

# AHP DB: a reference database of proteins in the human aqueous humor

Tae Jin Lee<sup>1</sup>, Arnav Goyal<sup>1</sup>, Garrett Jones<sup>1</sup>, Joshua Glass<sup>1</sup>, Vishal Doshi<sup>1</sup>, Kathryn Bollinger<sup>2</sup>, Lane Ulrich<sup>2</sup>, Saleh Ahmed<sup>1</sup>, Sai Karthik Kodeboyina<sup>1,3</sup>, Amy Estes<sup>2</sup>, Marc Töteberg-Harms<sup>2</sup>, Wenbo Zhi<sup>1</sup>, Shruti Sharma<sup>1,2</sup> and Ashok Sharma<sup>1,2,4,\*</sup>

<sup>1</sup>Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta University, 1120 15th Street, Augusta, GA 30912, USA

<sup>2</sup>Department of Ophthalmology, Medical College of Georgia, Augusta University, 1120 15th Street, Augusta, GA 30912, USA

<sup>3</sup>Mass General Brigham, 399 Revolution Drive, Somerville, MA 02145, USA

<sup>4</sup>Department of Population Health Sciences, Medical College of Georgia, Augusta University, 1120 15th Street, Augusta, GA 30912, USA

\*Corresponding author: Tel: +7067216335; Email: [assharma@augusta.edu](mailto:assharma@augusta.edu)

Citation details: Lee, T.J., Goyal, A., Jones, G. *et al.* AHP DB: a reference database of proteins in the human aqueous humor. *Database* (2024) Vol. 2024: article ID baae001; DOI: <https://doi.org/10.1093/database/baae001>

## Abstract

The aqueous humor (AH) is a low-viscosity biofluid that continuously circulates from the posterior chamber to the anterior chamber of the eye. Recent advances in high-resolution mass-spectrometry workflows have facilitated the study of proteomic content in small-volume biofluids like AH, highlighting the potential clinical implications of the AH proteome. Nevertheless, in-depth investigations into the role of AH proteins in ocular diseases have encountered challenges due to limited accessibility to these workflows, difficulties in large-scale AH sample collection and the absence of a reference AH proteomic database. In response to these obstacles, and to promote further research on the involvement of AH proteins in ocular physiology and pathology, we have developed the web-based Aqueous Humor Proteomics Database (AHP DB). The current version of AHP DB contains proteomic data from 307 human AH samples, which were analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS). The database offers comprehensive information on 1683 proteins identified in the AH samples. Furthermore, relevant clinical data are provided for each analyzed sample. Researchers also have the option to download these datasets individually for offline use, rendering it a valuable resource for the scientific community.

Database URL: <https://ahp.augusta.edu/>

## Introduction

The aqueous humor (AH) is a transparent, low-viscosity fluid produced by the ciliary body epithelium, circulating from the posterior to the anterior chamber of the eye (1). As a multi-functional biofluid, the AH plays essential roles in supplying nutrients, removing waste and providing oxygen to the avascular tissues of the ocular environment (2, 3). Homeostatic balance is achieved through continuous drainage and secretion, ensuring a steady flow that replenishes depleted nutrients and regulates intraocular pressure (IOP) within optimal physiological limits (4–6).

The proteomic composition of the AH is crucial for various cellular processes, including cell-to-cell communication, signal transduction, immunological modulation and cell proliferation (7–9). Consequently, researchers have investigated changes within the AH proteome associated with different pathophysiological conditions (7, 8, 10–26). These investigations have led to evidence linking specific AH proteins to ocular diseases, such as cataracts (12, 15, 19, 23), glaucoma (15, 18, 21, 27), age-related macular degeneration (20, 22), uveitis

(28, 29), retinoblastoma (30, 31) and diabetic retinopathy (17, 24). Moreover, the dynamics of the AH proteome have been observed to vary across demographic variables, such as sex (13) and race (32), with certain proteins involved in inflammatory protection and immune response found to be more abundant in AH samples from female and African American cohorts, respectively. These discoveries underscore the importance of further exploring the AH proteome and the potential to unveil novel molecular mechanisms.

Given that the AH is a small-volume biofluid with low protein concentration, reproducibility in proteomic profiling of the AH has been limited, as detecting low-abundance proteins can be inconsistent or potentially masked by more abundant molecules, such as transferrin, albumin and immunoglobulins (13, 33–36). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) has emerged as a valuable analytical platform for identifying and quantifying a wide range of proteins in complex, low-volume biofluids, including tears, cerebrospinal fluid, saliva and the AH (7, 8, 10–12, 15, 16, 27, 33, 37–42). The exceptional sensitivity, high-throughput

capabilities and reliable detection of lower-abundance proteins have significantly expanded the known AH proteome (33, 43).

In this study, we analyzed 307 AH samples using LC-MS/MS and developed a web-based resource known as the 'Aqueous Humor Proteomics Database' (AHP DB). This reference database of commonly identified AH proteins aims to provide valuable insights to the vision research community, helping them to better understand physiological and pathological proteomic signatures within the AH.

## Materials and methods

### Human subjects and sample collection

Participants in this study were recruited from patients undergoing cataract or glaucoma surgical procedures at the Augusta University Medical Center. Written, informed consent was obtained from all subjects recruited, and this study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Augusta University (IRB# 611480). Following paracentesis incision into the anterior chamber, the AH was drained using a cortical cleaving hydrodissector (585157, Beaver-Visitec International, Waltham, MA, USA). Typical AH sample volumes ranged from 50 to 200  $\mu$ L depending on the depth of the anterior chamber. The AH sample was immediately transferred to a 1.7 mL Eppendorf tube and stored at  $-80^{\circ}\text{C}$  until further analysis. These methods were safe for all subjects and did not affect the duration of the surgery. Following sample collection, a chart review was conducted to obtain demographic and clinical information pertaining to the study participants.

### AH sample preparation and LC-MS/MS analysis

The AH samples were lyophilized and reconstituted in 30  $\mu$ L of 8 M urea in 50 mM Tris-HCl buffer at pH 8.0. To minimize interference with proteomic mapping, cysteine residues in each sample were reduced and alkylated using 20 mM DTT and 55 mM iodoacetamide, respectively. Subsequently, 240  $\mu$ L of 50 mM ammonium bicarbonate buffer was added to each sample to lower the urea concentration to  $<1$  M. Total protein concentration was measured using a Bradford assay kit, following the manufacturer's instructions.

For LC/MS-MS preparation, the samples were digested with trypsin at a 1:20 (*w/w*) at  $37^{\circ}\text{C}$  overnight. After trypsin digestion, the samples were cleaned using a C18 spin plate and then separated and analyzed via the Ultimate 3000 nano-UPLC system and Orbitrap Fusion Tribrid mass spectrometer. The reconstituted peptide mixture (6  $\mu$ L) was trapped and washed on a Pepman100 C18 spin plate using a gradient of 2% acetonitrile in water and 0.1% formic acid for 10 min at a flow rate of 20  $\mu$ L/min. Subsequently, the samples were separated on a Pepman100 RSLC C18 column using a gradient of 2–40% acetonitrile in water and 0.1% formic acid at a flow rate of 300 nL/min for 120 min.

The eluted peptides were subjected to nano-electrospray ionization at  $275^{\circ}\text{C}$  and 2000 V. The peptides were then analyzed in the Orbitrap Fusion Tribrid mass spectrometer using data-dependent acquisition in positive mode. The parameters for analysis included 120 000 FWHM from 300 to 1500  $m/z$ , ion-trap MS analyzer in top speed mode at a 2 s cycle time

and dynamic exclusion settings with the repeat count set to 1, repeat duration set to 15 s and exclusion duration set to 30 s. Collision-induced dissociation (CID) fragmentation was used with normal collision energy set to 30%.

### Protein identification and quantification

Proteome Discoverer (version 2.2, Thermo Fisher Scientific, Waltham, MA, USA) software was utilized to process raw MS files for protein identification and quantification. The SequestHT algorithm was employed to align identified peptide sequences to the UniProt-SwissProt database, with parameters set for 10 ppm precursor mass tolerance, 0.6 Da product ion tolerance, static carbamidomethylation for cysteine (+57.021 Da), dynamic oxidation for methionine (+15.995 Da) and dynamic phosphorylation for serine, threonine and tyrosine (+79.966 Da). Proteins containing similar peptide sequences that could not be differentiated by LC-MS/MS were instead grouped based on the principles of parsimony. The peptide-spectrum match (PSM) count of each protein was used as a semi-quantitative measure of relative protein expression in AH.

### Database and website development

AHP DB was developed on a Microsoft platform using Asp.net for web development with JavaScript to create a highly interactive web interface. Microsoft SQL Server was used for data storage, with the website hosted on an Internet Information Services (IIS) web server. Data generated from LC-MS/MS analyses were uploaded to the server via the web interface. AHP DB was designed with internal programming to process the data files, generate appropriate tables and add data records into the database. User-friendly filtering and search options for each data column were also integrated. Clinical data corresponding to each sample were collected and updated in a Microsoft Excel file, which was subsequently uploaded to the database. The website was designed to automatically update all database pages and downloadable data sets following each upload.

## Results

### Database contents

The AHP DB serves as an accessible public resource for the collection and reference of AH proteomic data, which is shared with the research community. The user interface developed for the AHP DB allows users to access each component of the database via the following tabs in the navigation bar at the top of each page: 'Home', 'Protein Summary', 'Protein Data', 'Clinical Data' and 'Data Download'.

The 'Home' tab brings the user to the front page of the database, containing a brief description of both the importance of AH and the technologies used to collect and analyze AH proteins for the database. The current version of the database contains 1683 proteins detected in  $>5\%$  of AH samples from 307 subjects.

The 'Protein Summary' tab displays a table of all identified proteins accompanied by the total number and mean number of PSMs for each protein, as well as the percentage of AH samples in which each protein was detected (Figure 1). Based on the data collected to date, the 50 most abundant proteins detected in the AH are listed in Table 1. The top

Accession ID	Gene Symbol	Protein Name	Gene Names	Total PSMs	Mean PSMs	Detected in (%)
P02768	ALBU	Albumin	ALB	1311368	4271.56	100
P02787	TRFE	Serotransferrin	TF	260258	847.75	100
P01009	A1AT	Alpha-1-antitrypsin	SERPINA1	78179	254.65	100
P0DOX5	IGG1	Immunoglobulin gamma-1 heavy chain		75897	247.22	76.22
P01024	CO3	Complement C3	C3	69816	227.41	100
P02790	HEMO	Hemopexin	HPX	62826	204.64	100
P01859	IGHG2	Immunoglobulin heavy constant gamma 2	IGHG2	56495	184.02	100
P00450	CERU	Ceruloplasmin	CP	52563	171.21	100
P01834	IGKC	Immunoglobulin kappa constant	IGKC	51002	166.13	100
P41222	PTGDS	Prostaglandin-H2 D-isomerase	PTGDS	50657	165.01	100

**Figure 1.** Protein Summary page containing accession ID, gene symbol, protein name, gene names, PSM counts and detection level.

**Alt text:** Screen capture of the database's 'Protein Summary' page, which contains a table with columns titled 'Accession ID', 'Gene Symbol', 'Protein Name', 'Gene Names', 'Total PSMs', 'Mean PSMs' and 'Detected in (%)'.

50 most abundant proteins are defined based on their detection levels (in at least 50% of samples) and higher mean PSM counts.

In the database, proteomic data can be accessed through the 'Protein Data' tab, which provides a visual display of the mass spectrometry output data for all identified proteins (Figure 2A). These raw data are also available for download in a separate file. The data are stratified by sample ID and detail the UniProt ID (accession number), score, coverage, number of identified proteins in the protein group, number of unique peptides, number of peptides, PSM count, number of amino acids, protein molecular weight and calculated isoelectric point (pI).

The score of the associated protein reflects the sum of ion scores corresponding to the peptides identified, providing a measure of the confidence in protein identification. The coverage column refers to the proportion of the total protein sequence that could be matched by identified peptides. The PSM count refers to the number of PSMs for each peptide collected and identified by LC-MS/MS. Factors such as run time, protein concentration, and sample variation may influence the total PSM count for any individual protein. The number of proteins refers to the number of protein groups that include the identified protein. The number of unique peptides denotes the number of peptide fragments that uniquely belong to one protein group. The number of peptides indicates the total number of peptide sequences covering a protein group. The design of the database allows each of these column items to be filtered by a search bar and sorted in ascending or descending order (Figure 2B).

To complement the proteomic contents of the AHP DB, a reference clinical data set is also available for the analysis of the proteomic data in the context of clinical and demographic characteristics. The 'Clinical Data' tab displays information about the subjects from whom the AH samples were collected. This includes relevant ocular pathology (cataract or glaucoma), demographic information (sex, age, race and ethnicity), ocular characteristics (IOP, cup area, disc area,

etc.), comorbidities, current medications and other clinical characteristics (Figure 3).

To provide easy access to all of the data for secondary analysis, four separate data sets related to the compiled AH proteomic studies are available for download on the 'Data Download' page (Figure 4). To access the download contents, users need to create an account with a unique username and password through the 'sign-up' page. The 'PSM Matrix' link downloads a file containing the PSM counts of each protein detected via LC-MS/MS, stratified by de-identified sample ID. The UniProt ID (labeled as 'Accession'), total PSM count among all samples and sample-by-sample PSM count are available for each protein. The 'Clinical Data' link downloads a CSV file containing all of the clinical and demographic information corresponding to each sample; this is the same information that is found in the 'Clinical Data' tab. The 'Raw Protein Data' link downloads a file that includes the full mass spectrometry output generated by Proteome Discoverer using the SequestHT algorithm. The data on the database's 'Protein Summary' tab are available for download on the 'Data Download' page as a CSV file by clicking on the 'Protein Summary' link.

### Data search and retrieval

To query information on any of the data display pages, the user can enter a character string into the search bar directly beneath the title of the chosen column. The data set will automatically be filtered to display content containing the desired search term. Additionally, users can input search terms into multiple columns simultaneously. Users can also sort through all displayed data sets using the drop-down menu located on the far right of the associated column name. Each data set within the database can be sorted by column in either ascending or descending order. Furthermore, users can filter and sort the data simultaneously in any visual data set on the database webpage. An example of these functions on the 'Protein Data' page is shown in Figure 2B.

**Table 1.** Top 50 proteins identified in human aqueous humor

Serial no.	Uniprot ID	Gene symbol	Protein name	Mean PSM	Samples Detected In (%)
1	P02768	ALBU	Albumin	4271.56	100
2	P02787	TRFE	Serotransferrin	847.75	100
3	P01009	A1AT	Alpha-1-antitrypsin	254.65	100
4	P0DOX5	IGG1	Immunoglobulin gamma-1 heavy chain	247.22	76.22
5	P01024	CO3	Complement C3	227.41	100
6	P02790	HEMO	Hemopexin	204.64	100
7	P01859	IGHG2	Immunoglobulin heavy constant gamma 2	184.02	100
8	P00450	CERU	Ceruloplasmin	171.21	100
9	P01834	IGKC	Immunoglobulin kappa constant	166.13	100
10	P41222	PTGDS	Prostaglandin-H2 D-isomerase	165.01	100
11	P02766	TTHY	Transthyretin	160.76	100
12	P36955	PEDF	Pigment epithelium-derived factor	158.89	100
13	P10745	RET3	Retinol-binding protein 3	156.18	99.35
14	P02763	A1AG1	Alpha-1-acid glycoprotein 1	141.03	100
15	P01860	IGHG3	Immunoglobulin heavy constant gamma 3	140.51	100
16	P0COL5	CO4B	Complement C4-B	135.16	96.74
17	P0DOX7	IGK	Immunoglobulin kappa light chain	113.8	100
18	P0COL4	CO4A	Complement C4-A	113.22	80.13
19	P02774	VTDB	Vitamin D-binding protein	112.03	100
20	P01861	IGHG4	Immunoglobulin heavy constant gamma 4	105.76	97.39
21	P02647	APOA1	Apolipoprotein A-I	100.74	100
22	P10909	CLUS	Clusterin	96.09	100
23	P0DOY2	IGLC2	Immunoglobulin lambda constant 2	83.8	98.05
24	P01034	CYTC	Cystatin-C	82.54	99.67
25	P19652	A1AG2	Alpha-1-acid glycoprotein 2	82.43	100
26	P01023	A2MG	Alpha-2-macroglobulin	73.1	100
27	P02765	FETUA	Alpha-2-HS-glycoprotein	70	100
28	P04217	A1BG	Alpha-1B-glycoprotein	66.48	100
29	P68871	HBB	Hemoglobin subunit beta	63.86	52.44
30	P22352	GPX3	Glutathione peroxidase 3	62.06	99.67
31	P01008	ANT3	Antithrombin-III	61.99	100
32	P02749	APOH	Beta-2-glycoprotein 1	58	100
33	P06396	GELS	Gelsolin	57.78	100
34	P01876	IGHA1	Immunoglobulin heavy constant alpha 1	56.43	99.35
35	Q9UBP4	DKK3	Dickkopf-related protein 3	54.2	99.67
36	P00751	CFAB	Complement factor B	53	100
37	P01011	AACT	Alpha-1-antichymotrypsin	49.49	100
38	Q13822	ENPP2	Ectonucleotide pyrophosphatase/phosphodiesterase family member 2	49.38	99.67
39	P00738	HPT	Haptoglobin	47.52	97.07
40	Q12805	FBLN3	EGF-containing fibulin-like extracellular matrix protein 1	47.28	100
41	P0DOX8	IGL1	Immunoglobulin lambda-1 light chain	45.77	64.17
42	P00747	PLMN	Plasminogen	44.16	100
43	Q9UBM4	OPT	Opticin	38.5	99.67
44	P07339	CATD	Cathepsin D	36.4	99.35
45	P25311	ZA2G	Plasma protease C1 inhibitor	35.55	100
46	P04264	K2C1	Zinc-alpha-2-glycoprotein	33.56	100
47	P06727	APOA4	Keratin, type II cytoskeletal 1	33.24	71.34
48	P00734	THRB	Apolipoprotein A-IV	32.5	98.7
49	P07998	RNAS1	Prothrombin	31.88	100
50	P25311	ZA2G	Ribonuclease pancreatic	31.44	99.67

## Discussion

AH plays a critical role in maintaining IOP and ocular homeostasis, relying on its diverse array of nutrients, electrolytes and protein content to modulate these functions. Characterizing the human AH proteome will further our understanding of the roles of specific molecules in health and enable more in-depth studies of their relation to pathogenesis. Due to interpersonal variation and the presence of low-abundance proteins, corroborating results can be challenging. Providing a reference database of the human AH proteome can supplement and guide interpretations of AH proteomics in future investigations of ocular pathology and physiology.

The AHP DB provides a publicly accessible collection of proteins identified in human AH using well-validated and highly sensitive mass spectrometry methods. AHP DB stores and presents proteomic information on human AH while concurrently offering protein summaries and corresponding clinical data. Subsets of the data compiled in this database have been analyzed and presented in our previous publications (32, 44, 45), shedding light on inter-population differences in the human AH proteome that may contribute to variations in ocular disease risk factors and incidence. As additional samples are collected and analyzed, the proteomic and corresponding clinical data will be incorporated into the database

## A

Home Protein Summary **Protein Data** Clinical Data Data Download

**Protein Data**

Protein Data Statistics (as of January 02, 2024)

- Total 307 sample(s) in database
- 16275 unique proteins measured
- 167471 total proteins measured

Below is small sample of total protein data. Please go to "Data Download" page to download full dataset.

SampleID	Accession ID	Score	Coverage	Proteins	UniquePeptides	Peptides	PSMs	AAs	MW_kDa	Calc_pI
AH001	A0A075B6K5	17.1925201416	13.91	2	1	1	4	115	12.32407787466	7.38623046875
AH001	A0A075B6N3	2.47639369964	14.78	1	1	1	2	115	12.92035170466	7.84033203125
AH001	A0A075B6P5	3.57409572601	10.83	7	1	1	1	120	12.94845829466	5.93505859375
AH001	A0A087WSY6	18.44853091239	26.09	2	2	2	6	115	12.52629652466	5.18603515625
AH001	A0A0A0MRZ8	9.26722049713	7.83	2	1	1	4	115	12.61733132466	5.28759765625
AH001	A0A0A0MS15	2.71924304962	7.56	1	1	1	1	119	13.04753210466	8.61669921875
AH001	A0A0B4J1U7	32.34989964962	5.79	1	1	1	16	121	13.47283923466	9.20263671875
AH001	A0A0B4J2D5	5.44331049919	7.46	2	1	1	2	268	28.12470200466	8.26513671875
AH001	A0A0C4DH25	36.13084101676	13.79	2	1	1	9	116	12.50723434466	4.58935546875
AH001	A0A0C4DH35	5.53605532646	9.4	1	1	1	2	117	12.80231167466	7.88427734375

Go to page: 1 Show rows: 10 1-10 of 2000

## B

Home Protein Summary **Protein Data** Clinical Data Data Download

**Protein Data**

Protein Data Statistics (as of January 02, 2024)

- Total 307 sample(s) in database
- 16275 unique proteins measured
- 167471 total proteins measured

Below is small sample of total protein data. Please go to "Data Download" page to download full dataset.

SampleID	Accession ID	Score	Coverage	Proteins	UniquePeptides	Peptides	PSMs	AAs	MW_kDa	Calc_pI
AH001	A0A0C4DH25	36.13084101676	13.79	2	1	1	9	116	12.50723434466	4.58935546875
AH001	P01825	38.27573561668	13.79	9	1	1	9	116	12.92761258466	9.29052734375
AH001	P61769	35.48043918609	18.49	1	1	1	9	119	13.70591144466	6.52197265625
AH003	P01825	36.36272978782	13.79	9	1	1	9	116	12.92761258466	9.29052734375
AH005	P01825	32.96547937393	13.79	9	1	1	9	116	12.92761258466	9.29052734375
AH005	Q9NR64	0	5.35	1	1	1	9	748	82.62672227466	6.39208984375
AH003	P0DOX2	32.62929677963	15.82	2	2	4	9	455	48.90321775466	6.66845703125
AH005	P06312	23.40804719924	19.83	1	2	2	9	121	13.37159317466	5.24951171875
AH001	O43505	31.78552985191	12.77	1	3	3	9	415	47.08887467466	7.19580078125
AH003	P24592	33.39134860038	33.75	1	3	3	9	240	25.30616460466	7.81103515625

Go to page: 1 Show rows: 10 1-10 of 15

**Figure 2.** (A) Protein Data page providing visual display of LC-MS/MS analysis of proteins by sample ID and with (B) filters applied by search (right arrow) and by sorting ascending or descending (left arrow).

**Alt text for Figure 2A:** shows a screen capture of the database's 'Protein Data' page, which contains a table with columns titled 'SampleID', 'Accession ID', 'Score', 'Coverage', 'Proteins', 'UniquePeptides', 'Peptides', 'PSMs', 'AAs', 'MW\_kDa' and 'Calc\_pI'.

**Alt text for Figure 2B:** shows a screen capture of the database's 'Protein Data' page with a green arrow pointing to the tool that allows the user to select 'Sort Ascending', 'Sort Descending' or 'Remove Sort' for any column. There is also a blue arrow pointing to one of the boxes below each column name, which allows the user to apply search filters.

at regular intervals. In the future, we also plan to add functions that allow other users to upload their own AH proteomic data to the database.

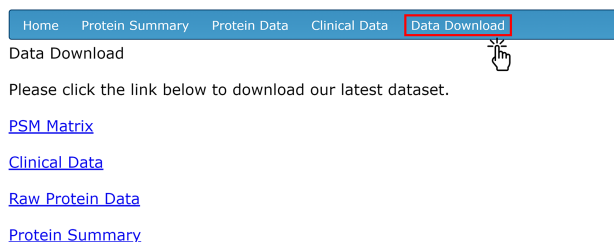
Protein families in the AH can be analyzed to identify differences associated with demographic variables. In one previous

study, the AH proteomic content of 88 subjects undergoing cataract surgery was analyzed to establish a preliminary overview of the human AH proteome. Notably, the levels of certain proteins, including SOD3 and TSPAN14, were found to differ significantly between African American and

Sample ID	C/G	Sex	Age	Race	Ethnicity	IOP	IOP Method	OD/OS	Incisional	Smoking	HTN	CVD	CVA	Collagen VD
AH001	Cataract	M	59	White	Not Hispanic		Tonopen	OS	Phaco	Y	N	Y	N	N
AH003	Cataract	F	48	White	Not Hispanic		Tonopen	OD	Phaco	Y	N	N	N	Y
AH005	Cataract	M	69	White	Not Hispanic		Tonopen	OD	Phaco	Y	N	N	N	N
AH007	Cataract	F	73	Black	Not Hispanic		Tonopen	OD	Phaco	Y	N	N	N	N
AH008	Glaucoma	M	75	Black	Not Hispanic		Tonopen	OS	TS	Y	N	N	N	N
AH009	Glaucoma	M	52	Black	Not Hispanic	24	Goldmann	OS	Phaco	N	N	N	N	N
AH010	Cataract	F	49	Black	Not Hispanic	24	Tonopen	OS	Phaco	N	Y	N	N	N
AH011	Cataract	F	62	Black	Not Hispanic			OD	Phaco	N	Y	N	N	Y
AH012	Cataract	M	62	White	Not Hispanic	14	Tonopen	OS	Phaco	N	Y	N	N	Y
AH013	Cataract	F	70	White	Not Hispanic	20	Tonopen	OS	Phaco	Y	Y	N	N	N

**Figure 3.** Clinical Data page detailing the clinical characteristics and demographic information for each sample.

**Alt text:** Screen capture of the database's 'Clinical Data' page, which contains a table with columns titled 'Sample ID', 'C/G', 'Sex', 'Age', 'Race', 'Ethnicity', 'IOP', 'IOP Method', 'OD/OS', 'Incisional', 'Smoking', 'HTN', 'CVD', 'CVA' and 'Collagen VD'. Additional columns extend beyond the width of the screen.



**Figure 4.** Data download options for protein and clinical data sets.

**Alt text:** Screen capture of the database's 'Data Download' page, which contains text stating 'Please click the link below to download our latest dataset', followed by links titled 'PSM Matrix', 'Clinical Data', 'Raw Protein Data' and 'Protein Summary'.

Caucasian subjects (32). Another prior study revealed sex- and race-specific differences among levels of apolipoproteins in the AH of subjects with primary open-angle glaucoma (45). Also, analysis of 49 samples from subjects with glaucoma revealed associations between the levels of several AH proteins and abnormal visual field parameters, demonstrating the clinical relevance of the AH proteome (44).

This database was generated using shotgun mass spectrometry, also referred to as bottom-up proteomics. Shotgun mass spectrometry represents a robust technique for the identification and quantification of proteins in complex biological samples. There are, however, inherent limitations to this methodology. Shotgun proteomics identifies peptides rather than intact proteins, potentially confounding protein identification in instances of multiple proteins sharing common peptide sequences. It is important to note that shotgun proteomics primarily offers relative quantification rather than absolute quantification. Additionally, shotgun proteomics using data-dependent acquisition prioritizes proteins present in higher abundance, thereby decreasing the reliability of detecting lower abundance proteins.

In conclusion, the current version of AHP DB offers a navigable interface featuring 1683 proteins detected in >5% of AH samples. Users can load and filter data according to clinical or proteomic profiles and subsequently download datasets for further analysis. AHP DB is a novel, freely accessible source of proteomic and corresponding clinical information about human AH. The database provides a user-friendly interface to enable researchers to analyze commonly expressed proteins from the human AH, thereby enhancing knowledge in the field of ophthalmology.

## Funding

National Institutes of Health (NIH) (grant #R01-EY029728 (Bethesda, MD, USA) to Dr A.S.); Center Core Grant for Vision Research (#P30-EY031631).

## Conflict of interest

None declared.

## References

- Goel, M., Picciani, R.G., Lee, R.K. *et al.* (2010) Aqueous humor dynamics: a review. *Open Ophthalmol. J.*, 4, 52–59.
- Sunderland, D.K. and Sapra, A. (2022) *Physiology, Aqueous Humor Circulation*. Treasure Island, FL, USA: StatPearls.
- Weissman, B., Fatt, I. and Rasson, J. (1981) Diffusion of oxygen in human corneas in vivo. *Invest. Ophthalmol. Visual Sci.*, 20, 123–125.
- Macknight, A.D., McLaughlin, C.W., Peart, D. *et al.* (2000) Formation of the aqueous humor. *Clin. Exp. Pharmacol. Physiol.*, 27, 100–106.
- To, C.H., Kong, C.W., Chan, C.Y. *et al.* (2002) The mechanism of aqueous humour formation. *Clin. Exp. Optom.*, 85, 335–349.
- Machiele, R., Motlagh, M. and Patel, B.C. (2022) *Intraocular Pressure*. StatPearls, Treasure Island, FL.

7. Murthy,K.R., Rajagopalan,P., Pinto,S.M. *et al.* (2015) Proteomics of human aqueous humor. *OMICS*, **19**, 283–293.
8. Chowdhury,U.R., Madden,B.J., Charlesworth,M.C. *et al.* (2010) Proteome analysis of human aqueous humor. *Invest. Ophthalmol. Vis. Sci.*, **51**, 4921–4931.
9. Jampel,H.D., Roche,N., Stark,W.J. *et al.* (1990) Transforming growth factor- $\beta$  in human aqueous humor. *Curr. Eye Res.*, **9**, 963–969.
10. Taube,A.B., Hardenborg,E., Wetterhall,M. *et al.* (2012) Proteins in aqueous humor from cataract patients with and without pseudoexfoliation syndrome. *Eur. J. Mass Spectrom. (Chichester)*, **18**, 531–541.
11. Pollreisz,A., Funk,M., Breitwieser,F.P. *et al.* (2013) Quantitative proteomics of aqueous and vitreous fluid from patients with idiopathic epiretinal membranes. *Exp. Eye Res.*, **108**, 48–58.
12. Bennett,K.L., Funk,M., Tschernutter,M. *et al.* (2011) Proteomic analysis of human cataract aqueous humour: comparison of one-dimensional gel LCMS with two-dimensional LCMS of unlabelled and iTRAQ(R)-labelled specimens. *J. Proteomics*, **74**, 151–166.
13. Perumal,N., Manicam,C., Steinicke,M. *et al.* (2017) Characterization of the human aqueous humour proteome: a comparison of the genders. *PLoS One*, **12**, e0172481.
14. Anshu,A., Price,M.O., Richardson,M.R. *et al.* (2011) Alterations in the aqueous humor proteome in patients with a glaucoma shunt device. *Mol. Vis.*, **17**, 1891–1900.
15. Kliuchnikova,A.A., Samokhina,N.I., Ilina,I.Y. *et al.* (2016) Human aqueous humor proteome in cataract, glaucoma, and pseudoexfoliation syndrome. *Proteomics*, **16**, 1938–1946.
16. Funke,S., Perumal,N., Bell,K. *et al.* (2017) The potential impact of recent insights into proteomic changes associated with glaucoma. *Expert Rev. Proteomics*, **14**, 311–334.
17. Chiang,S.Y., Tsai,M.L., Wang,C.Y. *et al.* (2012) Proteomic analysis and identification of aqueous humor proteins with a pathophysiological role in diabetic retinopathy. *J. Proteomics*, **75**, 2950–2959.
18. Hubens,W.H.G., Mohren,R.J.C., Liesenborghs,I. *et al.* (2020) The aqueous humor proteome of primary open angle glaucoma: an extensive review. *Exp. Eye Res.*, **197**, 108077.
19. Ji,Y., Rong,X., Ye,H. *et al.* (2015) Proteomic analysis of aqueous humor proteins associated with cataract development. *Clin. Biochem.*, **48**, 1304–1309.
20. Kim,T.W., Kang,J.W., Ahn,J. *et al.* (2012) Proteomic analysis of the aqueous humor in age-related macular degeneration (AMD) patients. *J. Proteome Res.*, **11**, 4034–4043.
21. Liu,X., Liu,X., Wang,Y. *et al.* (2021) Proteome characterization of glaucoma aqueous humor. *Mol. Cell. Proteomics*, **20**, 100117.
22. Rinsky,B., Beykin,G., Grunin,M. *et al.* (2021) Analysis of the aqueous humor proteome in patients with age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.*, **62**, 18.
23. Xiang,M., Zhang,X., Li,Q. *et al.* (2017) Identification of proteins in the aqueous humor associated with cataract development using iTRAQ methodology. *Mol. Med. Rep.*, **15**, 3111–3120.
24. Xiao,H., Xin,W., Sun,L.M. *et al.* (2021) Comprehensive proteomic profiling of aqueous humor proteins in proliferative diabetic retinopathy. *Transl. Vis. Sci. Technol.*, **10**, 3.
25. Yao,J., Chen,Z., Yang,Q. *et al.* (2013) Proteomic analysis of aqueous humor from patients with branch retinal vein occlusion-induced macular edema. *Int. J. Mol. Med.*, **32**, 1421–1434.
26. Yao,J., Liu,X., Yang,Q. *et al.* (2013) Proteomic analysis of the aqueous humor in patients with wet age-related macular degeneration. *Proteomics Clin. Appl.*, **7**, 550–560.
27. Yu,M., Xie,F., Liu,X. *et al.* (2020) Proteomic study of aqueous humor and its application in the treatment of neovascular glaucoma. *Front. Mol. Biosci.*, **7**, 587677.
28. Pessuti,C.L., Medley,Q.G., Li,N. *et al.* (2023) Differential proteins expression distinguished between patients with infectious and noninfectious uveitis. *Ocul. Immunol. Inflamm.*, 1–8.
29. Choi,J.A., Ju,H.H., Lee,J. *et al.* (2022) Increased complement-associated inflammation in cytomegalovirus-positive hypertensive anterior uveitis patients based on the aqueous humor proteomics analysis. *J. Clin. Med.*, **11**, 2337.
30. Galardi,A., Stathopoulos,C., Colletti,M. *et al.* (2022) Proteomics of aqueous humor as a source of disease biomarkers in retinoblastoma. *Int. J. Mol. Sci.*, **23**, 13458.
31. Busch,M.A., Haase,A., Miroshnikov,N. *et al.* (2022) TFF1 in aqueous humor - a potential new biomarker for retinoblastoma. *Cancers (Basel)*, **14**, 677.
32. Kodeboyina,S.K., Lee,T.J., Churchwell,L. *et al.* (2020) The constitutive proteome of human aqueous humor and race specific alterations. *Proteomes*, **8**: 34.
33. Qian,W.J., Jacobs,J.M., Liu,T. *et al.* (2006) Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol. Cell. Proteomics*, **5**, 1727–1744.
34. Clarke,W., Rhea,J.M. and Molinaro,R. (2013) Challenges in implementing clinical liquid chromatography-tandem mass spectrometry methods—the light at the end of the tunnel. *J. Mass Spectrom.*, **48**, 755–767.
35. Karpievitch,Y.V., Polpitiya,A.D., Anderson,G.A. *et al.* (2010) Liquid chromatography mass spectrometry-based proteomics: biological and technological aspects. *Ann. Appl. Stat.*, **4**, 1797–1823.
36. Yu,T. and Okamura,R. (1987) Comparative study of native proteins in aqueous humor and serum—detection of characteristic aqueous humor proteins. *Jpn. J. Ophthalmol.*, **31**, 235–248.
37. Hu,Y., Malone,J.P., Fagan,A.M. *et al.* (2005) Comparative proteomic analysis of intra- and interindividual variation in human cerebrospinal fluid. *Mol. Cell. Proteomics*, **4**, 2000–2009.
38. Jacobs,J.M., Adkins,J.N., Qian,W.-J. *et al.* (2005) Utilizing human blood plasma for proteomic biomarker discovery. *J. Proteome Res.*, **4**, 1073–1085.
39. Hu,S., Jiang,J. and Wong,D.T. (2010) Proteomic analysis of saliva: 2D gel electrophoresis, LC-MS/MS, and western blotting. *Methods Mol. Biol.*, **666**, 31–41.
40. Baldwin,M.A. (2004) Protein identification by mass spectrometry: issues to be considered. *Mol. Cell. Proteomics*, **3**, 1–9.
41. Di Girolamo,F., Lante,I., Muraca,M. *et al.* (2013) The role of mass spectrometry in the “Omics”. *Era. Curr. Org. Chem.*, **17**, 2891–2905.
42. Soria,J., Acera,A., Merayo-lloves,J. *et al.* (2017) Tear proteome analysis in ocular surface diseases using label-free LC-MS/MS and multiplexed-microarray biomarker validation. *Sci. Rep.*, **7**, 1–15.
43. Pitt,J.J. (2009) Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin. Biochem. Rev.*, **30**, 19–34.
44. Kodeboyina,S.K., Lee,T.J., Bollinger,K. *et al.* (2021) Aqueous humor proteomic alterations associated with visual field index parameters in glaucoma patients: a pilot study. *J. Clin. Med.*, **10**, 1180.
45. Patel,P.A., Lee,T.J., Kodeboyina,S.K. *et al.* (2021) Intra-population differences of apolipoproteins in the aqueous humor. *Lipids Health Dis.*, **20**, 128.

