

COMMENTARY

**OPEN ACCESS**  
Full open access to this and  
thousands of other papers at  
<http://www.la-press.com>.

## Prospects for Proteomics Directed Genomic and Genetic Analyses in Disease Discoveries

Sanjoy K. Bhattacharya

Bascom Palmer Eye Institute, 1638 NW 10th Avenue, Suite 706A, University of Miami, Miller School of Medicine, Miami, FL, 33136, USA. Email: [sbhattacharya@med.miami.edu](mailto:sbhattacharya@med.miami.edu)

---

**Abstract:** Proteomic discoveries are usually made using database searches for identification of proteins in a given protein sample derived from cells or tissues. High throughput searches leave a number of peptides not analyzed for a variety of reasons, such as posttranslational modification or a mutation that results changes in the peptide that is not present in databases. Such mutations may be critically important in causing disease conditions. Accounts from ocular diseases are presented where the search provided results often from non-conventional databases (such as structural database instead of protein database) due to the presence of information about a mutant peptide. We contemplate that better algorithms and the ability to determine probabilities of different amino acids in the available sequence may permit combinatorial analysis with genomics which may help identify new disease associated mutations directly from the sequence of the captured peptides. In addition, the *de novo* analysis of spectra of the unidentified peptides may provide mutation or polymorphism information enabling additional insight about the disease association of a mutation or posttranslational modification.

**Keywords:** proteomics, disease discoveries

---

*Proteomics Insights* 2009:2 23–26

This article is available from <http://www.la-press.com>.

© the authors, licensee Libertas Academica Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://www.creativecommons.org/licenses/by/2.0>) which permits unrestricted use, distribution and reproduction provided the original work is properly cited.

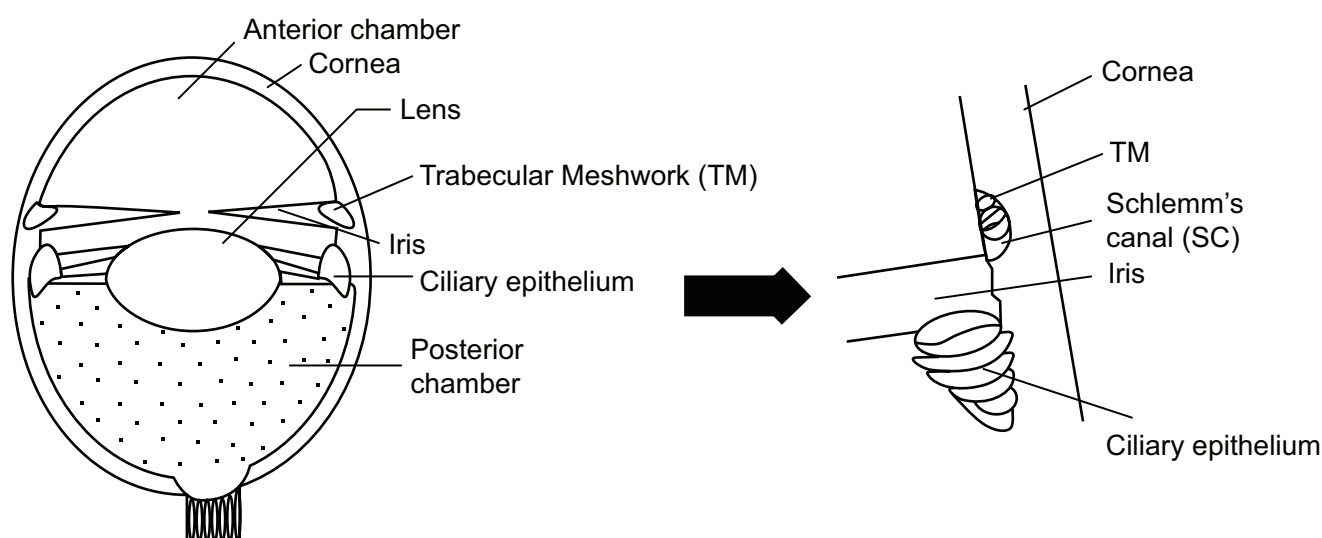
## Introduction

Since 1990s mass spectrometry has gained momentum for the analysis of cell and tissue proteomes.<sup>1</sup> Identification of proteins through mass spectrometry of trypsin digested peptides using a solid phase matrix assisted laser desorption/ionization (MALDI) device utilizes time of flight (TOF) as a measure of the mass to charge ratio. Calibration is performed using the mass of synthetic peptides and accurate masses of measured peptides are used to identify protein matches using databases. The lower and upper extremes of the spectrum of experimentally determined peptide masses utilized for such calibration are usually inaccurate, as calibration standards do not extrapolate in a linear fashion. This approach is often referred to MS<sup>1</sup> analysis (or simply as MS analysis) often leads to inaccurate and erroneous identification. In this type of analyses, even if the search is restricted to a given organism databases, a number of peptide masses are often left unidentified. That is, some peptide masses even within the well calibrated region of the spectra often cannot be identified, if the search is restricted to specific organisms or databases. This is due to posttranslational modification or unrecorded haplotypes or disease mutations. Despite explosive growth in mass spectrometry and the fast changing pace in which new technology is becoming available in this area, a large number of biologists still apply MALDI-TOF methodology where the

lack of identification becomes inherent and valuable information is left behind.

Superior mass spectrometers, based on matrix ionization or electrospray ionization, utilize MS-MS or MS<sup>2</sup> and of late often MS<sup>n</sup> where n is greater than 2. In MS-MS devices, the selected precursor is broken down in a collision cell with an inert gas in a chemically defined manner from which amino acid sequences can be determined.<sup>1,2</sup> Database searches which utilize these peptide sequence matches also suffer from some of these limitations. That is if there is no algorithm for possible mutation in one or more amino acids compared to existing information for the peptide and unknown posttranslational modifications of a peptide then that peptide is not ascribed to alignment against a sequence and is left aside as unanalyzed.

Unrecorded mutation in the identified peptides is one of the reasons and posttranslational modifications are another reason that leave behind a number of unidentified peptides obtained in mass spectrometry experiments. Often some program searches invariably look into structural databases (for example, some versions of ProteinLynx™ software) and find a match due to presence of the peptide sequence in structure databases instead of regular protein or gene sequence databases. However, improvement of search algorithms and incorporation of possibility of mutation may enable capture of a mutation directly from the peptide and enable further investigation at



**Figure 1.** A schematic diagram of the eye and an enlargement of the anterior chamber angle. The regions of the eye are as indicated.

**Table 1.** Mutant proteins identified glaucomatous TM.

Protein identification	Accession number*	Peptide matches	Donors <sup>†</sup>
Lysozyme (Human Mutant A96I, W109h)-Chain A	<b>1BB4</b>	1	1
Hemoglobin A Chain	<b>1C7D</b>	3	2
Hemoglobin beta chain variant Hb S-Wake	AAN11320	1	1
Keratin 9 (mutant)	O00109	1	2
Kerato-epithelin (R124C mutant)	Q15582/AAB88695	1	2
Hbbm fused globin	Q14477	4	2

<sup>†</sup>Number of donors exhibiting the indicated protein.

\*SwissProt, NCBI protein and NCBI structure accession numbers are in regular, italics and bold italic types respectively.

the genomic level and performing targeted genetic analysis.

## Identification of Mutant Proteins in Ocular Diseases

### Glaucoma

Proteomic analysis of trabecular meshwork (TM) tissue (Fig. 1), a region that offers increased resistance to aqueous outflow in the anterior eye chamber identified a number of mutant proteins (Table 1) in glaucomatous TM tissue compared to normal controls.<sup>3</sup> A number of the identified mutant proteins were associated with the circulatory system, consistent with the vascular theory of glaucoma. Two proteins were identified from structure database: human Lysozyme mutant A96I, W109h (Accession number: 1BB4) and Hemoglobin A Chain (1C7D). This is perhaps due to the fact that a match was readily identifiable in structure database but not in the protein database. The other identified mutant proteins are: hbbm fused globin (Accession number: Q14477), and Hb-S-Wake (Accession number: AAN11320), mutant keratin 9 (O00109) and kerato-epithelin (Q15582; Table 1). One interesting aspect of these mutations is that they occur with high frequency among African-Americans,<sup>3</sup> which is consistent with the clinical finding that African-American have relatively higher prevalence rates of glaucoma.<sup>4</sup>

### Pseudoexfoliation syndrome

The results of mass spectrometric identification of proteins in pseudoexfoliation syndrome were recently presented.<sup>5–7</sup> Antibody pattern in the aqueous humor from glaucoma and pseudoexfoliation syndrome patients also have been subjected to mass spectrometric

pattern determination.<sup>8</sup> The variant peptides of lysyl lysyl oxidase (LOXL1) were detected directly in the pseudoexfoliation syndrome material.<sup>7</sup>

## Concluding Remarks

Proteomics is applied to identify the proteins in cells and tissues. The peptide sequence of a given protein may differ from the gene sequence translated product in a database for a variety of reasons, such as gene mutation, gene splicing and posttranslational modifications. The *de novo* sequencing of unidentified peptide masses or better algorithms for incorporation of mutation or posttranslational modification may aid in searches. In the future, due to such efforts, direct identification of a mutant peptide may shorten the genomic and genetic discoveries linking the mutant product with disease or pathogenic conditions.

## Acknowledgements

I thank all individuals who participated in the original studies. The original work was partly supported by NIH grants EY015266 and EY016112.

## Disclosure

The author reports no conflicts of interest.

## References

1. Godovac-Zimmermann J, Brown LR. Perspectives for mass spectrometry and functional proteomics. *Mass Spectrom Rev.* 2001;20:1–57.
2. Godovac-Zimmermann J, Mulvey C, Konstantoulaki M, Sainsbury R, Brown LR. Promise of multiphoton detection in discovery and diagnostic proteomics. *Expert Rev Proteomics.* 2007;4:161–73.
3. Bhattacharya SK, Rockwood EJ, Smith SD, et al. Proteomics reveals cochlin deposits associated with glaucomatous trabecular meshwork. *J Biol Chem.* 2005;280:6080–4.
4. Wilson MR, Martone JF. Epidemiology of Chronic Open-Angle Glaucoma. In *The Glaucomas* (2nd edn) (Ritch R, et al. eds), Mosby 1996.



5. Lee RK. The molecular pathophysiology of pseudoexfoliation glaucoma. *Curr Opin Ophthalmol.* 2008;19:95–101.
6. Lee R, Bhattacharya SK. Molecular pathophysiology of pseudoexfoliation glaucoma-genes and proteins. In *XVIII International Congress for Eye Research.* 2008:114.
7. Sharma S, Chataway T, Burdon KP, et al. Identification of LOXL1 protein and Apolipoprotein E as components of surgically isolated pseudoexfoliation material by direct mass spectrometry. *Exp Eye Res.* 2009;89:479–85.
8. Joachim SC, Wuenschig D, Pfeiffer N, Grus FH. IgG antibody patterns in aqueous humor of patients with primary open angle glaucoma and pseudoexfoliation glaucoma. *Mol Vis.* 2007;13:1573–9.

**Publish with Libertas Academica and every scientist working in your field can read your article**

*“I would like to say that this is the most author-friendly editing process I have experienced in over 150 publications. Thank you most sincerely.”*

*“The communication between your staff and me has been terrific. Whenever progress is made with the manuscript, I receive notice. Quite honestly, I’ve never had such complete communication with a journal.”*

*“LA is different, and hopefully represents a kind of scientific publication machinery that removes the hurdles from free flow of scientific thought.”*

**Your paper will be:**

- Available to your entire community free of charge
- Fairly and quickly peer reviewed
- Yours! You retain copyright

**<http://www.la-press.com>**