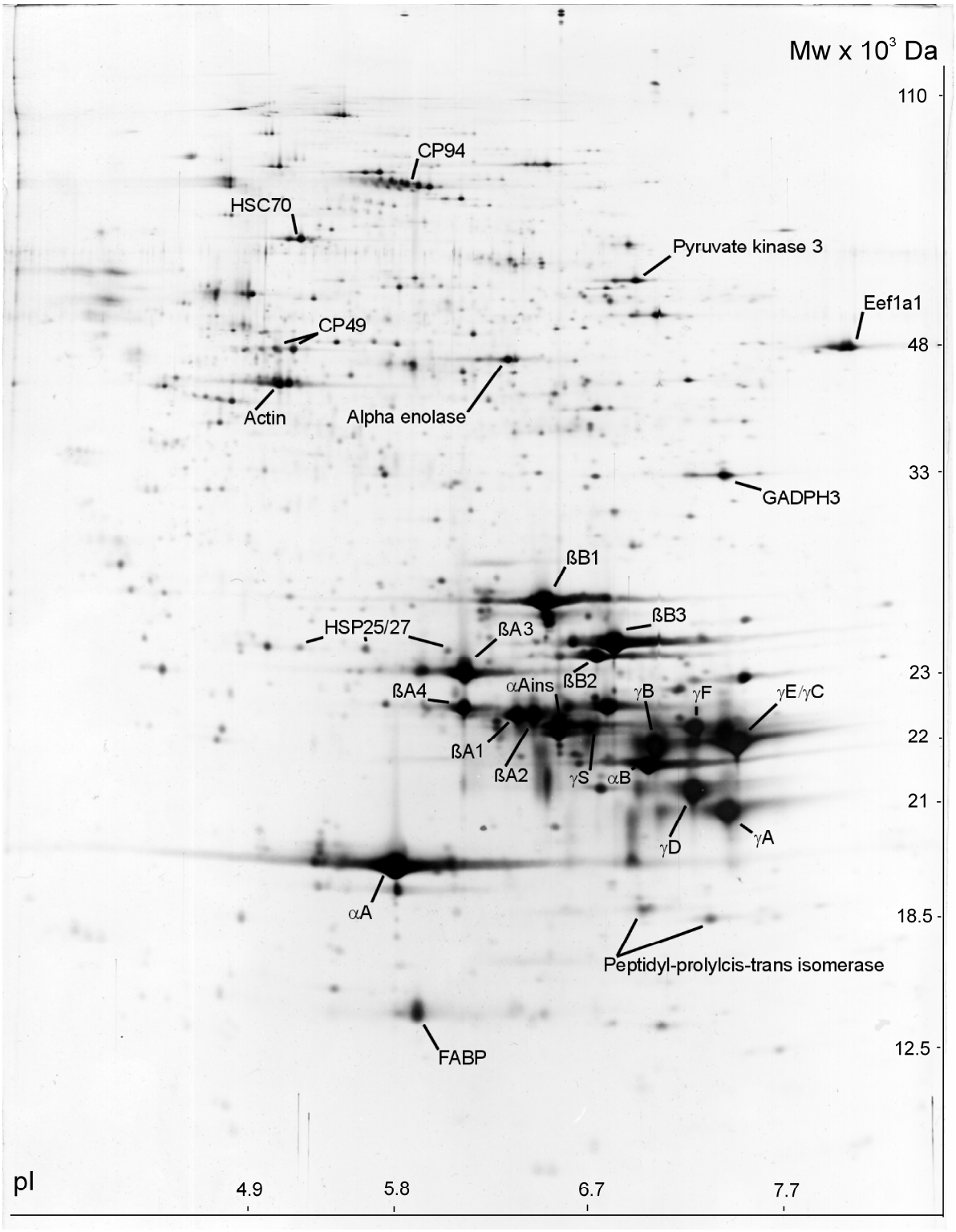


Figure 3. Large scale 2-DE protein spot pattern of the urea soluble fraction of the lens proteins of the *Mus musculus* C57BL/6J strain with some identified proteins. The crystallins are named as token only.

in humans and mice including mutations in the crystallin, connexin, and membrane protein genes (Graw, 2004).

Studies have demonstrated knockout of the Alpha 3 connexin gene *Gja3* results in dense nuclear cataracts in mouse lenses beginning at 2 – 3 weeks of age with 100% penetrance (Gong et al., 1997, Baruch et al., 2001). With the loss of the Alpha 3 connexin protein, gap junction architecture and consequently the entire gap junction network is perturbed. The lack of Alpha 3 connexin is compensated by Alpha 8 connexin in the lens cortex. In the lens nucleus this is not the case, probably due to early C-terminal Alpha 8 connexin proteolysis (Evans et al., 1993). Two effects are observed. The membranes of the fiber cells in the lens nucleus become porous facilitating calcium ion influx and the fiber cells of the nucleus are uncoupled from the gap junction network and the fiber cells of the lens cortex. As a result ion currents dissipate, calcium ion out flux is abated and intracellular calcium is accumulated (Robinson and Patterson, 1982-1983, Gong et al., 1997, 1998). Under these conditions the calcium dependant protease Lp82, a splice variant of muscle type calpain p94 and a member of the calpain protease family absent in humans (Ma et al., 1998, 1999; Fukiage et al., 2002, Reed et al., 2003) enjoys heightened activity and cleaves a number of crystallins in the lens nucleus including Gamma crystallin between asparagine residue 73 and serine residue 74. The onset of the nuclear cataract is definitively correlated to this cleavage event which produces 9 and 11 kDa Gamma crystallin truncation products (Gong et al., 1997; Baruch et al., 2001).

It was shown that *Mus musculus* C57BL/6J mice are less susceptible to cataract because of *Gja3* knockout and loss of gap junctions than *Mus musculus* 129/SvJ mice (Gong et al., 1999). Nuclear cataract begins to develop as early as postnatal day 12 in the homozygous *Gja3* knockout 129/SvJ mice while lenses of C57BL/6J mice homozygous for the knockout are still transparent at postnatal day 18 to 20. Indeed,

nuclear cataracts in transgenic mice are mild on the C57BL/6J but severe on the 129/SvJ background at two months of age. This suggests the presence of factors that modify the onset and severity of the pathology (Graw, 2003) which may be attributed to genetic variability, epigenetic variability or other causes in the two strains. However, the molecules that actually effect the phenotype are proteins and therefore a comprehensive proteomics study was undertaken.

1.4. Goals of this Study

The purpose of this work was the investigation of the variability in the proteins of the mouse eye lens and its effects on the differential onset and severity of cataract. Protein labeling and LC-MS techniques were developed and established for rapid realization of differential protein abundance and applied to the urea soluble lens proteins of the mouse species *Mus musculus* and *Mus spretus*. For the detection of modifying factors impacting cataract phenotype onset a 2-DE and MS based study of the protein abundance in the lenses of the wild type and the transgenic homozygous *Gja3* knockout mutants of the *Mus musculus* C57BL/6J and 129/SvJ strain was conducted. The lenses used in the study from both the wild type and mutant C57BL/6J animals showed no abnormalities while lenses from the mutant 129/SvJ animals showed nuclear cataracts at ten days of age. Otherwise the development of the animals was normal. Comparative analysis of the wild type and mutant C57BL/6J and 129/SvJ lens 2-DE protein spot patterns should clearly discriminate protein species involved in healthy lens development and cataract onset and development and protein species which modify cataract onset and severity in each strain.

2. Experimental Procedures

2.1. Chemicals and Reagents

Premium chemicals were purchased from various suppliers. Premium distilled water was either purchased from various suppliers or produced in-house with a Milli-Q Biocel system from Millipore (Billerica, MA, USA).

2.2. Mouse Eye Lenses

Fourteen weeks old male and female *Mus musculus* and *Mus spretus* animals were sacrificed. The lenses of the eyes were extracted and decapsulated immediately post mortem removing the outer epithelial cell layer (Toyofuku and Bentley, 1970) and frozen in liquid nitrogen and stored at -80°C until analysis. The animals were handled according to the guidelines for animal care of the Institute of Human Genetics of the Charité in Berlin. The lenses of ten days old *Mus musculus* C57BL/6J and 129/SvJ wild type and transgenic homozygous *Gja3* knockout animals with anterior epithelial cells intact (Dewey et al., 1995) were shipped on dry ice from the Department of Ophthalmology and Visual Sciences at the University of Illinois at Chicago and stored at -80°C until analysis. The animals were handled in accordance with the ARVO Statement for Use of Animals in Ophthalmic and Vision Research and the guidelines established by the Animal Care Committee of the University of Illinois at Chicago.

2.3. Lens Protein Preparation for Labeling with the ICAT Reagent

Two lenses each of *Mus musculus* and *Mus spretus* animals were ground down in a mortar and pestle under liquid nitrogen. The resulting powder was weighed and suspended in 5 to 6 times 10^{-3} weight per volume (w/v) of fresh buffer containing 9 M urea, 200 mM Tris, 5 mM EDTA, 5 mM tributylphosphine (TBP) and 0.1% sodium dodecyl sulfate (SDS), pH adjusted to 8.3 with HCl; i. e. 10 mg were suspended in 60 µl of buffer. Following agitation the suspension was centrifuged at 100000 g for 40 minutes. The supernatant containing the urea soluble lens protein fraction was removed and diluted to a final concentration of 6 M urea, 200 mM Tris, 5 mM EDTA, 5 mM TBP and 0.1% SDS, and the pH adjusted to 8.3 with HCl. This is the ICAT labeling buffer described by Smolka and coworkers (Smolka et al., 2001). The procedure yielded final volumes between 60 and 80 µl which were stored in 15 µl aliquots at -80°C until use.

2.4. Determination of Protein Concentration According to Lowry

The protein concentrations in the urea soluble protein fraction of the lenses were determined according to a modified protocol according to Lowry (Lowry et al., 1951, Peterson, 1977). A Bovine serum albumine (BSA) solution with a concentration of 0.5 mg/ml was employed as a standard for calibration. Zero, 10, 25, 50, 75, 100, 150 and 200 µl of BSA solution in a total volume of 1000 µl water were used for calibration. Two and five µl of sample in a total volume of 1000 µl water were used for protein concentration determination. One hundred µl of 0.15% sodium desoxycholate solution (DOC) were added to the diluted BSA standards and samples, which were then agitated and incubated at room temperature (RT) for ten minutes. One hundred µl of 72%

trichlorine acetate (TCA) were added, the solutions were then agitated and centrifuged at 3000 g at 4°C for 20 minutes. The supernatant was discarded and the remaining film was air dried for 30 minutes. The dried film was suspended in 1000 µl water. One thousand µl of solution A containing 0.05% potassium tartrate, 0.025% copper sulfate, 2.5% sodium carbonate, 0.2 N sodium hydroxide (NaOH) and 2.5% SDS in water were added. The mixture was agitated and incubated for ten minutes at room temperature (RT). Five hundred µl of 0.33 N Folin reagent were added and the mixture was agitated immediately. The mixture was then incubated for 30 minutes at RT. Following incubation the optical density (OD) of the BSA standards and samples were measured in a photometer at a transmission wavelength of 750 nm. ODs were recorded and sample concentrations calculated with the Excel software from Microsoft (Redmond, WA, USA).

2.5. Lens Protein Labeling with the ICAT Reagent

The ICAT reagent (Applied Biosystems, Foster City, CA, USA) reacts specifically with the side chains of the amino acid cysteine. One hundred µg of the *Mus musculus* lens proteins were labeled with the light (C12) form, 100 µg of the *Mus spretus* lens proteins with the heavy (C13) form of the cleavable ICAT reagent according to a protocol derived from an in-house protocol from Ruedi Aebersolds laboratory and Smolka and coworkers (Smolka et al., 2001).

Considering 90% of lens proteins are crystallins (Bloemendal, 1977) with a molecular weight ranging from about 20 to 30 kDa, an average molecular weight of 25 kDa for all lens proteins was estimated. With this calculation 100 µg constitute 4 nmol of protein. Four cysteine residues per lens protein was estimated taking into account that Alpha crystallins have only one cysteine residue and the presence of some higher molecular weight lens proteins. Therefore there is an amount of 16 nmol of cysteine

residues in the lens protein fraction. The cleavable ICAT reagent is supplied in tubes with 175 nmol reagent per tube. The amount in one tube is sufficient to establish a molar excess of ICAT reagent over free sulfhydryls. A total reaction volume of 100 µl was used for the labeling procedure. This amounts to a 1.75 mM ICAT reagent concentration sufficient for constitutive labeling.

One tube containing the light and one tube containing the heavy form of the reagent were briefly spun. Twenty µl of fresh ICAT labeling buffer were transferred to each tube, the tubes were then agitated and briefly spun. The solved ICAT reagents were transferred to standard reaction tubes from Eppendorf (Hamburg, Germany). One hundred µg of *Mus musculus* and *Mus spretus* lens proteins were transferred to the respective tubes. The reaction volumes were adjusted to 100 µl with the ICAT labeling buffer. The tubes were wrapped in aluminum foil to shield the reaction from light and agitated at 37°C for 2 hours. The reaction was quenched with a 10 fold molar excess of dithiothreitol (DTT) and the labeled proteins were combined bringing the total volume and protein amount to 200 µl and 200 µg, respectively. The combined reaction stock was diluted 6 fold with water lowering the urea concentration to 1 M and the proteins digested with 2% (w/v) Trypsin (Sequencing grade modified Trypsin, Promega, Madison, WI, United States) at 37°C overnight. Two independent differentially labeled reaction stocks based on separate preparations of lenses were produced.

2.6. Purification of ICAT Labeled Peptides

The ICAT labeled peptides were enriched using cation exchange and avidin batch chromatography. A column holder, cation exchange and avidin columns, fittings and buffers and cleaving reagents are available in kits from Applied Biosystems. Water and buffers employed in the purification procedures were warmed to and kept at 37°C. The

columns were placed into the column holder and connected to a 1 ml Hamilton (Reno, NV, USA) Teflon® luer lock (TLL) syringe with capillary tubing from GE Healthcare (Fairfield, CT, USA) with finger tight fittings and a luer lock adapter. The syringe was placed into a KDS 100 syringe pump from kdScientific (Holliston, MA, USA) for liquid phase application.

One hundred µl of the digested reaction stock containing 100 µg of ICAT labeled peptides were diluted with 2000 µl of Cation Exchange Buffer Load and then agitated. The cation exchange column was equilibrated with 2000 µl of Cation Exchange Buffer Load. Following equilibration the sample was loaded onto the column with a flow rate of 4 ml per hour. The column was washed with 2000 µl of Cation Exchange Buffer Load and the peptides subsequently eluted with 1000 µl of Cation Exchange Buffer Elute with a flow rate of 4 ml per hour. A 1 µl aliquot was removed for control.

One thousand µl of Affinity Buffer Load were added to the peptides eluted from the cation exchange chromatography. The affinity column was pre-cleaned with 4000 µl of Affinity Buffer Elute and equilibrated with 4000 µl of Affinity Buffer Load. The sample was loaded onto the column with a flow rate of 4 ml per hour and the column was washed with 1000 µl of Affinity Buffer Load followed by 2000 µl of Affinity Buffer Wash 1, 2000 µl of Affinity Buffer Wash 2 and 2000 µl of water with a flow rate of 6 ml per hour. The peptides were eluted with 800 µl of Affinity Buffer Elute with a flow rate of 4 ml per hour. Flow through from the column wash with Affinity Buffer Wash 2 and a 1µl aliquot of the eluent were kept as a control. Both columns were cleaned according to the instructions provided with the kits.

The eluted peptides from the affinity chromatography were cryo-lyophilized. Ninety-five µl of Cleaving Reagent A were combined with 5 µl of Cleaving Reagent B and then agitated. The combined cleaving reagent was added to the cryo-lyophilized peptides and then agitated. The mixture was incubated at 37℃ for 120 minutes and evaporated in

an Eppendorf Concentrator 5301 at 45°C. The enriched, fully evaporated cleaved differentially ICAT labeled peptides from the lens proteins of *Mus musculus* and *Mus spretus* lenses were solved in 100 µl 0.1% formic acid (FA) in water for analysis which amounted to a final protein concentration of 1µg or 40 pmol per µl.

2.7. Strong Cation Exchange HPLC

Strong cation exchange HPLC for enhanced separation of peptides was established on a SMART System with a Mono S^{TM} PC 1.6/5 strong cation exchange column (Amersham Biosciences, a division of GE Healthcare). Zero point zero five percent (0.05%) FA in 25% ACN in water was used as solvent A and 1M potassium chloride (KCl) in solvent A was used as solvent B.

Two µg of the enriched, differentially ICAT labeled peptides were diluted to a concentration of 0.01 µg per µl with solvent A in a total volume of 200 µl; the pH of the solution was adjusted to around 3 with FA. The flow rate was set to 100 µl per minute. The column was equilibrated for 10 minutes with solvent A. The peptides were loaded onto the column with a 100 µl Hamilton syringe using a 200 µl external sample loop. The peptides were eluted with a linear gradient inclination of 0 – 50% solvent B in solvent A over a duration of 20 minutes. Following elution the column was washed with 100% solvent B for 20 minutes and equilibrated with solvent A for 10 minutes.

The procedure was up-scaled for preparative separation of the ICAT labeled peptides. Twenty µg of the peptides were diluted to a concentration of 0.002 µg per µl with solvent A in a total volume of 10 ml and the pH adjusted. The peptides were sequentially loaded onto the column in five steps with a volume of 2 ml each using a 1ml Hamilton syringe and a 2 ml external sample loop. The flow rate was set to 200 µl per minute, each loading phase was set to proceed for twenty minutes so the sample was

completely purged from the sample loop and loaded onto the column. Pump refills were incorporated into the chromatographic program. The peptides were eluted with a linear gradient inclination of 0 – 50% solvent B in solvent A over a duration of 30 minutes. Fractions were collected with a fraction volume of 1 ml. The fraction collection was programmed to incorporate peak characteristics. The minimum peak width was set to 0.70 minutes in chromatographic duration, the baseline noise level was set to 0.00174 AU, the peak start and end was set to an absolute AU value of 0.005 and peak valley consideration was set to ON. The fraction collection was stopped after the elution gradient reached 50% solvent B in solvent A.

2.8. Reverse Phase LC and On-line ESI Ion Trap Mass Spectrometry of ICAT Labeled Peptides

Reverse phase liquid chromatography and on-line mass spectrometric analysis of differentially ICAT labeled peptides was performed on an UltiMate™ nano HPLC system from LC Packings (A Dionex Company, Sunnyvale, CA, USA) connected to an LCQ Deca XP ion trap instrument from Thermo Finnigan, a subsidiary of the Thermo Electron Corporation (Waltham, MA, USA) and on an 1100 series nano HPLC system connected to an XCT ion trap instrument both from Agilent Technologies (Palo Alto, CA, USA). For the system from Thermo Finnigan, fifty pmol or a little more than 2 µg of peptides in solvent A (0.1% FA in water) were loaded onto a C18 PepMap™ column with a diameter of 0.075 mm and a length of 150 mm from LC Packings over a desalting column with a flow rate of 20 µl per minute. The peptides were eluted from the column into the mass spectrometer with a flow rate of 200 nl per minute with a linear gradient inclination of 0-30% of solvent B (0.1% FA in ACN) in solvent A over a duration of 175 minutes. Mass

spectra were acquired throughout the chromatographic duration over a mass range of 350 to 2000 Da with a minimum ion signal intensity of 100000. Three spectra were acquired in series: a spectrum of all ions in the trap, a data dependant zoom spectrum plus and minus 5 Da around the highest intensity ion signal and an MS/MS spectrum of the highest intensity ion signal. The dynamic exclusion was enabled so an ion would not be subjected to renewed MS/MS spectral acquisition regardless of its signal intensity. The exclusion duration was set at three minutes with a tolerance of plus and minus 3 Da around an ion signal mass. Calibration with an external standard achieved a mass accuracy of better than 0.5 Da.

For the system from Agilent Technologies between 30 and 240 pmol of peptides in solvent A (0.1% FA in water) were loaded onto a ZORBAX stable bond C18 column with a diameter of 0.075 mm and a length of 150 mm from Agilent Technologies over a desalting column with a flow rate of 20 µl per minute in 5 separate analytical runs. The peptides were eluted from the column into the mass spectrometer with a flow rate of 200 nl per minute with a linear gradient inclination of 3-30% of solvent B (0.1% FA in ACN) in solvent A over a duration of 175 minutes. Mass spectra were acquired throughout the chromatographic duration over a mass range of 200 to 2200 Da. The two ion signals with the highest intensity were subjected to MS/MS spectral acquisition if their absolute signal intensity exceeded 10000 counts and their relative signal intensity exceeded 2% of baseline intensity. Active Exclusion was enabled. An ion would not be subjected to renewed MS/MS spectral acquisition regardless of its signal intensity after up to three MS/MS spectral acquisition events of the same ion for 1.5 minutes in chromatographic duration after the original MS/MS event. Calibration with an external standard achieved a mass accuracy of better than 0.4 Da.

2.9. Lens Protein Identification and Quantification with the ICAT Reagent

The raw data from the LCQ Deca XP ion trap instrument were evaluated and processed with the Bioworks Browser and SEQUEST software from Thermo Finnigan (Eng et al., 1994). The processed spectra were extracted from the raw data and deposited as scans for searches of a local version of the NCBI Protein sequence database (Bethesda, MD, USA) with the same software. The mass peak signals in an MS/MS spectrum had to surpass a signal to noise threshold of 1000 and the spectrum had to contain at least 15 mass peaks to be extracted into a scan. The database was searched with the extracted scans allowing a matching tolerance of 3 Da for the precursor mass and of 0.1 Da for matching of MS/MS fragment ion masses. Methylation of aspartate and glutamate residues and phosphorylation of serine and threonine residues were tolerated as variable modifications. In addition to an Xcorr higher than 1 from the identification software, scans had to contain four mass peaks from peptide fragment ions of the same ion series indicating three consecutive amino acid residues in the corresponding peptide sequence.

The differential protein abundance in the *Mus musculus* and *Mus spretus* lenses was determined with the XPRESS^{TM} software which performs the user defined extraction of the areas of the ion mass peak signals over the acquired ion current duration and the calculation of the ratio of the areas. The software can be downloaded from the Institute for Systems Biology in Seattle, WA, USA (Han et al., 2001). The ratios from all of the peptide pairs from a protein were averaged for relative quantification of the proteins abundance.

The raw data from the XCT ion trap instrument was evaluated and processed with the SpectrumMill software from Agilent Technologies which was also employed for database searches and relative protein quantification. For processing, the signal to noise

threshold was set to 5. MS/MS spectra with the same precursor mass with a tolerated error of 0.5 Da and a range of 120 MS/MS spectra in the acquired ion current and at least five fragment ion signals were extracted into scans. A local version of the NCBI Protein sequence database was searched with the extracted scans with the precursor mass tolerance for matches set to 0.8 Da and the tolerance for fragment ion matches set to 0.4 Da. The peptide and protein identification was only considered conclusive if a significant score was achieved. For relative protein quantification the software performs the parameter dependant extraction of the ion signal areas of an identified, ICAT labeled peptide pair from the acquired ion current and calculates the appropriate ratio of the areas.

2.10. Protein Standard and Lens Protein Preparation for Labeling with the iTRAQ Reagent

An equimolar mixture of five proteins, Bovine serum albumine (BSA), Carbonic anhydrase, Conalbumin, Cytochrome C and Myoglobin was prepared with a concentration of 10 µg per µl in water. Two lenses each of *Mus musculus* and *Mus spretus* animals were ground down in a mortar and pestle under liquid nitrogen. The resulting powder was weighed and suspended in 5 to 6 times 10^{-3} weight per volume (w/v) RapiGest^{TM} SF buffer from Waters (Milford, MA, USA) in 100 mM ammonium bicarbonate (ABC) at a concentration of 0.2% weight per volume (w/v), i. e. 10 mg were suspended in 60 µl buffer. The suspension was agitated and centrifuged at 100000 g in a microcentrifuge. The supernatant, containing the soluble protein fraction was kept and stored in 15 µl aliquots at -80°C until use. The protein concentration was determined as above.

2.11. Protein Standard and Lens Protein Labeling with the iTRAQ Reagent

The iTRAQ reagent reacts specifically with the amino groups of amino acids. The equimolar protein mixture was labeled with the iTRAQ reagent with the reporter group mass of 117 Da. The *Mus musculus* lens proteins were labeled with the iTRAQ reagent with the reporter group mass of 114 Da, the *Mus spretus* lens proteins were labeled with the iTRAQ reagent with the reporter group mass of 115 Da. Fifty µg of the protein standard mixture in 5 µl of water in 20 µl of RapiGest^{TM} SF buffer and one hundred µg of lens proteins of each species in 20 µl of RapiGest^{TM} SF buffer were reduced with Tris (2-carboxyethyl) phosphine (TCEP) at a concentration of 5 mM in the reaction volume at 60°C for 60 minutes and briefly spun. Free sulfhydryls were blocked with Methyl methanethiosulfonate (MMTS) at a concentration of 10 mM in the reaction volume. The proteins were digested with 2% (w/v) Trypsin at 37°C overnight. The vials containing the iTRAQ reagents were thawed at RT. Seventy µl of ethanol (EtOH) were added to the vials containing the iTRAQ reagents with the reporter group masses of 114, 115 and 117 Da. The iTRAQ reagents were added to the reaction stocks bringing the total volumes to about 100 µl. The reaction stocks were agitated and incubated at 60°C for 60 minutes and briefly spun. The reaction stocks from the *Mus musculus* and *Mus spretus* species were combined for a total mass of 200 µg in a total volume of around 200 µl.

2.12. Reverse Phase LC and Semi On-line MALDI/TOF/TOF Mass Spectrometry of iTRAQ Labeled Peptides

The iTRAQ labeled peptides from the protein standard and the *Mus musculus* and *Mus spretus* lens proteins were separated using reverse phase liquid chromatography,

spotted onto MALDI plates with an automated spotting machine and analyzed by MALDI/TOF/TOF mass spectrometry. Fifty µg of the protein standard constituting 60 pmol of each protein and 1.7 µg of the lens proteins considering an average protein molecular weight of 25 kDa and estimating 200 total proteins in the lens (MacCoss et al., 2001) constituting 340 fmol of each protein were loaded onto a C18 PepMap™ column with a diameter of 0.075 mm and a length of 150 mm from LC Packings with an UltiMate™ HPLC System. Solvent A was 0.05% trifluoric acetate (TFA) in 5% ACN in water, solvent B was 0.04% TFA in 80% ACN in water. Elution was performed with a gradient inclination of 8 – 30% of solvent B in solvent A over a chromatographic duration of 90 minutes at a flow rate of 200 nl per minute.

Fractions were spotted onto blank 2x2 MALDI plates from Applied Biosystems with a Probot™ micro fraction collector from LC Packings. Three hundred spots were spotted per plate. Three hundred fractions of the protein standard were collected over a duration of 100 minutes. Five hundred and seventy six fractions of the lens proteins were collected in ten second intervals. Five mg/ml α-cyano-4-hydroxycinamic acid (CHCA) MALDI matrix containing 2% (w/v) CHCA (Sigma Aldrich, Munich, Germany), in 0.3% TFA in 50% ACN in water was added to the fractions at 1 µl per minute for deposition of the sample matrix mixture onto the MALDI plates. PMF and MS/MS spectra of the crystallized samples were acquired with a 4700 Proteomics Analyzer MALDI/TOF/TOF instrument from Applied Biosystems. PMF spectra were acquired with 1200 laser shots per spot. Internal calibration was preferred over the instruments default calibration. MS/MS spectra were acquired by metastable dissociation without CID with at least 2500 and at most 10000 laser shots per spot position under stop conditions. The instruments default calibration was used. Five MS/MS spectra were acquired per spot with the dynamic exclusion enabled. This excluded an ion from renewed MS/MS spectral

acquisition in the 5 spots following its first acquisition. Internal calibration achieved a mass accuracy of better than 30 ppm.

2.13. Protein Standard Identification and Lens Protein Identification and Quantification with the iTRAQ Reagent

The raw data from the 4700 Proteomics Analyzer was processed and evaluated with the GPS Explorer version 2.0 software from Applied Biosystems. Searches of the NCBI Protein sequence database were also performed with this software supported by the in-house licensed Mascot software (Matrix Science, Boston, MA, USA) with a precursor mass matching tolerance of 60 ppm and a fragment ion mass matching tolerance of 0.4 Da. The variable modifications (N-term)_iTRAQ, Lysine (K)_iTRAQ, N-Acetyl (Protein), Oxidation (M), Pyro-Glu (N-term Q) and MMTS (C) were tolerated. The peptide and protein identification was only considered conclusive if a Mascot Score indicating protein identity with a p-value less than 0.05 was achieved. The peptide ion masses had four decimal places in the raw data, the masses of the corresponding peptide sequence suggestions identified with the Mascot software had two decimal places in the software output. These masses and the peptide sequence suggestions were exported into the Microsoft Excel software. The MALDI plate spot positions where the corresponding MS/MS spectra were acquired were displayed in the Mascot software output and were used to match the masses of the peptide sequence suggestions to the actual ion masses in the raw data. This was necessary to verify the correct assignment of the reporter group ion signals to the corresponding peptide sequence suggestions.

The GPS Explorer software was used to extract the areas of the reporter group ion signals with the masses of 114 and 115 Da from the raw data. The extracted areas

were also exported to Microsoft Excel and corrected if necessary. The peak area ratios of the reporter group ion signals were calculated for each peptide. The abundance ratios of all of the peptides of one protein were used to calculate the mean ratio and standard deviation for the protein and describe the relative abundance of the protein in the two mouse species.

2.14. Lens Protein Preparation for 2-DE

Two lenses each of same sex 14 weeks old *Mus musculus* and *Mus spretus* animals and of same sex ten days old wild type and transgenic homozygous *Gja3* knockout mutant *Mus musculus* C57BL/6J and 129/SvJ mice were ground in a mortar and pestle under liquid nitrogen. The resulting powder was weighed and suspended in 5 times 10^{-3} weight per volume (w/v) buffer containing 9 M urea, 70 mM DTT, 50 mM potassium chloride, 25 mM Tris pH adjusted to 7.3, 3 mM EDTA, 2.9 mM benzamidine, 1 mM PMSF, 2.1 µM leupeptin, 0.1 µM pepstatin, 2% w/v Servalyte 2-4 and 1% w/v CHAPS, i.e. 10 mg were suspended in 50 µl buffer. The suspension was agitated and centrifuged at 100000 g. The supernatant, containing the urea soluble protein fraction was kept and stored at -80°C until use. The protein concentration was determined as described above.

2.15. Two-dimensional Gel Electrophoresis of Lens Proteins

Large scale two-dimensional electrophoresis employing a gel size of 23 x 30 x 0.25 cm combining carrier ampholyte IEF and SDS-PAGE was performed according to a protocol of Klose and Kobalz (Klose and Kobalz, 1995). The 2-DE gels of the *Mus*

musculus and *Mus spretus* lens proteins were produced in the laboratory of Joachim Klose at the Institute of Human Genetics of the Charité in Berlin. The 2-DE gels of the three independent preparations of the *Mus musculus* wild type and *Gja3* knockout C57BL/6J and 129/SvJ lens proteins were produced in the laboratory of Peter Jungblut at the Max Planck Institute for Infection Biology in Berlin. Sixty µg of the urea soluble lens protein fractions were loaded onto the IEF tube gels for silver staining (Jungblut and Seifert, 1990), 300 µg for Coomassie Brilliant (CBB) Blue G250 staining (Doherty et al., 1998) of proteins.

2.16. Computer Supported Protein Spot Matching and Image Analysis

The 2-DE gels of the silver stained lens proteins of the *Mus musculus* C57BL/6J and 129/SvJ strain were scanned at 150 dpi with a Umax Mirage II SE laser scanner with the Magic Scan 4.5 software both from Umax Data Systems (Willich, Germany). The protein spot patterns were matched with the PDQuest 2-DE image analysis software version 7.1.0. from Bio-Rad (Hercules, CA, USA). The protein spot staining intensities were normalized and quantified. The software was used to select significantly variant protein spots. The approved candidates were excised manually from 2-DE gels of the CBB G250 stained lens proteins for in-gel digestion and mass spectrometry.

2.17. In-gel Digestion of Lens Proteins with Trypsin

The proteins were in-gel digested as described by Otto and co-workers (Otto et al., 1996) with some modifications. The procedures were performed on a clean surface under a hood in a basin under water to ensure low Keratin contamination. The protein

spots were manually excised from the 2-DE gels. The proteins were destained in 200 mM NH_4HCO_3 in 50% acetonitrile (ACN) in water for 30 minutes at 37 °C. The proteins were then incubated in 50 mM NH_4HCO_3 in 5% ACN in water for 30 minutes at 37 °C, dried in a Concentrator 5301 from Eppendorf and digested with 2% (w/v) Trypsin in 50 mM NH_4HCO_3 in 5% ACN in water at 37°C overnight.

The resulting peptides were extracted from the gel by passive diffusion in 0.5% trifluoric acid (TFA) in 60 % ACN in water and in 100% ACN. The extracted peptides were dried. The dried peptides were solved in 0.1% TFA in 33% ACN in water for MALDI mass spectrometry or in 0.1% formic acid (FA) in water for ESI mass spectrometry.

2.18. Peptide Sample Cleanup and Desalting with ZipTip® C18 Micro Pipette Tips

ZipTip® pipette tips from Millipore (Billerica, MA, USA) are micro pipette tips with a bed of chromatography media fixed to their ends. ZipTip® pipette tips with C-18 chromatographic media with chromatographic bed volumes of 0.2 and 0.6 µl were used. A ZipTip® was wet with 50% ACN and 0.1 %TFA and equilibrated with 0.1% TFA in water. Peptides in at least 3 µl of solvent volume were loaded onto the chromatographic material in the ZipTip® by holding the tip in the solution and pipetting up and down several times. The bound peptides were washed in aliquoted 0.1% TFA in water and eluted from the chromatographic material by holding the tip in 1 to 5 µl of aliquoted 0.1%TFA in 60% ACN in water and pipetting up and down 5 times. The eluted peptides were diluted to concentrations suitable for mass spectrometry, or dried and resolved for mass spectrometry.

2.19. Dephosphorylation of Peptides with Shrimp Alkaline Phosphatase

Peptides in one half µl of 0.1% TFA in 33% ACN in water were diluted with 5 µl of 50 mM NH_4HCO_3 in 5% ACN in water. One µl of shrimp Alkaline phosphatase from Roche (Mannheim, Germany) with a concentration of approximately 0.5 µg/µl in 50% glycerol containing buffer was added and the reaction stock was agitated and incubated at 37°C for 1 hour. Following incubation the reaction stock was lyophilized and solved in 2.5 µl of 0.1% TFA in 33% ACN in water for mass spectrometry.

2.20. Mass Spectrometry of In-gel Digested Lens Proteins

The peptides were loaded onto a C18 PepMap™ column (LC Packings) with a diameter of 0.075 mm and a length of 150 mm over a desalting column with an UltiMate™ nano HPLC system. The peptides were eluted from the column into an LCQ Deca XP ion trap mass spectrometer at a flow rate of 200 nl per minute with a linear gradient inclination of 0-30% solvent B (0.1% FA in ACN) in solvent A (0.1% FA in water) over a duration of 175 minutes. Mass spectra were acquired throughout chromatographic duration over a mass range of 350 to 2000 Da with a minimum signal intensity of 100000. Three spectra were acquired in series if the minimum signal intensity was met: a spectrum of all ions in the trap, a data dependant zoom spectrum plus and minus 5 Da around the highest intensity ion signal and an MS/MS spectrum of the highest intensity ion signal. Dynamic exclusion was enabled.

Alternatively, the peptides were loaded onto a ZORBAX stable bond C18 column with a diameter of 0.075 mm and a length of 150 mm (Agilent Technologies) with an 1100 series nano HPLC system (Agilent Technologies). The peptides were eluted from

the column into an XCT ion trap mass spectrometer (Agilent Technologies) at a flow rate of 200 nl per minute with a linear gradient inclination of 3-30% solvent B (0.1% FA in ACN) in solvent A (0.1% FA in water) over a duration of 175 minutes. Ion signals were acquired throughout chromatographic duration over a mass range of 200 to 2200 Da. The ions with the three most intense signals were subjected to MS/MS spectral acquisition if their signal intensity exceeded 10000 counts and 2% of baseline intensity. Active Exclusion was enabled.

Alternatively, the peptides were mixed with an equal amount of 4% (w/v) 2, 5-dihydroxybenzoicacid (DHB) (Bruker Daltonics, Billerica, MA, USA) in 0.5% TFA in 33% ACN in water for analysis with a Voyager Elite MALDI mass spectrometer (Perseptive Biosystems, Framingham, MA, USA). Peptide mass fingerprint (PMF) mass spectra were obtained with an accelerating voltage of 20 kV, a grid voltage of 70%, a guide wire voltage of 0.05% a delay of 200 ns and a low mass gate of 500. Internal calibration achieved a mass accuracy of better than 100 ppm.

Alternatively the peptides were mixed with a four times amount of 2% (w/v) α-cyano-4-hydroxycinamic acid (CHCA) in 0.3% TFA in 50% ACN in water for analysis with a 4700 Proteomics Analyzer MALDI mass spectrometer. PMF mass spectra were acquired with an accelerating voltage of 20 kV, 1200 laser shots per MALDI plate position and a low mass gate of 800. Tandem mass spectrometry (MS/MS) mass spectra were acquired without collision gas and with at least 2500 laser shots and at most 10000 laser shots per MALDI plate position and under stop conditions after the accumulated spectrum reached a signal to noise threshold (S/N) of 40 and if at least 15 peaks in the spectrum surpassed the S/N.

2.21. Identification of Lens Proteins

The peptide and peptide fragment ion masses recorded in mass spectra were used to search the NCBI protein sequence database with the Mascot mass spectrometry software suite version 2.0 (Matrix Science). The software matches ion masses to protein primary structure segment masses to identify proteins. The set of ion masses from a mass spectrum used as an input for the software (peak list) from PMF mass spectra acquired with the Voyager Elite mass spectrometer were generated by manual selection of ion signals with the Grams/386^{TM} software version 3.03 (Galactic Industries, Salem, NH, USA).

The peak lists from PMF mass spectra acquired with the 4700 Proteomics Explorer mass spectrometer were generated automatically with the 4000 Series ExplorerTM software version 3.0 with a minimum S/N of 10 or higher depending on spectral quality and a maximum number of 30 peaks. The peak lists from MS/MS mass spectra were generated depending on spectral quality with a minimum S/N of 2 or higher, a minimum peak area of 5 and a maximum number of 50 peaks.

The peak lists from mass spectra acquired with the XCT mass spectrometer were generated automatically with the DataAnalysis software version 2.0 (Agilent Technologies) with a minimum ion signal intensity threshold of $100x10^3$ counts.

The database searches were performed with the software species controller set to *Mus musculus*. A maximum peptide mass accuracy error of 100 ppm was tolerated in searches with the data from the Voyager Elite mass spectrometer with one exception where an error of 150 ppm was tolerated. A maximum peptide mass accuracy error of 60 ppm and peptide fragment mass accuracy error of 0.4 Da was tolerated for searches with the data from the 4700 Proteomics Explorer mass spectrometer. A maximum peptide mass and peptide fragment mass accuracy error of 0.4 Da was tolerated in searches

with the data from the XCT and the LCQ Deca XP mass spectrometers. A maximum of one trypsin missed cleavage was tolerated.

The protein identification was only considered conclusive if a Mascot Score indicating protein identity with a p-value less than 0.05 was achieved. Two MS/MS spectra which contained four consecutive mass peaks indicative of three amino acid residues in series in tryptic segments of the proteins primary structure had to be achieved for database searches with MS/MS data with one exception.

2.22. Detection of Modifications and Structure and Function Analysis of Lens Proteins

The PMF mass spectra of protein digests were comparatively analyzed with the MS-Screener© software developed in-house. Protein primary structures were *in silico* analyzed with the GPMAW version 5.02 software from Lighthouse data (Hanstholm, Denmark) as well as numerous software packages available on the ExPASy Proteomics Server (Geneva, Switzerland) and the NCBI server (Bethesda, MD, USA) in the world wide web.

The MS/MS spectra were evaluated to definitively characterize primary structure segments and identify modifications to database sequences in the lens proteins. The sequence tags found by evaluation of the MS/MS spectra were used to search the NCBI Protein sequence database, the Swiss-Prot Protein sequence database and the MSDB Protein sequence database with the Mascot software.

Ab initio protein tertiary structure predictions were performed with the HMMSTR/Rosetta server (Bystroff and Shao, 2002) available on the ExPASy Proteomics Server in the world wide web. Molecular modeling and visualization of PDB files was

done with the Swiss-PDB Viewer available on the ExPASy Proteomics Server in the world wide web.

An interaction network of the lens proteins was constructed with the Osprey Network Visualization System version 1.2.0 from the Tyers Lab available in the world wide web under the URL http://biodata.mshri.on.ca/osprey (Breitkreutz et al., 2003).

2.23. Results storage

The results were entered and are stored in the publicly available Proteome 2-D PAGE Database from the Max Planck Institute for Infection Biology (Berlin, Germany) under the URL http://www.mpiib-berlin.mpg.de/2D-PAGE/.

3. Results

3.1. Analysis of the Variability in the Lens Proteins in the *Mus musculus* and *Mus spretus* species with the Cleavable ICAT Reagent

The first of two independently prepared stocks of the mouse eye lens proteins were labeled with the ICAT reagent and analyzed with the LC/ESI-MS/MS system from LCPackings and Thermo Finnigan. The protein concentrations were 37.2 µg per µl in the urea soluble protein fraction of the *Mus musculus* and 43.3 µg per µl in the *Mus spretus* lenses. The acquired total ion current is shown in Figure 4.

The software extracted 1072 scans from the acquired total 8410 mass spectra in the raw data. The database search produced 757 protein identification suggestions. Multiple independent peptide sequence suggestions and therefore multiple protein identification suggestions were often assigned to one scan. The 757 protein identification suggestions were filtered according to criteria reflecting the quality of the scans and the corresponding peptide sequence suggestions. The top and second ranked peptide sequence suggestions for a given scan were retained in all cases. The scans were reevaluated manually and poor or ambiguous peptide sequence suggestions and the corresponding scans discarded. This resulted in 43 peptide sequence suggestions assigned to 54 high quality scans. (Supplementary Table 1, Appendix).

Thirty scans corresponding to twenty three peptide sequence suggestions showed labeling with either the light or heavy form of the cleavable ICAT reagent. Eighteen protein identification suggestions were quantified of which some were different database entries for the same protein. Ultimately, 6 proteins, Beta A2 crystallin, Beta B2 crystallin, Gamma B crystallin, Gamma C crystallin, Fatty acid binding protein and UDP-

glucuronosyltransferase were identified and quantified, some based on a single peptide sequence suggestion assigned to a single scan (Supplementary Table 1, Appendix).

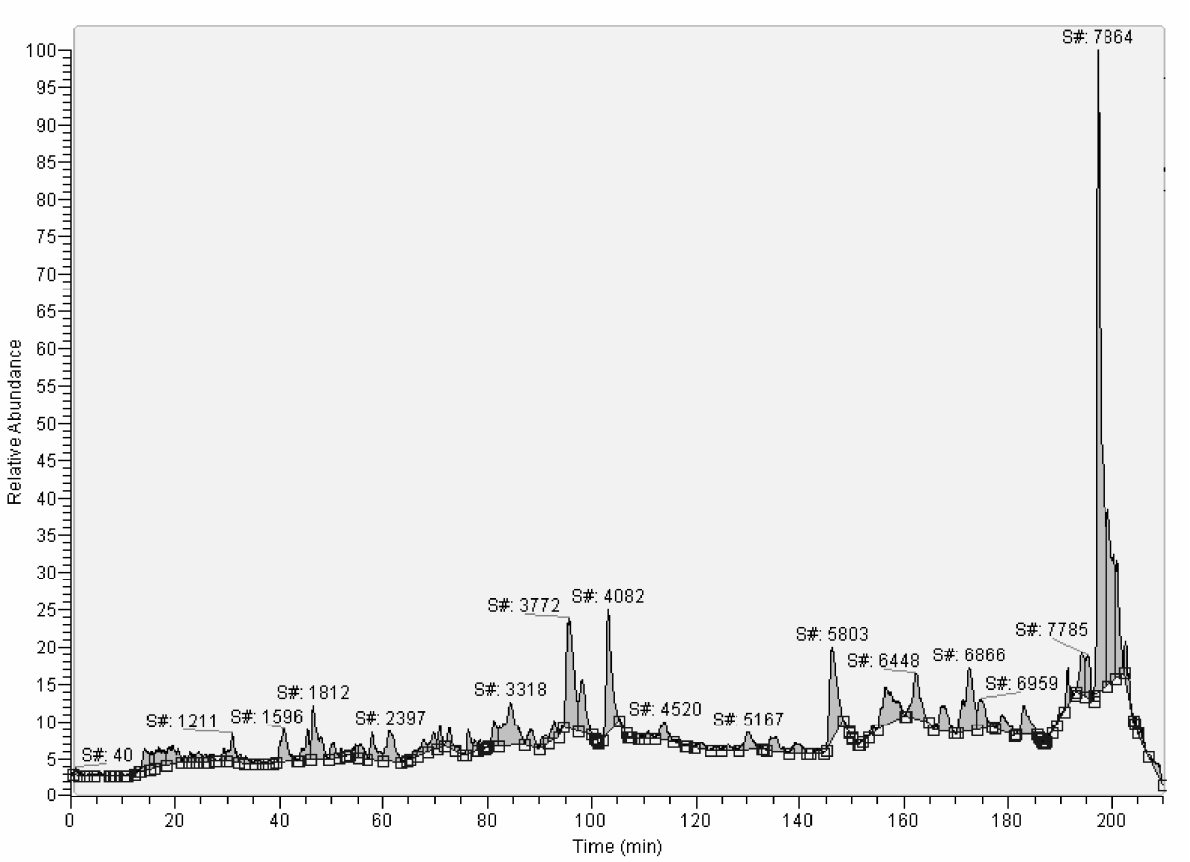


Figure 4. The total ion current acquired in the analysis of the first of two independently prepared lens protein stocks labeled with the ICAT reagent with the LC/ESI-MS/MS system from LCPackings and Thermo Finnigan.

The second of the independently prepared stocks was labeled with the ICAT reagent and analyzed separately 5 times with the Agilent Technologies System. The protein concentrations in the urea soluble lens protein fractions were 30.3 µg per µl in the *Mus musculus* and 29.5 µg per µl in the *Mus spretus* species. The acquired total ion current of analysis run 1 is shown in Figure 5.

The software extracted between 500 and 1000 scans from around 10000 raw MS/MS spectra each in all 5 separate analysis of the reaction stock. Stringent processing criteria as suggested in the softwares user guide were applied prior to the database

search. All scans were additionally manually validated and poor or ambiguous scans were discarded. For analysis 1, the search produced 25 peptide sequence suggestions assigned to 34 high quality scans, all labeled with either the light or heavy form of the cleavable ICAT reagent. Thirteen proteins (13) were identified and quantified, in some cases based on a single peptide sequence suggestion and a single scan (Supplementary Table 2, Appendix).

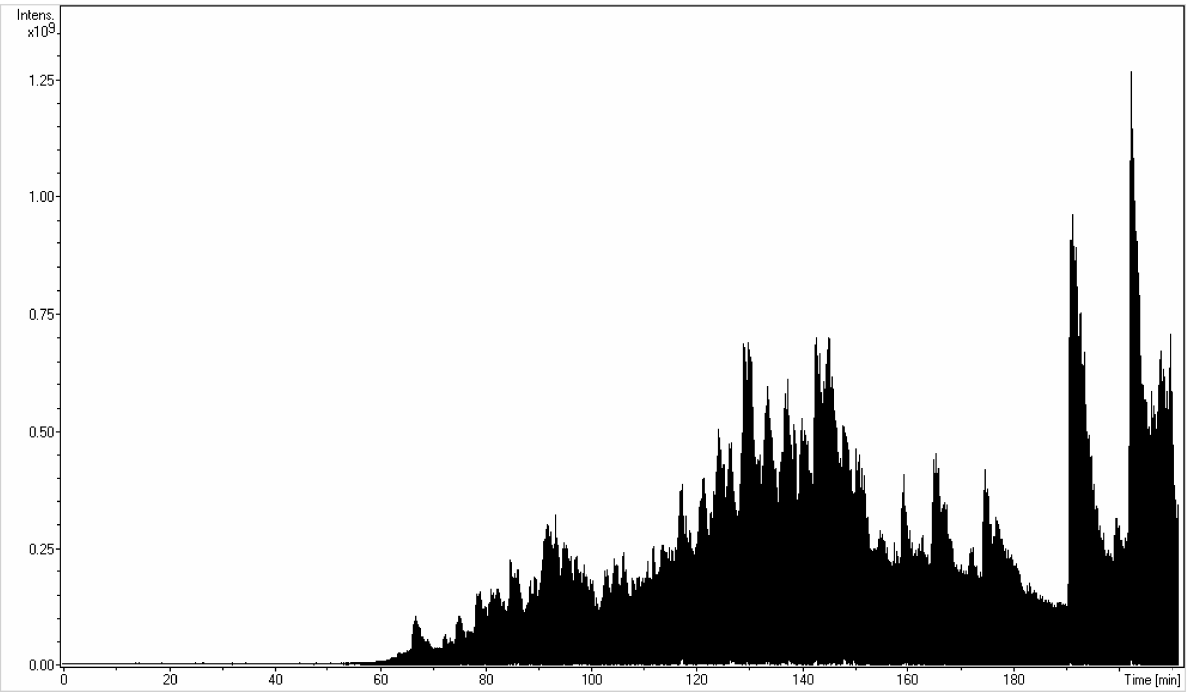


Figure 5. The total ion current acquired in the first of five analysis of the second of two independently prepared lens protein stocks labeled with the ICAT reagent with the LC/ESI-MS/MS system from Agilent Technologies.

Analysis 2 through 5 of the same reaction stock produced 94 high quality scans in total. Twenty nine peptide sequence suggestions were unambiguously assigned to these scans. All of the scans contained ion signals indicative of either the light or heavy form of the cleavable ICAT reagent. Thirteen (13) proteins were identified and quantified in total in some cases based on a single peptide sequence suggestion and a single scan. (Supplementary Table 3, Appendix).

The variability in the lens proteins in the two species is listed for the sixteen major crystallins ubiquitous to mammals for both reaction stocks (Table 1). The abundance ratios of the peptide sequence suggestions were summed and averaged to arrive at the final protein abundance ratios. The investigation of the first stock with the Thermo Finnigan system identified and quantified four (4) crystallin proteins, the investigation of the second stock with the Agilent Technologies system identified and quantified seven (7) crystallin proteins. The proteins Alpha A and Alpha A insert crystallin, Beta A1 and Beta A3 crystallin and the Gamma crystallins A and D, E and F were not discriminated and unambiguously quantified because of primary structure sequence homologies in the crystallin protein superfamily. The protein abundance ratios of the Beta A2, Beta A4, Beta B1 and Beta B2 and Gamma B and Gamma C crystallin proteins were accurately determined and are around 1. The variability of these proteins in the urea soluble fractions of the lenses of the *Mus musculus* and *Mus spretus* species therefore seems small.

3.2. Analysis of the Variability in the Lens Proteins in the *Mus musculus* and *Mus spretus* species with the iTRAQ Reagent

The iTRAQ reagent labeled protein standard and the lens proteins were analyzed using an UltiMate™ HPLC System, a Probot™ micro fraction collector and a 4700 Proteomics Analyzer MALDI/TOF/TOF mass spectrometer. The protein concentrations in the urea soluble protein fraction were 30.4 µg per µl in the *Mus musculus* and 23.8 µg per µl in the *Mus spretus* lenses.

Eight hundred and ninety three (893) MS/MS spectra were acquired in the analysis of the protein standard. All five components were identified with high Mascot

Table 1. Summary of protein variability in the Crystallin protein superfamily in the urea soluble lens protein fraction of the *Mus musculus* and *Mus spretus* species at 14 weeks of age determined with the ICAT reagent.

Crystallin Protein	Cysteine Content			Analysis with Thermo Finnigan System			Analysis with Agilent Technologies System		
	Cys Residues	Tryptic Cys Peptides[a] Total	Tryptic Cys Peptides[a] Discriminate	Tryptic Cys Peptides Detected[b]	Number of Scans	Protein Abundance Ratio[c] *Mus musculus*/ *Mus spretus*	Tryptic Cys Peptides Detected[b]	Number of Scans	Protein Abundance Ratio[c] *Mus musculus*/ *Mus spretus*
Alpha A crystallin 62201857	1	1	0	1	5	NA	1	14	NA
Alpha A insert crystalline 117332	1	1	0	1	5	NA	1	14	NA
Alpha B crystalline 6753530	0	0	0	0	0	NA	0	0	NA
Beta A3/A1 crystallin[d] 20304089	8	6	6	1	7	1.57 0.03	2	26	1 0.29
Beta A2 crystallin 10946978	7	5	5	2	7	1.18 0.05	2	14	1.18 0.14
Beta A4 crystallin 10946672	4	3	3	0	0	NA	2	6	0.92 0.53
Beta B1 crystallin 12963789	4	4	4	0	0	NA	2	11	1.22 0.10
Beta B2 crystallin 50401872	2	2	2	1	2	1.72 0	2	8	0.1 12.99
Beta B3 crystallin 10946674	3	1	1	0	0	NA	1	4	0.21 7.53
Gamma A crystallin 6724317	9	4	2	1[0]	2	NA	1[0]	1	NA
Gamma B crystallin 42733606	9	4	2	2[1]	1	1.45 0	2[1]	5	0.03 22.85
Gamma C crystallin 6681037	9	4	2	2[1]	1	1.23 0	1[0]	1	NA
Gamma D crystallin 6681039	7	3	2	0	0	NA	0	0	NA
Gamma E crystallin 34978370	6	3	0	2[0]	3	NA	2[0]	12	NA
Gamma F crystallin 21746155	6	4	1	2[0]	3	NA	2[0]	12	NA

[a] The number of tryptic cysteine containing peptides in the crystallin proteins was calculated considering complete enzymatic cleavage after every arginine and lysine residue and a molecular weight appropriate to the ion traps optimal detection range between masses 800 and 3000 Da.
[b] The number in brackets indicates the number of discriminate tryptic cysteine containing peptides detected for each protein.
[c] The mean and standard deviations of the abundance ratios of the peptide sequence suggestions identified in the analysis are listed for each protein.
[d] The proteins could not be discriminated because of protein sequence homology.
NA, not applicable indicates the protein could not be quantified.

scores. Bovine serum albumine had the highest score of 1945 followed by Conalbumin with a score of 1476, Myoglobin with a score of 537, Carbonic anhydrase with a score of 365 and Cytochrome C with a score of 270. The results for Bovine serum albumine were examined in detail. Sixty six peptide sequence suggestions were assigned to 66 MS/MS spectra. Several peptide sequence suggestions were assigned to the same MS/MS spectrum. MS/MS spectra with an ion score below 30 were inspected manually. Fifteen MS/MS spectra and the peptide sequence suggestions assigned to them were of insufficient quality and were discarded. Examination of the remaining peptide sequence suggestions showed that not all free amino groups on all peptides were labeled with the iTRAQ reagent and that some peptides showed no labeling with the iTRAQ reagent at all. Of the 51 peptide sequence suggestions based on high quality MS/MS spectra 24 (47%) had all amino groups labeled, 21 (41%) had some but not all amino groups labeled and 6 (12%) had no amino groups labeled. The 51 peptide sequence suggestions corresponded to 32 distinct peptide sequences. Nine of these had all amino groups constitutively labeled, 21 had only partially labeled amino groups and 2 were entirely unlabeled (Figure 6).

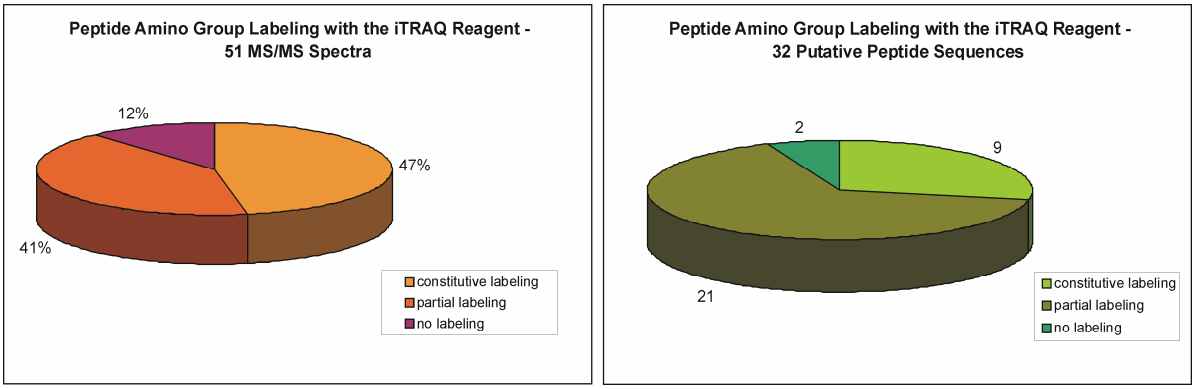


Figure 6. Frequency of amino group labeling of peptides of Bovine serum albumine with the iTRAQ reagent.

The chemical reactions involved in labeling peptide amino groups with the iTRAQ reagent are hampered by urea. The RapiGest^{TM} SF buffer from Waters, a buffer with

similar properties but lacking urea and compatible with the iTRAQ reagent chemistry was suggested by Applied Biosystems as an alternative. Two-dimensional electrophoresis gels of proteins from lenses of 14 weeks old *Mus musculus* and *Mus spretus* mice solved in this buffer were compared to 2-DE gels of proteins solved in urea to check the buffers overall performance. The 2-DE protein spot patterns of the 16 protein spots containing the highly abundant unmodified translation products of the major crystallin genes ubiquitous to mammals were highly similar to each other and to standard lens protein spot patterns (Jungblut et al., 1998; Ueda et al., 2002; Hoehenwarter et al., 2005, 2006b) in spot positions and staining intensity, so lens protein solubilization was satisfactory.

The analysis produced 2472 MS/MS spectra. Seventy five protein identification suggestions were produced with a significant Mascot score. Not all amino groups of all peptide sequence suggestions were labeled with the iTRAQ reagent. The peptide sequence suggestions and the corresponding MS/MS spectra with ion scores below 30 were inspected manually and discarded if they were poor. The peptide sequence suggestions with partial labeling of amino groups and the corresponding MS/MS spectra were also discarded. Two hundred and sixty five (265) peptide sequence suggestions with all amino groups labeled were unambiguously assigned to an equal number of high quality MS/MS spectra. Two hundred and four (204) peptide sequence suggestions were not redundant indicating that several peptide sequence suggestions were assigned to more than 1 MS/MS spectrum as would be the case if a prominent ion signal is detected and acquired again outside the dynamic exclusion range. The peptide sequence suggestions produced 36 protein identification suggestions. Due to extensive homology in crystallin protein primary structure Alpha A and Alpha A insert crystallin, Beta A1 and Beta A3 crystallin and Gamma D crystallin could not be discriminated. Thirty three (33) proteins were identified and accurately quantified including all sixteen crystallins ubiquitous to mammals with the exceptions noted above.

Results

This is a comparatively rapid, extensive realization of the variability in the urea soluble protein fraction of the lenses of the *Mus musculus* and *Mus spretus* species. It includes not only the crystallins but many of the predominant non crystallin proteins in the lens fiber cells. The major elements of the fiber cell cytoskeleton including both obligate beaded filament assembly partners CP49 (Phakinin) and CP94 (Filensin), Vimentin, Actin and Alpha tubulin and the Major intrinsic membrane protein MIP26; the molecular chaperones HSP27/25 and Hsp70; the Fatty acid binding protein FABP and a number of metabolic proteins were all identified and quantified. As in the ICAT based investigation, the protein abundance ratios are 1 in most cases. Indeed, only four proteins (Gamma F crystallin, Galectin, Glutathione-S-transferase, Ubiquitin/60S ribosomal fusion protein) showed a two-fold or more increase or decrease in abundance suggesting the protein variability is hardly significant. Some of the proteins were identified with one high quality MS/MS spectrum only (Table 2).

Six peptide sequence suggestions with all amino groups labeled were unambiguously assigned to high quality MS/MS spectra with ion signals with a mass of 114 Da and lacking ion signals with a mass of 115 Da. This suggests that the peptides were present in the lenses of the *Mus musculus* species and that the same peptides were not present or not labeled in the lenses of the *Mus spretus* species due to absence of the protein, polymorphism or modification. To determine whether the peptides were not labeled by accident the results were checked for corresponding unlabeled or partially labeled peptide sequence suggestions lacking the iTRAQ reagent reporter group ion mass of 115 Da in the MS/MS spectra. This was not the case for five of the six peptide sequence suggestions. Four possible protein sequence polymorphisms in the lenses of the *Mus musculus* and *Mus spretus* species were detected in four proteins, Alpha B crystallin, Gamma B crystallin, Gamma F crystallin and Gamma S crystallin (Table 3, Figure 7).

Table 2. Summary of protein variability in the urea soluble lens protein fraction of the *Mus musculus* and *Mus spretus* species at 14 weeks of age determined with the iTRAQ reagent.

Protein Number	Protein Name	Accession Number	Mascot Score[a]	Matched Peptides		Protein Abundance Ratio[b] *Mus musculus/Mus spretus*	
				Total	Valid and Discriminate	MEAN	STADD
1	Alpha A/A insert crystallin	30794510[c]	627	31	11	1.14	0.2
2	Alpha B crystallin	6753530	600	28	12[d]	1.51	0.46
3	Beta A3/A1 crystallin[e]	20304089	1095	34	11	1.14	0.25
4	Beta A2 crystallin	10946978	525	18	5	1.02	0.22
5	Beta A4 crystallin	10946672	567	15	7	1.22	0.2
6	Beta B1 crystallin	12963789	1003	36	13	0.92	0.27
7	Beta B2 crystallin	50401872	1472	52	20	1.23	0.25
8	Beta B3 crystallin	10946674	1078	33	13	1.02	0.19
9	Gamma A crystallin	6724317	550	25	5	1.08	0.51
10	Gamma B crystallin	42733606	761	46	8[d]	1.2	0.3
11	Gamma C crystallin	6681037	646	35	4	1.46	0.24
12	Gamma D crystallin[e]	6681039	404	25	0	0	0
13	Gamma E crystallin	31982854[f]	598	42	1	1.78	0
14	Gamma F crystallin	21746155	677	41	5[d]	2.07	0.49
16	Gamma N crystallin	23346485	60	3	1	0.77	0
15	Gamma S crystallin	6753532	557	22	10[d]	1.17	0.2
17	Cytoplasmic beta Actin	6671509	478	20	5	1.2	0.08
18	CP49, Bfsp2	51765089[g]	356	15	8	1.14	0.19
19	HSP27/25	7305173[g]	344	12	6	0.98	0.14
20	E-FABP	6754450	300	15	6	0.86	0.09
21	RIKEN cDNA 4732495G21	30425250	294	10	1	1.08	0
22	Vimentin	31982755	257	8	5	0.94	0.26
23	HSP70	31560686[f]	197	8	3	1.48	0.23
24	Peptidylprolyl isomerase A	6679439[f]	171	7	5	1.12	0.07
25	Acyl-CoA binding protein	6681137	100	5	2	1.96	0.12
26	Alpha 6 tubulin	6678469	92	3	2	1.45	0.01
27	MIP eye lens fiber	31543250	87	4	4	1.31	0.27
28	CP94, Bfsp1	6753188[g]	82	3	1	1.21	0
29	Glutathione-S-transferase	32401425	79	3	2	0.35	0.25
30	Ubiquitin/60S ribosomal fusion protein	9845265	72	2	2	2.28	0.4
31	Obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	859441	70	10	1	1.01	0
32	Galectin	15961	70	5	2	3.43	0
33	Bassoon, presynaptic cytomatrix protein	6671648	58	9	1	1.29	0
34	Gelsolin	28916693	55	4	2	0.93	0.18
35	Hypothetical XP_488995	51707653	53	4	1	1.12	0
36	Acetyl-Coenzym A carboxylase	269786	51	6	1	1.17	0

[a] A Mascot score of 31 or more indicates peptide sequence and protein identity with a P-value less than 0.05.
[b] The mean and standard deviations of the abundance ratios of the peptide sequence suggestions identified in the analysis are listed for each protein.
[c] Alpha A insert crystallin was discriminated, therefore its accession number is listed, Alpha A crystallin is also present.
[d] A putative protein sequence polymorphism or modification in this protein in *Mus musculus* and *Mus spretus* was identified by iTRAQ labeling and a quenched reporter ion signal in the MS/MS spectrum of the respective iTRAQ labeled peptide.
[e] The proteins were not discriminated.
[f] The protein sequence with this accession number is different from another protein sequence with another accession number for the same protein detected in the 2-DE analysis of the urea soluble lens proteins of the wild type C57BL/6J mice listed in Supplementary Table 4. The differences are covered by mass spectrometric data indicating a genuine polymorphism or other event.
[g] The protein sequence with this accession number is different from another protein sequence with another accession number for the same protein detected in the 2-DE analysis of the urea soluble lens proteins of wild type C57BL/6J mice listed in Supplementary Table 4. The polymorphism site is not covered by mass spectrometric data indicating that the sequence discrepancy could be due to software processing and scoring.

Table 3. Putative polymorphisms or post translational modifications in the urea soluble lens protein fraction of 14 weeks old *Mus musculus* and *Mus spretus* mice detected by the iTRAQ strategy

Protein	Sequence Coverage[a]	Polymorphic Peptide Sequence Suggestion	Mascot Score[b]
Alpha B crystallin	49%	EEKPAVAAAPK	66
Gamma B crystallin	61%	SCCLIPQHSGTYR	63
		SCCLIPQHSGTYR	76
Gamma F crystallin	52%	SCHLIPHSTSHR	47
Gamma S crystallin	40%	AVHLSSGGQAK	81

[a] The sequence coverage of the protein is based on all detected peptides listed in Table 2.
[b] A Mascot score of 31 or more indicates peptide sequence and protein identity with a P-value less than 0.05.

Peptide Sequence Suggestions	Area *Mus musculus*	Area *Mus spretus*	Ratio *Mus musculus*/ *Mus spretus*
EIHSCK + (N-term)_iTRAQ; Lysine(K)_iTRAQ	59753.4	66726.5	0.90
WMGLNDR + (N-term)_iTRAQ	1299132.9	988871.0	1.31
EIHSCK + (N-term)_iTRAQ; Lysine(K)_iTRAQ; MMTS (C)	1050398.0	1303734.5	0.81
WMGLNDR + Oxidation (M); (N-term)_iTRAQ	345214.2	267286.3	1.29
IQVFEK + (N-term)_iTRAQ; Lysine(K)_iTRAQ	1213181.5	0.0	
QYLLDK + (N-term)_iTRAQ; Lysine(K)_iTRAQ	2261642.3	2015925.1	1.12
ISFYEDR + (N-term)_iTRAQ	1667217.3	1415879.9	1.18
AVHLSSGGQAK + (N-term)_iTRAQ; Lysine(K)_iTRAQ	534073.0	0.0	
VEGGTWAVYER + (N-term)_iTRAQ	1281903.8	1273640.1	1.01
KPVDWGAASPAIQSFR + (N-term)_iTRAQ; Lysine(K)_iTRAQ	1429567.3	1181441.4	1.21

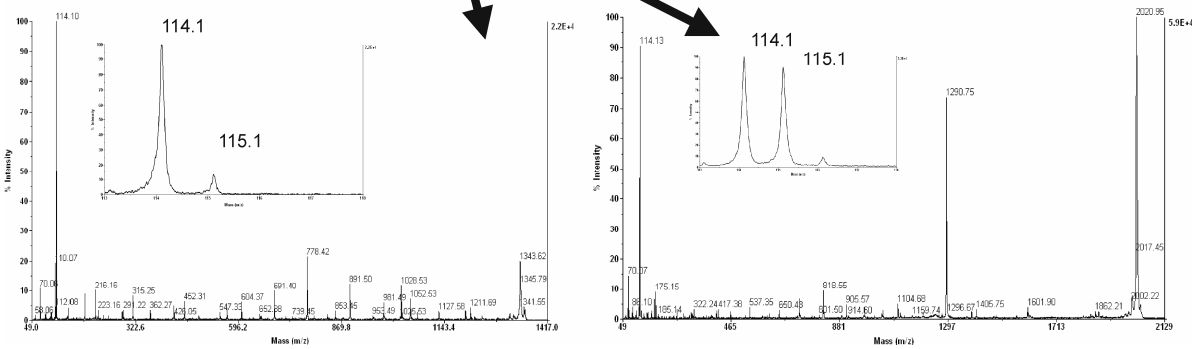


Figure 7. Detection of polymorphisms in Gamma S crystallin with the iTRAQ reagent. All peptide sequence suggestions assigned to Gamma S crystallin constitutively labeled with the iTRAQ reagent and the areas of the ion signals of the iTRAQ reagent reporter groups are shown. Two MS/MS spectra, one with the ion signals of both reporter groups and one lacking an ion signal of the reporter group with the mass of 115 Da corresponding to a possible polymorphic peptide are also shown.

3.3. Analysis of the Variability in the Lens Proteins in the *Mus musculus* C57BL/6J and 129/SvJ Strain and the Connection to Cataract Onset and Severity with 2-DE, Quantitative Image Analysis and MS

3.3.1. Proteomics Achieves a Comprehensive Characterization of Lens Proteins and Strain Specific Lens Protein Variability

The lens proteins of the wild types and homozygous *Gja3* knockout mutants of the *Mus musculus* C57BL/6J and 129/SvJ strain were studied with 2-DE for protein species

separation and visualization and mass spectrometry for protein identification and protein species characterization (Figure 8). The mutant 129/SvJ mice had developed the cataract phenotype while the lenses of the mutant C57BL/6J mice were unaffected at ten days of age. Otherwise the development of the animals was normal.

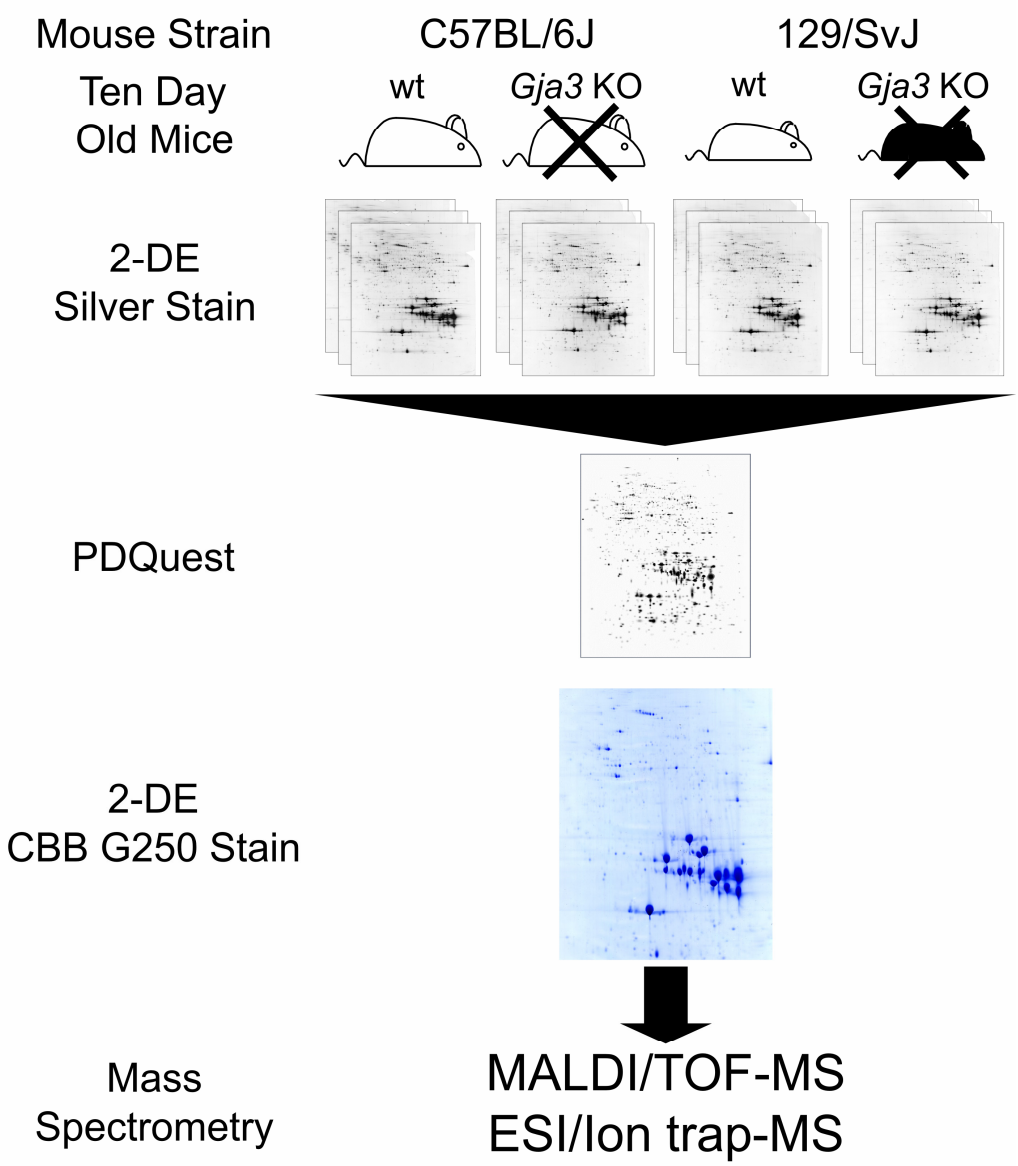


Figure 8. Schematic of the experimental procedures. The urea soluble protein fraction of two lenses each of wild type and homozygous *Gja3* knockout mutant C57BL/6J and 129/SvJ mice was independently prepared three times. The samples were applied to large scale ampholyte carrier based 2-DE. The separated protein species were silver stained and the protein spot patterns scanned and incorporated into a matchset in the PDQuest software. The protein spot patterns were matched to each other and the protein spot staining intensities were quantified and normalized. The protein spots were selected for excision and subsequent mass spectrometric analysis. The sample preparations were applied to large scale carrier ampholyte based 2-DE and the separated protein species were CBB G250 stained. The selected protein spots were excised from these 2-DE gels and analyzed by mass spectrometry.

Results

The urea soluble protein fraction of lenses of ten day old C57BL/6J and 129/SvJ wild type and *Gja3* knockout mutant mice was independently prepared three times and applied to large scale ampholyte carrier based 2-DE (Klose and Kobalz, 1995). The proteins in the 2-DE gels were silver stained and the resulting 2-DE protein spot patterns were scanned and analyzed with the PDQuest image analysis software. The lens protein fractions were separated into around 1000 protein spots each. All twelve protein spot patterns were matched to each other thereby clearly and reproducibly discriminating not only unique protein species from the individual lens protein fractions but also differential protein species abundance.

The staining intensity of a protein spot on a 2-DE gel is proportional to the abundance of the protein species in the spot. The protein spot staining intensities were first quantified and normalized. The correlation coefficient determined by the PDQuest software was 87% for the protein spot patterns of the three independent preparations of the C57BL/6J wild type, 92% for the C57BL/6J *Gja3* KO, 90% for the 129/SvJ wild type and 90% for the 129/SvJ *Gja3* KO lens proteins. This demonstrates the exceptional reproducibility of the technique and its suitability for a comparative investigation of protein abundance.

The proteins in many congruent spots in the protein spot patterns of some or all of the lens protein fractions were identified based on mass spectrometric analysis of a protein spot from 2-DE gels of the wild type C57BL/6J mice. The molecular weight and pI values of the identified proteins were calculated from their positions in the protein spot patterns in relation to their theoretical molecular weight and pI values (Aksu et al., 2002).

All protein spots in the protein spot patterns of the strain wild types that had a 2-fold or greater staining intensity variation and that were present on 2-DE gels of two of the three independent preparations and that met good 2-DE protein spot criteria in a

visual examination i.e. clearly defined spot borders and little to no streaking or smearing were considered significant. These were 94 protein spots.

Sixty seven of these 94 protein species were also visible when stained with Coomassie Brilliant Blue G250 (CBB G250). This dye is less sensitive than silver staining but is compatible with mass spectrometry. All of the variant protein species were characterized with mass spectrometry except three where no conclusive results were obtained and one where the protein staining intensity variation exceeded 100% among the three independent preparations of the same sample. The mean staining intensities of the protein species were calculated and normalized to the C57BL/6J strain wild type and are listed (Table 4) The variant protein species were classified into arbitrary taxa refelecting their function (Figure 9). All of the characterized protein species are central to the major mechanisms of healthy lens and cataract development.

A substantial amount of protein species was characterized in addition to the variant protein species described above. In total, this is the largest and most comprehensive investigation of mouse lens proteins to date. The protein spot patterns with the identified protein spots marked by a mouse sensitive cross linked to an information page can be seen in the publicly available Proteome 2-D PAGE Database of the Max Planck Institute for Infection Biology (Berlin, Germany) under the URL http://www.mpiib-berlin.mpg.de/2D-PAGE/. All of the protein species characterized by mass spectrometry in the urea soluble lens protein fraction of the C57BL/6J strain wild type are listed (Supplementary Table 4, Appendix).

Table 4. Summary of protein variability in the urea soluble lens protein fraction of the *Mus musculus* C57BL/6J and 129/SvJ strains at ten days of age

Increased in the C57BL/6J Wild Type					Mw and pI		Relative Protein Abundance[a]			
Protein Spot Denominator	Protein Name	Accession Number NCBI	Mascot Score[b]	Protein Identification Method	Theoretical	2-DE	C57BL/6J wt	C57BL/6J *Gja3* KO	129/SvJ wt	129/SvJ *Gja3* KO
SSP 1422	CP49, Phakinin	50400333	193	MALDI/TOF/TOF-PMF	45.71 / 5.2	43.30 / 5.1	1	0.3	0	0
SSP 1528	CP49, Phakinin	50400333	94	MALDI/TOF/TOF-PMF	45.71 / 5.2	49.21 / 5.1	1	0.4	0	0
SSP 1531	Gamma-actin	74187644	98	MALDI/TOF/TOF-PMF	41.71 / 5.2	46.67 / 5.2	1	1.1	0	0
SSP 1536	CP49, Phakinin	50400333	78	MALDI/TOF/TOF-PMF	45.71 / 5.2	45.44 / 5.0	1	0.2	0	0
SSP 1537	CP49, Phakinin	50400333	66	MALDI/TOF/TOF-PMF	45.71 / 5.2	45.53 / 4.9	1	0	0	0
SSP 1538	CP49, Phakinin	50400333	129	MALDI/TOF/TOF-PMF	45.71 / 5.2	49.44 / 5.2	1	0.7	0	0
SSP 2315	HSP27/25	547679	126	MALDI/TOF-PMF	23.01 / 6.1	29.17 / 5.4	1	0.8	0.3	0.2
SSP 2521	CP49, Phakinin	50400333	230	MALDI/TOF/TOF-PMF	45.71 / 5.2	49.51 / 5.4	1	1.1	0	0
SSP 2522	CP49, Phakinin	50400333	177	MALDI/TOF/TOF-PMF	45.71 / 5.2	49.45 / 5.2	1	0.8	0	0
SSP 2523	CP49, Phakinin	50400333	170	MALDI/TOF/TOF-PMF	45.71 / 5.2	49.41 / 5.3	1	0.8	0	0
SSP 2527	CP49, Phakinin	50400333	201	MALDI/TOF/TOF-PMF	45.71 / 5.2	47.07 / 5.2	1	1.1	0	0
SSP 2617	CP49, Phakinin	50400333	111	MALDI/TOF/TOF-PMF	45.71 / 5.2	51.22 / 5.4	1	0.7	0	0
SSP 3101	Alpha A crystallin	62001857	103	MALDI/TOF/TOF-PMF	19.78 / 5.8	19.42 / 5.5	1	0.4	0.4	0.3
SSP 3313	HSP27/25	547679	186	MALDI/TOF-PMF	23.01 / 6.1	29.24 / 5.7	1	0.9	0.2	0.2
SSP 3617	CP94, Filensin	26343715	80	MALDI/TOF/TOF-PMF	71.46 / 5.6	50.71 / 5.7	1	0.7	0.3	0.2
SSP 3804	DnaK-type molecular chaperone Hsc70 homolog	74205924	67	MALDI/TOF/TOF-PMF	63.92 / 6.1	61.91 / 5.5	1	1.1	0.1	0.1
SSP 3827	CP94, Filensin	26343715	70	MALDI/TOF/TOF-PMF	71.46 / 5.6	67.97 / 5.8	1	1.2	0.2	0.2
SSP 4316	HSP27/25	547679	138	MALDI/TOF/TOF-PMF	23.01 / 6.1	29.30 / 6.1	1	0.9	0.3	0.3
SSP 4326	CP94, Filensin	26343715	250	MALDI/TOF/TOF-MS/MS	71.46 / 5.6	39.86 / 5.8	1	0.2	0	0
SSP 4329	Endoplasmatic reticulum protein ERp29, precursor	16877776	87	MALDI/TOF/TOF-MS/MS	28.81 / 5.9	31.60 / 5.9	1	1.2	0	0
SSP 4501	CP94, Filensin	26343715	139	MALDI/TOF/TOF-PMF	71.46 / 5.6	50.57 / 5.8	1	0.9	0.2	0.2
SSP 4806	CP94, Filensin	26343715	77	MALDI/TOF/TOF-PMF	71.46 / 5.6	67.48 / 5.9	1	1.2	0.2	0.2
SSP 4811	CP94, Filensin	26343715	95	MALDI/TOF/TOF-PMF	71.46 / 5.6	67.63 / 5.9	1	1.2	0.3	0.3
SSP 4812	CP94, Filensin	26343715	143	MALDI/TOF/TOF-PMF	71.46 / 5.6	67.40 / 5.9	1	0.9	0.4	0.4
SSP 4818	CP94, Filensin	26343715	195	MALDI/TOF/TOF-MS/MS	71.46 / 5.6	67.36 / 6.0	1	1.3	0.2	0.2
SSP 5522	Beta B1 crystallin	12963789	74	MALDI/TOF/TOF-PMF	27.99 / 6.8	48.30 / 6.2	1	0.2	0	0
SSP 6125	Gamma B crystallin	42733606	132	MALDI/TOF/TOF-MS/MS	21.07 / 6.9	23.96 / 6.7	1	0.5	0.4	0.4
SSP 6132	Gamma N crystallin	23346485	118	MALDI/TOF/TOF-PMF	21.39 / 6.3	23.38 / 6.6	1	1.1	0	0
SSP 6325	Annexin A1	70912321	187	MALDI/TOF-PMF	38.71 / 7.7	40.06 / 6.5	1	1.4	0	0
SSP 6717	Chaperonin subunit 6A	74204595	195	MALDI/TOF/TOF-MS/MS	58.01 / 6.5	58.46 / 6.5	1	1.0	0	0
SSP 7223	Gamma S crystallin	6753532	76	MALDI/TOF/TOF-PMF	20.84 / 6.9	27.00 / 6.8	1	0.8	0.5	0.4
Increased in the 129/SvJ Wild Type					Mw and pI		Relative Protein Abundance[a]			
Protein Spot Denominator	Protein Name	Accession Number NCBI	Mascot Score[b]	Protein Identification Method	Theoretical	2-DE	C57BL/6J wt	C57BL/6J *Gja3* KO	129/SvJ wt	129/SvJ *Gja3* KO
SSP 1013	Beta A3 crystallin	20304089	76	MALDI/TOF/TOF-PMF	25.19 / 6.0	14.62 / 5.1	1	6.1	12.6	11.1
SSP 1018	Beta B3 crystallin	10946674	114	MALDI/TOF/TOF-PMF	24.27 / 6.7	14.91 / 5.2	1	6.0	8.4	6.5
SSP 2010	Alpha A crystallin	62201857	83	MALDI/TOF/TOF-PMF	19.78 / 5.8	17.56 / 5.4	1	4.7	7.1	6.5
SSP 2104	Beta A3 crystallin	20304089	91	MALDI/TOF/TOF-PMF	25.19 / 6.0	19.54 / 5.3	1	4.6	10.1	8.4
SSP 3004	Alpha A crystallin	62201857	102	MALDI/TOF/TOF-PMF	19.78 / 5.8	17.18 / 5.6	1	5.6	8.4	6.3
SSP 3007	Alpha A crystallin	62201857	130	MALDI/TOF/TOF-PMF	19.78 / 5.8	17.48 / 5.6	1	3.3	8.3	6.3
SSP 3010	Beta B3 crystallin	10946674	95	MALDI/TOF/TOF-PMF	24.27 / 6.7	16.12 / 5.8	1	6.6	7.3	6.1
SSP 3014	Alpha A crystallin	62201857	105	MALDI/TOF/TOF-PMF	19.78 / 5.8	18.23 / 5.6	1	4.1	7.3	3.3
SSP 3109	Alpha A crystallin	62201857	126	MALDI/TOF/TOF-PMF	19.78 / 5.8	20.89 / 5.7	0	+	+	+
SSP 3809	DnaK-type molecular chaperone Hsc70 homolog	74205924	95	MALDI/TOF/TOF-MS/MS	63.92 / 6.1	66.16 / 5.6	0	0	+	+
SSP 4005	Alpha A crystallin	62201857	103	MALDI/TOF/TOF-PMF	19.78 / 5.8	16.96 / 5.9	1	1.8	2.0	3.1
SSP 4008	Alpha A crystallin	62201857	102	MALDI/TOF/TOF-PMF	19.78 / 5.8	16.15 / 5.9	1	7.0	7.9	9.3
SSP 4106	Alpha A crystallin	62201857	102	MALDI/TOF/TOF-PMF	19.78 / 5.8	18.97 / 6.0	1	6.4	10.8	9.1
SSP 5002	Alpha A crystallin	62201857	84	MALDI/TOF/TOF-PMF	19.78 / 5.8	16.77 / 6.2	1	5.0	7.0	6.2
SSP 5015	Beta A3/A1 crystallin	20304089	314	MALDI/TOF/TOF-MS/MS	25.19 / 6.0	17.02 / 6.5	0	+	+	+
SSP 5409	Annexin A1	NA	104	MALDI/TOF/TOF-PMF	38.71 / 7.7	37.86 / 6.4	1	1.5	6.7	6.8
SSP 6003	Alpha A crystallin	62201857	200	MALDI/TOF/TOF-MS/MS	19.78 / 5.8	16.55 / 6.6	1	5.9	8.3	8.1
SSP 6010	Gamma C crystallin	6681037	113	MALDI/TOF/TOF-PMF	20.9 / 7.6	16.03 / 6.6	1	1.5	6.5	6.8
SSP 6011	Alpha A crystallin	62201857	77	ESI-MS/MS	19.78 / 5.8	16.57 / 6.6	0	+	++	++
SSP 6107	Alpha B crystallin	6753530	76	MALDI/TOF/TOF-PMF	20.06 / 6.8	18.30 / 6.6	1	7.7	8.4	7.0
SSP 6122	Gamma N crystallin	23346485	117	MALDI/TOF/TOF-PMF	21.39 / 6.3	20.83 / 6.7	0	0	+	+
SSP 6127	Gamma D crystallin; Alpha A crystallin	34784220; 62201857	102; 67	MALDI/TOF/TOF-PMF	21.1 / 7.0; 19.78 / 5.8	20.38 / 6.8	0	+	+	+
SSP 6704	Chaperonin subunit 6A	6753324	66	MALDI/TOF/TOF-PMF	57.97 / 6.6	59.56 / 6.6	1	0.7	5.9	4.9
SSP 7001	Alpha A crystallin	62201857	337	MALDI/TOF/TOF-MS/MS	19.78 / 5.8	17.37 / 6.8	1	4.1	6.8	5.8
SSP 7017	Beta A3/A1 crystallin gene product	26350979	76	MALDI/TOF/TOF-PMF	12.43 / 6.5	13.85 / 7.0	1	6.5	12.9	10.1
SSP 7106	Syntaxin binding protein	19353298	72	MALDI/TOF/TOF-MS/MS	23.66 / 9.3	18.02 / 6.9	1	3.2	4.5	4.1
SSP 8006	Alpha B crystallin	6753530	128	MALDI/TOF/TOF-PMF	20.06 / 6.8	16.86 / 7.2	1	3.3	8.4	7.2
SSP 8014	Gamma E crystallin	34978370	96	MALDI/TOF/TOF-PMF	21.18 / 7.1	17.15 / 7.3	1	3.6	4.3	4.2
SSP 8118	Gamma A crystallin	6724317	120	MALDI/TOF/TOF-PMF	21.4 / 7.5	19.68 / 7.5	1	1.5	2.4	1.8
SSP 9013	Beta B3 crystallin	10946674	72	MALDI/TOF/TOF-PMF	24.27 / 6.7	15.25 / 8.0	1	5.8	8.7	7.3

SSP 9014	Beta B3 crystallin	10946674	208	MALDI/TOF/TOF-MS/MS	24.27 / 6.7	15.18 / 8.0	1	1.5	2.4	1.8
SSP 9017	Gamma B crystallin	2507570	79	MALDI/TOF/TOF-PMF	21.12 / 7.55	14.23 / 8.2	1	4.8	7.0	6.5

[a] The mean of the normalized staining intensity of the silver stained proteins of the three independent preparations of the urea soluble proteins of the lens is listed as the ratio of the wild type and the mutant of both strains and the wild type of the C57BL/6J strain. When the protein is absent in the wild type of the C57BL/6J strain the protein abundance is indicated in relative terms as increased (+) or decreased (-).
[b] A Mascot score of 63 or more indicates protein identity with a P-value less than 0.05.

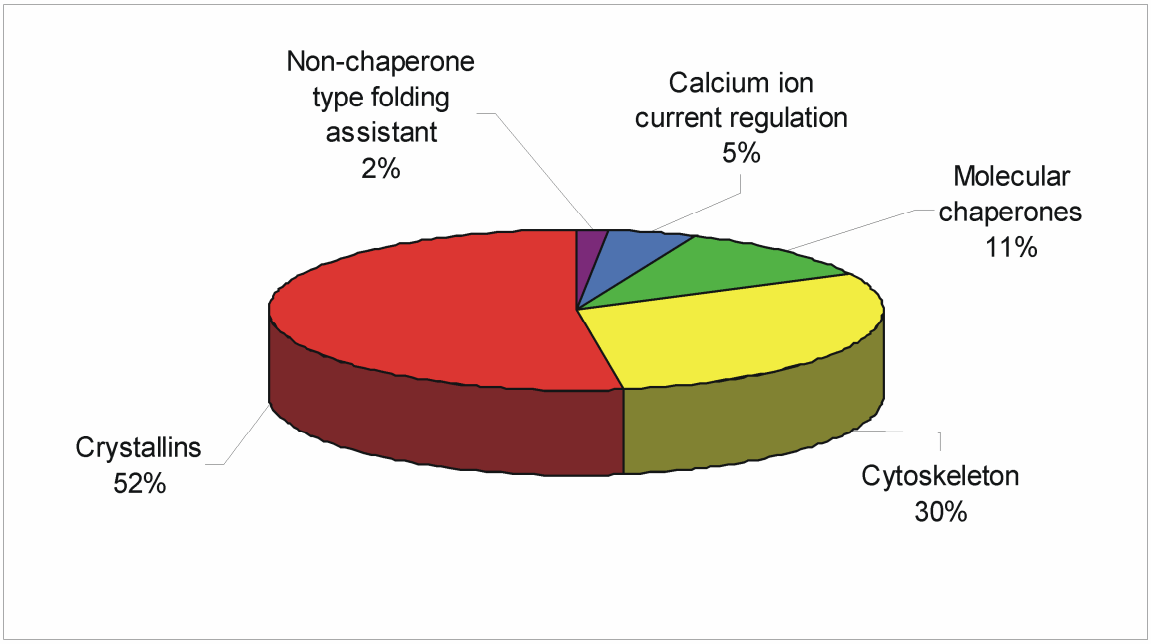


Figure 9. Classification of the significant protein variability in the urea soluble protein fractions of the lenses of the C57BL/6J and 129/SvJ wild type mice. The characterized variant protein species were entered into arbitrary classes reflecting protein function. The Alpha crystallin protein species were entered into the class crystallins although they are also molecular chaperones.

The investigation describes in detail all significant protein variability in the lenses of the two strains up to the limits of the applied techniques. In addition, the investigation identifies several factors that directly effect differential cataract development and severity. A previously described point mutation in the acceptor splice site of exon 2 in the CP49 encoding gene *Bfsp2* was long suspected to modify the onset and severity of cataract in the 129/SvJ mouse strain (Prescott et al., 2001, Sandilands et al., 2004). This was confirmed. Molecular chaperones and other protein folding assistants are highly abundant in the lens and involved in achieving and maintaining protein solubility and lens transparency. Two proteins which modulate intracellular calcium levels, Syntaxin binding

protein and Annexin A1 are also modifying factors that help to sustain calcium ion currents in the face of gap junction loss in the lens nucleus. In total, the results certainly show that healthy lens and cataract development cannot be attributed to the presence or absence of a few factors alone, but that distinct, precisely regulated networks of protein species and their interactions are responsible for these phenomena in each of the investigated inbred mouse strains.

3.3.2. The Loss of CP49 and the Beaded Filament is a Modifying Factor for Cataract Onset and Severity in the *Mus musculus* 129/SvJ strain

The proteins CP49, also known as Phakinin (Merdes et al., 1993; Hess et al., 1996; Binkley et al., 2002; Hoehenwarter et al., 2005) and CP94, also known as Filensin (Merdes et al., 1991; Hess et al., 1998), are related to the intermediate filament protein family (Quinlan et al., 1996). They are obligate assembly partners forming the backbone of the primary cytoskeletal element in the differentiated lens fiber cells, the beaded filament and are unique to the lens (Maisel and Perry, 1972; Ireland and Maisel, 1984, 1989). Alpha A crystallin and Alpha B crystallin associate with the backbone constituting the beads on the beaded filament and complete this cellular structure (Nicholl and Quinlan, 1994).

The beaded filament is integral to the structure and morphology of lens fiber cells and the lens as a whole. Mutations in the CP49 gene in humans that result in autosomal dominant congenital cataract (Jakobs et al., 2000) and juvenile-onset progressive cataract (Conley et al., 2000) have been described. Targeted disruption of the CP49 gene in C57BL/6J mice predictably led to (Alizadeh et al., 2002; Sandilands et al., 2003) radical changes in fiber cell morphology and membrane organization resulting in

decreased lens optical quality and light scatter but not cataract. It seems disruption of the beaded filament is compensated by other, possibly vimentin dependent mechanisms, creating an alternative albeit reduced cytoskeleton which maintains lens architecture and transparency (Sandilands et al., 2003).

Analysis of the 129/SvJ mouse genomic DNA detected a naturally occurring point mutation in the acceptor splice site of exon 2 in the CP49 gene *Bfsp2* in this mouse strain but not in C57BL/6J. Consequently, it was reported that the CP49 protein could not be detected in the lenses of 129/SvJ mice with polyclonal antibodies (Sandilands et al., 2004). The proteomics results confirm this. The CP49 protein was identified in 10 protein spots on a 2-DE gel of the urea soluble lens proteins of C57BL/6J wild type mice. The protein species were also present in lenses of the C57BL/6J mutants however were absent in both wild type and mutants of the 129/SvJ strain (Figure 10). As previously reported, there is some truncation of the full length protein species (Sandilands et al., 1995). Most of the total CP49 protein however is unmodified and equally abundant in the lenses of the C57BL/6J wild type and mutants. This indicates CP49 protein abundance is primarily unaffected by the *Gja3* knockout.

The obligate assembly partner of CP49, CP94 was identified in 8 protein spots on a 2-DE gel of the urea soluble lens proteins of C57BL/6J wild type mice (Figure 11). The protein species were equally abundant in lenses of C57BL/6J wild type and mutants and were markedly decreased in lenses of wild type and mutant 129/SvJ mice. As previously reported, CP49 abundance affects the stoichiometry of its obligate assembly partner CP94 and of the beaded filament as a whole (Alizadeh et al., 2002; Sandilands et al.,2003).

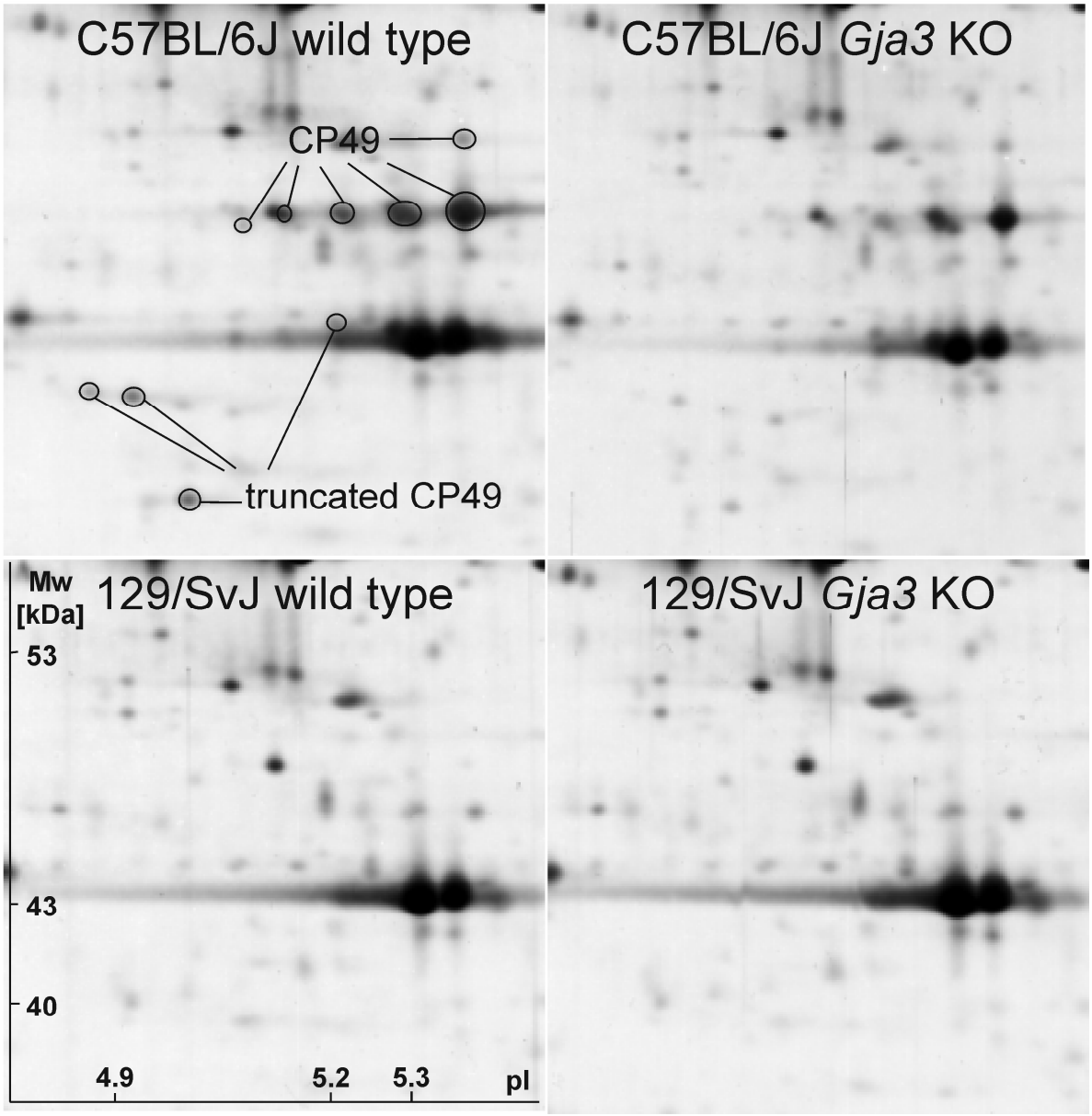


Figure 10. Section of the 2-DE protein spot pattern that shows the CP49 protein species. The protein spots are circled and absent in the patterns of the urea soluble protein fractions of the lenses of the 129/SvJ strain.

Taken together, it seems the atrophied cytoskeleton and altered membrane organization resulting from the disruption of the beaded filament in the 129/SvJ strain leads to increased cell fragility. This phenomenon is well documented for other intermediate filament defects and diseases (Fuchs and Cleveland, 1998). More fragile and porous cell membranes promote calcium ion influx as has been shown in many cataract model systems and as is also the case in the *Gja3* knockout mutant lenses

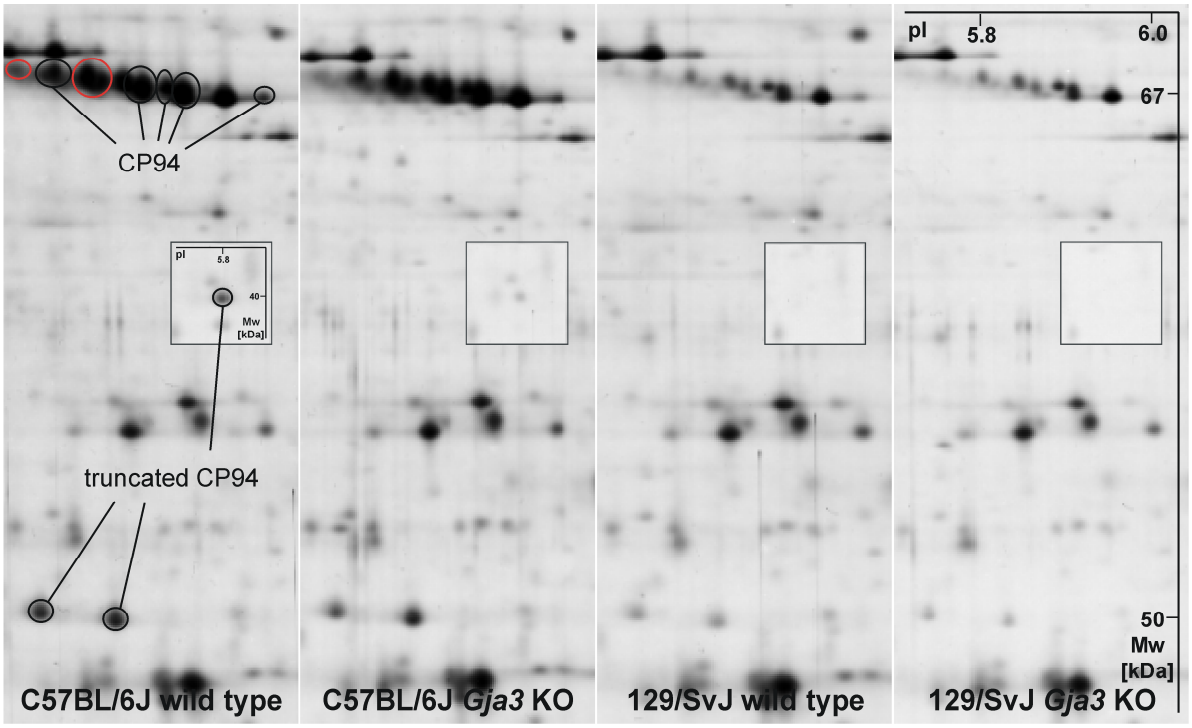


Figure 11. Section of the 2-DE protein spot pattern that shows the CP94 protein species. The protein spots are circled and less intense in the patterns of the urea soluble protein fractions of the lenses of the 129/SvJ strain. Protein spots which were significantly variant and visible on CBB G250 stained gels but for which no identification was achieved are circled red.

because of the loss of the Alpha 3 connexin integral membrane proteins. Thus the increased intracellular calcium levels and in turn Lp82 protease activity as a result of differential fiber cell morphology is a feature of the healthy physiology in the lens of the 129/SvJ mouse strain yet at the same time constitutes a predisposition towards the cataract disease. The combined lack of the CP49 and Alpha 3 connexin proteins leads to pathological alterations.

3.3.3. Protein Folding Assistants are Instrumental to Achieve and Maintain Lens Transparency and Help to Inhibit Cataract Onset and Development

3.3.3.1. The ATP Independent Lower Molecular Weight Molecular Chaperones

The unmodified, phosphorylated and doubly phosphorylated HSP27/25 (Thiede et al., 2005) protein was identified in three protein spots from 2-DE gels of the urea soluble lens proteins of the C57BL/6J wild type mice. The protein species are significantly less abundant in lenses of the 129/SvJ strain in the wild type and the mutant (Figure 12).

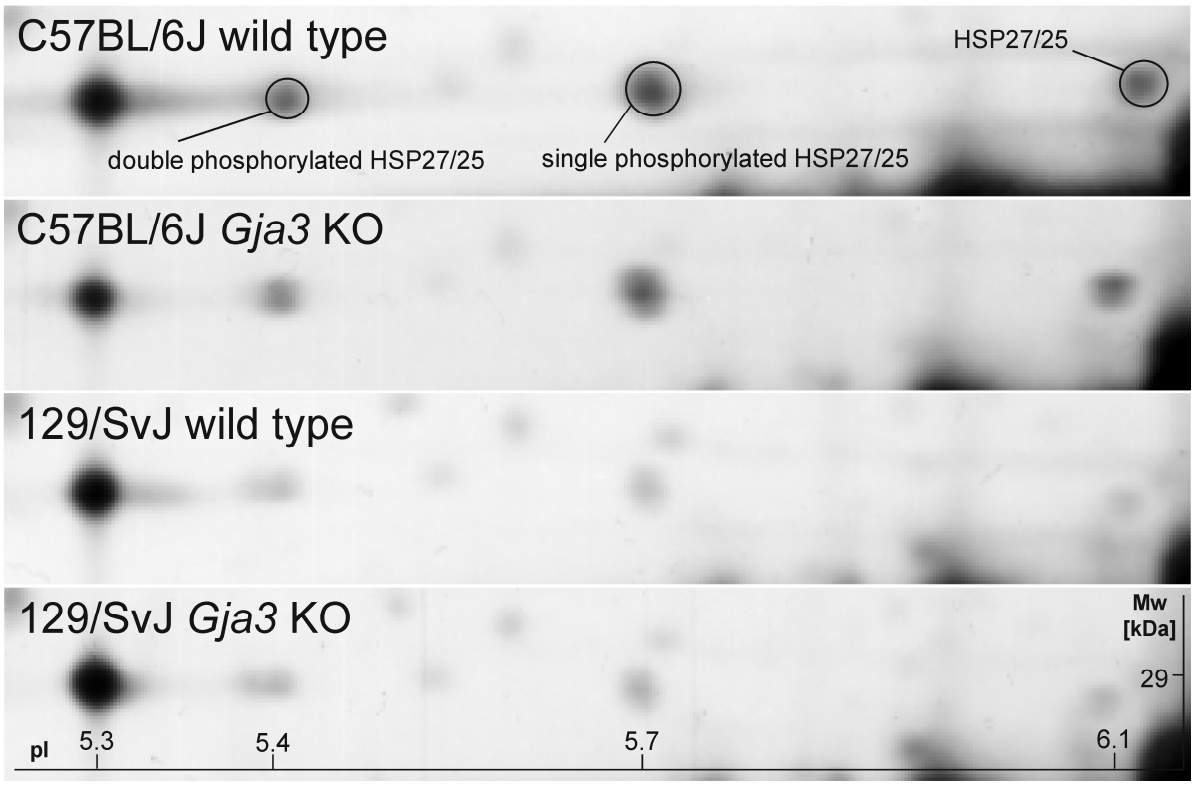


Figure 12. Section of the 2-DE protein spot pattern that shows the HSP27/25 protein species. The protein spots are circled and less intense in the patterns of the urea soluble protein fractions of the lenses of the 129/SvJ strain.

The HSP27/25 protein is a member of the small heat shock protein family (sHSP) (Ingolia et al., 1982) and very similar in structure and function to two other prominent, highly abundant small heat shock proteins in the lens, Alpha A and Alpha B crystallin

(Klemenz et al., 1991; Merck et al., 1993; Leroux et al., 1997a). All three protein monomers act as subunits in obligate functional higher order assembly (van den Oetelaar et al., 1990; Gesierich and Pfeil, 1996; Leroux et al., 1997b; Bova et al., 2000; Haley et al., 2000) and can dynamically exchange to form homo- or hetero-oligomeric complexes (Bova et al., 2000; Fu and Liang, 2003).

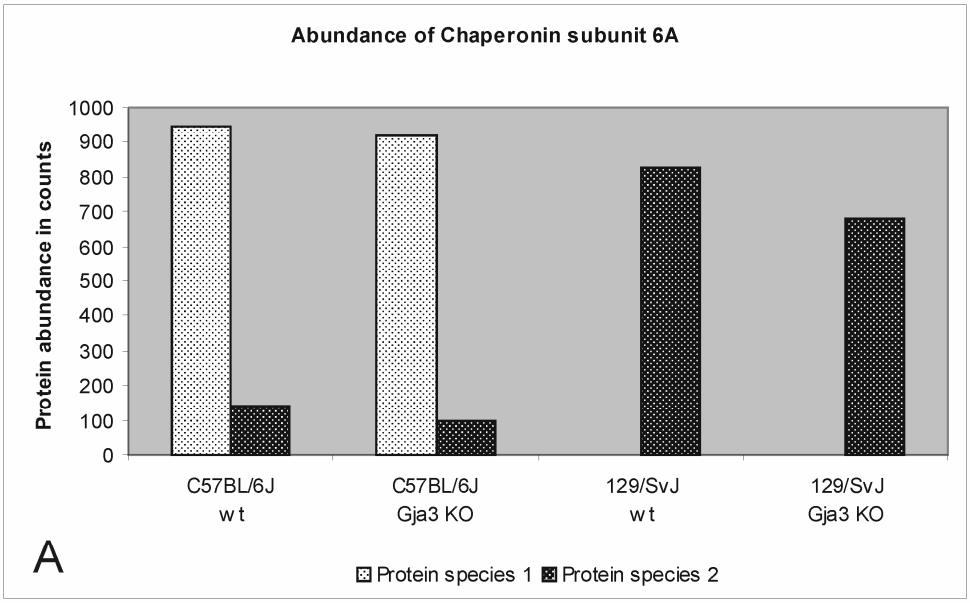
Alpha A and Alpha B crystallin and HSP27/25 are molecular chaperones that influence the folding properties of substrate proteins preventing unspecific protein aggregation and insolubilization independently of ATP and in some cases they may prevent cataract (Horwitz 1992; Jakob et al., 1993, Brady et al., 1997; Horwitz et al., 1998, Litt et al., 1998). The Alpha A and Alpha B crystallin proteins constitute around 30% of the lens protein mass and were similarly abundant in the lenses of wild type and mutants of both strains. This suggests an involvement to assist in the initial proper conformational arrangement of other crystallins beyond their intrinsic function as soluble transparent material.

It has been shown that Alpha B crystallin and HSP27/25 associate with intermediate filament proteins particularly Vimentin and modulate assembly of the cytoskeletal network (Nicholl and Quinlan, 1994; Perng et al., 1999) and that HSP27/25 inhibits the polymerization of Actin filaments (Miron et al., 1988, 1991), another major component of lens fiber cell architecture (Ramaekers et al., 1981; Rafferty, 1985). The comparatively low abundance of HSP27/25 to the Alpha A and Alpha B crystallin proteins and its differential abundance in the strains indicate a specialized function connected to the characteristic differences in the C57BL/6J and 129/SvJ cytoskeleton. The evidence also indicates that the stochiometry of the Alpha crystallin subunits affects the function of the high molecular weight oligomer.

3.3.3.2. The ATP Dependent Higher Molecular Weight Molecular Chaperones

The proteins Chaperonin subunit 6A and DnaK-type molecular chaperone hsc70 homolog are ATP dependant higher molecular weight molecular chaperones belonging to the Chaperonin or GroEL (Sigler et al., 1998) and Hsp70 or DnaK type (Bukau and Horwich, 1998) protein families. Two distinct Chaperonin subunit 6A protein species are differentially abundant in the lenses of the two strains. Protein species 1 is exclusively present in the lenses of C57BL/6J mice where it is similarly abundant in the wild type and the mutant. Protein species 2 is significantly more abundant in the wild type and mutant lenses of the 129/SvJ strain. There is a similar situation for two DnaK-type molecular chaperone hsc70 homolog protein species. Protein species A is present in the lenses of both strains but more abundant in the C57BL/6J mice, protein species B is only present in the lenses of the 129/SvJ mice. Both protein species are also similarly abundant in the wild type and the mutant (Figure 13).

The ATP dependent higher molecular weight molecular chaperones, particularly Hsp70, can adopt sequestered denatured proteins from the low molecular weight ATP independent sHSP molecular chaperones and return them to their native fold (Wickner et al., 1999; Wang and Spector, 2000). Via this repair mechanism proteins regain their solubility and avoid destruction by cytosolic ATP dependent proteases. It has also been reported that chaperonins modulate the assembly of Actin filaments (Gao et al., 1992) and that Hsp70 interacts with intermediate filaments (Green and Liem, 1989). Clearly questions of protein solubility and proper arrangement are a primary concern in the lens and the high abundance of molecular chaperones is evident. The results indicate molecular chaperones are imminently connected to upholding protein solubility and are instrumental in maintaining the transparency of the lens and in some protein species specific instances help to prevent cataract.



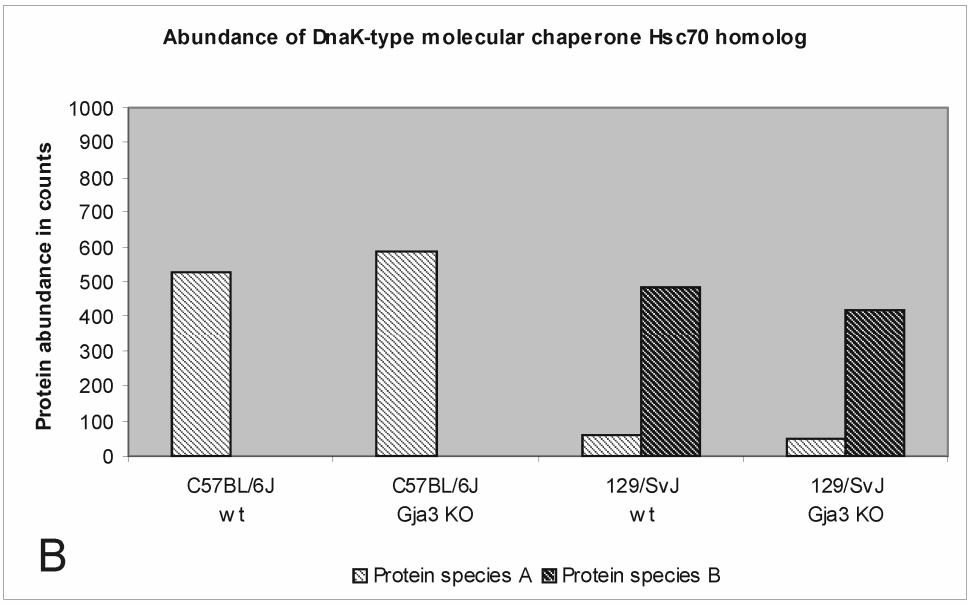


Figure 13. Differential abundance of the A) Chaperonin subunit 6A protein species and the B) DnaK-type molecular chaperone Hsc70 homolog protein species in the urea soluble protein fractions of the lenses of the wild type and the mutant of the C57BL/6J and 129/SvJ strain. Protein abundance is indicated in protein spot staining intensity counts.

3.3.3.3. The Non-chaperone Type Folding Assistants

An Endoplasmatic reticulum protein 29 (ERp29) protein species is abundant in the lenses of wild type and mutant C57BL/6J mice and absent in the lenses of the 129/SvJ

strain (Figure 14). This protein has been characterized only recently and is found in the lumen of the endoplasmatic reticulum (Demmer et al., 1997; Mkrtchian et al., 1998). It is ubiquitously present and highly conserved in animals however it remains a functional orphan. Studies on the pure native protein isolated from rat liver cells have disallowed calcium binding, classical disulphide isomerase and molecular chaperone activity. A hypothesis of non-chaperone type folding and protein escort properties is emerging (Hubbard et al., 2004; Mkrtchian and Sandalova, 2006).

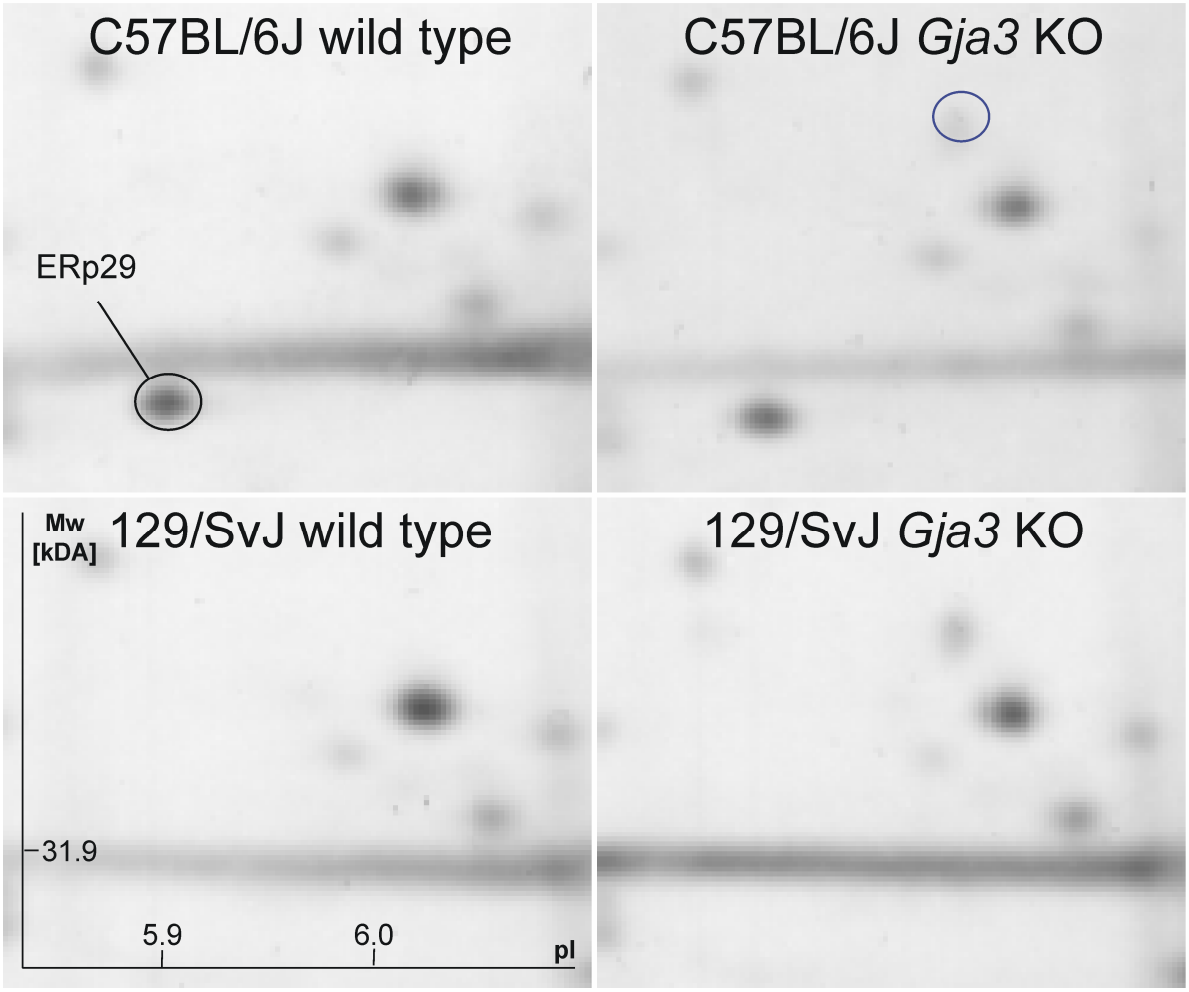


Figure 14. Section of the 2-DE protein spot pattern that shows the Erp29 protein species. The protein spot is circled and absent in the patterns of the urea soluble protein fractions of the lenses of the 129/SvJ strain. A protein spot which is only present in the 2-DE protein spot patterns of the *Gja3* KO mutants and which presumably contains a protein species resulting exclusively from the mutation is circled in blue.

The fiber cells of the lens lack organelles including the endoplasmatic reticulum. The results indicate ERp29 may be involved in escorting the crystallins at the point of translation during terminal differentiation at the transition from lens epithelial to lens fiber cell. In this regard the protein could be regarded as the archetype chaperone in the lens. A host of other functions possibly connected to cataract suppression are also conceivable.

3.3.4. Syntaxin Binding Protein and an Alternative Calcium Ion Channel Network Presumably Effect Calcium Ion Currents in the Lens

Syntaxin binding protein also called Amisyn (Scales et al., 2002) was identified in a protein spot from 2-DE gels of the urea soluble lens proteins of the 129/SvJ wild type mice. The characterized protein species is more abundant in the lenses of the mutant than the wild type C57BL/6J mice and even more abundant in the lenses of the mutant and the wild type 129/SvJ mice (Figure 15). This indicates Syntaxin binding protein abundance is connected to the knockout of the *Gja3* gene as well as the strain specific variability. The common denominator is the reduced membrane integrity and the accompanying elevated intracellular calcium ion levels in the fiber cells. The results suggest that Syntaxin binding protein abundance is elevated in a stress response type manner.

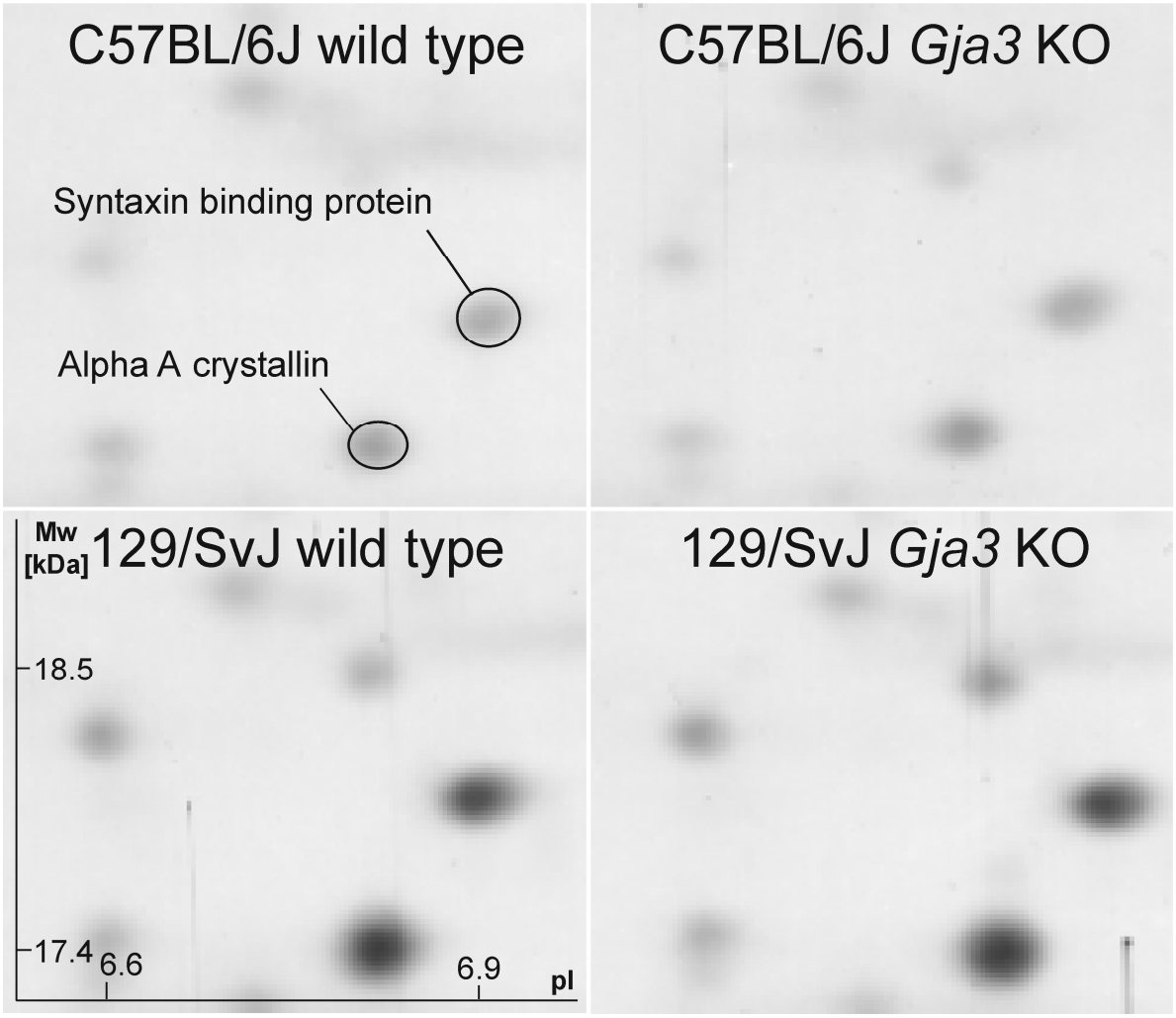


Figure 15. Section of the 2-DE protein spot pattern that shows the Syntaxin binding protein species. The protein spot is circled and less intense in the patterns of the urea soluble protein fractions of the lenses of the C57BL/6J strain. A protein spot containing a significantly variant Alpha A crystallin protein species (SSP 7001, Table 4) is also circled.

Syntaxin or HPC-1 is an integral membrane protein with a C-terminal transmembrane domain (Yoshida et al., 1992; Bennett et al., 1992) and binds to a protein interaction site on the cytoplasmic region connecting the second and third of four transmembrane domains of the presynaptic calcium channels particularly of the N-type (Wheeler et al., 1994). It has been shown that Syntaxin has an inhibitory effect on these channels both directly and as a second messenger via G-protein pathways and leaves them in a closed state preventing calcium ion current (Stanley and Mirotnzik, 1997; Jarvis et al., 2000; 2002). Syntaxin binding protein is not membrane associated and binds to Syntaxin (Scales et al. 2002). The results indicate it acts on Syntaxin in a

competitive inhibitor type manner, sequestering Syntaxin from binding to sites on the presynaptic calcium channels and G-protein (Jarvis et al., 2000) thereby keeping the channels open and facilitating calcium ion current.

The implications for the lens are profound. A voltage gated calcium channel network consisting of L-type calcium channels has been suspected (Srivastava et al., 1997; Baruch et al., 2001) but its components have not been clearly identified. Other evidence attests that Syntaxin associates with N-type and P/Q-type channels but not with L-type channels (El Far et al., 1995; Pupier et al., 1997). Together this suggests the presence of at least one type of voltage gated calcium channel network in the lens. This network may be as essential to directing the overall pattern of current flow as the gap junction network. The G-protein beta subunit like protein identified in the lens proteins of the wild type and the mutant animals of both strains is a further indicator. The increased Syntaxin binding protein levels presumably modulate calcium ion currents via this alternative pathway and counteract the effects of the loss of the gap junctions in the lens nucleus and the loss of the beaded filament.

3.3.5. Two Annexin A1 Protein Species Specifically and Differentially Regulate Calcium Ion Currents in the Lenses of the two Strains

Two protein species of the anti-inflammatory protein Annexin A1 are differentially abundant in the urea soluble fractions of the C57BL/6J and the 129/SvJ lenses. Protein species X is exclusively present in the lenses of the wild type and the mutant C57BL/6J strain. Protein species Y is twofold more abundant in the mutant than in the wild type C57BL/6J lenses however its total abundance in this strain is minimal. It is substantially more abundant in the lenses of the wild type and the mutant 129/SvJ strain where its

total abundance is comparable to the abundance of protein species X in the C57BL/6J strain (Figure 16).

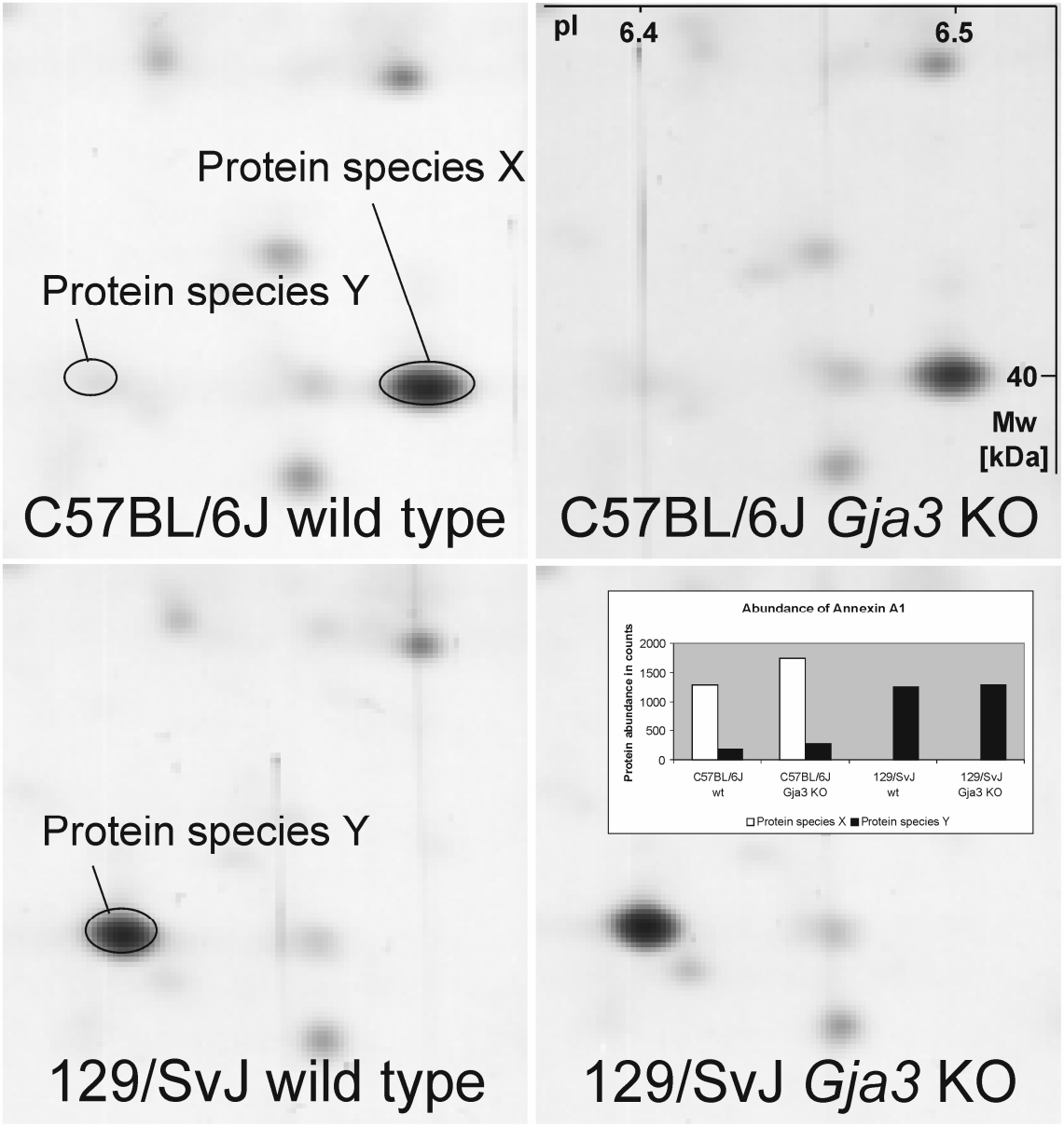


Figure 16. Section of the 2-DE protein spot pattern that shows the Annexin A1 protein species. The protein spots are circled. The diagram in the upper right hand panel shows the differential abundance of the protein species in the urea soluble protein fractions of the lenses of the wild type and the mutant of the C57BL/6J and the 129/SvJ strain in protein spot staining intensity counts.

Annexin A1 is a member of the Annexin protein family, which ubiquitously binds calcium and binds to plasma membranes in a reversible, calcium ion dependant manner (Gerke and Moss, 2002). It activates calcium ion transport in a concentration dependent

manner in human neutrophils (Solito et al., 2003). Particularly the proteins N-terminal domain is involved in this process. There is also evidence that Annexins themselves are calcium ion channels however there is some speculation as to the exact molecular mechanisms involved (Pollard et al., 1992; Burger et al., 1996).

Annexins are known for post translational modifications which affect protein function (Gerke and Moss, 2002). MALDI/TOF/TOF-MS/MS analysis of a tryptic peptide of the protein species Y from amino acid residues 82 to 97 of the Annexin A1 primary structure with the mass 1816.91 Da carrying a single charge produced y10 and y11 fragment ion signals with a 1 Da mass increase over the calculated masses for the corresponding unmodified peptide fragment ions (Figure 17). This indicates deamidation at amino acid residue 88. The ion signals had low intensity mass peaks corresponding to the unmodified peptide fragment ions possibly due to the partial loss of the modification during experimental procedures. The modification is also present in protein species X. The shift in the positions of the protein spots in the protein patterns indicates there is an additional alteration in the primary structure of protein species Y.

The results indicate the distinct Annexin A1 protein species regulate calcium ion transport activity specifically in the lenses of the two strains. Annexin A1 binds to cell membranes under high calcium concentrations and induces calcium ion current. The protein species could act on the N- or L- type calcium ion channels described above or its intrinsic properties as a calcium ion channel could come to bear. Whatever the exact mechanism, the results indicate that the protein species acts in the nucleus of the lenses of the C57BL/6J strain reducing calcium ion concentrations in a rescue type effect.

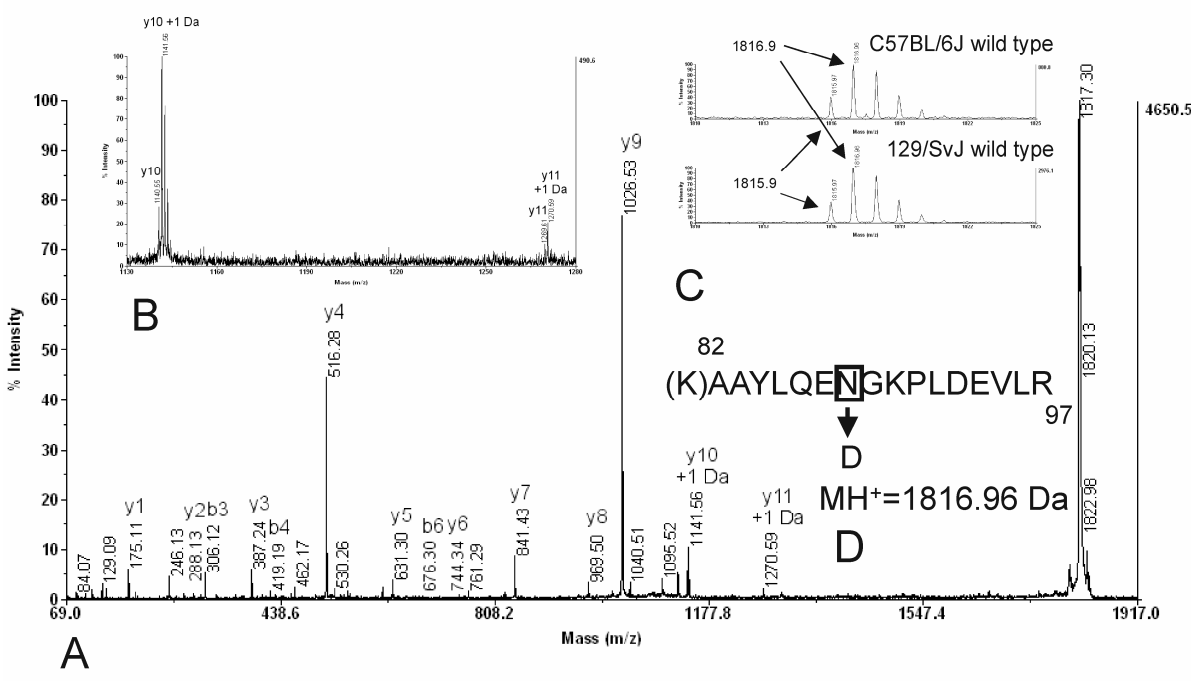


Figure 17. Mass spectrometric analysis indicating deamidation of amino acid residue 88 of the Annexin A1 primary structure. A) MS/MS spectrum of the single charged ion with the monoisotopic mass 1816.91 Da. The fragment ion signals that delineate the amino acid sequence of the analyzed peptide are indicated (b and y ion series). The y10 and y11 fragment ion signals with a 1 Da mass increase over the calculated masses for the corresponding unmodified peptide fragment ions of the tryptic peptide from amino acid residues 82 to 97 of the Annexin A1 primary structure are also indicated. B) Detail of the y10 and y11 fragment ion signals. The low intensity mass peaks corresponding to the respective unmodified peptide fragment ions can be seen. C) Detail of the ion signals with the single charged monoisotopic masses of the unmodified (1815.9) and deamidated (1816.9) tryptic peptides from amino acid residues 82 to 97 of the Annexin A1 primary structure in PMF spectra from tryptic digests of protein spots from 2-DE gels of the urea soluble protein fractions of the lenses of the wild type C57BL/6J and 129/SvJ mice. D) Amino acid sequence of the tryptic peptide from amino acid residues 82 to 97 of the Annexin A1 primary structure. The deamidation of amino acid residue 88 and the single charged monoisotopic mass of the deamidated peptide are indicated. All masses are in Daltons.

3.3.6. Many Truncated Crystallin Proteins are Differentially Abundant in the Lenses of the Wild Type and Mutant Mice of Both Strains

Most of the protein species that are differentially abundant in the urea soluble protein fraction of the lenses of the two strains belong to the crystallin protein superfamily. Many of them are significantly more abundant in the lenses of the mutant C57BL/6J and the wild type and the mutant 129/SvJ strain. The protein species, primarily the Alpha A crystallins, are at positions less than the theoretical molecular weight of the

full length proteins in the protein spot patterns. This indicates extensive protein truncation. This is the result of the elevated Lp82 protease activity because of the high intracellular calcium ion levels in the lenses of the mutant mice because of the loss of the Alpha 3 connexin and in the lenses of the 129/SvJ mice because of the loss of the beaded filament. Also N-terminally truncated Beta B1 crystallin is highly abundant exclusively in the mutant lenses of both strains (data not shown). This protein species seems to be directly correlated with the onset and development of the nuclear cataract together with the previously reported gamma crystallin protein species. Other investigations have linked calcium ion levels, calpain protease activity and truncated Alpha and Beta crystallins to cataract onset and development in mice (Piatigorsky et al., 1978; Garber et al., 1984; David et al., 1994; Nakamura et al., 2000; Descamps et al., 2005).

In total, the investigation comprehensively realized the variability in the urea soluble protein fraction of the lenses of the *Mus musculus* C57BL/6J and 129/SvJ strain. The abundance of a large number of the protein species was significantly altered indicating the total variability is considerable.

4. Discussion

The investigations described here are a systematic analysis of the protein variability in the lens of the eye of the mouse which led to the detection of factors that modify the onset of an induced genetic disorder, the knockout of the Alpha 3 connexin gene *Gja3* leading to nuclear cataract. This simple animal model reflects the clinical variability of monogenic congenital pathologies following a Mendelian inheritance pattern in humans (Slavotinek and Tifft, 2002; Kyriakou et al., 2004; Hartmann and Kulozik, 2006). In the course of the investigations it has become clear that the transition from health to cataract is a matter of surpassing and perpetuating thresholds in healthy physiological pathways in the lens. The impact of the protein species and the need for a general understanding of every protein species as an individual effector with a distinct abundance, structure and function also became apparent.

Indeed, the abundance of protein species is as essential to regulatory networks as their presence and methods to rapidly and accurately quantify large numbers of protein species in complex samples are necessary. This demand is primarily being met by the development of stable isotope coded protein reagents and respective ion signal quantification in mass spectrometry but also by the development and improvement of 2-DE spot staining intensity quantification softwares (Patton, 2002). Both strategies were employed in the investigations with different results.

4.1. Analysis of the Variability in the Lens Proteins in the *Mus musculus* and *Mus spretus* Species with one-step Protein Labeling Strategies and LC-MS

4.1.1. The ICAT Strategy

The investigation with the cleavable ICAT reagent employing the LC Packings and Thermo Finnigan system identified and quantified six (6) proteins. The investigation with the Agilent Technologies system identified and quantified thirteen (13) proteins.

With the exception of the Alpha, Beta B2 and Beta B3 crystallins, the crystallins contain a suitable number of cysteine residues for application of the ICAT strategy. The number of cysteine containing tryptic peptide sequences in the ion traps optimal detection range, masses between 800 and 3000 Da, however is limited. The number of discriminate tryptic peptide sequences unique to a particular crystallin protein and useful for unambiguous identification and quantification is further reduced among the Gamma crystallins because of extensive protein sequence homology. From these calculations, the Alpha crystallins and Gamma E crystallin cannot be identified at all and the number of peptides *de facto* available for analysis of the other crystallin proteins are, except for Beta A3 and Beta A1 crystallin, which cannot be discriminated outright because of protein sequence homology, 1, 2 and at most 5 (Table 1). Therefore, conditions for the study of lens proteins with the ICAT strategy are not ideal. There were some exceptions, for instance the sole cysteine containing tryptic peptide from Alpha A and Alpha A insert crystallin, which has a monoisotopic mass in excess of 3000 Da was detected. The proteins however could not be unambiguously identified and quantified because the peptide is not discriminate.

Discussion

The liquid chromatography and mass spectrometry machines used in the investigations both performed comparably well. The quantification softwares namely the XPRESS^{TM} software used with the Thermo Finnigan system and the SpectrumMill software from Agilent Technologies performed differently. The XPRESS^{TM} software features a user interface which allows manual evaluation and if necessary correction of suggested ion signal area ratios and accordingly peptide and protein abundance ratios. The SpectrumMill software lacks such an interface so errors cannot be amended. The software calculates the area of an ion signal over a parameter defined software generated duration in the acquired ion current. In addition the software searches for and if available calculates the areas of ion signals corresponding to same peptides labeled with the other form of the cleavable ICAT reagent or other user specified reagents over the same duration as the original ion signal. Thus the software acquires the areas of the ion signals of all possible combinations of a cleavable ICAT labeled peptide pair. Following database searching and assignment of a peptide sequence suggestion the appropriate ratio of areas is calculated from these previously acquired data. The major inconsistencies in peptide abundance ratios which are evident in the Supplementary Tables 2 and 3 and the standard deviations in some of the protein abundance ratios listed in Table 1 are indicative of errors in the original quantification of the ion signal areas of the differentially labeled peptides with the same primary structure from the *Mus musculus* and *Mus spretus* lenses. An example is shown (Figure 18).

The range between 125 and 175 minutes of the acquired ion current of the unfragmented ions with the masses of 1066 Da +/- 0.4 Da (top) and 1063 Da +/- 0.4 Da (bottom) from the third analysis of the second independently prepared reaction stock are shown. The scan numbers corresponding to the beginning and end of the ion signal peak at around 150 minutes are also shown. The corresponding mass spectra were extracted from the raw data into scans and are listed with scan numbers and peptide abundance

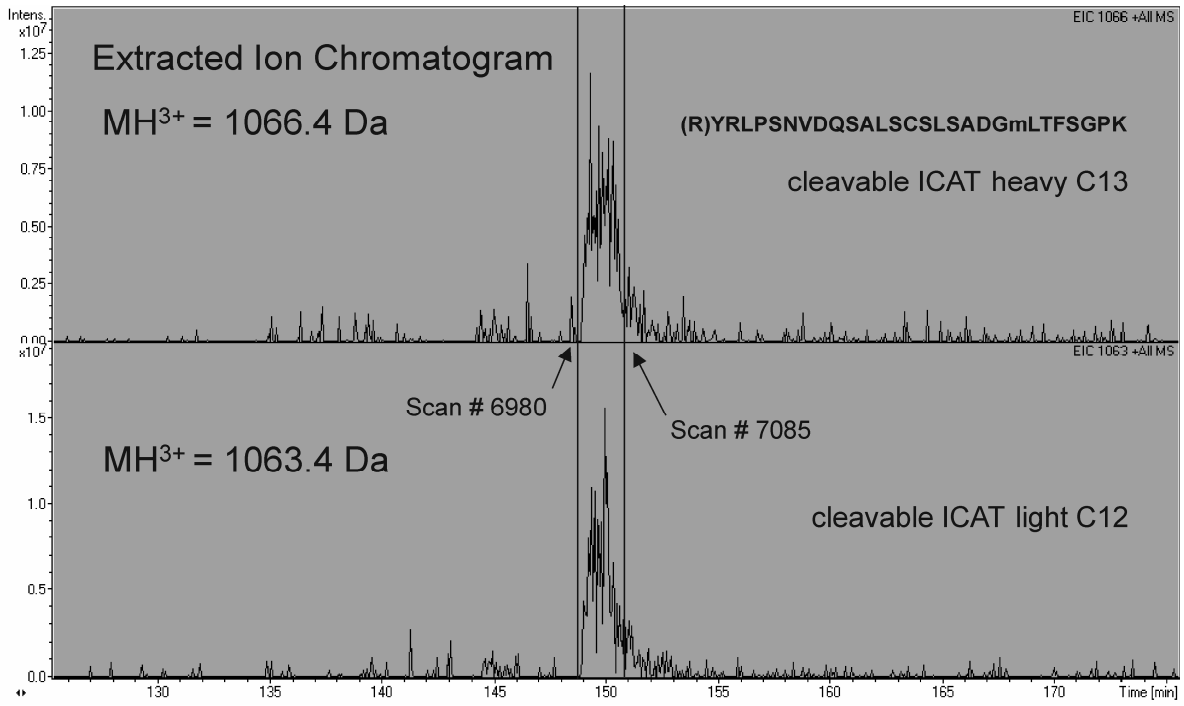


Figure 18. Extracted ion chromatogram of the ion signals with the masses of 1066.4 Da and 1063.4 Da from an analysis of the ICAT labeled lens proteins of the *Mus musculus* and *Mus spretus* species. The ion signals are from the differentially labeled peptides with the indicated sequence and were eluted from chromatography and acquired simultaneously in the mass spectrometer.

ratios in Supplementary Table 3 in lines 9 through 12 under Protein Number 1. The peptide sequence for the only cysteine containing tryptic peptide in Alpha A and Alpha A insert crystallin YRLPSNVDQSALSCSLSADGMLTFSGPK with the methionine oxidated was unambiguously assigned to the scans. The peptide abundance ratios suggested by the software for these scans are 0.027, 6.539, 0.008 and 7.380 respectively. Not only is the range between the four values roughly 1000 but the actual ratio of the ion signal peak areas is around 1 as is evident from the figure. The peptide abundance ratios suggested by the SpectrumMill software are clearly wrong.

This error could easily be corrected with the XPRESS^{TM} software which allows arbitrary definition of the beginning and end points for the calculation of the ion signal peak areas and their ratio. It seems both softwares perform well when the mass spectrometric raw data is of high quality (Moulder et al., 2004), the SpectrumMill software

however quickly encounters problems. Therefore, the abundance ratios of the peptide sequence suggestions for the Beta A2 crystallin, Beta B2 crystallin, Gamma B crystallin and Gamma C crystallin proteins produced with the XPRESS™ software are correct, the low number of identified and quantified peptides however makes statistically significant protein quantification impossible.

In summary, the ICAT strategy for rapid comprehensive identification and quantification of proteins is not ideal for the crystallins because of their nature. Indeed, as shown by Schmidt and co-workers (Schmidt et al., 2004) the ICAT strategy has difficulties even in discriminating proteins let alone protein species in higher eukaryotes due to the inherent limitation to cysteine residues and the high amount of paralogs and post translational modifications. Additionally in the lens, the abundance of peptides from the non-crystallin proteins is minuscule compared to the shear mass of the crystallins which constitute around 90% of total lens protein mass. Therefore, these peptides are difficult to detect in high numbers. Implementation of the established strong cation exchange HPLC could improve results however the ICAT strategy should not be applied to the eye lens.

4.1.2. The iTRAQ Strategy

The quantification of the lens proteins in the *Mus musculus* and *Mus spretus* species was also attempted with the iTRAQ reagent and a 4700 Proteomics Analyzer MALDI/TOF/TOF mass spectrometer from Applied Biosystems. The RapiGest™ SF buffer was shown to be responsible for the phenomenon of non-constitutive labeling of peptide amino groups in an investigation of the human 20S proteasome subunits where constitutive labeling with the iTRAQ reagent was achieved in water under otherwise

identical conditions (Janek et al., 2006). However, this should not adversely effect protein quantification because the phenomenon should be evenly distributed throughout.

Thirty six (36) protein identification suggestions were produced and thirty three (33) proteins including most of the crystallins ubiquitous to mammals identified and quantified. A comprehensive investigation of human lens proteins using high end MudPit with 15 successive strong cation exchange LC steps for extended peptide separation before reverse phase LC and mass spectrometry identified 270 proteins (MacCoss et al., 2002). Considering the investigation described here used only reverse phase LC for peptide separation before mass spectrometry and the mitotically active cells of the anterior epithelium were removed, the number of identified proteins is high. Of course it must be noted that many of the proteins were only identified with one or two discriminate peptides which makes statistically significant quantification impossible.

The investigation detected four putative sequence polymorphisms in four proteins, Alpha B crystallin, Gamma B crystallin, Gamma F crystallin and Gamma S crystallin. The lack of a reporter ion signal in the MS/MS spectra of peptides with the same primary structure from the lenses of the two species labeled with the iTRAQ reagent is indicative of a lack of a specific peptide in the respective sample which barring errors in experimental procedures must be due to polymorphism or modification of the peptide. The NCBI Protein sequence database which contains a comprehensive collection of *Mus musculus* protein sequences was used for identification of peptides and proteins. Consequently, polymorphic peptides from the *Mus spretus* species could not be detected as there was no adequate reference database. The mass spectra of these polymorphic peptides must be present in the raw data and lack the reporter ion signal characteristic of the iTRAQ reagent with which the *Mus musculus* lens protein samples were labeled. Theoretically, as all peptides in up to four samples can be differentially labeled with the iTRAQ reagent and individually discriminated in one mass spectrum this technique

combined with a powerful *de novo* peptide sequencing software could prove ideal for rapid und reliable screening of polymorphisms and modifications on a proteome wide scale.

Indeed, the question of polymorphisms and protein species must always be addressed when analyzing complex samples without extensive prior prefractionation. As an example from the iTRAQ investigation, the peptide sequence suggestion GQMSEITDDCLSLQDR from the Gamma B crystallin protein sequence with the NCBI accession number 42733606 with N-terminal iTRAQ labeling was assigned to the MS/MS spectrum number 2197 with a Mascot score of 109. Another peptide sequence suggestion GQMVEITDDCPHLQDR from the Gamma E crystallin protein sequence with the NCBI accession number 31982854 with N-terminal iTRAQ labeling was assigned to the same MS/MS spectrum with a lower but still acceptable Mascot score of 22. Both peptide sequence suggestions and the corresponding MS/MS spectra were manually reviewed and validated.

The Gamma E crystallin sequence with the above accession number is distinct from the *Mus musculus* Gamma E crystallin sequence 34978370 which was detected in the 2-DE based analysis of the ten day old *Mus musculus* C57BL/6J proteome and which was shown to perfectly match all codons of the *Mus musculus* Gamma E crystallin gene *CrygE* open reading frame (Hoehenwarter et al., 2005). The peptide sequence suggestion assigned to the MS/MS spectrum in the iTRAQ investigation covers a polymorphic site between these two sequences although the lenses are from the *Mus musculus* C57BL/6J strain. Thus, while there is no doubt that both Gamma B crystallin and Gamma E crystallin proteins are present, the discrepancy in the Gamma E crystallin sequences presents a conundrum. Several similar situations are footnoted in Table 2 and Supplementary Table 1.

Discussion

In comparing the ICAT and iTRAQ strategies for one-step relative quantification of proteins in two samples, the iTRAQ strategy seems to be more advantageous. It features less complex reaction chemistry and a simpler and more robust quantification procedure where every peptide is individually quantified in a single MS/MS spectrum. The ICAT strategy also has an additional affinity chromatography step for enrichment of labeled peptides which presents an additional avenue for error. The ICAT strategies premier advantage is that it reduces sample complexity by specifically labeling cysteine amino acid residues.

Considering that 90% of the lens protein mass belongs to the crystallin protein superfamily and the very high protein sequence homology in this protein superfamily the rationale of applying any LC-MS strategy must be carefully assessed. Many lower abundant proteins will be overshadowed by the crystallins and remain undetected unless a very high degree of separation is achieved. Also, polymorphic peptides and proteins often cannot be distinguished unless a very high mass accuracy in mass spectrometry is achieved as is possible with FTICR instruments.

In both the ICAT and iTRAQ based LC-MS investigations of the variability in the proteins of the eye lenses in the *Mus musculus* and *Mus spretus* species the protein abundance ratio for most of the crystallins and for many of the non-crystallin proteins is 1. This indicates the variability is marginal which at first is not surprising for two closely related species in a terminally differentiated, more or less static organ. The relative standard deviations of the mean peptide abundance ratios are around 20% or lower for most of the proteins. This is in the range of previous experiments employing both the ICAT and iTRAQ strategy for relative protein abundance quantification (Han et al., 2001; Ross et al., 2004).

4.2. Analysis of the Variability in the Lens Proteins in the *Mus musculus* C57BL/6J and 129/SvJ Strain with 2-DE, Quantative Image Analysis and Mass Spectrometry

Two-dimensional electrophoresis, the PDQuest image analysis software and a MALDI/TOF/TOF mass spectrometer were used to analyze the variability in the lens proteins of the *Mus musculus* strains C57BL/6J and 129/SvJ. Both the wild type and homozygous Alpha 3 connexin gene *Gja3* knockout mutants were analyzed to define in detail the effects of the variability on the onset and severity of cataract. Several factors that directly impact the development of the pathology were determined; however it became clear, that a similar yet specific network of protein interactions is responsible for the healthy and pathological development of the lens of the eye in the two inbred strains.

The loss of CP49 and in turn the beaded filament and the accompanying radical changes in fiber cell and lens morphology as a whole are clearly powerful modifying factors resulting in a predisposition for cataract in the 129/SvJ strain. The increased abundance of some protein folding assistants in the C57BL/6J strain may help to keep proteins in solution and to moderate the effects of the mutation. The Syntaxin binding protein and the Annexin A1 protein species seem to promote calcium ion current flow and to counteract calcium ion accumulation in the nucleus of the lenses of the C57BL/6J strain.

The recently described non-chaperone type folding assistant ERp29 is a novelty in the eye lens. The Syntaxin binding protein and the G-protein beta subunit protein (Supplementary Table 4) which were identified in the lenses of the wild type and the *Gja3* knockout mutants of both strains are part of the N-type calcium ion channel network in neurons. Other components of this network were identified in both the ICAT (Syntaxin 4,

listed in Supplementary Table 2, Appendix, Protein Number 16) and iTRAQ (Bassoon, presynaptic cytomatrix protein, listed in Table 2, Protein Number 33) investigations of the *Mus musculus* and *Mus spretus* lens proteins. These results indicate that there is at least one calcium ion channel network present in the lens in addition to the gap junction network which could be equally important in the regulation of lenticular calcium ion transport.

The total protein interactions which are apparent from the results are shown in an interaction network (Figure 19). The proteins are grouped into the arbitrary taxa Protease, Protein folding assistants, N- / L-type calcium ion channel, Gap junction calcium ion channel, Crystallins and Cytoskeleton. Physical interactions between proteins are indicated by green colored arrows connecting nodes. Proximal interactions were defined as interactions between proteins in a pathway with the interaction partners removed from direct physical interaction by at least one pathway component and are indicated by red colored arrows connecting nodes. Distal interactions were defined as interactions between proteins outside of a defined pathway and are indicated by black colored arrows connecting nodes. Proposed interactions were defined as interactions between proteins based on interactions of homologous proteins or protein family members of the interaction partners or on protein function and are indicated by blue colored arrows connecting nodes.

The crystallin proteins in solution constitute the bulk of the transparent lenticular material (Delaye and Tardieu, 1983). Increased modification and truncation of the crystallin proteins accompanies the normal ageing process and the cataract disease (Garber et al., 1984; Miesbauer et al., 1994; Lampi et al., 1998; Ma et al., 1998; Ueda et al., 2002). An overall organ specific architecture must be maintained for proper light diffraction (Kuszak et al., 1994). Considering the results in light of these observations it is clear that each protein species has a function in the supra-molecular order which can be

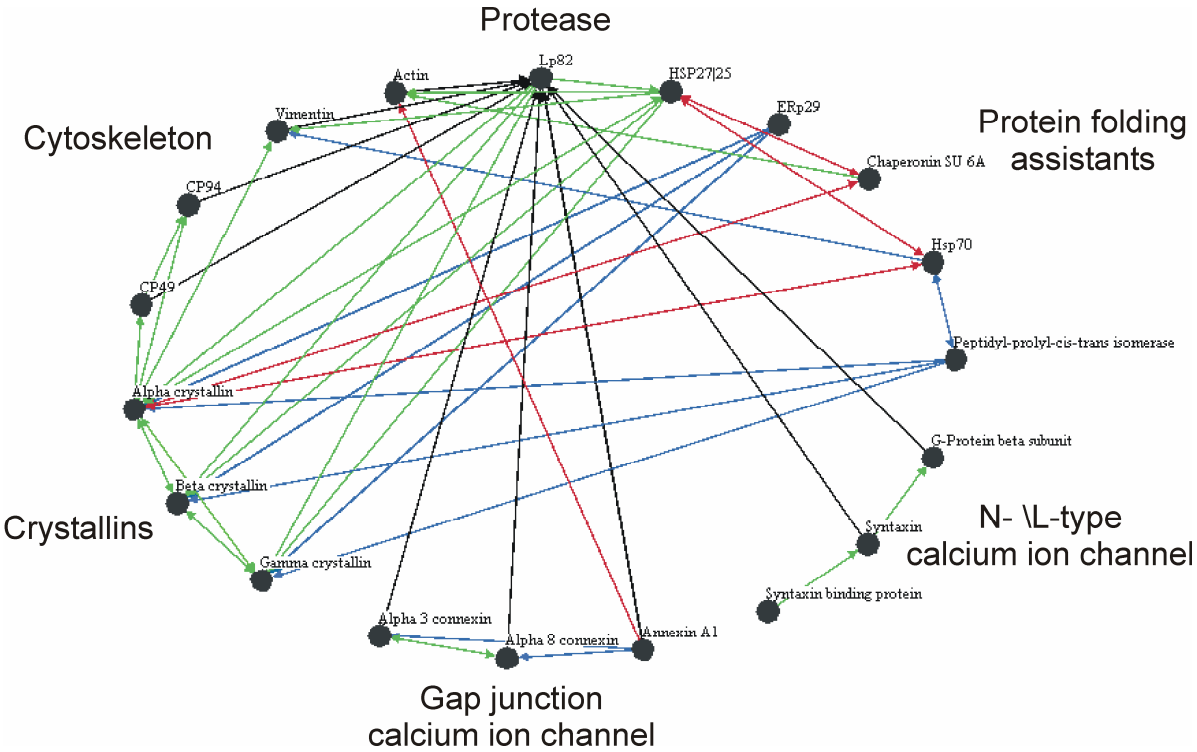


Figure 19. Network of protein interactions central to health and pathology in the lenses of mice. Physical interactions are indicated by green colored arrows, proximal interactions are indicated by red colored arrows, distal interactions are indicated by black colored arrows and proposed interactions are indicated by blue colored arrows. The proteins are arbitrarily classified reflecting protein function.

defined. The truncated crystallin proteins are essential components of the healthy lens just as any other protein species. The disruption of the calcium ion currents and the subsequent increase in calcium ion levels elevates the activity of the calpain proteases. Perpetuated elevated protease activity leads to an improper stochiometry of the truncated crystallin proteins which subsequently aggregate and become insoluble. This process cannot be checked by mitigating factors such as molecular chaperone response or calcium ion transport activation after a certain threshold is surpassed. Factors that promote increased calcium levels in their own right lower this critical threshold. Therefore a network of proteins is responsible for healthy and pathological development in the lens of the eye. This network is similar but distinct in the two investigated strains and presumably similar in other inbred mouse strains yet unique in its own right in every mouse strain.

4.3. Conclusions

The diversity of the protein species and their functional implications become clear when considering the results of the described investigations together. In the investigation of the *Mus musculus* C57BL/6J and 129/SvJ strain the protein spot containing the full length Alpha A crystallin gene product is large and saturated (SSP 4121, Supplementary Table 4). For this protein species 100% sequence coverage was achieved, i.e. it was actually identified as a protein species. It is equally highly abundant in the lenses of both the wild type and the mutants of both *Mus musculus* strains. Alpha A crystallin was also identified as the only component of 13 other protein spots in the protein spot patterns with significant variation in staining intensity (Table 4). One of these protein species was analyzed in greater detail and a functional hypothesis was formulated based on molecular modeling (Hoehenwarter et al., 2006a). The exact functions of the other 12 Alpha A crystallin protein species remain unknown.

The investigation of the *Mus musculus* and *Mus spretus* species with the iTRAQ strategy also assessed the abundance ratio of Alpha A crystallin. As for most of the other lens proteins it is around 1 (Table 2). This indicates the variability in the lens proteins is marginal in the two closely related species which does not seem surprising for a terminally differentiated, quiescent organ. However, the abundance of the individual types of Alpha A crystallin molecules, the protein species, goes completely undetected. Thus the necessary resolution was not reached and many of the functional molecular interactions are entirely overlooked, a limitation of many proteomics techniques (Schmidt et al., 2004; Hoehenwarter et al., 2006a). The variability in lens proteins is indeed present but manifested on the protein species and not the protein level. This is also true

for many other biological phenomena and makes the comprehensive investigation of the proteome on the protein species level necessary for true understanding to be achieved.

The ability to discriminate the individual components and to simultaneously grasp the total components of a protein sample demonstrates the power of a proteomics investigation. Which of the results from the mouse model beyond the detection of individually distinct modifying factors for monogenic congenital pathologies and the establishment of functional networks in health and pathology are applicable to humans is difficult to assess. Increased intracellular calcium ion concentration is definitely connected to cataract onset and development in humans (Sanderson et al., 2000) however other parallels are not yet established. Also, a mutation implicated in the onset of zonular pulverulant cataract connected to a loss of gap junction coupling (Pal et al., 2000) was mapped to the Alpha 3 connexin gene *GJA3* in two families (Mackay et al., 1999). An assessment of proteomics for directed investigations of human samples which go beyond routine protein identification could prove insightful. Applied systematically and by an experienced professional the proteomics methodology can rapidly produce clear clinically relevant results.

A study of the lens integral membrane proteins by solubilization of the urea insoluble protein fraction with high SDS concentrations (Kibbelaar and Bloemendal, 1979) and subsequent SDS-PAGE and mass spectrometry would give a clearer picture of fiber cell membrane spanning calcium ion channels and calcium ion homeostasis. A backcross of the homozygous C57BL/6J with the 129/SvJ *Gja3* knockout mutants should confirm the effects of the modifying factors described above.

5. References

Aarts HJ, Lubsen NH, Schoenmakers JG (1989) Crystallin gene expression during rat lens development. Eur J Biochem 183: 31-36

Abgar S, Backmann J, Aerts T, Vanhoudt J, Clauwaert J (2000) The structural differences between bovine lens alpha A- and alphaB-crystallin. Eur J Biochem 267: 5916-5925

Aksu S, Scheler C, Focks N, Leenders F, Theuring F, Salnikow J, Jungblut PR (2002) An iterative calibration method with prediction of post-translational modifications for the construction of a two-dimensional electrophoresis database of mouse mammary gland proteins. Proteomics 2: 1452-1463

Alizadeh A, Clark Ji, Seeberger T, Hess J, Blankenship T, Spicer A, FitzGerald PG (2002) Targeted genomic deletion of the lens-specific intermediate filament protein CP49. Invest Ophthalmol Vis Sci 43: 3722-3727

Baruch A, Greenbaum D, Levy E, Nielsen P, Gilula N, Kumar N, Bogyo M (2001) Defining a link between gap junction communication, proteolysis, and cataract formation*. J Biol Chem 276: 28999-29006

Bassnett S (2002) Lens organelle degradation. Exp Eye Res 74: 639-649

References

Bateman OA, Lubsen NH, Slingsby C (2001) Association behaviour of human betaB1-crystallin and its truncated forms. Exp Eye Res 73: 321-331

Bateman OA, Sarra R, van Genesen ST, Kappe G, Lubsen NH, Slingsby C (2003) The stability of human acidic beta-crystallin oligomers and hetero-oligomers. Exp Eye Res 77: 409-422

Bennett MJ, Choe S, Eisenberg D (1994) Domain swapping: entangling alliances between proteins. Proc Natl Acad Sci USA 91: 3127-3131

Bennett MJ, Schlunegger MP, Eisenberg D (1995) 3D domain swapping: a mechanism for oligomer assembly. Protein Sci 4:2455-2468

Bennett MK, Calakos N, Scheller RH (1992) Syntaxin: a synaptic protein implicated in docking of synaptic vesicles at presynaptic active zones. Science 257: 255-259

Bera S, Thampi P, Cho WJ, Abraham EC (2002) A positive charge preservation at position 116 of alpha A-crystallin is critical for its structural and functional integrity. Biochemistry 41: 12421-12426

Berbers GA, Hoekman WA, Bloemendal H, de Jong WW, Kleinschmidt T, Braunitzer G (1984) Homology between the primary structures of the major bovine beta-crystallin chains. 139: 467-479

References

Berengian AR, Bova MP, Mchaourab HS (1997) Structure and function of the conserved domain in alphaA-crystallin. Site-directed spin labeling identifies a beta-strand located near a subunit interface. Biochemistry 36: 9951-9957

Bhat SP (2003) Crystallins, genes and cataract. Prog Drug Res 60: 205-262

Bhat SP (2004) Transparency and non-refractive functions of crystallins--a proposal. Exp Eye Res 79: 809-816

Biemann K, Cone C, Webster BR, Arsenault GP (1966) Determination of the amino acid sequence in oligopeptides by computer interpretation of their high-resolution mass spectra. J Am Chem Soc 88: 5598-5606

Binkley PA, Hess J, Casselman J, FitzGerald P (2002) Unexpected variation in unique features of the lens-specific type I cytokeratin CP49. Invest Ophthalmol Vis Sci 43: 225-235

Bloemendal H (1977) The vertebrate eye lens. Science 197: 127-138

Bloemendal H, de Jong WW (1991) Lens proteins and their genes. Prog Nucleic Acid Res Mol Biol 41:259-281

Bova MP, Ding LL, Horwitz J, Fung BK (1997) Subunit exchange of alphaA-crystallin. J Biol Chem 272: 29511-29517

References

Bova MP, McHaourab HS, Han Y, Fung BK (2000) Subunit exchange of small heat shock proteins. Analysis of oligomer formation of alphaA-crystallin and Hsp27 by fluorescence resonance energy transfer and site-directed truncations. J Biol Chem 275: 1035-1042

Bradley RH, Ireland M, Maisel H (1979) The cytoskeleton of chick lens cells. Exp Eye Res 28: 441-453

Brady JP, Garland D, Duglas-Tabor Y, Robison WG Jr, Groome A, Wawrousek EF (1997) Targeted disruption of the mouse alpha A-crystallin gene induces cataract and cytoplasmatic inclusion bodies containing the small heat shock protein alpha B-crystallin. Proc Natl Acad Sci USA 94: 884-889

Breci LA, Tabb DL, Yates 3rd JR, Wysocki VH (2003) Cleavage N-terminal to proline: analysis of a database of peptide tandem mass spectra. Anal Chem 75: 1963-1971

Breitkreutz BJ, Stark C, Tyers M (2003) Osprey: a network visualization system. Genome Biol 4: R22

Breitman ML, Lok S, Wistow G, Piatigorsky J, Treton JA, Gold RJ, Tsui LC (1984) Gamma-crystallin family of the mouse lens: structural and evolutionary relationships. Proc Natl Acad Sci USA 81: 7762-7766

Breuker K, Oh H, Lin C, Carpenter BK, McLafferty FW (2004) Nonergodic and conformational control of the electron capture dissociation of protein cations. Proc Natl Acad Sci USA 101: 14011-14016

References

Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 chaperone machines. Cell 92: 351-366

Burger A, Berendes R, Liemann S, Benz J, Hofman A, Gottig P, Huber R, Gerke V, Thiel C, Romisch J, Weber K (1996) The crystal structure and ion channel activity of human annexin II, a peripheral membrane protein. J Mol Biol 257: 839-847

Bystroff C, Shao Y (2002) Fully automated ab initio protein structure prediction using I-SITES, HMMSTR and ROSETTA. Bioinformatics 18 Suppl 1: S54-61

Carrell RW, Lomas DA (1997) Conformational disease. Lancet 350: 134-138

Carter JM, Hutcheson AM, Quinlan RA (1995) In vitro studies on the assembly properties of the lens proteins CP49, CP115: coassembly with alpha-crystallin but not with vimentin. Exp Eye Res 60: 181-192

Carver JA (1999) Probing the structure and interactions of crystallin proteins by NMR spectroscopy. Prog Retin Eye Res 18: 431-462

Carver JA, Cooper, PG, Truscott, RJ (1993) 1H-NMR spectroscopy of beta B2-crystallin from bovine eye lens. Eur J Biochem 213: 313-320

Caspers GJ, Leunissen JA, de Jong WW (1995) The expanding small heat-shock protein family, and structure predictions of the conserved "alpha-crystallin domain". J Mol Evol 40: 238-248

References

Castellanos-Serra LR, Fernandez-Patron C, Hardy E, Huerta V (1996) A procedure for protein elution from reverse-stained polyacrylamide gels applicable at the low picomole level: An alternative route to the preparation of low abundance proteins for microanalysis. Electrophoresis 17: 1564-1572

Chakraborty A, Regnier FE (2002) Global internal standard technology for comparative proteomics. J Chromatogr A 949: 173-184

Chambers C, Russell P (1991) Deletion mutation in an eye lens beta-crystallin. An animal model for inherited cataracts. J Biol Chem 266: 6742-6746

Clout NJ, Basak A, Wieligman K, Bateman OA, Jaenicke R, Slingsby C (2000) The N-terminal domain of betaB2-crystallin resembles the putative ancestral homodimer. J Mol Biol 304: 253-257

Comisarow MB, Marshall AG (1974) Fourier transform ion cyclotron resonance spectroscopy. Chem Phys Lett 25: 282-283

Conley YP, Erturk D, Keverline A, Mah TS, Keravala A, Barnes LR, Bruchis A, Hess JF, FitzGerald PG, Weeks DE, Ferrell RE, Gorin MB (2000) A juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with a missense mutation in the beaded filament structural protein-2. Am J Hum Genet 66: 1426-1431

Cooper PG, Aquilina JA, Truscott RJ, Carver JA (1993a) Supramolecular order within the lens: 1H NMR spectroscopic evidence for specific crystallin-crystallin interactions. Exp Eye Res 59: 607-616

References

Cooper PG, Carver JA, Truscott RJ (1993b) 1H-NMR spectroscopy of bovine lens beta-crystallin. The role of the beta B2-crystallin C-terminal extension in aggregation. Eur J Biochem 213: 321-328

David LL, Calvin HI, Fu SC (1994) Buthionine sulfoximine induced cataracts in mice contain insolubilized crystallins with calpain II cleavage sites. Exp Eye Res 59: 501-504

Delaye M, Tardieu A (1983) Short-range order of crystallin proteins accounts for eye lens transparency. Nature 302: 415-417

Demmer J, Zhou C, Hubbard MJ (1997) Molecular cloning of ERp29, a novel and widely expressed resident of the endoplasmatic reticulum. FEBS Lett 402: 145-150

Dewey J, Bartling C, Rae JL (1995) A non-enzymatic method for lens decapsulation which leaves the epithelial cells attached to the fibers. Curr Eye Res 14: 357-362

Descamps FJ, Martens E, Proost P, Starckx S, Van den Steen, PE, Van Damme J, Opdenakker G (2005) Gelatinase B/matrix metalloproteinase-9 provokes cataract by cleaving lens betaB1 crystallin. FASEB J 19: 29-35

Doherty NS, Littman BH, Reilly K, Swindell AC, Buss JM, Anderson NL (1998) Analysis of changes in acute-phase plasma proteins in an acute inflammatory response and in rheumatoid arthritis using two-dimensional gel electrophoresis. Electrophoresis 19: 355-363

References

Driessen HP, Herbrink P, Bloemendal H, de Jong WW (1981) Primary structure of the bovine beta-crystallin Bp Chain. Internal duplication and homology with gamma-crystallin. Eur J Biochem 121: 83-91

Edman P (1949) A method for the determination of the amino acid sequence in peptides. Arch Biochem 22: 475-476

el Far O, Charvin N, Leveque C, Martin-Moutot N, Takahashi M, Seagar MJ (1995) Interaction of a synaptobrevin (VAMP)-syntaxin complex with presynaptic calcium channels. FEBS Lett 361: 101-105

Ellis M, Alousi S, Lawniczak J, Maisel H, Welsh M (1984) Studies on lens vimentin. Exp Eye Res 38: 195-202

Eng JK, McCormack AL, Yates 3rd JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5: 976-989

Evans CW, Eastwood S, Rains J, Gruijters WT, Bullivant S, Kistler J (1993) Gap junction formation during development of the mouse lens. Eur J Cell Biol 60: 243-249

Farnsworth P, Singh K (2004) Structure function relationship among alpha-crystallin related small heat shock proteins. Exp Eye Res 79: 787-794

Farnsworth PN, Frauwirth H, Groth-Vasselli B, Singh K (1998) Refinement of 3D structure of bovine lens alpha A-crystallin. Int J Biol Macromol 22: 175-185

References

Farnsworth PN, Groth-Vasselli B, Greenfield NJ, Singh K (1997) Effects of temperature and concentration on bovine lens alpha-crystallin secondary structure: a circular dichroism spectroscopic study. Int J Biol Macromol 20: 283-291

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246: 64-71

Fernald RD, Wright SE (1983) Maintenance of optical quality during crystalline lens growth. Nature 301: 618-620

FitzGerald PG (1988) Age-related changes in a fiber cell-specific extrinsic membrane protein. Curr Eye Res 7: 1255-1262

Fu L, Liang JJ (2003) Enhanced stability of alpha B-crystallin in the presence of small heat shock protein Hsp27. Biochem Biophys Res Commun 302: 710-714

Fuchs E, Cleveland DW (1998) A structural scaffolding of intermediate filaments in health and disease. Science 279: 514-519

Fukiage C, Nakajima E, Ma H, Azuma M, Shearer T (2002) Characterization and regulation of lens-specific calpain Lp82. J Biol Chem 277: 20678-20685

Gao Y, Thomas JO, Chow RL, Lee GH, Cowan NJ (1992) A cytoplasmic chaperonin that catalyzes beta-actin folding. Cell 69: 1043-1050

References

Garber AT, Goring D, Gold RJ (1984) Characterization of abnormal proteins in the soluble lens proteins of CatFraser mice. J Biol Chem 259: 10376-10379

Gerke V, Moss SE (2002) Annexins: from structure to function. Physiol Rev 82: 331-371

Gesierich U, Pfeil W (1996) The conformational stability of alpha-crystallin is rather low: calorimetric results. FEBS Lett 393: 151-154

Gong X, Agopian K, Kumar NM, Gilula NB (1999) Genetic factors influence cataract formation in alpha 3 connexin knockout mice. Dev Genet 24: 27-32

Gong X, Baldo G, Kumar N, Gilula N, Mathias R (1998) Gap junctional coupling in lenses lacking $\alpha 3$ connexin. Proc Natl Acad Sci USA 95: 15303-15308

Gong X, Li E, Klier G, Huang Q, Wu Y, Lei H, Kumar NM, Horwitz J, Gilula NB (1997) Disruption of alpha3 connexin gene leads to proteolysis and cataractogenesis in mice. Cell 91: 833-843

Goodenough DA (1992) The crystalline lens. A system networked by gap junctional intercellular communication. Semin Cell Biol 3: 49-58

Goring DR, Breitman, ML, Tsui LC (1992) Temporal regulation of six crystallin transcripts during mouse lens development. Exp Eye Res 54: 785-795

Graw J (2003) The genetic and molecular basis of congenital eye defects. Nat Rev Genet 4: 876-888

References

Graw J (2004) Congenital hereditary cataracts. Int J Dev Biol 48: 1031-1044

Graw J, Liebstein A, Pietrowski D, Schmitt-John T, Werner T (1993) Genomic sequences of murine gamma B- and gamma C-crystallin-encoding genes: promoter analysis and complete evolutionary pattern of mouse, rat and human gamma-crystallins. Gene 136: 145-156

Green LA, Liem RK (1989) Beta-internexin is a microtubule-associated protein identical to the 70-kDa heat-shock cognate protein and the clathrin uncoating ATPase. J Biol Chem 264: 15210-15215

Groenen PJ, Merck KB, de Jong WW, Bloemendal H (1994) Structure and modifications of the junior chaperone alpha-crystallin. From lens transparency to molecular pathology. Eur J Biochem 225: 1-19

Gygi SP, Rist B, Gerber SA, Turecek, F, Gelb MH, Aebersold R (1999) Quantative analysis of complex protein mixtures using isotope-coded affinity tags. Nat. Biotechnol. 17: 994-999

Haley DA, Bova MP, Huang QL, Mchaourab HS, Stewart PL (2000) Small heat-shock protein structures reveal a continuum from symmetric to variable assemblies. J Mol Biol 298: 261-272

Haley DA, Horwitz J, Stewart PL (1998) The small heat-shock protein, alphaB-crystallin, has a variable quaternary structure. J Mol Biol 277: 27-35

References

Han DK, Eng J, Zhou H, Aebersold R (2001) Quantative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol 19: 946-951

Harding JJ (1972) Conformational changes in human lens proteins in cataract. Biochem J 129: 97-100

Harding JJ (2002) Viewing molecular mechanisms of ageing through a lens. Ageing Res Rev 1: 465-479

Harrington V, McCall S, Huynh S, Srivastava K, Srivastava OP (2004) Crystallins in water soluble-high molecular weight protein fractions and water insoluble protein fractions in aging and cataractous human lenses. Mol Vis 10: 476-489

Hartmann K, Kulozik AE (2006) [Genetic modifiers of homozygous sickle cell disease]. Klin Padiatr 218: 170-173

Hejtmancik JF, Wingfield PT, Chambers C, Russel P, Chen HC, Sergeev YV, Hope JN (1997) Association properties of beta B2- and beta A3-crystallin: ability to form dimers. Protein Eng 10: 1347-1352

Hejtmancik JF, Wingfield PT, Sergeev YV (2004) Beta-crystallin association. Exp Eye Res 79: 377-383

References

Hendriks W, Weetnik H, Voorter CE, Sanders J, Bloemendal H, de Jong WW (1990) The alternative splicing product alpha Ains-crystallin is structurally equivalent to alpha A and alpha B subunits in the rat alpha crystallin aggregate. Biochim Biophys Acta 1037: 58-65

Hess JF, Casselman JT, FitzGerald PG (1996) Gene structure and cDNA sequence identify the beaded filament protein CP49 as a highly divergent type I intermediate filament protein. J Biol Chem 271: 6729-6735

Hess JF, Casselman JT, Kong AP, FitzGerald PG (1998) Primary sequence, secondary structure, gene structure, and assembly properties suggests that the lens-specific cytoskeletal protein filensin represents a novel class of intermediate filament protein. Exp Eye Res 66: 625-644

Hoehenwarter W, Ackermann R, Zimny-Arndt U, Kumar NM, Jungblut PR (2006a) The necessity of functional proteomics: Protein species and molecular function elucidation exemplified by in vivo alpha A crystallin N-terminal truncation. Amino Acids 31: 317-323

Hoehenwarter W, Klose J, Jungblut PR (2006b) Eye Lens Proteomics. Amino Acids 30: 369-389

Hoehenwarter W, Kumar NM, Wacker M, Zimny-Arndt U, Klose J, Jungblut PR (2005) Eye lens proteomics: from global approach to detailed information about phakinin and gamma E and F crystallin genes. Proteomics 5: 245-257

Horwitz J (1992) Alpha-crystallin can function as a molecular chaperone. Proc Natl Acad Sci USA 89: 10449-10453

References

Horwitz J, Huang QL, Ding L, Bova MP (1998) Lens alpha-crystallin: chaperone like properties. Methods Enzymol 290: 365-383

Hubbard MJ, Mangum JE, McHugh NJ (2004) Purification and biochemical characterization of native ERp29 from rat liver. Biochem J 383: 589-597

Ingolia TD, Craig EA (1982) Four small Drosophila heat shock proteins are related to each other and to mammalian alpha-crystallin. Proc Natl Acad Sci USA 79: 2360-2364

Ireland M, Maisel H (1984) A cytoskeletal protein unique to lens fiber cell differentiation. Exp Eye Res 38: 637-645

Ireland M, Maisel H (1989) A family of lens fiber cell specific proteins. Lens Eye Toxic Res 6: 623-638

Iwaki T, Kume-Iwaki A, Goldman JE (1990) Cellular distribution of alpha B-crystallin in non-lenticular tissues. J Histochem Cytochem 38: 31-39

Jakob U, Gaestel M, Engel K, Buchner J (1993) Small heat shock proteins are molecular chaperones. J Biol Chem 268: 1517-1520

Jakobs PM, Hess JF, FitzGerald PG, Kramer P, Weleber RG, Litt M (2000) Autosomal-dominant congenital cataract associated with a deletion mutation in the human beaded filament protein gene BFSP2. Am J Hum Genet 66: 1432-1436

References

Janek K, Schmid M, Jungblut P, Kuckelkorn U, Kloetzel PM (2006) Relative Protein Quantification of 20S Proteasomes using iTRAQ Label. Poster at DGMS Meeting 2006

Jarvis SE, Barr W, Feng ZP, Hamid J, Zampouni GW (2002) Molecular determinants of syntaxin 1 modulation of N-type calcium channels. J Biol Chem 277: 44399-44407

Jarvis SE, Magga JM, Beedle AM, Braun JE, Zamponi GW (2000) G protein modulation of N-type calcium channels is facilitated by physical interactions between syntaxin 1A and Gbetagamma. J Biol Chem 275: 6388-6394

Jonscher KR, Yates 3rd JR (1997) The quadrupole ion trap mass spectrometer--a small solution to a big challenge. Anal Biochem 244: 1-15

Jungblut PR, Otto A, Favor J, Lowe M, Muller EC, Kastner M, Sperling K, Klose J (1998) Identification of mouse crystallins in 2D protein patterns by sequencing and mass spectrometry. Application to cataract mutants. FEBS Lett 435: 131-137

Jungblut P, Thiede B (1997) Protein identification from 2-DE gels by MALDI mass spectrometry. Mass Spectrom Rev 16: 145-162

Jungblut PR, Seifert R (1990) Analysis by high-resolution two-dimensional electrophoresis of differentiation-dependant alterations in cytosolic protein pattern of HL-60 leukemic cells. J Biochem Biophys Methods 21:47-58

References

Kamei A, Iwase H, Masuda K (1997) Cleavage of amino acid residue(s) from the N-terminal region of alpha A- and alpha B-crystallins in human crystalline lens during aging. Biochem Biophys Res Commun 231: 373-378

Kapphahn RJ, Ethen CM, Peters EA, Higgins L, Ferrington DA (2003) Modified alpha A crystallin in the retina: altered expression and truncation with aging. Biochemistry 42: 15310-15325

Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60: 2299-2301

Kibbelaar M, Bloemendal H (1975) The topography of lens proteins based on chromatography and two-dimensional gel electrophoresis. Exp Eye Res 21: 25-36

Kibbelaar MA, Bloemendal H (1979) Fractionation of the water-insoluble proteins from calf lens. Exp Eye Res 29: 679-688

Klemenz R, Frohli E, Steiger RH, Schafer R, Aoyama A (1991) Alpha B-crystallin is a small heat shock protein. Proc Natl Acad Sci USA 88: 3652-3655

Klose J, Kobalz U (1995) Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis 16: 1034.1059

References

Klose J, Nock C, Herrmann M, Stühler K, Marcus K, Blüggel M, Krause E, Schalwyk LC, Rastan S, Brown SDM, Büssow K, Himmelbauer H, Lehrach H (2002) Genetic analysis of the mouse brain proteome. Nat Genet 30: 385-393

Koretz JF, Handelman GH (1988) How the human eye focuses. Sci Am 259: 92-99

Krah A, Miehlke S, Pleissner KP, Zimny-Arndt U, Kirsch C, Lehn N, Meyer TF, Jungblut PR, Aebischer T (2004) Identification of candidate antigens for serologic detection of Helicobacter pylori-infected patients with gastric carcinoma. Int J Cancer 108: 456-463

Kröger RH, Campbell MC, Munger R, Fernald RD (1994) Refractive index distribution and spherical aberration in the crystalline lens of the African cichlid fish *Haplochromis burtoni*. Vis Res 34:1815-1822

Kumar NM, Gilula NB (1996) The gap junction communication channel. Cell 84: 381-388

Kuszak, JR, Peterson KL, Sivak JG, Herbert KL (1994) The interrelationship of lens anatomy and optical quality. II. Primate lenses. Exp Eye Res 59: 521-535

Kuwabara T (1975) The maturation of the lens cell: a morphologic study. Exp eye Res 20: 427-443

Kyriakou T, Hodgkinson C, Pontefract DE, Iyengar S, Howell WM, Wong YK, Eriksson P, Ye S (2004) Genotypic effect of the -565C>T polymorphism in the ABCA1 gene promoter on ABCA1 expression and severity of atherosclerosis. Arterioscler Thromb Vasc Biol 25: 418-423

References

Lambert S (1994) Lens In: Albert D, Jakobiec F (Eds.) Principles and Practices of Ophthalmology. Saunders, Philadelphia PA 461

Lampi KJ, Ma Z, Hanson SR, Azuma M, Shih M, Shearer TR, Smith DL, Smith JB, David LL (1998) Age-related changes in human lens crystallins identified by two-dimensional electrophoresis and mass spectrometry. Exp Eye Res 67: 31-43

Lampi KJ, Oxford JT, Bachinger HP, Shearer TR, David LL, Kapfer DM (2001) Deamidation of human beta B1 alters the elongated structure of the dimer. Exp Eye Res 72: 279-288

Lapko VN, Smith DL, Smith JB (2002) S-methylated cysteines in human lens gamma S-crystallins. Biochemistry 41: 14645-14651

Lapko VN, Smith DL, Smith JB (2003) Methylation and carbamylation of human gamma-crystallins. Protein Sci 12: 1762-1774

Leroux MR, Ma BJ, Batelier G, Melki R, Candido EP (1997a) Unique structural features of a novel class of small heat shock proteins. J Biol Chem 272: 12847-12853

Leroux MR, Melki R, Gordon B, Batelier G, Candido EP (1997b) Structure-function studies on small heat shock protein oligomeric assembly and interaction with unfolded polypeptides. J Biol Chem 272: 24646-24656

References

Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR 3rd (1999) Direct analysis of protein complexes using mass spectrometry. Nat Biotechnology 17: 676-682

Litt M, Kramer P, LaMorticella DM, Murphey W, Lovrien EW, Weleber RG (1998) Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. Hum Mol Genet 7: 471-474

Lok S, Tsui LC, Shinohara T, Piatigorsky J, Gold R, Breitman M (1984) Analysis of the mouse gamma-crystallin gene family: assignment of multiple cDNAs to discrete genomic sequences and characterization of a representative gene. Nucleic Acids Res 12: 4517-4529

Louris JN, Cooks GR, Syka JEP, Kelley PE, Stafford GC, Todd JFJ (1987) Instrumentation, applications, and energy deposition in quadrupole ion-trap tandem mass spectrometry. Anal Chem 59: 1677-1685

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem 193: 265-275

Lubsen NH, Aarts HJ, Schoenmakers JG (1988) The evolution of lenticular proteins: the beta- and gamma-crystallin super gene family. Prog Biophys Mol Biol 51: 47-76

Ma H, Fukiage C, Azuma M, Shearer TR (1998) Cloning and expression of mRNA for calpain Lp82 from rat lens: splice variant of p94. Invest Ophthalmol Vis Sci 39: 454-461

References

Ma H, Hata I, Shih M, Fukiage C, Nakamura Y, Azuma M, Shearer T (1999) Lp82 is the dominant form of calpain in young mouse lens*. Exp Eye Res 68: 447-456

MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, Tasto JJ, Gould KL, Wolters D, Washburn M, Weiss A, Clark JI Yates JR 3rd (2002) Shotgun identification of protein modifications from protein complexes and lens tissue. Proc Natl Acad Sci USA 99: 7900-7905

MacKay D, Ionides A, Kibar Z, Rouleau G, Berry V, Moore A, Shiels A, Bhattacharya S (1999) Connexin46 mutations in autosomal dominant congenital cataract. Am J Hum Genet 64: 1357-1364

Maisel H, Perry MM (1972) Electron microscope observations on some structural proteins of the chick lens. Exp Eye Res 14: 7-12

Mathias RT, Rae JL, Baldo GJ (1997) Physiological properties of the normal lens. Physiol Rev 77: 21-50

Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL (2000) The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. Anal Chem 72: 552-558

Merck KB, Groenen PJ, Voorter CE, de Haard-Hoekman WA, Horwitz J, Bloemendal H, de Jong WW (1993) Structural and functional similarities of bovine alpha-crystallin and mouse small heat-shock protein. A family of chaperones. J Biol Chem 268: 1046-1052

References

Merdes A, Brunkener M, Horstman H, Georgatos SD (1991) Filensin: a new vimentin-binding, polymerization-competent, and membrane-associated protein of the lens fiber cell. J Cell Biol 115: 397-410

Merdes A, Gounari F, Georgatos SD (1993) The 47-kD lens-specific protein phakinin is a tailless intermediate filament protein and an assembly partner of filensin. J Cell Biol 123: 1507-1516

Miesbauer LR, Zhou X, Yang Z, Yang Z, Sun Y, Smith DL, Smith JB (1994) Post-translational modifications of water-soluble human lens crystallins from young adults. J Biol Chem 269: 12494-12502

Minassian DC, Mehra V (1990) 3.8 million blinded by cataract each year: projections from the first epidemiological study of incidence of cataract blindness in India. Br J Ophthalmol 74: 341-343

Miron T, Vancompernolle K, Vandekerckhove J, Wilchek M, Geiger B (1991) A 25-kD inhibitor of actin polymerization is a low molecular mass heat shock protein J Cell Biol 114: 255-261

Miron T, Wilchek M, Geiger B (1988) Characterization of an inhibitor of actin polymerization in vinculin-rich fraction of turkey gizzard smooth muscle. Eur J Biochem 178: 543-553

References

Mkrtchian S, Fang C, Hellman U, Ingelman-Sundberg M (1998) A stress-inducible rat liver endoplasmatic reticulum protein ERp29. Eur J Biochem 251: 304-313

Mkrtchian S, Sandalova T (2006) ERp29, an unusual redox-inactive member of the thioredoxin family. Antioxid Redox Signal 8: 325-337

Mörner CT (1894) Untersuchung der Proteinsubstanzen in den leichtbrechenden Medien des Auges. Z Physiol Chem 18: 61-106

Moulder R, Filen JJ, Salmi J, Katajamaa M, Nevalainen OS, Oresic M, Aittokallio T, Lahesmaa R, Nyman TA (2005) A comparative evaluation of software for the analysis of liquid chromatography-tandem mass spectrometry data from isotope coded affinity tag experiments. Proteomics 5: 2748-2760

Nakamura Y, Fukiage C, Shih M, Ma H, David LL, Azuma M, Shearer TR (2000) Contribution of calpain Lp82-induced proteolysis to experimental cataractogenesis in mice. Invest Ophthalmol Vis Sci 41: 1460-1466

Nicholl JP, Quinlan RA (1994) Chaperone activity of alpha-crystallins modulates intermediate filament assembly. EMBO J 13: 945-953

Norledge BV, Mayr EM, Glockshuber R, Bateman OA; Slingsby C, Jaenicke R, Driessen HP (1996) The X-ray structures of two mutant crystallin domains shed light on the evolution of multi-domain proteins. Nat Struct Biol 3: 267-274

References

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250: 4007-4021

Olthoff JK, Ly IA, Cotter RJ (1988) A pulsed time-of-flight mass spectrometer for liquid secondary ion mass spectrometry. Rapid Commun Mass Spectrom 2: 171-175

Otto A, Thiede B, Muller EC, Scheler C, Wittmann-Liebold B, Jungblut P. (1996) Identification of human myocardial proteins separated by two-dimensional electrophoresis using an effective sample preparation for mass spectrometry. Electrophoresis 17: 1643-1650

Pal J, Liu X, MacKay D, Shiels A, Berthoud V, Beyer E, Ebihara L (2000) Connexin46 mutations linked to congenital cataract show loss of gap junction channel function. Am J Physiol Cell Physiol 279: C596-C602

Papaconstantinou J (1967) Molecular aspects of lens cell differentiation. Science 156: 338-346

Pasta SY, Raman B, Ramakrishna T, Rao ChM (2003) Role of the conserved SRLFDQFFG region of alpha-crystallin, a small heat shock protein. Effect on oligomeric size, subunit exchange and chaperone-like activity. J Biol Chem 278: 51159-51166

Patton WF (2002) Detection technologies in proteome analysis. J Chromatogr B Analyt Technol Biomed Life Sci 771: 3-31

References

Perng MD, Cairns L, van den Ijssel P, Prescott A, Hutcheson AM, Quinlan RA (1999) Intermediate filament interactions can be altered by HSP27 and alphaB-crystallin. J Cell Sci 112: 2099-2112

Peterson GL (1977) A simplification of the protein assay method of Lowry et al. which is more generally applicable. Anal Biochem 83: 346-356

Piatigorsky J, Fukui HN, Kinoshita JH (1978) Differential metabolism and leakage of protein in an inherited cataract and a normal lens cultured with ouabain. Nature 274: 558-562

Pollard HB, Guy HR, Arispe N, de la Fuente M, Lee G, Rojas EM, Pollard JR, Srivastava M, Zhang-Keck ZY, Merezhinskaya N, Cao-Huy H, Burns AL, Rojas E (1992) Calcium channel and membrane fusion activity of synexin and other members of the Annexin gene family. 62: 15-18

Prescott AR, Sandilands A, Quinlan RA, Wang X, Rong P, Chang B, Gong X (2001) The lens specific intermediate filament protein CP49 as a potential genetic modifier in cataractogenesis. Poster at ARVO Meeting 2001

Pupier S, Leveque C, Marqueze B, Kataoka M, Takahashi M, Seagar MJ (1997) Cysteine string proteins associated with secretory granules of the rat neurohypophysis. J Neurosci 17: 2722-2727

Puri N, Augusteyn RC, Owen EA, Siezen RJ (1983) Immunochemical properties of vertebrate alpha-crystallins. Eur J Biochem 134: 321-326

References

Quinlan RA, Carte JM, Sandilands A, Prescott AR (1996) The beaded filament of the eye lens: an unexpected key to intermediate filament structure and function. Trends Cell Biol 6: 123-126

Rae JL, Bartling C, Rae J, Mathias RT (1996) Dye transfer between cells of the lens. J Membr Biol 150: 89-103

Rafferty NS (1985) Lens morphology, in The ocular lens: structure, function & pathology (Maisel H ed.) Marcel Dekker, Inc., New York 1-60

Ramaekers FC, Boomkens TR, Bloemendal H (1981) Cytoskeletal and contractile structures in bovine lens cell differentiation. Exp Cell Res 35: 454-461

Reed N, Castellini M, Ma H, Shearer T, Duncan M (2003) Protein expression patterns for ubiquitous and tissue specific calpains in the developing mouse lens. Exp Eye Res 76: 433-443

Robinson KR, Patterson JW (1982-83) Localization of steady currents in the lens. Curr Eye Res 2: 843-847

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3: 1154-1169

References

Sanderson J, Marcantonio JM, Duncan G (2000) A human lens model of cortical cataract: Ca2+ induced protein loss, vimentin cleavage and opacification. Invest Ophthalmol Vis Sci 41: 2255-2261

Sandilands A, Prescott AR, Carter JM, Hutcheson AM, Quinlan RA, Richards J, FitzGerald PG (1995) Vimentin and CP49/filensin form distinct networks in the lens which are independently modulated during lens fiber cell differentiation. J Cell Sci 108: 1397-1406

Sandilands A, Prescott AR, Wegener A, Zoltoski RK, Hutcheson AM, Masaki S, Kuszak JR, Quinlan RA (2003) Knockout of the intermediate filament protein CP49 destabilises the lens fibre cell cytoskeleton and decreases lens optical quality, but does not induce cataract. Exp Eye Res 76: 385-391

Sandilands A, Wang X, Hutcheson AM, James J, Prescott AR, Wegener A, Pekny M, Gong X, Quinlan RA (2004) Bfsp2 mutation found in mouse 129 strains causes the loss of CP49´and induces vimentin-dependent changes in the lens fibre cell cytoskeleton. Exp Eye Res 78: 875-889

Scales SJ, Hesser BA, Masuda ES, Scheller RH (2002) Amisyn, a novel syntaxin-binding protein that may regulate SNARE complex assembly. J Biol Chem 277: 28271-28279

References

Schmidt F, Donahoe S, Hagens K, Mattow J, Schaible UE, Kaufmann SH, Aebersold R, Junfblut PR (2004) Complementary analysis of the Mycobacterium tuberculosis proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. Mol Cell Proteomics 3: 24-42

Schoenmakers JG, Gerding JJ, Bloemendal H (1969) The subunit structure of alpha-crystallin. Isolation and characterization of the S-carboxymethylated acidic subunits from adult and embryonic origin. Eur J Biochem 11: 472-481

Senn M, Venkataraghavan R, McLafferty FW (1966) Mass spectrometric studies of peptides. 3. Automated determination of amino acid sequences. J Am Chem Soc 88: 5593-5597

Shearer TR, Ma H, Shih M, Fukiage C, Azuma M (1999) Calpains in the lens of the eye. Wang KKW, Yeun P–W, eds. *CALPAIN: Pharmacology and Toxicology of Calcium-Dependent Protease*, 331-347 Taylor & Francis Philadelphia

Shinohara T, Robinson EA, Appella E, Piatigorsky J (1982) Multiple gamma-crystallins of the mouse lens: fractionation of mRNAs by cDNA cloning. Proc Natl Acad Sci USA 79: 2783-2787

Sigler PB, Xu Z, Rye HS, Burston, SG, Fenton WA, Horwich AL (1998) Structure and function in GroEL-mediated protein folding. Annu Rev Biochem 67: 581-608

Singh K, Groth-Vasselli B, Farnsworth PN (1998) Interaction of DNA with bovine lens alpha-crystallin: its functional implications. Int J Biol Macromol 22: 315-320

Slavotinek AM, Tifft CJ (2002) Fraser syndrome and cryptophthalmos: review of the diagnostic criteria and evidence for phenotypic modules in complex malformation syndromes. J Med Genet 39: 623-633

Slingsby C (1985) Structural variation in lens crystallins. Trends Biochem Sci 10: 281-284

Slingsby C, Bateman OA (1990) Quaternary interactions in eye lens beta-crystallins: basic and acidic subunits of beta-crystallins favor heterologous association. Biochemistry 29: 6592-6599

Slingsby C, Norledge B, Simpson A, Bateman OA, Wright G, Driessen HPC, Lindley PF, Moss DS, Bax B (1996) X-ray diffraction and structure of crystallins. Prog Retin Eye Res 16: 3-29

Smith JB, Liu Y, Smith DL (1996) Identification of possible regions of chaperone activity in lens alpha-crystallin. Exp Eye Res 63: 125-128

Smolka MB, Zhou H, Purkayastha S, Aebersold R (2001) Optimization of the isotope-coded affinity tag-labeling procedure for quantative proteome analysis. Anal Biochem 297: 25-31

Solito E, Kamal A, Russo-Marie F, Buckingham JC, Marullo S, Perretti M (2003) A novel calcium dependent proapoptotic effect of annexin 1 on human neutrophils. FASEB J 17: 1544-1546

References

Srinivasan AN, Nagineni CN, Bhat SP (1992) alpha A-crystallin is expressed in non-ocular tissues. J Biol Chem 267:23337-23341

Srivastava SK, Wang LF, Ansari NH, Bhatnagar A (1997) Calcium homeostasis of isolated single cortical fibers of rat lens. Invest Ophthalmol Vis Sci 38: 2300-2312

Stanley EF, Mirotznik RR (1997) Cleavage of syntaxin prevents G-protein regulation of presynaptic calcium channels. Nature 385: 340-343

Swanson SK, Washburn MP (2005) The continuing evolution of shotgun proteomics. Drug Discov Today 10: 719-725

Sze SK, Ge Y, Oh H, McLafferty FW (2002) Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue. Proc Natl Acad Sci USA 99: 1774-1779

Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T (1988) Protein and polymer analysis up to m/z 100,000 by laser ionization time of flight mass spectrometry. Rapid Commun Mass Spectrometry 2: 151-153

Tardieu A, Laporte D, Licinio P, Krop B, Delaye M (1986) Calf lens alpha-crystallin quaternary structure. A three-layer tetrahedral model. J Mol Biol 192:711-724

Thiede B, Hohenwarter W, Krah A, Mattow J, Schmid M, Schmidt F, Jungblut PR (2005) Peptide mass fingerprinting. Methods 35: 237-247

Toyofuku H, Bentley PJ (1970) The effects of decapsulation on ion movements across the lens of the toad, Bufo marinus. Influence on drug actions. Invest Ophthalmol 9: 959-965

Ueda Y, Duncan MK, David LL (2002) Lens proteomics: the accumulation of crystallin modifications in the mouse lens with age. Invest Ophthalmol Vis Sci 43: 205-215

van den Oetelaar PJ, van Someren PF, Thomson JA, Siezen RJ, Hoenders HJ (1990) A dynamic quaternary structure of bovine alpha-crystallin as indicated from intermolecular exchange of subunits. Biochemistry 29: 3488-3493

van der Ouderaa FJ, de Jong WW, Bloemendal H (1973) The amino-acid sequence of the alphaA2 chain of bovine alpha-crystallin. Eur J Biochem 39: 207-222

van Rens GL, Raats JM, Driessen HP, Oldenburg m, Wijnen JT, Khan PM, de Jong WW, Bloemendal H (1989) Structure of the bovine eye lens gamma s-crystallin gene (formerly beta s). Gene 78: 225-233

Wang K, Spector A (2000) alpha-crystallin prevents irreversible protein denaturation and acts cooperatively with other heat-shock proteins to renature the stabilized partially denatured protein in an ATP-dependent manner. Eur J Biochem 267: 4705-4712

Washburn MP, Wolters D, Yates JR 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19: 242-247

References

Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphrey-Smith I (1995) Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis 16: 1090-1094

Wheeler DB, Randall A, Tsien RW (1994) Roles of N-type and Q-type Ca2+ channels in supporting hippocampal synaptic transmission. Science 264: 197-111

Wickner S, Maurizi MR, Gottesman S (1999) Posttranslational quality control: folding, refolding, and degrading proteins. Science 286: 1888-1893

Wistow G, Slingsby C, Blundell T, Driessen H, DeJong W, Bloemendal H (1981) Eye-lens proteins: the three-dimensional structure of beta-crystallin predicted from monomeric gamma-crystallin. FEBS Lett 133: 9-16

Wistow GJ, Piatigorsky J (1988) Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. Annu Rev Biochem 57: 479-504

Wolters DA; Washburn MP, Yates JR 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 73: 5683-5690

Yates JR 3rd (1998) Mass spectrometry and the age of the proteome. J Mass Spectrom 33: 1-19

References

Yoshida A, Oho C, Omori A, Kuwahara R, Ito T, Takahashi M (1992) HPC-1 is associated with synaptotagmin and omega-conotoxin receptor. J Biol Chem 267: 24925-24928

Yu NT, DeNagel DC, Pruett PL, Kuck Jr, JFR (1985) Disulphide bond formation in the eye lens. Proc Natl Acad Sci USA 82: 207-214

Zarina S, Slingsby C, Jaenicke R, Zaidi ZH, Driessen H, Srinivasan N (1994) Three-dimensional model and quaternary structure of the human eye lens protein gamma S-crystallin based on beta- and gamma-crystallin X-ray coordinates and ultracentrifugation. Protein Sci 3: 1840-1846

6. Appendix

Supplementary Table 1. Summary of peptide sequence suggestions, protein identification suggestions and ICAT quantification in the urea soluble lens protein fraction of the *Mus musculus* and *Mus spretus* species at 14 weeks of age with the Thermo Finnigan system

Protein Identification Suggestion Number	Protein Name and Matched Scans	SEQUEST Score	Accession Number	Peptide Sequence Suggestion	MH^+ in Da	Charge	XC	ΔCn	Sp	RSp	Ions	ICAT *Mus musculus/ Mus spretus*
#1	Alpha A crystallin	50.24	62201857[a]									
	3867 - 3870			R.LPSNVDQSALSCSLSADGM#LTFSGPK.V	2868.38	3	4.37	0.38	870.8	1	3/10	0.94
	3873 - 3879			R.LPSNVDQSALSC*SLSADGM#LTFSGPK.V	2877.38	3	4.30	0.43	813.3	1	31/100	0.94
	4377 - 4385			R.LPSNVDQSALSCSLSADGMLTFSGPK.V	2852.38	3	4.87	0.45	1264.1	1	31/100	1.05
	4394			R.LPSNVDQSALSC*SLSADGMLTFSGPK.V	2861.38	3	4.75	0.48	1267.3	1	29/100	1.05
	3760			R.YRLPSNVDQSALSCSLSADGM#LTFSGPK.V	3187.55	3	2.38	0.27	552.7	1	13/54	1.05
#2	Crystallin, alpha A [Mus musculus]	50.24	20899400[b]									
	3867 - 3870			R.LPSNVDQSALSCSLSADGM#LTFSGPK.V	2868.38	3	4.37	0.38	870.8	1	3/10	0.94
	3873 - 3879			R.LPSNVDQSALSC*SLSADGM#LTFSGPK.V	2877.38	3	4.30	0.43	813.3	1	31/100	0.94
	4377 - 4385			R.LPSNVDQSALSCSLSADGMLTFSGPK.V	2852.38	3	4.87	0.45	1264.1	1	31/100	1.05
	4394			R.LPSNVDQSALSC*SLSADGMLTFSGPK.V	2861.38	3	4.75	0.48	1267.3	1	29/100	1.05
	3760			R.YRLPSNVDQSALSCSLSADGM#LTFSGPK.V	3187.55	3	2.38	0.27	552.7	1	13/54	1.05
#3	Beta A3/A1crystallin	70.23	117406[c]									
	1562			R.M#EFTSSCPNVSER.N	1729.76	2	4.21	0.64	982.5	1	17/24	1.54
	1568 - 1571			R.M#EFTSSCPNVSER.N	1729.76	2	4.61	0.60	1169.2	1	3/4	1.54
	1574 - 1577			R.M#EFTSSC*PNVSER.N	1738.76	2	3.56	0.53	791.7	1	5/8	1.54
	1952			R.MEFTSSC*PNVSER.N	1722.76	2	3.19	0.63	374.4	1	13/24	1.59
	1955 - 1961			R.MEFTSSCPNVSER.N	1713.76	2	4.13	0.61	696.1	1	17/24	1.59
	1964			R.MEFTSSC*PNVSER.N	1722.76	2	3.64	0.66	322.0	1	1/2	1.59
	1553			R.M#EFTSSC*PNVSER.N	1738.76	2	2.67	0.31	313.1	1	1/2	1.59
#4	Beta A2 crystallin	70.20	10946978									
	2763			R.VTLFEGENFQGC*K.F	1707.82	2	2.84	0.39	685.8	1	7/12	1.18
	2769			R.VTLFEGENFQGCK.F	1698.82	2	3.76	0.38	1344.1	1	3/4	1.18
	2772			R.VTLFEGENFQGC*K.F	1707.82	2	3.76	0.51	941.1	1	2/3	1.18
	2775			R.VTLFEGENFQGCK.F	1698.82	2	3.95	0.35	1156.1	1	17/24	1.18
	2778			R.VTLFEGENFQGC*K.F	1707.82	2	3.49	0.51	614.5	1	13/24	1.18
	2781			R.VTLFEGENFQGCK.F	1698.82	2	3.76	0.34	1010.4	1	2/3	1.18
	1838 - 1844			R.LLSDCANVCER.G	1676.82	2	2.87	0.39	573.0	1	7/10	1.04
#5	Beta B2 crystallin	10.15	50401872									
	3421			K.AGSVLVQAGPWVGYEQANCK.G	2304.15	3	3.10	0.52	835.0	1	13/38	1.72
#6	Crystallin, beta B2; betaB2-crystallin; Philly cataract [Mus musculus]	10.15	6681035[b]									
	3421			K.AGSVLVQAGPWVGYEQANCK.G	2304.15	3	3.10	0.52	835.0	1	13/38	1.72
#7	Gamma A crystallin	20.19	6724317									
	3305 - 3311			R.CYECSSDCPNLQTYFSR.C	2697.18	3	3.73	0.43	1030.5	1	7/16	1.08
	3314			R.C*YEC*SSDC*PNLQTYFSR.C	2724.18	3	3.41	0.41	1000.9	1	25/64	1.08
#8	I48353 gamma 4-crystallin - mouse (fragment)	20.19	2137312[b]									
	3305 - 3311			R.CYECSSDCPNLQTYFSR.C	2697.18	3	3.73	0.43	1030.5	1	7/16	1.08
	3314			R.C*YEC*SSDC*PNLQTYFSR.C	2724.18	3	3.41	0.41	1000.9	1	25/64	1.08
#9	Gamma B crystallin	30.23	2507570									
	3305 - 3311			R.CYECSSDCPNLQTYFSR.C	2697.18	3	3.73	0.43	1030.5	1	7/16	1.08
	3314			R.C*YEC*SSDC*PNLQTYFSR.C	2724.18	3	3.41	0.41	1000.9	1	25/64	1.08
	2429 - 2440			R.GQM#SEITDDCLSLQDR.F	2053.93	2	4.53	0.51	1222.6	1	2/3	1.45
#10	Gamma C crystallin	30.19	6681037									
	3305 - 3311			R.CYECSSDCPNLQTYFSR.C	2697.18	3	3.73	0.43	1030.5	1	7/16	1.08
	3314			R.C*YEC*SSDC*PNLQTYFSR.C	2724.18	3	3.41	0.41	1000.9	1	25/64	1.08
	2286			K.GVM#M#ELSEDCSCIQDR.F	2302.00	3	2.61	0.38	690.5	1	1/3	1.23
#11	I49614 gamma-C-crystallin - mouse (fragment)	30.19	2137313[b]									

Continued...

	3305 - 3311			R.CYECSSDCPNLQTYFSR.C	2697.18	3	3.73	0.43	1030.5	1	7/16	1.08
	3314			R.C*YEC*SSDC*PNLQTYFSR.C	2724.18	3	3.41	0.41	1000.9	1	25/64	1.08
	2286			K.GVM#M#ELSEDCSCIQDR.F	2302.00	3	2.61	0.38	690.5	1	1/3	1.23
#12	I49617 gamma-F-crystallin - mouse (fragment)	20.21	2137316[d]									
	1823 - 1826			R.HYEC*STDHSNLQPYFSR.C	2320.02	3	3.24	0.41	543.1	1	23/64	1.92
	1802 - 1808			R.HYECSTDHSNLQPYFSR.C	2311.02	3	4.12	0.62	1101.1	1	27/64	1.92
#13	Crystallin, gamma E; Crystallin, gamma polypeptide 5 [Rattus norvegicus]	20.21	27545356[d]									
	1823 - 1826			R.HYEC*STDHSNLQPYFSR.C	2320.02	3	3.24	0.41	543.1	1	23/64	1.92
	1802 - 1808			R.HYECSTDHSNLQPYFSR.C	2311.02	3	4.12	0.62	1101.1	1	27/64	1.92
#14	CRGF_MOUSE Gamma crystallin F	30.21	2493884[d]									
	1802 - 1808			R.HYECSTDHSNLQPYFSR.C	2311.02	3	4.12	0.62	1101.1	1	27/64	1.92
	1823 - 1826			R.HYEC*STDHSNLQPYFSR.C	2320.02	3	3.24	0.41	543.1	1	23/64	1.92
	1652			R.GQM#VEITDDCSHLQDR.F	2089.94	3	2.30	0.07	508.7	1	1/3	3.13
#15	JS0596 gamma-E-crystallin - mouse	30.21	90505[d]									
	1802 - 1808			R.HYECSTDHSNLQPYFSR.C	2311.02	3	4.12	0.62	1101.1	1	27/64	1.92
	1823 - 1826			R.HYEC*STDHSNLQPYFSR.C	2320.02	3	3.24	0.41	543.1	1	23/64	1.92
	1652			R.GQM#VEITDDCSHLQDR.F	2089.94	3	2.30	0.07	508.7	1	1/3	3.13
#16	Crystallin, gamma E [Mus musculus]	30.21	6681041[d]									
	1802 - 1808			R.HYECSTDHSNLQPYFSR.C	2311.02	3	4.12	0.62	1101.1	1	27/64	1.92
	1823 - 1826			R.HYEC*STDHSNLQPYFSR.C	2320.02	3	3.24	0.41	543.1	1	23/64	1.92
	1652			R.GQM#VEITDDCSHLQDR.F	2089.94	3	2.30	0.07	508.7	1	1/3	3.13
#17	Fatty acid-binding protein, epidermal (E-FABP)	10.12	6754450									
	2037 - 2040			K.TTVFSC*NLGEK.F	1434.71	2	2.35	0.38	486.8	1	3/5	0.8
#18	UDP-glucuronosyltransferase 2 family, member 5 [Mus musculus]	10.07	6678501									
	4607			K.C*LLFIYR.F	1163.64	2	1.48	0.04	247.3	2	7/12	0.24
#19	Similar to zinc finger protein Peg3 - mouse [Rattus norvegicus]	10.06	27675518									
	4095			K.PFECGSEM#RQAMSM#GNLR.N	2303.02	3	1.23	0.08	414.8	1	19/68	not valid
#20	B42026 cyclic AMP response element DNA-binding protein isoform 1 - mouse	34.07	284734									
	4119 - 4130			K.QLLLAHK.D	822.52	2	1.17	0.13	694.8	1	11/12	
#21	Activating transcription factor 2 [Mus musculus]	34.07	20837284									
	4119 - 4130			K.QLLLAHK.D	822.52	2	1.17	0.13	694.8	1	11/12	
#22	C42026 cyclic AMP response element DNA-binding protein isoform 2 - mouse	34.07	284735									
	4119 - 4130			K.QLLLAHK.D	822.52	2	1.17	0.13	694.8	1	11/12	
#23	A34785 DNA-binding protein mXBP - mouse (fragment)	34.07	109793									
	4119 - 4130			K.QLLLAHK.D	822.52	2	1.17	0.13	694.8	1	11/12	
#24	A42026 cAMP response element-binding protein 3 - mouse	34.07	284736									
	4119 - 4130			K.QLLLAHK.D	822.52	2	1.17	0.13	694.8	1	11/12	
#25	Carbonyl reductase 1; carbonyl reductase [Mus musculus]	20.36	6671690									
	3644			K.ELLPLIK.P	825.55	2	1.46	0.06	397.8	11	3/4	
	3638			K.ELLPLIK.P	825.55	2	1.23	0.09	280.2	13	2/3	
	3650			K.ELLPLIK.P	825.55	2	1.31	0.08	234.6	17	2/3	
#26	ANR5_MOUSE Ankyrin repeat domain protein 5	20.28	20137465									
	4673			R.LENLQIYR.V	1048.58	2	1.36	0.23	286.7	4	4/7	
	4308 - 4313			R.LENLQIYR.V	1048.58	2	1.40	0.15	343.8	9	9/14	
#27	Cofilin 1, non-muscle [Mus musculus]	20.11	6680924									
	2280			K.ESKKEDLVFIFWAPENAPLK.S	2361.25	3	2.22	0.01	394.8	1	5/19	
#28	I49528 hypothetical protein - mouse (fragment)	18.47	2137411									
	4262 - 4268			K.IKPQAR.T	712.45	2	1.01	0.05	169.8	7	3/5	
#29	Platelet-activating factor acetylhydrolase, isoform 1b, alpha1 subunit	18.10	6679201									
	4262 - 4268			K.LQPQAR.V	712.41	2	1.01	0.08	169.8	7	3/5	
#30	Myelocytomatosis oncogene [Mus musculus]	16.57	27545183									
	4435			K.RSESGSSPSR.G	1049.50	2	1.42	0.10	272.9	2	1/2	

Continued...

	4441			K.RSESGSSPSR.G	1049.50	2	1.22	0.21	272.6	3	1/2
#31	Toll-like receptor 4; lipopolysaccharide response [Mus musculus]	16.35	10946594								
	4694			R.QQVELYR.L	935.50	2	1.04	0.05	208.8	1	2/3
	4397 - 4403			R.QQVELYR.L	935.50	2	1.30	0.11	175.7	16	1/2
#32	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 9	12.06	24429590								
	4572			R.ELLPVKK.F	826.54	2	1.24	0.04	232.1	7	2/3
	4253 - 4259			R.ELLPVKK.F	826.54	2	1.20	0.23	222.6	16	3/4
#33	COQ6_MOUSE Ubiquinone biosynthesis monooxgenase COQ6 r,	10.24	26006706								
	4615			K.LLIGADGHK.S	923.53	2	1.08	0.15	382.4	1	5/8
#34	WDRD_MOUSE WD-repeat protein 13	10.14	20140638								
	1140 - 1146			R.AAKAAVNK.L	772.47	2	1.18	0.20	290.2	1	9/14
#35	Cytoplasmic linker 2; cytoplasmic linker protein 1, 115 kDa [Mus musculus]	10.09	6753562								
	3180 - 3186			K.EIALLK.A	686.45	1	1.12	0.04	338.2	2	7/10
#36	1-acylglycerol-3-phosphate O-acyltransferase 1	10.08	9256517								
	4708			R.LLLLHAK.Y	807.55	2	1.52	0.23	194.6	8	7/12
#37	General transcription factor II H, polypeptide 1 (62kD subunit)	10.07	6680125								
	4406			K.DLLQQLLPK.F	1067.65	2	1.43	0.01	350.2	3	11/16
#38	Hormonally upregulated Neu-associated kinase [Mus musculus]	10.06	7657216								
	4288			K.LDKNLPSHK.Q	1051.59	2	1.21	0.06	225.6	1	5/8
#39	Kinesin family member 9; kinesin 9 [Mus musculus]	10.06	6754442								
	3180 - 3186			K.ELALLK.Q	686.45	1	1.12	0.00	338.2	2	7/10
#40	Ligase III, DNA, ATP-dependent [Mus musculus]	10.05	6754546								
	4359			R.GGIKPIPK.H	809.52	2	1.06	0.11	295.4	3	9/14
#41	Similar to endoplasmic reticulum resident protein 58	6.20	27672320								
	1750 - 1756			K.HAILVKSNLSDLLEKLK.W	1921.15	3	1.58	0.14	539.3	3	23/64
#42	Telomeric repeat binding factor 1 [Mus musculus]	4.16	6678287								
	2639			K.SSTFLM#KAATK.V	1200.63	2	1.29	0.16	257.5	20	11/20
#43	ITN2_MOUSE Intersectin 2 (SH3 domain-containing protein 1B)	4.06	20138801								
	3969 - 3974			K.ELIMSNTK.L	935.49	2	1.18	0.06	214.1	13	1/2

[a] The proteins Alpha A and Alpha A insert crystallin could not be discriminated as they have the same single cysteine containing peptide. The Alpha A crystallin accession number is listed.

[b] The protein sequence with this accession number is different from another protein sequence with another accession number for the same protein detected in the 2-DE analysis of the urea soluble lens proteins of wild type C57BL/6J mice listed in Supplementary Table 4. The polymorphism site is not covered by mass spectrometric data indicating that the sequence discrepancy could be due to software processing and scoring.

[c] The proteins could not be discriminated.

[d] These Gamma E and Gamma F crystallin sequences are different from the Gamma E and Gamma F crystallin sequences identified in the 2-DE analysis of the urea soluble lens proteins of wild type C57BL/6J mice listed in Supplementary Table 4. These protein sequences and ultimately Gamma E and Gamma F crystallin could not be discriminated.

* Indicates the cysteine residue is labeled with the C13 or heavy version of the cleavable ICAT reagent.

# Indicates the methionine residue is oxidated.

Supplementary Table 2. Summary of peptide sequence suggestions, protein identification suggestions and ICAT quantification in the urea soluble lens protein fraction of the *Mus musculus* and *Mus spretus* species at 14 weeks of with the Agilent Technologies system (first of five analysis)

Protein Number	LC-MS/MS Scan	Peptide Sequence Suggestion	M/Z in Da	Charge	MH^+ in Da[a]	ΔM in Da[a]	Accession Number	Protein Name	Cys	ICAT *Mus musculus*/ *Mus spretus*
1	001-0102.7768.7768.0	(R)YRLPSNVDQSALSCSLSADGMLTFSGPK(V)	1057.75	3	3171.546	-0.311	NA	Alpha A/A insert crystallin[b]	cICAT-C12	4.939
	001-0102.6969.7011.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1062.97	3	3171.546	15.349	NA	Alpha A/A insert crystallin[b]	cICAT-C12	8.029
2	001-0102.3925.3969.2	(R)MEFTSSCPNVSER(N)	862.06	2	1722.788	0.324	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	2.054
	001-0102.3288.3304.2	(R)mEFTSSCPNVSER(N)	865.43	2	1713.758	16.095	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	1.063
	001-0102.3292.3321.3	(R)mEFTSSCPNVSER(N)	577.26	3	1713.758	16.007	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	0.967
	001-0102.3297.3309.2	(R)mEFTSSCPNVSER(N)	869.92	2	1722.788	16.044	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.024
	001-0102.3049.3105.3	(K)RmEFTSSCPNVSER(N)	632.27	3	1878.889	15.905	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.078
	001-0102.3061.3103.3	(K)RmEFTSSCPNVSER(N)	629.02	3	1869.859	15.186	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	1.056
	001-0102.4300.4324.3	(R)GYQYILECDHHGGDYK(H)	712.18	3	2133.975	0.549	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	0.790
3	001-0102.3622.3649.2	(R)LLSDCANVCER(G)	838.75	2	1676.810	-0.318	10946978	Beta A2 crystallin	cICAT-C12	0.797
	001-0102.5266.5299.2	(R)VTLFEGENFQGCK(F)	850.12	2	1698.816	0.416	10946978	Beta A2 crystallin	cICAT-C12	1.084
	001-0102.5278.5329.3	(R)VTLFEGENFQGCK(F)	570.00	3	1707.846	0.138	10946978	Beta A2 crystallin	cICAT-C13	0.852
4	001-0102.7576.7576.0	(R)HEFTAECPSVLELGFETVR(S)	800.95	3	2400.196	0.639	10946672	Beta A4 crystallin	cICAT-C13	0.583
5	001-0102.5941.5946.0	(K)ICLFEGANFK(G)	689.28	2	1377.729	-0.177	12963789	Beta B1 crystallin	cICAT-C13	0.774
	001-0102.4987.5022.3	(R)RVEFSGECLNLGDR(G)	607.75	3	1821.892	-0.657	12963789	Beta B1 crystallin	cICAT-C12	0.847
	001-0102.4977.5004.3	(R)RVEFSGECLNLGDR(G)	610.80	3	1830.922	-0.538	12963789	Beta B1 crystallin	cICAT-C13	0.848
6	001-0102.6142.6153.3	(K)AGSVLVQAGPWVGYEQANCK(G)	768.53	3	2304.145	-0.570	50401872	Beta B2 crystallin	cICAT-C12	0.950
	001-0102.5917.5976.0	(K)IIIFEQENFQGHSHELSGPCPNLK(E)	989.03	3	2964.468	0.607	50401872	Beta B2 crystallin	cICAT-C12	36.812
	001-0102.5920.5920.0	(K)IIIFEQENFQGHSHELSGPCPNLK(E)	991.96	3	2973.498	0.366	50401872	Beta B2 crystallin	cICAT-C13	0.041
	001-0102.5991.6015.0	(K)IIIFEQENFQGHSHELSGPCPNLK(E)	988.88	3	2964.468	0.157	50401872	Beta B2 crystallin	cICAT-C12	14.409
7	001-0102.5776.5796.3	(K)RCELSAECPNLTDSLLEK(V)	831.88	3	2493.283	0.342	10946674	Beta B3 crystallin	cICAT-C13	0.072
8	001-0102.3996.4050.3	(R)SCCLIPQHSGTYR(M)	640.31	3	1918.927	-0.012	NA	Gamma A/B/C crystallin[b]	cICAT-C12	50.910
9	001-0102.3555.3574.3	(R)GQmVEITDDCSHLQDR(F)	697.06	3	2073.933	15.231	NA	Gamma E/F crystallin[b]	cICAT-C12	1.480
	001-0102.4086.4120.3	(R)HYECSTDHSNLQPYFSR(C)	770.77	3	2311.020	-0.726	NA	Gamma E/F crystallin[b]	cICAT-C12	1.071
	001-0102.4137.4194.4	(R)HYECSTDHSNLQPYFSR(C)	578.47	4	2311.020	-0.164	NA	Gamma E/F crystallin[b]	cICAT-C12	1.181
	001-0102.4159.4215.3	(R)HYECSTDHSNLQPYFSR(C)	774.00	3	2320.050	-0.066	NA	Gamma E/F crystallin[b]	cICAT-C13	1.212
10	001-0102.4233.4237.0	(K)TTVFSCNLGEK(F)	475.94	3	1425.705	0.100	6754450	Fatty acid-binding protein, epidermal (E-FABP)	cICAT-C12	7.523
11	001-0102.7560.7560.0	(K)LPRPRDLQPFPVCQALVYRGHSDLVR(C)	1090.53	3	3268.794	0.780	15215181	Block of proliferation 1	cICAT-C13	0.038
12	001-0102.3256.3268.0	(K)NEAIQAAHDSVAQEGQCR(V)	718.76	3	2154.000	0.265	18203410	Ubiquitin carboxyl-terminal hydrolase isozyme L1	cICAT-C12	12.250
13	001-0102.4953.4971.3	(R)VLQNFTVQPCK(E)	504.97	3	1512.830	0.065	6681115	Cytochrome P450 3A13 (CYPIIIA13)	cICAT-C13	0.069
14	001-0102.3742.3747.0	(K)SVVKVLQtCQRTR(Q)	611.84	3	1754.016	79.489	28893277	Unnamed protein product	cICAT-C13	0.231
15	001-0102.4561.4561.0	(K)QmFTSLSFmGyMTKCLKK(Q)	828.39	3	2371.169	111.986	25052332	Hypothetical protein XP_196974	cICAT-C12	1.431
16	001-0102.5139.5160.2	(K)CNsMQsEYR(E)	757.49	2	1353.598	160.374	6678177	Syntaxin 4	cICAT-C13	2.416
17	001-0102.5221.5221.0	(K)VmESNCSK(L)	575.46	2	1133.538	16.374	26349145	Unnamed protein product	cICAT-C13	1.165

[a] The mass M refers to the mass of the unmodified cleavable ICAT labeled peptide sequence.
[b] The proteins could not be discriminated.
m Indicates the methionine residue is oxidated.
s Indicates the serine residue is phosphorylated.
t Indicates the threonine residue is phosphorylated.
y Indicates the tryptophane residue is phosphorylated.

Supplementary Table 3. Summary of peptide sequence suggestions, protein identification suggestions and ICAT quantification in the urea soluble lens protein fraction of the *Mus musculus* and *Mus spretus* species at 14 weeks of age with the Agilent Technologies system (analysis 2 through 5 of the same reaction stock)

Protein Number	LC-MS/MS Run	LC-MS/MS Scan	Peptide Sequence Suggestion	M/Z in Da	Charge	MH^+ in Da[a]	ΔM in Da[a]	Accession Number	Protein Name	Cys	ICAT *Mus musculus*/ *Mus spretus*
	2	001-0101.7559.7571.0	(R)LPSNVDQSALSCSLSADGMLTFSGPK(V)	951.44	3	2852.381	-0.077	NA	Alpha A/A insert crystallin[b]	cICAT-C12	3.428
	2	001-0101.6670.6695.3	(R)LPSNVDQSALSCSLSADGmLTFSGPK(V)	957.07	3	2852.381	16.813	NA	Alpha A/A insert crystallin[b]	cICAT-C12	1.102
	3	001-0103.6765.6805.0	(R)LPSNVDQSALSCSLSADGmLTFSGPK(V)	960.04	3	2861.411	16.693	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.021
	3	001-0103.6826.6855.0	(R)LPSNVDQSALSCSLSADGmLTFSGPK(V)	960.03	3	2861.411	16.663	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.009
	3	001-0103.7803.7803.0	(R)YRLPSNVDQSALSCSLSADGMLTFSGPK(V)	1060.61	3	3180.576	-0.761	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.041
	5	40804010.8772.8772.0	(R)YRLPSNVDQSALSCSLSADGMLTFSGPK(V)	1060.70	3	3180.576	-0.491	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.139
1	2	001-0101.6859.6895.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1063.40	3	3171.546	16.639	NA	Alpha A/A insert crystallin[b]	cICAT-C12	11.270
	2	001-0101.6890.6917.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1066.34	3	3180.576	16.429	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.016
	3	001-0103.6982.7032.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1066.40	3	3180.576	16.609	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.027
	3	001-0103.6999.7048.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1063.26	3	3171.546	16.219	NA	Alpha A/A insert crystallin[b]	cICAT-C12	6.539
	3	001-0103.7057.7057.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1066.05	3	3180.576	15.559	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.008
	3	001-0103.7087.7087.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1063.02	3	3171.546	15.499	NA	Alpha A/A insert crystallin[b]	cICAT-C12	7.380
	5	40804010.7869.7878.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1066.25	3	3180.576	16.159	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.069
	5	40804010.7933.7981.0	(R)YRLPSNVDQSALSCSLSADGmLTFSGPK(V)	1066.00	3	3180.576	15.409	NA	Alpha A/A insert crystallin[b]	cICAT-C13	0.054
	2	001-0101.3827.3883.3	(R)MEFTSSCPNVSER(N)	572.05	3	1713.758	0.377	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	0.377
	2	001-0101.3850.3868.2	(R)MEFTSSCPNVSER(N)	862.18	2	1722.788	0.564	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.199
	2	001-0101.3856.3869.2	(R)MEFTSSCPNVSER(N)	857.57	2	1713.758	0.375	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	1.159
	3	001-0103.3933.3946.2	(R)MEFTSSCPNVSER(N)	857.32	2	1713.758	-0.126	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	1.017
	3	001-0103.3942.3964.2	(R)MEFTSSCPNVSER(N)	861.57	2	1722.788	-0.656	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.009
	5	40804010.4355.4390.2	(R)MEFTSSCPNVSER(N)	857.45	2	1713.758	0.135	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	1.107
	5	40804010.4362.4372.2	(R)MEFTSSCPNVSER(N)	862.22	2	1722.788	0.644	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.182
	2	001-0101.3070.3100.3	(K)RmEFTSSCPNVSER(N)	632.25	3	1878.889	15.845	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	0.952
	3	001-0103.3067.3073.0	(K)RmEFTSSCPNVSER(N)	629.18	3	1869.859	15.666	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	16.722
2	4	40803005.3509.3513.3	(K)RmEFTSSCPNVSER(N)	632.32	3	1878.889	16.055	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	0.988
	4	40803005.3542.3549.0	(K)RmEFTSSCPNVSER(N)	629.36	3	1869.859	16.206	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	5.286
	2	001-0101.3286.3286.2	(R)mEFTSSCPNVSER(N)	869.70	2	1722.788	15.604	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.035
	2	001-0101.3302.3302.2	(R)mEFTSSCPNVSER(N)	865.58	2	1713.758	16.395	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	0.979
	3	001-0103.3265.3301.2	(R)mEFTSSCPNVSER(N)	869.95	2	1722.788	16.104	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	0.933
	3	001-0103.3277.3283.2	(R)mEFTSSCPNVSER(N)	864.94	2	1713.758	15.115	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	1.035
	4	40803005.3746.3750.3	(R)mEFTSSCPNVSER(N)	580.30	3	1722.788	16.096	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	0.904
	5	40804010.3776.3776.3	(R)mEFTSSCPNVSER(N)	580.56	3	1722.788	16.877	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	1.338
	2	001-0101.4142.4142.3	(R)GYQYILECDHHGGDYK(H)	708.71	3	2124.945	-0.831	20304089	Beta A3/A1 crystallin[b]	cICAT-C12	0.687
	4	40803005.4756.4771.3	(R)GYQYILECDHHGGDYK(H)	712.29	3	2133.975	0.879	20304089	Beta A3/A1 crystallin[b]	cICAT-C13	0.677
	2	001-0101.3580.3584.2	(R)LLSDCANVCER(G)	838.76	2	1676.810	-0.298	10946978	Beta A2 crystallin	cICAT-C12	0.772
	2	001-0101.3587.3587.2	(R)LLSDCANVCER(G)	839.26	2	1676.810	0.702	10946978	Beta A2 crystallin	cICAT-C12	0.905
	2	001-0101.3563.3568.2	(R)LLSDCANVCER(G)	848.12	2	1694.870	0.362	10946978	Beta A2 crystallin	cICAT-C13	0.047
	3	001-0103.3606.3636.2	(R)LLSDCANVCER(G)	847.49	2	1694.870	-0.898	10946978	Beta A2 crystallin	cICAT-C13	0.024
	4	40803005.4079.4079.2	(R)LLSDCANVCER(G)	847.92	2	1694.870	-0.038	10946978	Beta A2 crystallin	cICAT-C13	0.026
	4	40803005.4081.4125.2	(R)LLSDCANVCER(G)	838.91	2	1676.810	0.002	10946978	Beta A2 crystallin	cICAT-C12	0.585
	4	40803005.4146.4173.2	(R)LLSDCANVCER(G)	838.97	2	1676.810	0.122	10946978	Beta A2 crystallin	cICAT-C12	0.702
3	5	40804010.4114.4158.2	(R)LLSDCANVCER(G)	847.96	2	1694.870	0.042	10946978	Beta A2 crystallin	cICAT-C13	0.034
	5	40804010.4176.4180.2	(R)LLSDCANVCER(G)	848.06	2	1694.870	0.242	10946978	Beta A2 crystallin	cICAT-C13	0.029
	2	001-0101.5060.5096.2	(R)VTLFEGENFQGCK(F)	850.20	2	1698.816	0.576	10946978	Beta A2 crystallin	cICAT-C12	0.891
	2	001-0101.5063.5113.3	(R)VTLFEGENFQGCK(F)	570.05	3	1707.846	0.288	10946978	Beta A2 crystallin	cICAT-C13	1.006
	2	001-0101.5072.5125.2	(R)VTLFEGENFQGCK(F)	854.40	2	1707.846	-0.054	10946978	Beta A2 crystallin	cICAT-C13	0.986
	2	001-0101.5074.5114.0	(R)VTLFEGENFQGCK(F)	849.67	2	1698.816	-0.484	10946978	Beta A2 crystallin	cICAT-C12	0.993
	4	40803005.5749.5787.2	(R)VTLFEGENFQGCK(F)	854.54	2	1707.846	0.226	10946978	Beta A2 crystallin	cICAT-C13	0.803

Continued...

	5	40804010.5745.5749.2	(R)VTLFEGENFQGCK(F)	849.91	2	1698.816	-0.004	10946978	Beta A2 crystallin	cICAT-C12	0.732
	5	40804010.5795.5805.2	(R)VTLFEGENFQGCK(F)	854.82	2	1707.846	0.786	10946978	Beta A2 crystallin	cICAT-C13	0.825
	2	001-0101.3167.3226.3	(R)LTSFRPVACANHR(D)	570.27	3	1707.916	0.878	10946672	Beta A4 crystallin	cICAT-C13	1.039
	4	40803005.3697.3717.3	(R)LTSFRPVACANHR(D)	566.94	3	1698.886	-0.082	10946672	Beta A4 crystallin	cICAT-C12	0.859
4	5	40804010.3591.3620.3	(R)LTSFRPVACANHR(D)	567.01	3	1698.886	0.128	10946672	Beta A4 crystallin	cICAT-C12	0.879
	5	40804010.3690.3744.3	(R)LTSFRPVACANHR(D)	567.16	3	1698.886	0.578	10946672	Beta A4 crystallin	cICAT-C12	1.083
	5	40804010.7657.7674.0	(R)RHEFTAECPSVLELGFETVR(S)	849.91	3	2547.267	0.448	10946672	Beta A4 crystallin	cICAT-C12	2.118
	2	001-0101.5659.5686.2	(K)ICLFEGANFK(G)	689.62	2	1377.729	0.503	12963789	Beta B1 crystallin	cICAT-C13	0.967
	3	001-0103.6090.6102.0	(K)ICLFEGANFK(G)	689.61	2	1377.729	0.483	12963789	Beta B1 crystallin	cICAT-C13	0.766
	4	40803005.6376.6405.2	(K)ICLFEGANFK(G)	685.01	2	1368.699	0.314	12963789	Beta B1 crystallin	cICAT-C12	0.772
5	5	40804010.6313.6317.0	(K)ICLFEGANFK(G)	689.80	2	1377.729	0.863	12963789	Beta B1 crystallin	cICAT-C13	0.715
	2	001-0101.4798.4822.3	(R)RVEFSGECLNLGDR(G)	607.93	3	1821.892	-0.117	12963789	Beta B1 crystallin	cICAT-C12	0.740
	3	001-0103.5031.5038.3	(R)RVEFSGECLNLGDR(G)	607.86	3	1821.892	-0.327	12963789	Beta B1 crystallin	cICAT-C12	1.036
	5	40804010.5413.5442.3	(R)RVEFSGECLNLGDR(G)	608.17	3	1821.892	0.603	12963789	Beta B1 crystallin	cICAT-C12	0.806
		40804010.5425.5441.3	(R)RVEFSGECLNLGDR(G)	611.07	3	1830.922	0.272	12963789	Beta B1 crystallin	cICAT-C13	0.799
	2	001-0101.5941.5974.3	(K)AGSVLVQAGPWVGYEQANCK(G)	768.86	3	2304.145	0.420	50401872	Beta B2 crystallin	cICAT-C12	1.142
	4	40803005.6720.6731.0	(K)AGSVLVQAGPWVGYEQANCK(G)	771.85	3	2313.175	0.359	50401872	Beta B2 crystallin	cICAT-C13	0.115
6	2	001-0101.5852.5852.0	(K)IIIFEQENFQGHSHELSGPCPNLK(E)	989.11	3	2964.468	0.847	50401872	Beta B2 crystallin	cICAT-C12	14.896
	3	001-0103.6061.6096.0	(K)IIIFEQENFQGHSHELSGPCPNLK(E)	988.93	3	2964.468	0.306	50401872	Beta B2 crystallin	cICAT-C12	7.552
	5	40804010.6469.6503.0	(K)IIIFEQENFQGHSHELSGPCPNLK(E)	989.09	3	2964.468	0.786	50401872	Beta B2 crystallin	cICAT-C12	7.694
	2	001-0101.5624.5624.0	(K)RCELSAECPNLTDSLLEK(V)	832.06	3	2493.283	0.882	10946674	Beta B3 crystallin	cICAT-C13	0.110
7	3	001-0103.5554.5554.0	(K)RCELSAECPNLTDSLLEK(V)	825.89	3	2475.222	0.432	10946674	Beta B3 crystallin	cICAT-C12	3.038
		001-0103.5895.5907.0	(K)RCELSAECPNLTDSLLEK(V)	825.59	3	2475.222	-0.468	10946674	Beta B3 crystallin	cICAT-C12	15.936
	4	40803005.6320.6320.0	(K)RCELSAECPNLTDSLLEK(V)	831.89	3	2493.283	0.372	10946674	Beta B3 crystallin	cICAT-C13	0.231
	5	40804010.6425.6458.0	(R)CYECSSDCPNLQTYFSR(C)	899.96	3	2697.175	0.690	NA	Gamma A/B/C crystallin[b]	cICAT-C12	-1.000
	2	001-0101.3907.3958.3	(R)SCCLIPQHSGTYR(M)	640.26	3	1918.927	-0.162	NA	Gamma A/B/C crystallin[b]	cICAT-C12	27.522
8		001-0101.3917.3934.2	(R)SCCLIPQHSGTYR(M)	959.95	2	1918.927	-0.035	NA	Gamma A/B/C crystallin[b]	cICAT-C12	57.144
	4	40803005.4419.4429.3	(R)SCCLIPQHSGTYR(M)	640.37	3	1918.927	0.168	NA	Gamma A/B/C crystallin[b]	cICAT-C12	21.175
	5	40804010.4328.4386.3	(R)SCCLIPQHSGTYR(M)	640.42	3	1918.927	0.318	NA	Gamma A/B/C crystallin[b]	cICAT-C12	0.977
	2	001-0101.3499.3512.3	(R)GQmVEITDDCSHLQDR(F)	697.24	3	2073.933	15.771	NA	Gamma E/F crystallin[b]	cICAT-C12	1.551
	2	001-0101.3991.4025.3	(R)HYECSTDHSNLQPYFSR(C)	773.91	3	2320.050	-0.336	NA	Gamma E/F crystallin[b]	cICAT-C13	1.040
	3	001-0103.4110.4146.3	(R)HYECSTDHSNLQPYFSR(C)	770.87	3	2311.020	-0.426	NA	Gamma E/F crystallin[b]	cICAT-C12	1.274
		001-0103.4119.4119.0	(R)HYECSTDHSNLQPYFSR(C)	774.13	3	2320.050	0.324	NA	Gamma E/F crystallin[b]	cICAT-C13	0.043
9		40803005.4507.4535.3	(R)HYECSTDHSNLQPYFSR(C)	774.28	3	2320.050	0.774	NA	Gamma E/F crystallin[b]	cICAT-C13	0.979
	4	40803005.4523.4579.3	(R)HYECSTDHSNLQPYFSR(C)	773.74	3	2320.050	-0.846	NA	Gamma E/F crystallin[b]	cICAT-C13	0.668
		40803005.4628.4639.0	(R)HYECSTDHSNLQPYFSR(C)	771.07	3	2311.020	0.174	NA	Gamma E/F crystallin[b]	cICAT-C12	8.831
		40804010.4481.4505.0	(R)HYECSTDHSNLQPYFSR(C)	771.14	3	2311.020	0.384	NA	Gamma E/F crystallin[b]	cICAT-C12	9.800
	5	40804010.4529.4529.0	(R)HYECSTDHSNLQPYFSR(C)	774.02	3	2320.050	-0.006	NA	Gamma E/F crystallin[b]	cICAT-C13	0.053
		40804010.4586.4642.3	(R)HYECSTDHSNLQPYFSR(C)	771.02	3	2311.020	0.024	NA	Gamma E/F crystallin[b]	cICAT-C12	1.312
	2	001-0101.4048.4070.2	(K)TTVFSCNLGEK(F)	713.38	2	1425.705	0.047	6754450	Fatty acid-binding protein, epidermal (E-FABP)	cICAT-C12	0.817
	4	40803005.4565.4605.2	(K)TTVFSCNLGEK(F)	718.20	2	1434.735	0.657	6754450	Fatty acid-binding protein, epidermal (E-FABP)	cICAT-C13	0.978
10	5	40804010.4555.4597.2	(K)TTVFSCNLGEK(F)	718.10	2	1434.735	0.457	6754450	Fatty acid-binding protein, epidermal (E-FABP)	cICAT-C13	0.687
	2	001-0103.5592.5616.0	(K)TTVFSCNLGEKFDETTADGR(K)	806.61	3	2418.125	-0.310	6754450	Fatty acid-binding protein, epidermal (E-FABP)	cICAT-C12	3.351
	3	40803005.6003.6031.3	(K)TTVFSCNLGEKFDETTADGR(K)	806.86	3	2418.125	0.440	6754450	Fatty acid-binding protein, epidermal (E-FABP)	cICAT-C12	1.118
11	4	40803005.4583.4635.3	(K)TQmAHMTALACLQEQQHK(Q)	774.25	3	2305.130	15.604	21314852	kinesin family member 18A	cICAT-C13	1.038
12	4	40803005.5363.5419.4	(K)KTQCVPATFPLVFNAR(M)	505.42	4	2019.085	-0.428	26343271	unnamed protein product	cICAT-C12	3.814
13	4	40803005.3030.3030.0	(K)DNEHKCSLTK(T)	468.15	3	1401.680	0.755	26353384	unnamed protein product	cICAT-C12	4.999
14	2	001-0101.4925.4925.2	(R)CPGWINIRYR(N)	757.68	2	1513.815	0.537	28495170	hypothetical protein XP_160778	cICAT-C13	2.525
15	2	001-0101.2797.2842.0	(R)CSRPPALPPPK(F)	466.83	3	1398.798	-0.324	30354325	2900002H16Rik protein	cICAT-C13	0.014
16	2	001-0101.4750.4804.3	(R)VLQNFTVQPCK(E)	504.98	3	1512.830	0.095	6681115	Cytochrome P450 3A13 (CYPIIIA13)	cICAT-C13	0.010

[a] The mass M refers to the mass of the unmodified cleavable ICAT labeled peptide sequence.
[b] The proteins could not be discriminated.
m Indicates the methionine residue is oxidated.

Appendix

Supplementary Table 4. Summary of proteins identified in the urea soluble lens protein fraction of the tens day old *Mus musculus* C57BL/6J strain with 2-DE and mass spectrometry

Protein Spot Denominator	ORF	Accession Number NCBI	Accession Number Swiss-Prot	Protein Name	Protein Identification Method	Mascot Score[a]	% Sequence Coverage
SSP 1010	NE	49259325	NE	Coactosin-like protein	MALDI/TOF/TOF-PMF	139	38
SSP 1013	12957	20304089	P02525	Beta A3 crystallin	MALDI/TOF/TOF-PMF	76	29
SSP 1018	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF/TOF-PMF	114	34
SSP 1422	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	193	38
SSP 1528	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	94	23
SSP 1531	87906	74187644	Q3TSB7	Gamma-actin	MALDI/TOF/TOF-PMF	98	29
SSP 1536	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	78	20
SSP 1537	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	66	25
SSP 1538	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	129	28
SSP 1632	NA	NA	NA	Vimentin	MALDI/TOF/TOF-MS/MS	554	NA
SSP 2010	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	83	35
SSP 2104	12957	20304089	P02525	Beta A3 crystallin	MALDI/TOF/TOF-PMF	91	35
SSP 2307	22223	18203410	Q9R0P9	Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCH-L1)	MALDI/TOF-PMF	177	68
SSP 2315	15507	547679	P14602	HSP27/25	MALDI/TOF-PMF	126	65
SSP 2521	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	230	42
SSP 2522	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	177	35
SSP 2523	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	170	35
SSP 2527	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	201	38
SSP 2617	1333828	50400333	Q6NVD9	CP49, Phakinin	MALDI/TOF/TOF-PMF	111	26
SSP 2716	NA	NA	NA	DnaK-type molecular chaperone Hsc70	ESI-MS/MS	491[b]	NA
SSP 2811	269523	74192715	Q3TXN9	Similar to Translational endoplasmatic reticulum containing ATPase	MALDI/TOF-PMF	135	20
SSP 3004	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	102	40
SSP 3007	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	130	73
SSP 3010	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF/TOF-PMF	95	32
SSP 3014	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	105	35
SSP 3101	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	103	42
SSP 3313	15507	547679	P14602	HSP27/25	MALDI/TOF-PMF	186	69
SSP 3605	14854	13277789	P51855	Glutathione synthetase	ESI-MS/MS	358[b]	31
SSP 3617	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-PMF	80	13
SSP 3804	96245	74205924	Q3UVN1	DnaK-type molecular chaperone Hsc70 homolog	MALDI/TOF/TOF-PMF	67	20
SSP 3827	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-PMF	70	15
SSP 4005	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	103	42
SSP 4008	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	102	35
SSP 4011	101790	74151396	Q3TLH6	Epidermal-type fatty acid binding protein (E-FABP)	MALDI/TOF/TOF-MS/MS	81	31
SSP 4103	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF-PMF	162	54
SSP 4106	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	102	36
SSP 4114	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF-PMF	155	58
SSP 4121	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF-PMF	202[c]	100
SSP 4216	12959	10946672	Q9JJV0	Beta A4 crystallin	MALDI/TOF-PMF	157	64
SSP 4219	12957	20304089	P02525	Beta A3 crystallin	MALDI/TOF/TOF-MS/MS	501	56
SSP 4316	15507	547679	P14602	HSP27/25	MALDI/TOF/TOF-PMF	138	41
SSP 4326	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-MS/MS	250	13
SSP 4329	67397	16877776	P57759	Endoplasmatic reticulum protein ERp29, precursor	MALDI/TOF/TOF-MS/MS	87	26
SSP 4501	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-PMF	139	19
SSP 4507	NA	NA	NA	(GDP) dissociation inhibitor 2	ESI-MS/MS	256[b]	NA
SSP 4806	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-PMF	77	18
SSP 4811	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-PMF	95	14
SSP 4812	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-PMF	143	23
SSP 4818	12075	26343715	Q8BKB1	CP94, Filensin	MALDI/TOF/TOF-MS/MS	195	11
SSP 5002	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-PMF	84	35
SSP 5014	NA	NA	NA	Gamma crystallin	ESI-MS/MS	74[b]	NA
SSP 5215	12957	20304089	P02525	N-acetylated Beta A1 crystallin	MALDI/TOF/TOF-MS/MS	417	61
SSP 5323	12960	12963789	Q9WVJ5	Beta B1 crystallin	MALDI/TOF/TOF-PMF	156	46
SSP 5409	NA	NA	NA	Annexin A1	MALDI/TOF/TOF-PMF	104	26
SSP 5519	NA	NA	NA	Alpha enolase 1	ESI-MS/MS	534[b]	NA
SSP 5522	12960	12963789	Q9WVJ5	Beta B1 crystallin	MALDI/TOF/TOF-PMF	74	28
SSP 6003	12954	62201857	P02490	N-terminally truncated Alpha A crystallin	ESI-MS/MS, MALDI/TOF/TOF-MS/MS	200	50
SSP 6010	12966	6681037	Q61597	Gamma C crystallin	MALDI/TOF-PMF	113	51
SSP 6107	12955	6753530	P23927	Alpha B crystallin	MALDI/TOF/TOF-PMF	76	29
SSP 6125	12965	42733606	P04344	Gamma B crystalllin	MALDI/TOF/TOF-MS/MS	132	47
SSP 6132	214301	23346485	NE	Gamma N crystallin	MALDI/TOF/TOF-PMF	118	46
SSP 6216	12958	10946978	Q9JJV1	Beta A2 crystallin	MALDI/TOF-PMF	175	87
SSP 6225	12954	117332	P24622	Alpha Ains crystallin	MALDI/TOF-PMF	200	75
SSP 6319	12960	12963789	Q9WVJ5	Beta B1 crystallin	MALDI/TOF-PMF	309	92
SSP 6325	96819	70912321	Q4FK88	Annexin A1	MALDI/TOF-PMF	187	55
SSP 6704	12466	6753324	P80317	Chaperonin subunit 6a	MALDI/TOF/TOF-PMF	66	15
SSP 6717	107943	74204595	Q3TW97	Chaperonin subunit 6a	MALDI/TOF/TOF-MS/MS	195	15
SSP 7001	12954	62201857	P02490	Alpha A crystallin	MALDI/TOF/TOF-MS/MS	337	24
SSP 7017	12957	26350979	Q8BNB5	Beta A3/A1 crystallin gene product	MALDI/TOF/TOF-PMF	76	72
SSP 7106	217517	19353298	Q8R3T5	Syntaxin binding protein	MALDI/TOF/TOF-MS/MS	72	52
SSP 7122	NA	NA	NA	Peptidyl-prolyl-cis-trans isomerase	ESI-MS/MS	158[b]	25
SSP 7127	12955	6753530	P23927	Alpha B crystallin	MALDI/TOF-PMF	187	53
SSP 7219	12961	50401872	P62696	Beta B2 crystallin	MALDI/TOF-PMF	213	78
SSP 7220	12970	6753532	O35486	Gamma S crystallin	MALDI/TOF-PMF	240	87
SSP 7221	12957	20304089	P02525	Beta A1 crystallin	MALDI/TOF/TOF-MS/MS	315	34
SSP 7223	12970	6753532	O35486	Gamma S crystallin	MALDI/TOF/TOF-PMF	76	38
SSP 7307	16828	13529599	Q99K20	Ldha protein	ESI-MS/MS	131[b]	10
SSP 7312	NA	NA	NA	G Protein beta subunit like	ESI-MS/MS	111[b]	NA
SSP 7318	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF-PMF	279	86
SSP 7403	74091	21311855	Q9DCJ9	N-acetylneuraminate pyruvate lyase	ESI-MS/MS	212[b]	17

SSP 7404	18655	80477474	Q5XJE7	Phosphoglycerate kinase 1	MALDI/TOF-PMF	253	58
SSP 7632	NA	NA	NA	Pyruvate Kinase 3	ESI-MS/MS	637[b]	30
SSP 7722	NA	NA	NA	Transketolase	ESI-MS/MS	135[b]	NA
SSP 8006	12955	6753530	P23927	Alpha B crystallin	MALDI/TOF/TOF-PMF	128	53
SSP 8014	12968	34978370	Q03740	Gamma E crystallin	MALDI/TOF/TOF-PMF	96	47
SSP 8108	NA	NA	NA	Peptidyl-prolyl-cis-trans isomerase	ESI-MS/MS	119[b]	19
SSP 8116	12967	34784220	Q6PGI0	Gamma D crystallin	MALDI/TOF/TOF-MS/MS	230	35
SSP 8117	12964	6724317	P04345	Gamma A crystallin	MALDI/TOF-PMF	273	90
SSP 8118	12964	6724317	P04345	Gamma A crystallin	MALDI/TOF/TOF-PMF	120	66
SSP 8120	12969	21746155	Q9CXV3	Gamma F crystallin	MALDI/TOF-PMF	204	89
SSP 8121	12968	34978370	Q03740	Gamma E crystallin	MALDI/TOF-PMF	193	74
SSP 8216	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF-PMF	109	61
SSP 8220	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF-PMF	202	73
SSP 8227	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF-PMF	225	77
SSP 8230	12966	6681037	Q61597	Gamma C crystallin	MALDI/TOF-PMF	191	77
	12968	34978370	Q03740	Gamma E crystallin	MALDI/TOF-PMF	172	74
SSP 8231	12969	21746155	Q9CXV3	Gamma F crystallin	MALDI/TOF-PMF	270	94
SSP 8232	12965	42733606	P04344	Gamma B crystallin	MALDI/TOF-PMF	206	75
SSP 8310	14862	6754084	P10649	Glutathione transferase mu1	MALDI/TOF-PMF	153	60
SSP 8314	56012	9256624	O70250	Phosphoglycerate mutase 2	MALDI/TOF-PMF	106	44
SSP 8432	NA	NA	NA	Glyceraldehyde-3-phosphate dehydrogenase	ESI-MS/MS	148[b]	12
SSP 8508	NA	NA	NA	Argininosuccinate synthetase 1	ESI-MS/MS	165[b]	12
SSP 8601	NA	NA	NA	Acetaldehyde dehydrogenase	ESI-MS/MS	427[b]	25
SSP 8905	12965	42733606	P04344	Gamma B crystallin	MALDI/TOF-PMF	173	53
SSP 9004	12955	6753530	P23927	Alpha B crystallin	MALDI/TOF-PMF	63	25
SSP 9013	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF/TOF-PMF	72	41
SSP 9014	12962	10946674	Q9JJV9	Beta B3 crystallin	MALDI/TOF/TOF-MS/MS	208	32
SSP 9017	12965	42733606	P04344	Gamma B crystallin	MALDI/TOF/TOF-PMF	79	49
SSP 9516	NA	NA	NA	Eef1a1 protein	MALDI/TOF-PMF	175	34
SSP 9703	18458	53754	P29341	Poly (A) binding protein	MALDI/TOF-PMF	189	35

[a] A Mascot score of 63 or more indicates protein identity with a P-value less than 0.05. The exceptions are additionally footnoted.

[b] A Mascot score of 36 or more indicates peptide sequence and protein identity with a P-value less than 0.05.

[c] A peptide mass accuracy error of 150 ppm was tolerated.

ORF. Open reading frame. The GenBank at NCBI gene ID was preferred, if this was unavailable the MGI (Mouse Genome Informatics) at The Jackson Laboratory gene ID is listed.

NA. Not applicable because the exact protein sequence could not be unambiguously determined or the % sequence coverage could not be calculated.

NE. Not entered because there was no listing for the sequence in the respective database.

Appendix

List of Non-Standard Abbreviations

ARVO	Association for Vision and Research in Ophthalmology
ABC	Ammonium hydrogencarbonate
ACN	Acetonitrile
BSA	Bovine serum albumine
CBB	Coomassie brilliant blue
CHAPS	3-[3-(Cholamidopropyl)dimethylammonio]-1-proanesulfonate
CHCA	α-cyano-4-hydroxycinamic acid
CID	Collision induced dissociation
DHB	2, 5-dihydroxybenzoicacid
DOC	Sodium desoxycholate
DTT	Dithiothreitol
ECD	Electron capture dissociation
EDTA	Ethylenediamine tetraacetic acid
ESI	Electrospray ionization
EtOH	Ethanol
FA	Formic acid
FTICR	Fourrier transform ion cyclotron resonance
GIST	Global internal standard technology
HPLC	High performance liquid chromatography
ICR	Ion cyclotron resonance
IEF	Isoelectric focussing
LC	Liquid chromatography
MALDI	Matrix assisted laser desorption ionisation
MMTS	Methylmethane-thiosulfonate
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MudPIT	Multi-dimensional protein identification technology
Mw	Molecular weight
N-term	N-terminal
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PMF	Peptide mass fingerprint
PMM	Peptide mass map
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post translational modification
Pyro-Glu	Pyro-glutamic acid
RT	Room temperature
S/N	Signal to noise
SDS	Sodium dodecyl sulfonate
SDS-PAGE	Sodium dodecyl sulfonate polyacrylamide gel electrophoresis
TBP	Tributylphosphine
TCA	Trichlorine acetate
TCEP	Tris (2-carboxyethyl) phosphine
TFA	Trifluoric acid
TOF	Time of flight
TOF/TOF	Tandem time of flight
w/v	Weight per volume
w/w	Percent by mass
2-DE	Two dimensional electrophoresis

Appendix

Curriculum Vitae

Personal Facts

Name	Wolfgang Hoehenwarter
Address	Bergmannstr. 27, 10961 Berlin, Germany
Telephone	+49 (0)178 352 6130
E-Mail	wolfbl21@yahoo.com
Date of Birth	4th June 1973
Place of Birth	Lake Forest, ILL, USA
Nationality	American, Austrian
Marital Status	unmarried

Primary Education

1979-1992	Primary School and Gymnasium, German School Washington D.C., USA
1986-1987	Mittelstufe, German School Riyadh, Kingdom of Saudi Arabia
1992	Abitur, German School Washington D.C., USA
1994	SAT Test, score 1500, top 1% of college bound seniors in USA

Higher Education

1992-2001	Graduate studies of Biology at the University of Innsbruck, Austria
2001	Masters Degree in Biology
2001-2006	Postgraduate studies at the TU Berlin, Max Planck Institute for Infection Biology and Institute of Human Genetics of the Charité, Berlin, Germany

Administrative Training and Experience

1993, 1994	Summer internship as an Administrative Assistant at the International Monetary Fund, Washington D.C., USA
1995	Member of the committee for administration of the conference "Students: From Revolution to Stagnation", an international conference for students in Innsbruck, Austria

Scientific Achievements

Scientific publications:
Hoehenwarter et al., 2005; Proteomics, 5: 245-257; PMID: 16583312
Hoehenwarter et al., 2006; Amino Acids, 30: 369-389; PMID: 15744838
Hoehenwarter et al., 2006; Amino Acids, 31: 317-323; PMID: 16964561
Mattow et al., 2004; Proteomics, 4: 2927-2941; PMID: 15378758
Thiede et al., 2005; Methods, 35: 237-247; PMID: 15722220

Oral presentations:
Sigma Aldrich Seminar Tour "New Tools and Techniques for Protein Research", Berlin, Germany, 26th May 2004
Shimadzu Seminar Tour "Neues aus der Proteomforschung", Munich, Germany, 5th Oktober 2004
SFB577 Annual Meeting, Kremmen, Germany, 29th to 30th April 2005
Applied Biosystems MALDI/TOF/TOF User Meeting, Berlin-Buch, Germany, 2nd June 2005
9th International Congress on Amino Acids and Proteins, Vienna, Austria, 8th to 12th August 2005
SFB577 Annual Meeting, Kremmen, Germany, 21st to 22nd April 2006

Poster presentations:
Proteomics Forum 2003, Munich, Germany, 13th to 16th September 2003