A combination of Edman sequence analysis and mass spectrometry identified the major proteins of the young human lens as αA, αB, βA1, βA3, βA4, βB1, βB2, βB3, γS, γC, and γD-crystallins and mapped their positions on two-dimensional electrophoretic gels. The primary structures of human βA1, βA3, βA4, and βB3-crystallin subunits were predicted by determining cDNA sequences. Mass spectrometric analyses of each intact protein as well as the peptides from tryptic-digested proteins confirmed the predicted amino acid sequences and detected a partially degraded form of βA3/A1 missing either 22 or 4 amino acid residues from its N-terminal extension. These studies were a prerequisite for future studies to determine how human lens proteins are altered during aging and cataract formation.

Elucidating the structure and function of crystallins, the major proteins of human lens, is important, because alterations in these proteins may contribute to cataract formation. The goal of our laboratories is to determine what changes occur in human crystallins during aging and cataract formation. To accomplish this goal it is first necessary to perform the following in young, normal human lenses: 1) determine which crystallin subunits are present in the young lens; 2) map the positions where these crystallins migrate on two-dimensional electrophoretic gels; and 3) deduce and confirm the amino acid sequences of these proteins.

The sequences of human αA, αB, βB1, βB2, γS, γC, and γD-crystallins have already been determined and their presence in the human lens has been demonstrated (1–11). Furthermore, the positions where human αA, αB, βB1, βB2, and γS-crystallins migrate on two-dimensional electrophoretic gels have been determined (12–14). In the present study, we completed the identification of all major crystallins resolved by two-dimensional electrophoresis, and add βA1, βA3, βB3, and βA4 to the list of major β-crystallins in young, normal human lens. We also correct the sequence of human βA3, complete the sequence determination of human βB3, and for the first time, report the sequence of human βA4. These deduced amino acid sequences were confirmed by mass spectrometry.

### Materials and Methods

**Identification of Major Human Crystallins**—The posterior poles of human eyes from organ donors 7-months of age or less were obtained from the Lions EyeBank of Oregon within 48 h post-mortem. Following decapsulation, the lenses were homogenized in 1.0 ml of 20 mM phosphate buffer (pH 7.0), and 0.1 mM EGTA. Water-soluble proteins were isolated by centrifugation at 10,000 × g for 30 min. Protein content was assayed by the BCA assay (Pierce) according to the manufacturer’s protocol using bovine serum albumin as a standard. Water-soluble fractions were then dried by vacuum centrifugation and stored at −70 °C prior to electrophoretic or chromatographic separation.

Two-dimensional electrophoresis of water-soluble lens proteins, transfer to polyvinylidene difluoride membranes, and direct sequence analysis were carried out as described previously (15). Non-equilibrium pH gradient electrophoresis using pH 3.5–10 ampholine was used in the first dimension and SDS-polyacrylamide electrophoresis in the second dimension. However, most crystallins were blocked on the N terminus and could not be identified by direct sequence analysis. Therefore, electroblots of two-dimensional gels were reversibly stained with Poncetan S, individual proteins digested with trypsin, and peptides isolated from the resulting mixtures for sequencing as described previously (16), except that peptides were separated using a 250 × 2.1-mm C18 column (Vydac, Hesperia, CA), and linear 100-min 0–35% acetonitrile gradient containing 0.1% trifluorocetic acid. The tryptic peptides derived from regions of electrophrams thought to contain more than one crystallin species were additionally analyzed by mass spectrometry using either FAB-MS or ESIMS as described below.

The identified crystallins were quantified after staining two-dimensional gels with Coomassie Brilliant Blue R-250, or colloidal Coomassie Brilliant Blue G-250 (Instaview Universal Stain, Gallard-Schlesinger Industries, Inc., Carle Place, NY). Each of the stained gels contained 100 μg of total protein. The density of each species was determined from images obtained with a Gel Doc 1000 camera and analysis using Molecular Analyst Software (Bio-Rad).

**Cloning and Sequencing of Human βA4, βB3, and βA3/A1 cDNAs—**Lenses from organ donors 19-months of age or less were obtained as described above. Total lens RNA was isolated by homogenization of dissected lenses in TRIzol reagent according to the manufacturer’s protocol (Life Technologies, Inc.). To amplify the 3’ end of the human lens βA4 cDNA, reverse transcription was performed on total RNA from human lens using an oligo(dT) containing adapter primer as described previously (3). The resulting cDNA was then subjected to 3’-RACE PCR according to the manufacturer’s protocol (3′-RACE System for Rapid Amplification of cDNA Ends, Life Technologies Inc.), using a 3’-RACE universal adapter primer, and the gene-specific sense primer AGGGCTGACCATCTTCGGGAGCA matching residues 390–410 of bovine βA4 cDNA (GenBank accession number M60328). Cycling conditions for PCR were as described previously (3), except an annealing temperature of 61 °C was used. The amplification resulted in a single PCR product of approximately 550 base pairs.

To amplify the 5’ end of the human lens βA4 cDNA, reverse transcription was performed on total RNA from human lens using the gene-specific antisense primer GTGATTCCTGAGTCGCTCTCCCT complementsary to residues 392–411 of the final human βA4 cDNA sequence.

---

**Abbreviations**

- FAB-MS: Fast atom bombardment mass spectrometry
- RACE: Rapid amplification of cDNA ends
- PCR: Polymerase chain reaction
- HPLC: High performance liquid chromatography
- ESIMS: Electrospray ionization mass spectrometry
Identification of Human Lens Crystallins

The cDNA was then homopolymer tailed according to the manufacturer’s protocol (5’–RACE System for Rapid Amplification of cDNA Ends, Life Technologies Inc.). The resulting cDNA was then PCR amplified using a second gene-specific antisense primer, TCTGGGCCAGGTCTCTCTCT, complementary to residues 571–591 of the final human βA4 cDNA, and a 5’–RACE anchor primer supplied by the manufacturer (Life Technologies, Inc.). Cycling conditions for PCR were as described previously (3), except an annealing temperature of 55 °C was used. The amplification resulted in a single PCR product of approximately 450 base pairs.

To amplify the unknown 5’ end of human βB3 cDNA, the same procedure described above for 5’–RACE of human βA4 cDNA was followed. The gene-specific antisense primer, CAGACACACAGCTGTGCTTGTGCTGCTGCTG, complementary to nucleotides 8–28 in exon 5 of human βB3 (GenBank accession number X15145), was used for reverse transcription of total lens RNA. A second gene-specific antisense primer, CCTC-GGCTCTCTGATAATT, complementary to nucleotides 115–133 in exon 4 of human βB3 (GenBank accession number X15144), was used to perform 5’–RACE of human βB3 cDNA. Cycling conditions were as described previously (3), except an annealing temperature of 60 °C was used. A PCR product of approximately 400 base pairs was obtained.

Preliminary mass spectrometric analysis of human βA3 protein suggested that a portion of the reported genomic sequence of human βA3 contained several errors in nucleotide identification in exons 4–6 (GenBank accession numbers M14304, M14305, and M14306). Therefore, the corresponding region of human βA3 cDNA was amplified and sequenced for comparison with the earlier genomic sequence. Total human lens RNA was reversed transcribed using the gene-specific antisense primer, GCAAGGTCTCATGTTGAGG, complementary to residues 185–204 of exon 6 of human βA3 (GenBank accession number M14306). After treating with RNase H, the cDNA was PCR amplified using the gene-specific sense primer, TGATCAGGAGAACTTTCAGG, and the antisense primer used above in the reverse transcription reaction. Cycling conditions for PCR were as described previously (3), except an annealing temperature of 55 °C was used. A PCR product near the expected size of 589 base pairs was produced.

PCR products were cloned with either the Prime PCR Cloner™ Cloning System (5 Prime → 3 Prime, Inc., Boulder, CO) or with the Original TA Cloning® Kit (Invitrogen, San Diego, CA). Plasmid DNA was isolated using either a FlexiPrep™ Kit (Pharmacia-Biotech, Inc., Piscataway, NJ), or Quantum Prep™ Kit (Bio-Rad). Following screening to confirm the presence of the correct sized insert, sequencing was performed using either the AutoRead sequencing kit and automated laser fluorescence DNA analysis system (Pharmacia Biotech), or a Cycle Sequencing Kit and Model 373A DNA Sequencer (Applied Biosystems, Inc., Foster City, CA). Depending on the plasmid utilized, sequencing was performed using either M13/pUC forward and reverse primers or T7 and Sp6 primers. Three clones of each of the 3’- and 5’-RACE products of the βA4 cDNA, and 2 clones of the PCR product of the βA3 cDNA were sequenced in both sense and antisense directions. The sequence from 4 antisense strands of the 5’-RACE product of βB3 cDNA were sequenced. DNA sequences were edited and theoretical pI values of proteins calculated using Geneworks 2.5 software (IntelliGenetics, Mountain View, CA).

Confirmation of Deduced βA3/A1, βB3, and βA4 Protein Sequences using Mass Spectrometry—The deduced amino acid sequence of βA3/A1, βB3, and βA4 crystallins were confirmed by: 1) determining the molecular mass of intact proteins; 2) determining the molecular masses of tryptic peptides; and 3) confirming the sequences of tryptic peptides. The types of mass spectrometry employed were ESI-MS, FAB-MS, and tandem mass spectrometry (MS/MS).

Soluble lens protein was isolated from 32-week gestation, 3–7-day-old, and 42-year-old human donors as described above. Proteins were dissolved in gel filtration buffer containing 50 mM phosphate (pH 7.0), 150 mM NaCl, and 5–mg portions injected onto a Superose 6 HR 10/30 gel filtration column (Pharmacia Biotech) at a 0.2 ml/min flow rate. Protein elution was monitored at 280 nm and β-crystallin aggregates of approximately 160,000 and 50,000 molecular weight (β1r and β1-crys- tallin fractions, respectively) were collected. These aggregates were then concentrated and desalted using Centricron™10 microconcentrators (Amicon, Inc., Beverly, MA) in preparation for mass analysis. β-Crystallin subunits were partially purified from β1r and β1 aggregates using a Vydac 4.6 × 150-mm C4 reversed phase column and 25–60% acetonitrile gradient containing 0.1% trifluoroacetic acid over 35 min and 1 ml/min flow rate. Reversed phase HPLC purified β-crystallins were then injected into the mass spectrometer with a 50:50 solution of acetonitrile and water at a flow rate of 5 μl/min. The masses of the isolated β-crystallins were determined using a Micromass Platform II electrospray ionization mass spectrometer with a quadrupole analyzer and Mass Lynx software (Micromass, Manchester, UK). For protein analysis, the instrument was calibrated with horse skeletal muscle myoglobin over the range of 700-1800 Da. Accuracy of protein molecular mass determinations was ±3 Da.

As an alternative to reversed phase HPLC, two-dimensional electrophoresis was used as a preparative tool to isolate the truncated βA3/A1 species. Total soluble proteins from a 4-day-old human donor were separated by two-dimensional electrophoresis, proteins visualized by precipitation of SDS-protein complexes with ice-cold 0.25 M KCl (17), and the truncated βA3 species excised and electroeluted from 12 gels. Proteins were electroeluted into a Centricron™ 10 microconcentrator using the method recommended by the manufacturer (Amicon, Inc.). Electroelution into this device facilitated the concentration and desalting of the sample for analysis by ESI-MS as described above.

The deduced amino acid sequences of βA3/A1, βB3, and βA4 were confirmed from the masses of peptides in tryptic digests of the proteins (50:1 substrate:trypsin, pH 8.2, 4 h). Lenses of 32-week-gestation, 4-day-old, and 42-year-old donors were used to isolate βB3, βA3, and βA4 for tryptic digestion, respectively. Tryptic peptides of βA4 and βB3 were prepared from proteins isolated by gel filtration and reversed-phase HPLC. Peptides of βA3 were prepared from protein separated by two-dimensional electrophoresis, transfer to polyvinylidene difluoride membrane, and digestion from the membrane surface with trypsin. The peptides were fractionated by C18, reversed-phase HPLC with a gradient of 2–50% acetonitrile in water, containing 0.1% trifluoroacetic acid. For peptide analysis, the mass spectrometers used included a Kratos MS-50 fast atom bombardment mass spectrometer (Kratos Analytical, Manchester, UK), a Micromass Platform II electrospray ionization mass spectrometer and a Micromass Autospec mass spectrometer with an orthogonal TOF analyzer for the MS/MS analyses. Some analyses employed a microbore column on-line to the electrospray ionization mass spectrometer (with a flow rate of 50 μl/min, with 5 μl/min entering the mass spectrometer and 45 μl/min monitored by a UV detector and collected for further analysis. The instruments were calibrated over the range of 200-3000 with Na2, the accuracy of the determinations was ±0.3 Da.

Sequences of tryptic peptides were also confirmed by MS/MS. This technique consists of one mass spectrometer to isolate the peptide of interest, a chamber where the peptide is fragmented by collision with xenon, and a second mass analyzer which determines the masses of the resulting fragments (18). The tandem mass spectrometer used in this investigation (Micromass Autospec oa-TOP) consisted of a conventional magnetic sector instrument as the first mass spectrometer and an orthogonal acceleration time-of-flight analyzer as the second mass spectrometer (19). These analyses were performed in the fast atom bombardment mode of ionization.

**RESULTS**

Identification of Crystallin Subunits of Young Lens Separated by Two-dimensional Electrophoresis—Two-dimensional Electrophoresis of Soluble Lens Proteins from 3-day-old Human Donor. Identification of major crystallin species was performed by both Edman sequencing and mass spectro- metric analysis as summarized in Table I and Figs. 2 and 3. The identification of βB1-crystallin and its degradation product βB1 (16–251) was previously reported (3). Note that γS and βA1, as well as βA3- and βB3-crystallins co-migrate. The left side of the gel is basic and the right side is acidic. Only the region of the two-dimensional gel containing crystallins is shown.

FIG. 1. Two-dimensional electrophoresis of soluble lens protein from 3-day-old human donor.
Identification of Human Lens Crystallins

**Table I**  
Identification of peptides from major crystallins from young human donors

| Crystallin<sup>a</sup> | Amino acid sequence | Residue numbers |
|----------------------|---------------------|-----------------|
| αA                   | HFSPEDLTVK<sup>b</sup> | 79–88           |
| αB                   | TLGFFYPSR           | 13–21           |
| βB2                  | APSWFDQLS           | 57–66           |
| βA3 (23–215)         | EWSHAQTQIQSIR       | 197–211         |
| βA3 (23–215)         | PTPGLGP             | 23–30           |
| βA4                  | MVVNEDEGFQGR        | 13–24           |
| γS                   | GKYTVLEXHDHSGDYK    | 158–173         |
| γC                   | EWSHATAQFQOSIR      | 177–191         |
| γD                   | GITLYEDRG           | 1–10            |

<sup>a</sup> The following crystallin species labeled in Fig. 1 were identified by comparing the results of Edman sequencing to reported amino acid sequences of crystallins. The sequences of αA, αB, βB2, βA3, βA4, and γS-crystallins were determined from tryptic fragments, and the sequences of βA3/α1 (23–215), γC, and γD-crystallins by direct Edman sequencing of unfragmented proteins.

<sup>b</sup> Several analyzed fractions of the reverse-phase purified tryptic fragments contained more than one peptide.

**Fig. 2.** Fast-atom bombardment mass spectra of tryptic peptides isolated from the region of two-dimensional gels containing both γS- and βA1-crystallin. A. mass spectrum of two tryptic peptides corresponding to the expected molecular masses of residues 84–94 (MH<sup>+</sup> 1156.6) and 148–153 (MH<sup>+</sup> 779.4) of human γS-crystallin. B, mass spectrum of a single tryptic fragment corresponding to the expected molecular mass of the acetylated N-terminal tryptic peptide of human βA1-crystallin (residues 1–14, MH<sup>+</sup> 1495.7).

**Table II**  
Relative amounts of each crystallin subunit in the soluble protein of young human lens

| Crystallin subunit | % of total |
|--------------------|------------|
| αA                 | 21.4 ± 1.0<sup>a</sup> |
| γS/α1<sup>b</sup>  | 15.4 ± 1.2  |
| βB2                | 14.3 ± 1.2  |
| γC                 | 14.3 ± 0.8  |
| βB1                | 8.8 ± 2.4   |
| αB                 | 6.3 ± 0.8   |
| βA3/βB3<sup>b</sup>| 6.2 ± 0.8   |
| βA4                | 4.7 ± 1.4   |
| γD                 | 2.5 ± 0.9   |
| βB1 (16–251)       | 2.2 ± 0.8   |
| βA3 (23–215)       | 1.6 ± 0.6   |
| Total              | 97.7        |

<sup>a</sup> Mean ± 1 S.D. (n = 4), resulting from the densitometric analysis of individual two-dimensional electrophoretic gels of soluble protein of newborn, 3-, 4-, and 7-day old human donors.

<sup>b</sup> The individual amounts of γS, βA1, βA3, and βB3 could not be estimated because these subunits migrated to similar positions during two-dimensional electrophoresis.

Due to their lack of acetylated N termini, the positions of γC and γD, as well as a truncated β-crystallin, could be determined by direct sequence analysis of electrophoretic proteins (Fig. 1, Table I). The truncated subunit contained a sequence matching that of βA3/α1 and could have arisen from truncation of the first 22 amino acids of βA3 or the first 4 amino acids of βA1. βA1 and βA3 are identical, except βA3 contains a longer N-terminal extension due to the use of an alternate start codon in the single βA3/α1 transcript (20). For simplicity, the truncated β-crystallin is referred to as βA3 (23–215). The presence of βA3 (23–215) in human lens was recently reported in another laboratory (21).

Although mRNAs coding for γA, γB, γC, and γD have been reported in fetal and neonatal human lenses, transcripts for γC and γD were more abundant (7). In the present study, γC and γD crystallins were detected, but no γA or γB crystallins were observed (Fig. 1, Table I). Earlier mass spectrometric and chromatographic analysis also detected γC and γD in human lenses, and found either small or undetectable amounts of γA and γB (10, 11). This suggested that little translation of either...
Aor
Aor
C
May occur in human lens. The lower migration position of D was not expected (Fig. 1), since its calculated molecular weight was nearly identical to C. This is likely explained by anomalous migration of \( g \)-crystallins during SDS-PAGE and not by post-translational modification.

Many of the crystallins described above were identified from the sequence of one peptide in each tryptic digest. This did not exclude the possibility that some spots observed on two-dimensional gels actually contained more than one protein species. Therefore, mass spectrometry was used to determine the masses of all peptides in tryptic digests from each of the two regions previously thought to contain only yS or \( \beta \)A3. The region containing yS (as determined by Edman sequencing) yielded peptides with masses matching peptides expected from both yS and \( \beta \)A1. Representative mass spectra showing two peptides with masses matching residues 84–94 and 148–153 of yS (Fig. 2A), and a peptide with a mass matching the acetylated N terminus of \( \beta \)A1 (residues 1–14) (Fig. 2B) are shown.

Densitometric Analysis of Two-dimensional Gels—Once the major crystallin subunits in young human lenses were identified, the relative amount of each subunit was determined by densitometric analysis of two-dimensional electrophoretic gels of the total soluble lens protein from newborn, 3-, 4-, and 7-day old human donors (Table II). The relative amounts of \( \beta \)A1 and yS, as well as \( \beta \)A3 and \( \beta \)B3, could not be determined from two-dimensional gels of whole soluble protein, because these proteins did not resolve from one another. However, two-dimensional electrophoresis of isolated \( \beta \)H- and \( \beta \)L-crystallin aggregates from 3- and 7-day-old human lenses removed yS and allowed estimation of the relative amount of \( \beta \)A1. \( \beta \)A1 and \( \beta \)A4 were found in nearly equal quantities in the young human lens.

\( \gamma \)A or \( \gamma \)B mRNA may occur in human lens. The lower migration position of \( \gamma \)D was not expected (Fig. 1), since its calculated molecular weight was nearly identical to \( \gamma \)C. This is likely explained by anomalous migration of \( \gamma \)-crystallins during SDS-PAGE and not by post-translational modification.

Many of the crystallins described above were identified from the sequence of one peptide in each tryptic digest. This did not exclude the possibility that some spots observed on two-dimensional gels actually contained more than one protein species.
tides 788–793. Human βA4 cDNA encoded a protein of 195 amino acids, excluding the N-terminal methionine. This gave a calculated molecular mass of 22,285 including the acetylation of their N terminus (see mass spectrometric analysis below). The predicted pI was 5.63. Excluding the N-terminal extensions, the amino acid sequence of human βA4 shared 92–94% sequence identity with rat2 and bovine (22) βA4, and 65% identity with chicken βA4 (23). The N-terminal extensions of human, rat, and chicken βA4 were all 10 residues in length, supporting the previous suggestion that the N-terminal extension of bovine βA4 is also 10 residues in length (23).

Since only exons 4–6 of human βB3 were previously reported (24), the 5’ end of human βB3 cDNA was amplified by 5’-RACE PCR and sequenced (Fig. 4B). The resulting sequence was 379 nucleotides in length, contained a 71-nucleotide untranslated region at its 5’ end, and coded for amino acids 1–102 of human βB3 (GenBank accession number U71216). This sequence overlapped with the first 114 base pairs of the previously reported sequence for human βB3 exon 4 (Genbank accession number X15144). Combining this sequence with the previously reported sequences of human βB3 exons 4–6 (GenBank accession numbers X15144, X15145, and X15146) resulted in a deduced sequence for a 211-amino acid protein with a calculated mass of 24,224, including the acetylated N-terminal methionine. The calculated pI was 5.77.

Because the mass of β3A determined by ESIMS (25,192) did not agree with the mass calculated from the published sequence (20), a portion of the human β3A cDNA containing the 3’ coding region was amplified by PCR and sequenced. The 549-base pair sequence with a stop codon at nucleotide 517 is shown in Fig. 4C (GenBank accession number U59058). This sequence differs from the earlier sequence at two positions within exon four (Genbank accession number M14304), and 4 positions within exon six (Genbank accession number M14306). Each of these nucleotide differences resulted in a change in the deduced amino acid sequence. The calculated masses for the new sequences of β3A and βA1 were 25,193 and 23,102, respectively, including the acetylation of their N termini. The pI values of β3A and βA1 were 5.58 and 6.17, respectively.

**Confirmation of the Deduced βA4, βB3, and βA3/A1 Amino Acid Sequences**—The molecular weights of the subunits in lens β4H- and βL-crystallin aggregates from a 3-day-old donor were determined by ESIMS (Table III). Two of the components had masses corresponding to βB1 and βB2. Two other components had masses corresponding to previously identified partially degraded βB1-crystallins (3). A spot corresponding to βB1 (40–251) was not identified during Edman sequencing of proteins separated by two-dimensional electrophoresis (Fig. 1). This protein, as well as other partially degraded forms of βB1 (3), may not have been detected by electrophoresis because they may co-migrate with other crystallins.

Experimentally determined masses of other β-crystallins agreed with the calculated masses of βA1, βA3, βA3 (23–215), βA4, and βB3 (Table III), confirming the deduced amino acid sequences of βA4, βB3, and βA3/A1 (Fig. 4). A spot corresponding to βA3 (23–215) was consistently observed by two-dimensional electrophoresis of soluble protein from newborn lenses (Table II). However, analysis of β-crystallin aggregates by reversed-phase HPLC and ESIMS barely detected a component with the mass of βA3 (23–215) in lenses less than 3 years old. The presence of βA3 (23–215) was confirmed in newborn lenses after soluble protein from a 4-day-old lens was separated by two-dimensional electrophoresis and the region containing βA3 (23–215) eluted and analyzed by ESIMS. A molecular mass of 23,001 was expected, because acrylamide reacts with cysteinyl residues, adding 71 Da to each of 5 cysteines (25). The ESIMS determined mass of eluted βA3 (23–215) was 22,999 (Fig. 5), supporting the Edman sequencing results (Table I) and confirming the presence of βA3 (23–215).

Further confirmation of the deduced sequences of human βA4, βB3, and βA3 was obtained from the molecular weights of the peptides in tryptic digests of the proteins. Peptides with masses corresponding to all portions of the sequences of βA4, βB3, and βA3 were found (Fig. 6). Analysis of the tryptic peptides also confirmed that the N termini of βA4, βB3, βA3 (Fig. 6), and βA1 (Fig. 2B) were all acetylated, and showed that the N-terminal methionine was removed from βA4 and βA1, but retained on βB3 and βA3. The peptides marked with an *asterisk* in Fig. 6 were additionally analyzed by MS/MS to confirm the proposed sequence. An example of an MS/MS spectrum of a peptide from βA4 is shown in Fig. 7. The spectrum shows the fragments formed by collisional activation of βA4 peptide 106–117, which has the sequence, LTI FEEQENFLGK. Collisional activation causes the peptide to fragment, primarily along the backbone. Fragments formed with the charge remaining with the C terminus are called the a, b, and c series; fragments formed with the charge remaining with the N terminus are called x, y, and z series (26). Masses of expected fragments due to the b and y* series (the * indicates an additional H on the fragment) are shown in the diagram at the top. The presence of many peaks in the spectrum with masses corresponding to the expected fragments confirmed that the peptide analyzed has the given sequence.

**DISCUSSION**

This investigation: 1) identified and quantified the major protein species in the young human lenses following separation by two-dimensional electrophoresis; 2) determined and confirmed the sequences of human βA4, βB3, and βA3/A1 crystallins; and 3) demonstrated the presence of a truncated βA3/A1 crystallin. The present report completes the mapping of the major human crystallin species separated by two-dimensional electrophoresis first initiated by De Vries et al. (12) and Dutiles et al. (13). Furthermore, the work demonstrates how an analysis of human crystallinns combining two-dimensional electrophoresis, Edman sequencing, and mass spectrometry can rapidly identify and characterize the primary structure of crystallins and detect their post-translational modifications. The information in this report will be used as a reference point.

---

2 K. J. Lampi and L. L. David, unpublished results.
for a thorough characterization of modifications occurring during lens maturation, aging, and opacification. The relative amounts of the total $\alpha$-, $\beta$-, and $\gamma$-crystallin subunits determined from two-dimensional gels were 27.7, 42.5, and 27.5%, respectively. This corresponds to previously published amounts of $\alpha$-, $\beta$-, and $\gamma$-crystallin aggregates isolated by gel filtration from lenses of similar age to the newborn lenses used in this study (27, 28). The percent contribution of individual $\alpha$- and $\gamma$-crystallin subunits to the total lens protein was also consistent with the literature (8, 10). The present

The relative amounts of the total $\alpha$-, $\beta$-, and $\gamma$-crystallin subunits determined from two-dimensional gels were 27.7, 42.5, and 27.5%, respectively. This corresponds to previously published amounts of $\alpha$-, $\beta$-, and $\gamma$-crystallin aggregates isolated by gel filtration from lenses of similar age to the newborn lenses used in this study (27, 28). The percent contribution of individual $\alpha$- and $\gamma$-crystallin subunits to the total lens protein was also consistent with the literature (8, 10). The present
A1-crystallin. N-terminal degradation of an Asn22-Pro23 of human lens was detected by mass spectrometric analysis. The co-migration of these cases, the presence of co-migrating crystallins was demonstrated. Sequencing identified just one of the crystallins because only a single tryptic peptide from a complex mixture was analyzed. In future studies, the majority of human crystallins, several crystallins did not resolve completely from one another. Initially Edman sequencing identified the major protein of rat lens (30), but its relative amount in the human lens is much lower. Changes in the relative concentrations of various crystallins in lens may alter its properties. Therefore, future studies will further characterize the relative amounts of crystallins and how this changes during lens maturation.

While two-dimensional electrophoresis was capable of resolving the majority of human crystallins, several crystallins did not resolve completely from one another. Initially Edman sequencing identified just one of the crystallins because only a single tryptic peptide from a complex mixture was analyzed. In these cases, the presence of co-migrating crystallins was detected by mass spectrometric analysis. The co-migration of human B3 and B3 was unexpected, since these two proteins migrate to different positions during two-dimensional electrophoresis of rat or bovine crystallins (16, 29), and have theoretical pI values which differ by 0.2 pH units.

Independent confirmation of the cDNA determined sequences of human B4, B3, and B3/B1 crystallins by mass spectrometric analysis can detect errors in the deduced sequences and characterize post-translational modifications. A major finding of the present paper was the truncation of the N-terminal extension of B1 has previously been reported (3). Even in the lenses from donors of less than 1-week-old, partially degraded forms of both B1 and B3/A1 were present. B1 and B3/A1 may be the most proteolytically labile crystallin subunits in the human lens. The major truncated forms of B3/A1 and B1 result from cleavage between an Asn22-Pro23 of B3 and Asn15-Pro16 of B1. A similar site is also present in B2 at Asn15-Pro16. However, no cleavage of B2 was seen in the lenses examined in this study. The asparagine-proline sequence in B2 occurs at the interface between the N-terminal extension and the first “Greek key” containing motif (22). 1H NMR spectroscopy has confirmed that the N-terminal extensions of crystallins have relatively greater flexibility than the Greek key containing motifs (31). Therefore, the proteolytically resistant Asn22-Pro23 region of B2 may be relatively inaccessible compared to the Asn-Pro regions of B1 and B3 which are cleaved. Of interest is the additional observation that B1 and B3 have longer N-terminal extensions than the proteolytically resistant B2 and B4 (57 and 30 residues versus 15 and 10 residues, respectively). The protease(s) responsible for these cleavages in newborn human lens remain unknown.

In conclusion, two-dimensional electrophoresis combined with Edman sequencing and mass spectrometry successfully identified the major crystallins in the young human lens and confirmed their predicted sequences and masses. This combination of techniques permitted separation and identification of proteins and modifications not otherwise possible. Knowledge of the sequences of all the human crystallins and their exact molecular weights obtained in this investigation will facilitate identification of the various modified forms which become more numerous with age. The ability to electroelute a modified crystallin species from two-dimensional gels and carry out mass spectrometric measurements on the recovered protein was noteworthy (Fig. 5). This technique should speed progress in identifying the numerous protein modifications occurring in human lenses with age. Post-translational modifications of crystallins may alter the way they interact with one another and affect the highly ordered protein structures required to maintain lens clarity. Determining the modifications of crystallins in human lenses from donors of different ages and stages of cataract will contribute to our understanding of cataractogenesis. Such information is important, because cataract is still a major cause of blindness.

---

3 L. L. David and M. K. Duncan, unpublished results.
Acknowledgments—Mass spectral analyses were performed by the Nebraska Center for Mass Spectrometry. DNA sequencing was performed by the Department of Microbiology Core Facility at the Oregon Health Sciences University. GenBank data was accessed through the National Center for Biotechnology Information, Bethesda, MD.

REFERENCES

1. de Jong, W. W., Terwindt, E. C., and Bloemendal, H. (1975) FEBS Lett. 58, 310–313
2. Kramps, J. A., de Man, B. M., and de Jong, W. W. (1977) FEBS Lett. 74, 82–84
3. David, L. L., Lampi, K. J., Lund, A. L., and Smith, J. B. (1996) J. Biol. Chem. 271, 4273–4279
4. Chambers, C., and Russell, P. (1993) Gene (Amst.) 133, 295–299
5. Miesbauer, L. R., Smith, J. B., and Smith, D. L. (1993) Protein Sci. 2, 290–291
6. Smith, J. B., Yang, Z., Lin, P., Zaidi, Z., Abbasi, A., and Russell, P. (1995) Biochem. J. 307, 407–410
7. Brakenhoff, R. H., Aarts, H. J. M., Reek, F. H., Lubsen, N. H., and Schoenmakers, J. G. G. (1990) J. Mol. Biol. 216, 519–532
8. Swamy, M. S., and Abraham, E. C. (1991) Curr. Eye Res. 10, 213–219
9. Aka, J., Katar, M., Rudner, G., and Maisel, H. (1988) Curr. Eye Res. 7, 353–359
10. Siezen, R. J., Thomson, J. A., Kaplan, E. D., and Benedek, G. B. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6086–6092
11. Miesbauer, L. R., Zhou, X., Yang, Z., Yang, Z., Sun, Y., Smith, D. L., and Smith, J. B. (1994) J. Biol. Chem. 269, 12494–12502
12. De Vries, A. C. J., Vermeer, M. A., Hendriks, L. A. M., Bloemendal, H., and Cohen, L. H. (1991) Exp. Eye Res. 53, 519–524
13. Datiles, M. B., Schumer, D. J., Zigler, J. S., Jr., Russell, P., Anderson, L., and Garland, D. (1992) Curr. Eye Res. 11, 669–677
14. Garland, D. L., Duglas-Tabor, Y., Jimenez-Acensio, J., Datiles, M. B., and Magno, B. (1996) Exp. Eye Res. 62, 285–291
15. David, L. L., Shearer, T. R., and Shih, M. (1993) J. Biol. Chem. 268, 1937–1940
16. David, L. L., and Shearer, T. R. (1993) FEBS Lett. 324, 265–270
17. Hager, D. A., and Burgess, R. R. (1980) Anal. Biochem. 109, 76–86
18. Biemann, K. (1990) in Methods in Enzymology (McCloskey, J., ed) Vol. 193, pp. 455–479, Academic Press, San Diego, CA
19. Medzhibrody, K. F., Adams, G. W., Burlingame, A. L., Bateman, R. H., and Green, M. R. (1996) J. Am. Soc. Mass Spectrom. 1, 1–10
20. Hogg, D., Toul, L-C., Gerin, M., and Breitman, M. L. (1986) J. Biol. Chem. 261, 12420–12427
21. Srivastava, O. P., and Srivatava, K. (1996) Invest. Ophthal. Vis. Sci. 37, S421
22. van Rens, G. L. M., Driessen, H. P. C., Nalini, V., Slingsby, C., de Jong, W. W., and Bloemendal, H. (1991) Gene (Amst.) 102, 179–188
23. Duncan, M. K., Haynes, J. I., II, and Piatigorsky, J. (1995) Gene (Amst.) 162, 189–196
24. Aarts, H. J. M., Jacobs, E. H. M., van Willigen, G., Lubsen, N. H., and Schoenmakers, J. G. G. (1989) J. Mol. Biol. 216, 407–410
25. Ringena, P. J., Hoenders, H. J., and Bloemendal, H. (1982) Exp. Eye Res. 34, 201–207
26. Roepstorff, P., and Fohlman, J. (1984) Biomed. Mass Spectrom. 11, 601
27. Thomson, J. A., and Augusteyn, R. C. (1985) Exp. Eye Res. 40, 393–410
28. Berbers, G. A. M., Hoekman, W. A., Bloemendal, H., de Jong, W. W., Kleinschmidt, T., and Braunitzer, G. (1984) Eur. J. Biochem. 139, 467–479
29. David, L. L., Azuma, M., and Shearer, T. R. (1994) Invest. Ophthal. Vis. Sci. 35, 785–793
30. Cooper, P. G., Carver, J. A., and Truscott, R. J. (1993) Eur. J. Biochem. 213, 321–328