PROTEOMICS: A TOOL FOR BETTER UNDERSTANDING OF CELIAC DISEASE

Muhammad Asif Rasheed
University of Tampere
Institute of Biomedical Technology
Faculty of Medicine
Master's thesis
June 2011

DEDICATION

ALLAH (*Subhanahu wa-ta'ala*), the most merciful and kind. The dedication goes to Allah, who always helped me during my problems and showed me the right path.

ACKNOWLEDGEMENT

The practical and written part of the thesis was done in Institute of Biomedical Technology (IBT), University of Tampere. First, I owe special thanks to Professor Mauno Vihinen for discussion and important counsel. I am thankful to Institute of Biomedical Technology (IBT) and celiac disease research group for giving me a great possibility to work and learn here.

I am most grateful to my family, specially my parents, for their ethical and financial support. Without their precious prayers and encouragement, I had not been able to accomplish my goals.
I am very gratified to Professor Markku Mäki, Katri Kaukinen and Katri Lindfors to give me the opportunity to work in celiac disease research group as well as for their support during this time.

I direct my deep appreciations to Sergio Caja (supervisor), Johanna M Salonen, Laura Airaksinen, Suvi Kalliokoski, Cristina Nadalutti, Tiina Rauhavirta and Boglarka Toth for their help, motivation and guidance.

MATER'S THESIS

Place:	UNIVERSITY OF TAMPERE Faculty of Medicine Institute of Biomedical Technology (IBT)
Author:	RASHEED, MUHAMMAD ASIF
Title:	Proteomics: A tool for better understanding of celiac disease
Pages:	67 pp.
Supervisors:	Professor Mauno Vihinen, Sergio Caja
Reviewers:	Professor Mauno Vihinen, Sergio Caja
Date:	June 2011

ABSTRACT

Background and aims: Celiac disease (CD) is a small intestine autoimmune disorder trigger by the ingestion of gluten containing cereals (wheat, barley and rye) in genetically predispose individuals (HLA-DQ2 or DQ8). The disease course with malabsorption, diarrhea, fail in child growth and extra intestinal manifestation such as ataxia, bone diseases, heart diseases, skin diseases. In addition, the diagnosis is based on the study of villous atrophy together with the titre of the antibodies against transglutaminase 2 (TG2), anti deamideated gliadine antibodies and endomysium antibodies (EMA) in both serum and duodenal biopsy of the expected person. On the other hand, proteomic analysis of the serum samples gives an idea regarding diagnosis and the treatment of this disease as well as the development of new markers. We aim to find out the difference between two methods of protein depletion in order to find more proteins and more reliable results during proteomic analysis of CD samples. Moreover, we aim to setup a protocol to study the circulating proteins in the blood of CD patients in order to better understand the diseases pathogenesis.

Methods: Samples from celiac patients' serum during active disease, same patients' serum after one year of gluten free diet and healthy donors were prepared for proteomic analysis. We used 10 samples in each group. The majority of proteins were depleted by two different methods (A and B) while separated by isoelectric focusing and 2-D gel electrophoresis and identified by Malditof-tof.

Results: The spots found by using method A were 348 while in method B, the amount was considerably increased to 689. Method B was selected for further analysis and statistically significant proteins were found. Some proteins were up or downregulated in celiac disease patients compared with the other groups.

Conclusions: We setup a suitable protocol for the proteomic analysis of serum samples derived from CD. Depletion of major proteins by method B compare to method A allows us to find more proteins and more reliable results. Most of the identified proteins were related to vascular damage, suggesting their participation in the angiogenesis defect during the disease. Hence, protein profile during active CD and GFD can help in the study of the pathogenesis of the disease.

Table of Contents

ABBREVIATIONS

2DE	Two-dimensional gel electrophoresis
APS	Ammonium pursulfate
BLAST	Basic local alignment search tool
CD	Celiac disease
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DTT	Dithiothrietol
EMA	Anti-endomysium antibody
ESI	Electrospray ionisation
IAA	Iodoacetamide
IEF	Isoelectric focusing
IEL	Intraepithelial leukocyte
IgG-tTG	IgG antibodies against transglutaminase
IL-15	Interleukin-15
KD	kilo dalton
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
NCBI	National center for biotechnology information
NMR	Nuclear magnetic resonance
PFF	Peptide fragmentation fingerprinting

pI	Isoelectric point
PMF	Peptide mass fingerprinting
RP-HPLC	Reverse phase high pressure liquid chromatography
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TEMED	Tetramethylethylenediamine
TG2	Transglutaminase 2
tTGA	Tissue transglutaminase antibody
TOF	Time of flight

1 Proteomics

1.1 Introduction and history

Proteomics is the study of structures and functions of proteins which are the main components of the physiological and metabolic pathways of cells. In fact, genome expression can be described by the proteins which are responsible for regulation and normal functioning of the cell[1]. Different cells and tissues of the body contain different proteins which may vary due to certain environmental and disease conditions. Most of the functions in an organism are performed by the macromolecular protein complexes for example DNA replication machinery, transcription machinery, protein-folding apparatus and many other protein complexes[2]. The cell processes, and the effect of proteins on the cell processes can be well understood through proteomics studies. It can also be helpful in studying the effect of the enviroment and the cell processes on the proteins. The differences between the diferent conditions of an organism can be highlighted and analyzed by the proteomics study. Moreover, serum proteomics lead to discover the new classes of biomarkers[3].

In order to get the protein expression status, different techniques from the field of proteomics must be used. The classic proteomics platforms include the two-dimensional gel electrophoresis (2DE) and mass spectrometry. In the past decades, the two-dimensional gel electrophoresis (2DE) and usual protein staining techniques were used as primary means to conduct comparative experiments. This technique was first introduces in 1975 by O'Farrel[4]. After the many improvements in 2DE technology, its popularization in the 1980s, and its use in conjunction with MS technology, it definitely became a major tool in a wide range of proteomics research[5 6]. The main advantages of 2DE include its ability to separate a wide number of proteins and visualize them[7]. The 2DE process is based on two independent separation methods, the first of which is isoelectric focusing (IEF). Through this process, differently charged proteins are separated on the basis of their isoelectric points by using a pH gradient strip. In the second step, the strip is transferred to a large

polyacrylamide gel, and the proteins are separated on the basis of their molecular masses. One instrument for second dimension is shown in Figure 2. So each 2DE gel generates a protein profile which can not only be visualized as fluorescent labelled spots but also stained by other methods e.g. silver staining and coomasie etc that represent the proteins. One common used method to label the proteins is to use three different CyDye dyes. The proteins are fluorescently labelled by three different CyDye DIGE fluor minimal dyes i.e. Cy2, Cy3 and Cy5. However it is not necessary that a single spot represents a single protein on the gel. In fact, Gygi et al. have shown that as many as six yeast proteins can co-migrate as a single spot on a 2D gel[8]. Similarly one protein can be represented in different spots as well for example the unmodified or post translational modified proteins. The technique has been used for over 30 years, and its reproducibility was clearly improved with the introduction of immobilized pH gradient gel strips and bioinformatics[7]. This technique is productive in providing relevant data about biological systems. Several authors have utilized this strategy to investigate plant protein expression in organelles[9 10 11 12]. Due to the high sensitivity and reproducibility of this method, it can be used in comparative proteomics studies. Computational and the physical components of the proteomics can be studied through mass spectrometry (MS) which is essential for protein identification, and is commonly associated with electrophoresis techniques. Mass spectrometry (MS) is an analytical technique that measures the mass-to-charge ratio of charged particles and the principle consists of ionizing chemical compounds to generate charged molecules or molecule fragments and measurement of their mass-to-charge ratios[13]. For this purpose, many techniques have been utilized, including Matrix-Assisted Laser Desorption/Ionization (MALDI) and Electrospray Ionisation (ESI)[14 15]. Ion trap and triple-quadruple tandem MS (MS⁄MS) spectrometers have also improved the sensitivity and mass accuracy[16].

Figure 1: Ettan IPGphor. Differently charged proteins can be separated by this machine on the basis of their isoelectric points by using an immobilized pH gradient strip

Figure 2: Ettan DALTsix. Proteins are separated on the basis of their molecular masses using polyacrylamide gel by this machine

Serum can be considered as a tissue containing proteins derived from cells as well in suspension[17]. The fluid part consist of the macromolecules such as immunoglobin, complement cascade and acute phase reactants while the cellular fraction consist of red blood cells, white blood cells and platelets. In most of proteomic analysis, separation of these components is required prior to analyse them to determine the basic set of proteins produced in a cell which is also helpful in determining the pathogenesis and diagnosis of a disease on the basis of proteins present in different body tissues and discovery of new markers of that disease.

Through applying the proteomics tool (Figure 3), and identifying the proteins in disease and the healthy individuals' samples, we can identify the biomarkers to differentiate the both classes. Such discoveries can lead to protein-based diagnostic tools and better understanding of a disease state. For example, protein expression pattern or profile during cancer is different compared to a normal healthy situation. Such unique proteins which are present in the diseased condition and not in the healthy cells can be used as markers for the disease study as well as target during the disease treatment.

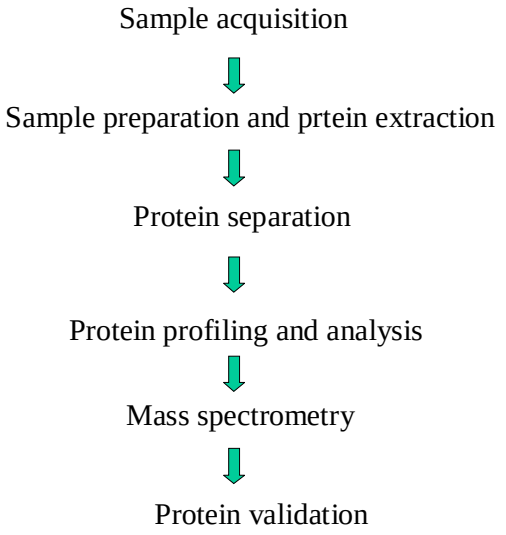

Figure 3: A workflow of proteomic analysis

1.2 Application of proteomics

The analysis of the proteins using a proteomic approach is very useful and has different applications. For example, proteins isolated from body tissues gives a complete idea of the tissue situation as well as provided a good basis for to study other processes such as post translational modifications, protein functionality or protein complexes[18]. In the next paragraphs, some of the most important applications have been described.

1.2.1 Drug discovery

Proteomics facilitated the ever increasing discovery of new, novel drug targets, their respective modes of action mechanistically, and their biological toxicology. During a drug discovery process, the determination of exact cause of the disease and bring it to the normal level is a big challenge. Studies have shown that there is a poor correlation between the regulation of transcripts and actual protein quantities because genome analysis does not account for post-translational processes e.g. protein modifications and protein degradation. So the methods in the drug-discovery process started to shift from genomics to proteomics[19 20]. Hence, the analysis of proteome has a much more accurate approach to identifying not only applicable biomarkers that will aid in diagnosis, but also effective remedies for diseases of varying origins[21 22 23].

1.2.2 Protein chips

After the success of DNA microarray, the researchers were inspired to develop similar kind of technique for the proteomics experiment which refers to the protein microarray. A quick and easy survey of the entire proteome of a cell within an organism is enabled through the protein chips. They are first developed by the researchers at Harvard university in the year 2000[24]. Protein chip technology is essential for high throughput functional proteomics[25]. In order to detect the concentration of the antigen-antibody interactions in biological samples, an expansive creation was immunoassays, the precursor to protein chips available since the 1980s. The researchers at Harvard

University combined the technologies of immunoassays and DNA microarrays to develop the protein chip[24]. During the analysis of the chip, there is the reading of the protein levels off the chip, and then the use of computer software to analyze the collected data. While only a small amount of protein is required for high throughput screening[26]. Protein chip experiments have many applications include identifying biomarkers for diseases, investigating protein-protein interactions, cancer research and medical diagnosis and testing for the presence of antibodies in a sample. Many companies, e.g. Biacore, Invitrogen, and Sigma-Aldrich, have started the production of industrial level protein array systems within the past five years. It can also be used for drug discovery and basic biological research.

1.2.3 Biomarker characterization

A biomarker is a characteristic that can be objectively measured and evaluated as an indicator of normal and disease processes[27]. The biomarker can be used to possibly predict and prevent the related disease. To validate a biomarker successfully, series of six steps should be accomplished which includes discovery, qualification, verification, assay optimization, validation and commercialization[28]. With the passage of time, as the technology advances, we will be able to get the different detections through a single measurement. So quantification of proteins and peptides for personalized medicine is one of the main goals. Through a single assay, possible diagnosis of many diseases can be done by multiplex the biomarkers[29].

1.2.4 Functional genomics and peptidomics

Functional genomics describes gene functions and interactions while in peptidomics, we analyse the peptide content within an organism, tissue, or even cell. The availability of genome sequences from different organisms was a breakthrough of the year 2000[30]. Such sequencing gives the information about evolutionary relationships and the diseases which are communicable through genetics. As the mRNA level often does not correlate with the proteins expression, similarly, the protein expression

level does not correlate with the functioning of those proteins[31] [32]. So in order to get better understanding of the proteins, their functions should be determined. There can be some modifications in the proteins i.e. post translational modifications for example phosphorylation, glycosylation and sulphation. These modifications as well as many other modifications are extremely important for protein function as they can determine activity, stability, localization and turnover[33]. Continuing progress has been made in this field, especially in the case of phosphorylation. Such events can be studied by generic strategies, because phosphopeptides are 80 Da heavier than their unmodified counterparts[34] [35] [36] [37] [38].

1.2.5 Quantitative proteomics

During classical proteomics approach, the total protein complement of a cell or tissue is electrophoresed on a two-dimensional gel to separate them on the basis of their charge and mass. By this approach, several hundred proteins can typically be visualized by radioactive or fluorescent labeling or silver staining. While the differences between the reference and altered states are measured by quantifying the ratios of spot intensities between independent twodimensional gels[39]. So the determination of the presence or absence of one protein sometimes is not sufficcient but to know the abundance in different conditions is more critical. Quantitative proteomics provides quantitative information about the proteins in the sample[40]. Determination of relative abundance of different proteins in two different conditions is more crucial. Sample preparation is one of the most important processes in quantification which influences the sensitivity very much. Depletion of highly abundant proteins is another very important step which increases the analytical dynamic range. The most important may be to know how the composition of protein complex changes in different cell types under different conditions, and after post translational modifications in order to study cellular signaling in that particular conditions. The changes in the protein abundance should be quantified to compare two different conditions for example disease and healthy one. Quantitative proteomics not only identify the proteins, but also their abundance in a sample is determined[41]. This

information can be used to determine the statistically different proteins on the basis of their abundance.

1.3 Limitations of proteomics

Despite the fact that proteomics over the time have demonstrated a very good method in the analysis of proteins, but it has some limiting factors that should be considered before to run any assay. For example, the 2D gel electrophoresis is although a standard technology, but it has also some limitations as well. The issues are related to its reproducibility and the separation of non soluble hydrophobic proteins[42]. However, by using the reagents e.g. thiourea, this issue can be partially resolved. Another problem related to 2D gel electrophoresis is the detection of low abundance proteins present in the sample[42 43]. If the major proteins are present in the sample and are not treated previously, then the quantity of florescence will be high in the major proteins which will decrease the florescence for low abundance proteins and they will not be detected during the scanning and staining processes. So this process is not suitable for the proteins having a size less than 8KD. During proteomics studies, the issue of proteins modifications should be considered because due to the modifications e.g. post translational modifications and glycosilation there can be different results with the same methodology. Similarly different methodologies in proteomics may lead to different results in the same samples.

1.4 Proteomics and bioinformatics

There are continuous improvements in the field of proteomics with the introduction of new technologies and significant improvements have been made in the last decade. For example, high throughput acquirement of proteome data is possible now a days. Bioinformatics introducing new algorithms in the field of proteomics to handle large and heterogeneous data sets. With the integrated use of "-omics" disciplines, the identification and the characterization of candidate genes, proteins and molecules involved in a given disease will probably represent one of the milestones of

future health care[44]. Proteomics deals with the identification of the proteins produced by cells in normal and diseased conditions, while metabolomics monitors the role of small molecules (lipids, sugars and amino acids) involved in daily cellular function[45]. Since the complete sequencing of the human genome, the development of mass scale research tools in this field, such as DNA microarrays, has greatly improved[45]. There are approximately 30,000 genes in the human genome and the number of proteins is likely at least three times higher, as resulting from alternative splicing and posttranslational modifications[46]. Microarray technique allows a complete profiling of the expression of a very high number of genes in a given tissue in a single experiment[47]. DNA microarrays can be used to measure the relative gene expression in different conditions, e.g. normal versus pathological tissues[47].

In the last decade, the protein separations techniques (e.g. 2DE) and protein analysis (e.g. mass spectrometry) were seen as the key technologies developed by Klose and O' Farrel for protein separation and imaging[48 49]. New algorithms for image analysis of two dimensional gels have been developed. Moreover, mass spectrometry data analysis algorithms for peptide mass fingerprinting (PMF) and peptide fragmentation fingerprinting (PFF) have been developed. After the completion of human genome project, more information is required in the field of proteomics. Other proteomics technologies include protein biochips, detection of protein-protein interaction by yeast two-hybrid system and the technologies for the determination of protein structure e.g. X-ray crystallography and NMR spectroscopy.

Proteomics is a vast field which can be subdivided into functional proteomics, structural proteomics, expression proteomics and interaction proteomics. Functional proteomics address to protein functions e.g. differential protein expression analysis and protein-protein interaction. In structural proteomics, there is analysis of protein structures with different techniques for example, X-ray crystallography and NMR spectroscopy[50]. This structural identification and analysis can help to identify the functions of newly discovered genes, protein interactions and the binding sites of a

protein. Expression proteomics helps to analyse the protein expressions in a sample. In interaction proteomics, there is analysis of different proteins interactions[50]. Such analysis is helpful to determine the proteins functions and their assembly in the large complexes. For this purpose, technologies e.g. yeast two-hybrid system, affinity purification, and mass spectrometry are very helpful. The analysis of the complex network of connections between proteins may allow the identification of potential molecular markers and targets[44 51]. Gene interaction networks have recently been demonstrated and described in several biological processes[47 52 53 54 55]. Originally developed for the analysis of gene sequences, bioinformatics is now widely applied to the study of gene and protein expression and to the analysis of genes and protein networks[56].

Many disciplins are part of the bioinformatics field, and for example biostatistics, biomathematis, computacional biology and biometrics are importnat parts in the field. Bioinformatics is also involve in different informatics disciplines e.g. databases, knowledge representation, software development methods and algorithm development. In the analysis of proteome data, these disciplines are needed for the knowledge discovery process. Lots of promising bioinformatics developments and the potential for knowledge discovery in biochemistry draws a lot of attention towards the field of bioinformatics in proteomics.

1.4.1 2D gel image analysis

After 2DE, the protein spots can be seen as stained protein spots on polyacrylamide gel. The spots can be detected and analyzed in differential proteome studies by sophisticated software packages. Some of these software are presented in the Table 1[50]. The software solutions differ in the applied algorithm and workflow. Some software performs image registration of whole images before spot detection[57]. Hence, each program analyzes the gel image in one way or other. These approaches potentially serve the image comparison in a more sophisticated way. The proteomics study contains several steps. The sample is prepared for protein separation studies. This step is very significant as it may influence the proteome and therefore, the exact protocol should be followed. To separate the

proteins several separation techniques are applied in proteomics. 2DE is the most widely spread technique[58] [49]. Other methods include chromatographic methods e.g. RP-HPLC, affinity chromatography, ion exchange chromatography or free-flow electrophoresis[59]. After 2DE, the proteins are cleaved into peptides by enzymatic digestion (e.g. by trypsin). This maybe also applied within a multidimensional chromatography. Finally the proteomics data should be stored and published to ensure the ability to query project specific questions and also to build knowledge libraries.

Table 1: Overview of 2D gel image analysis software

Software	Company	Website
Delta 2D	Decodon	www.decodon.com
ImagepIQ	Proteome Systems Ltd.	www.proteomesystems.com
Kepler	Large Scale Biology	www.lsbc.com
Melanie	Geneva Bioinformatics SA	www.genebio.com
Phoretix 2D /Progenesis	Nonlinear Dynamics	www.nonlinear.com
Proteomweaver	Definiens	www.definiens.com

1.4.2 MS data interpretation

There are exponentially growing protein databases and the improvements in accuracy and resolution of mass spectrometry (MS). The technologies e.g. mass spectrometry and related interpretation software can be combined for identification and characterization of proteins in proteomics[60]. During a proteomic experiment, the proteins are first separated on gel by 2DE followed by the digestion of the proteins by an enzyme that digests certain amino acids. The peptides are then analysed by mass spectrometry in order to identify the proteins[61] [49]. Examples of such mass spectrometers are given in figure 4 and 5. Sometimes, in a shotgun approach, the protein separation is not performed while

only the protein complex is digested by the enzyme and resulting peptides are separated by liquid chromatography (LC) followed by the analysis of peptides by mass spectrometry[62 63].

Figure 4: ESI-hybrid quadrupole-TOF. A type of mass spectrometer to analyse the peptide fragments in order to identify the protein (http://www.btk.fi/proteomics/instrumentation/).

Figure 5: MALDI TOF-TOF. Another type of mass spectrometer to analyse the peptide fragments in order to identify the protein (http://www.btk.fi/proteomics/instrumentation/).

With the development of proteomics, there is generation of thousands of MS spectra per day. It makes the manual methods inadequate for the analysis. So there is need to transfer the advanced capabilities of an expert human user into a sophisticated MS interpretation software which can supersede the requirement of human expert[64 65 66]. Therefore, such algorithms and software already have been, and still must be developed. In proteomics, the responsibilities for MS interpretation software can be MS data pre-processing, peptide mass fingerprinting, peptide fragmentation fingerprinting and de novo sequencing. A short description of all of them is as follow.

1.4.3 MS data pre-processing

Peak lists are the collection of signals of a spectrum which are used by most MS software tools for MS based protein identification and characterization. Signal extraction and calibration are the most common first steps in the MS data interpretation process. Due to high throughput technology, manual peak list generation has become out of date. There are some software tools available which automatically extract signals from acquired MS data[67 68 69]. However, during automatic peak extraction, there can be some hurdles e.g. noise and signal overlapping which make the expert mass spectrometrist superior to computational tools. Although, now a days in proteomics, manual peak extraction is rarely used compare to automatic peak list generation. Proteins are detected in higher charge states by using electrospray ionisation mass spectrometry (ESI-MS). Algorithms used for this purpose are based on maximum entropy, ZScore and heuristic models[70 71 72]. There are different characteristics of different spectrometers e.g. to get mass accuracy, calibration of the spectra is necessary in time of flight (TOF) mass spectrometer measurements. While, for protein identification from large sequence databases, a mass accuracy better than 50 ppm is required in case of PMF[73 74 75 76]. There can be different problems during automatic calibration for example, lacking internal calibrants, poor mass accuracy of calibrant signals and ambiguity in matching signals to calibrants. To overcome these problems, the MS data analysis tool ScoreBooster is included in the proteome

bioinformatics platform[77] with the characteristics of automatic multi-step calibration that is rather based on a collection of potential calibrants than on specific calibrants.

1.4.4 Peptide Mass Fingerprinting (PMF)

Protein identification using peptide mass fingerprinting is a widespread technique for protein analysis after 2DE. Five different research groups published the concept of peptide mass fingerprinting (PMF) in 1993[78 79 80 81 82]. In PMF, protein sequence database can be automatically searched by using a software tool. It reports the best matching proteins of the database. The constituents of simple protein mixtures can also be identified by peptide mapping, although PMF is usually applied to pure proteins[83 84]. There are many software tools for PMF protein identification. Some of them are presented in the Table 2.

Table 2: Peptide mass fingerprinting software and corresponding URL addresses

Program name	URL
Mascot	http://www.matrixscience.com
MS-Fit	http://prospector.ucsf.edu/
Mowse	http://srs.hgmp.mrc.ac.uk/cgi-bin/mowse
PepMapper	http://wolf.bms.umist.ac.uk/mapper/
Peptident	http://us.expasy.org/tools/peptident.html
PeptideSearch	http://www.mann.embl-idelberg.de/GroupPages/PageLink/peptidesearchpage.html

In the algorithm, proteins of a protein sequence database are digested *in silico* according to the specificity of the chosen enzyme. Then a theoretical MS spectrum is constructed and the calculated masses are then compared to the measured PMF data. A score for each matching database entry is calculated according to the comparison which qualifies the search result. The top ranked proteins are probably the correct identifications. To restrict the research and to make the searches more specific, PMF software provides various parameters. For example species, molecular weight and pI

of a protein can be useful to make the search more specific. Some algorithms simply rank possible protein candidates according to the number of matching masses[85 81]. MOWSE, was developed from the observed distribution frequency of peptides in the source database[79]. It is a more sophisticated scoring algorithm for PMF data. The MOWSE score is implemented in the algorithms MS-Fit and Mascot[86 87].

ProFound is another algorithm, which uses Bayesian theory to rank the protein sequences in the database according to their probability of occurrence[83]. By using simulations, the quality of results of PMF algorithms can be stated[88 89]. Direct calculation of the probability that the measured masses correspond to the sequence in the database can be another option[83 87]. Another method includes simultaneous activating of several search engines and combining their results to a single scoring scheme called "meta-score"[77]. The search engines may report different proteins and apply different ranking due to different scoring approaches. Combination of the results by a meta-score can reduce false positives. Because it evaluates the protein, either reported by another search engine or not.

1.4.5 Peptide Fragmentation Fingerprinting (PFF)

In PFF, software tools perform theoretical fragmentation of peptides and generate theoretical spectra by in silico digestion of database proteins. The approach of PFF is similar to peptide mass fingerprinting (PMF) and therefore it is referred as peptide fragmentation fingerprinting (PFF). Both strategies PMF and PFF are complementary and can be applied subsequently to obtain extensive data for unambiguous protein identification. The spectra of peptide fragmentation can be correlated to protein sequences in a database[90 91]. Peptide fragmentation provides primary structure information on a single peptide. A sequence tag of about eight amino acid lengths is required to enable unambiguous protein identification from a sequence database. By the use of peptide fragmentation, identification of the constituents of protein mixtures or posttranslational modifications can be performed. Different software tools for the identification of peptides and proteins by PFF are given in the Table 3.

SEQUEST algorithm is one of the most frequently used algorithms now a days which identifies peptides from their fragmentation pattern. It was published as early as 1994[91]. Sequest Browser provides summarization of LC-ESI MS/MS runs in addition to various other tools. Other summarizing tools for Sequest are DTASelect[92], Protein Prophet[93], Qscore[65], Final Sore and Support Vector Machine Classification[94] [38]. The Mascot algorithm employs the same probability based scoring approach for both, PMF and PFF. It identifies proteins by using peptide fragmentation data[87]. PepFrag and MSSeq are similar programs for the elucidation of peptide fragmentation data using a less sophisticated scoring approach[95 86]. SONAR includes a probability based scoring approach[96].

Table 3: Peptide fragmentation fingerprinting software and corresponding URL addresses

Program name	URL
MS-Seq	http://prospector.ucsf.edu/
MS-Tag	http://prospector.ucsf.edu/
PepFrag	http://prowl.rockefeller.edu/prowl/pepfragch.html
SEQUEST	http://fields.scripps.edu/sequest/
Sonar	http://65.219.84.5/service/prowl/sonar.html
Mascot	http://www.matrixscience.com/

1.4.6 *De novo* sequencing

If the genome from which the protein is derived is not yet sequenced or in case of posttranslational modifications, de novo sequencing from peptide fragmentation spectra can become inevitable. Because PMF or PFF data for protein identification are only appropriate if the sequence of the unknown protein is contained in a sequence database. There can be global approach or local approach in de novo sequencing. In global approach, there is generation of all theoretically possible amino acid sequences and their corresponding theoretical spectra. They are matched with the

measured spectra to find the best matching sequence[97 98]. Different techniques have been developed to reduce the space of possible solutions because the number of sequences grows exponentially with the length of the peptide. But there is a possibility to miss the correct peptide. Therefore, the local approach is commonly used in newer de novo sequencing algorithms[99 100 101]. The local approach generates a spectrum graph representing the measured spectrum. The peaks in the spectrum serve as vertices while the edges of the graph correspond to linking vertices differing by the mass of an amino acid and find the correct path through the spectrum graph. The main problems arise from incomplete fragmentation and random noise. Therefore, the interpretation often leads to uncertain results. Commonly used de novo algorithms are DeNovoX, Lutefisk[100], Rapido DeNovo, SEQPEP[97], SeqMS[99], Sherenga[101].

1.4.7 Proteome databases

There are different protein sequence databases for MS data interpretation. The choice of the applied database is influenced by update frequency, redundancy, organism, linkage to protein function and support by the search algorithm. Most commonly used databases are Swiss-Prot and NCBInr, which are also used for BLAST search[102 103]. Swiss-Prot database is a manually curated high quality database which covers post-translational modifications, domain structures, citations and protein functions. The NCBI protein database is a comprehensive database with the collection of several databases and is frequently updated. During the interpretation of mass spectrometry data with PMF or PFF algorithm, the user has to decide about the specificity of the search. Either it should be limited to the sequences of a distinct organism or open to taxonomy of higher level or even to search all known protein sequences. Such limitations e.g. taxonomy limitation, decrease CPU time and increase sensitivity of the search. Additionally, it may decrease identification rate because of incomplete sequence databases, limitations in gene prediction, incomplete support of taxonomy classification or the search algorithms and sample cross contamination. So high quality multiple PFF data sets of one given protein should be searched without taxonomy restriction. The efforts

have been made for high quality protein database. The Uniprot database (www.uniprot.org) combines Swiss-Prot, TrEMBL and PIR to provide a new stable and comprehensive resource for information on proteins, their sequences and their functions. The International Protein Index (www.ebi.ac.uk/IPI/IPIhelp.html) combines Swiss-Prot, TrEMBL, Ensembl and RefSeq. Human protein reference database (www.hprd.org) concentrate on all human proteins, their protein interaction, their post-translational modification and ontologies based on manual curation of publications. Web tools e.g. (www.ebi.ac.uk/proteome/) provide the possibilities to view a proteome in an in silico 2DE pattern style. Although the proteomics databases lack information for protein expression and protein quantification but they serve as valuable information sources for comparing theoretical proteomes and as sources for protein identifications. Swiss 2D PAGE database is a web based database (www.expasy.org/ch2d/) which is linked to the SwissProt protein database. The Swiss 2D PAGE current version (Release 18.6, September 2006 and updates up to 31-January-2008) contains 1265 entries in 36 reference maps from human, mouse, *Arabidopsis thaliana, Dictyostelium discoideum, Escherichia coli, Saccharomyces cerevisiae, and Staphylococcus aureus* (N315).

2 Celiac disease

2.1 Introduction

Celiac disease (CD) is a disorder of small intestine in genetically susceptible individuals due to the sensitivity of gluten containing cereals (wheat, rye, and barley). Thus, the presence of gluten in food products leads to massive immune reactions and damages the small intestine, affecting the absorption of nutrients.Willem Karel Dicke (late 1940s) was the first who observed that the ingestion of certain grains was harmful in the persons with the CD[104 105]. In addition to the gastrointestinal symptoms, the disease expresses itself with extraintestinal manifestations, such as osteoporosis[106], dermatitis herpetiformis[107], neurological problems[108], liver disorders[109], arthritis[110] and obstetric problems[111]. Traditionally CD has been considered as a fairly uncommon gastrointestinal disorder affecting mainly children but according to current knowledge the disease has evolved to a common systemic condition affecting individuals of all age groups.

CD is associated with major histocompatibility (MHC) class II genes, the alleles encoding the human leukocyte antigen molecules (HLA)-DQ2 and HLA-DQ8 and almost all patients with CD carry these HLA alleles[112]. However, DQ2 and DQ8 are also common (in 30-40%) among healthy individuals, the majority of whom never develop the disorder during their life time. The diagnosis of the disease is done by bowel biopsy or the detection of the certain antibodies in the blood e.g. endomysial antibodies, gliadin antibodies, reticulin antibodies. The only treatment is the gluten free diet although there can be different other methods to diagnose and treat this disease.

2.2 Signs and symptoms

CD is just not specifically related to intestine but it also affects the other body systems which leads to other disorders e.g. osteoporosis, malignancy, and infertility. So this is a multisystem disorder which also associated with other autoimmune disorders[113], such as dermatitis herpetiforms, type 1 diabetes, or autoimmune thyroiditis, and with neurological and genetic disorders[114 115 116]. The major symptoms are related to the digestive system although it varies from person to person. Some

common symptoms include diarrhea, vomiting, weight loss, abdominal pain, irritability, fatigue, arthritis, anxiety, seizures and infertility. The diarrhea becomes pale and malodorous with abdominal pain and cramping. CD can lead to an increased risk of both adenocarcinoma and lymphoma of the small bowel.

Some people may not have symptoms of the disease but there may be the complications of the disease e.g. malnutrition, anaemia, liver disease, osteoporosis, intestine cancer and miscarriage. The variation of the symptoms also depends on the intake of gluten. If a person is not taking gluten or less gluten in feed, there will be less symptoms of the disease. Age is another factor in symptoms variation. Due to autoimmune disorder, immune system may attack the healthy tissues of the body which leads to other diseases for example type 1 diabetes, autoimmune thyroid disease, autoimmune liver disease, rheumatoid arthritis, Addison's disease (hormonal glands are destroyed), Sjögren's syndrome (tear and saliva producing glands are destroyed). Moreover ulcerative jejunitis and structuring can also be there if this disease remains untreated[117]. Due to chronic malabsorption of the nutrients, there can be weight loss, fatigue and lack of energy. Folic acid and vitamin B_{12} deficiencies can lead to megaloblastic anaemia. Similarly calcium and vitamin D deficiencies can cause osteoporosis. CD is also associated with bacterial overgrowth of the small intestine which can worsen the malabsorption[118].

2.3 Epidemiology

In many Western countries the disease affects approximately 1% of the populations but it has recently been shown that the true prevalence of CD is increasing over time[119].

In general, the disease affects around 1.5-3% of the general population in both developed and developing countries while the prevalence is increasing in United States and Europe[120 121 122]. CD abundance is more in Caucasians especially in the Western Europe where it affects 1 per 130–300 individuals[123 124 125]. The statistics of the prevalence of this disease in USA is as follow.

Table 4: Statistics showing the prevalence of celiac disease among different categories in USA

Category	Prevalence (%)
Not-at-risk groups	0.75
Symptomatic patients	1.8
Second-degree relatives of a patient	2.6
First-degree relatives	4.5

This profile is similar to the prevalence in Europe[126]. The disease is one of the most common immunemediated disorders with the considerable prevalence of undetected cases[104]. It is more common in females than in males at a ratio of 3 to 1[127]. The studies in Europe and United States, suggest that its prevalence is much higher than previous estimates, raising the concern that the disease is widely under diagnosed. The increase may be due to the reason of better diagnostic approaches. However, the cases of the disease are increasing so a better understanding is necessary in order to properly diagnose and treat the disease.

2.4 Pathophysiology

Gluten-containing cereal prolamins (gliadin in wheat, secalin in rye and hordein in barley) have a high amount of repetitive glutamine- and proline-rich sequences making them highly resistant to proteolytic degradation by human gastric, pancreatic and intestinal brush-border enzymes, even in healthy individuals[128 129].

The gliadin peptidases provoke two types of immune reactions i.e. innate response and the adaptive response. Gliadin-peptides, the so called immunogenic peptides (peptides within the α-gliadin 33-mer peptide 56-89) activate the adaptive immune response. First, these peptides are post-translationally modified by a ubiquitously expressed multifunctional enzyme transglutaminase (TG) 2, which catalyzes the deamidation of distinct glutamine amino acids to glutamic acid residues[130 131]. Such deamidation greatly enhances their binding to HLA-DQ2 and thereby also potentiates

celiac patient T cell stimulation[130] [132]. As a result pro-inflammatory cytokines are secreted in parallel to the small-bowel mucosal tissue remodelling and damage characterized by villous atrophy, crypt hyperplasia and inflammation. Pathogenesis of CD is shown in figure 6. In CD patients the gluten-induced small-bowel mucosal deterioration occurs gradually from normal morphology (shown in the left) to villous atrophy with crypt hyperplasia and inflammation (shown in the right).

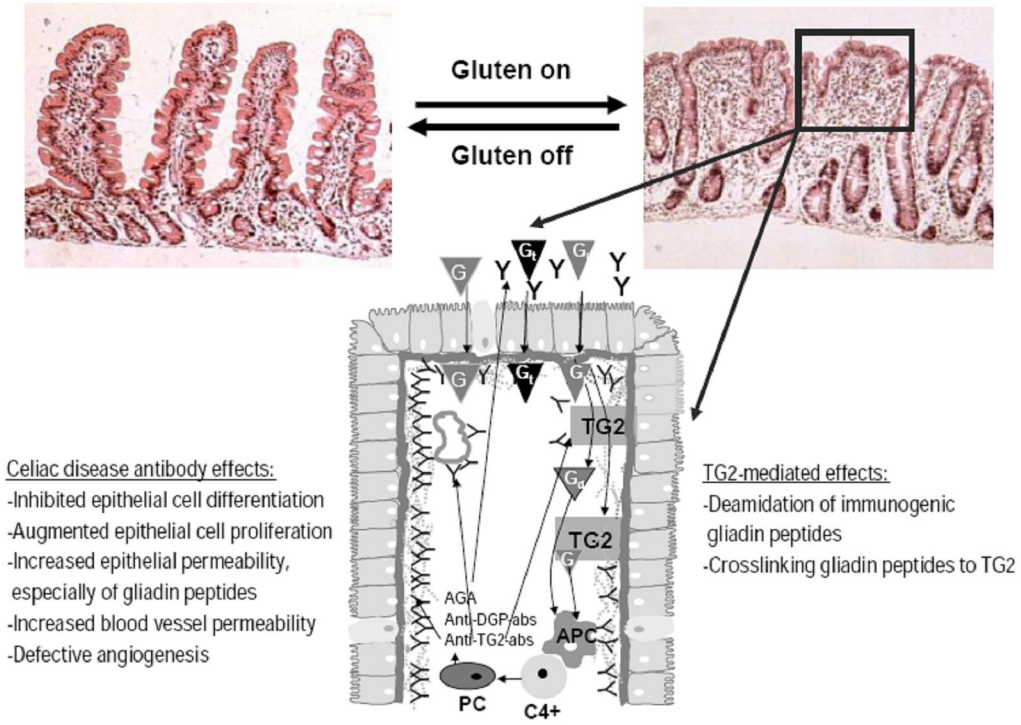

Figure 6: Pathogenesis of celiac disease (CD): Normal intestinal morphology (shown in the left) to villous atrophy with crypt hyperplasia and inflammation (shown in the right). The schematic presentation (below) shows what happens in the small-bowel mucosa at cellular level in untreated celiac patients. (Modified from Caja et al 2011)

Toxic gliadin peptides (G_t) evoke an innate immunity response eventually leading to epithelial cell death and increased epithelial permeability. This enables the immunogenic gliadin peptides (G_i) to enter the lamina propria where transglutaminase 2 (TG2) either deamidates or crosslinks them to itself. Deamidated gliadin peptides (G_d) or gliadin-TG2 complexes are taken up by antigen

presenting cells (APC). This is followed by DQ2/DQ8-dependent activation of CD4+ T cells and subsequent secretion of antibodies by mucosal plasma cells (PC). TG2-autoantibodies are found to be deposited in the small-intestinal mucosa below the epithelial basement membrane and around capillaries but antibodies are also present in the lumen (indicated by Y in the figure 6).

Along the process, B cells start to secrete antibodies against the trigger, gliadin, but also against various self-antigens[133]. The pathology change of the disease can be done by Marsh classification as follow[134].

Table 5: Marsh classification. The study of pathological changes of the disease at various stages.

Classification	Symptoms
Marsh stage 0	Normal mucosa
Marsh stage 1	Increased number of intra-epithelial lymphocytes, usually exceeding 20 per 100 enterocytes
Marsh stage 2	Proliferation of the crypts of Lieberkuhn
Marsh stage 3	Partial or complete villous atrophy
Marsh stage 4	Hypoplasia of the small bowel architecture

2.5 Diagnosis

Most of the time, CD is under diagnosed or misdiagnosed hence a good test is necessary to diagnose this disease correctly. This is important to know that a celiac patient should eat the gluten containing diet before the diagnosis otherwise the diagnostic tests will be negative. CD can be confused with some other diseases e.g. irritable bowel syndrome, iron-deficiency anaemia caused by menstrual blood loss, inflammatory bowel disease, diverticulitis, intestinal infections, and chronic fatigue syndrome. There can be different options to diagnose this disease.

In the diagnosis, celiac specific antibodies at the time of diagnosis and their disappearance after taking the gluten free diet also supports the diagnosis. For the confirmation, a biopsy is mandatory in asymptomatic cases. The small intestine mucosa damages over years or even decades. In the initial cases, the mucosa turns from normal morphology to villous atrophy and crypt hyperplasia[134]. It has also been suggested that in the cases with high celiac antibodies in serological screening, the small bowel biopsy might not be needed[135]. In the early stage cases, there might be mild mucosal changes which are not sufficient to fulfil the diagnosis criteria.

2.5.1 Serum diagnosis

The most popular method and first line investigation is blood testing in which the level of certain antibodies is determined e.g. anti-tissue transglutaminase antibodies (tTGA) or anti-endomysium antibodies (EMA). For the diagnosis of CD, the most sensitive antibody test is immunoglobulin A (IgA) class, and antireticulin and anti gliadin tests were introduced first in early1970s[136]. After a decade, highly valid IgA class endomysial antibodies (EMA) were detected by indirect immunofluorescence method by using the monkey oesophagus as antigen[137]. So a new era of serological testing of CD opened up. The testing of tTG should be done first because it is an easier test to perform moreover total serum IgA level is also checked in parallel. IgA deficiency in celiac patients may be unable to produce the required antibodies on which these tests depend which can lead to false negative results. In those patients, IgG antibodies against transglutaminase (IgG-tTG) may be used as diagnosis. However the tests for EMA and tTG antibodies are considered as the best source to screen the CD as they have also clear correlation with carriage of HLA DQ2 and DQ8[138]. Now a days, the rapid tests for tTG antibodies are developed. Only the sample of fingertip blood is required which makes the testing more simple and quick[139].

2.5.2 Small intestinal biopsy

Intestinal biopsy is another routinely test for identification of CD in which the sample is taken through endoscope directly from the small intestine. The characteristic features of CD e.g. villous atrophy and crypt hyperplasia were described first in an operative specimen of the jejunum and by peroral biopsies[140]. The upper endoscopy and biopsy of the duodenum is performed i.e. distal to the duodenal bulb. Multiple samples from different areas of intestine should be taken because there can be some healthy tissues from where you are taking the sample which can lead to false negative result so at least four samples should be taken[141]. There are three main types of mucosal changes i.e. infiltrative (type 1), hyperplastic (type 2) and destructive (type 3)[134]. The most serious change is type 3 destructive in which the villi are rudimentary or even absent while in hyperplastic (type 2), there is an increase number of IELs (intraepithelial lymphocytes) with enlarged crypts. The least dangerous situation is infiltrative (type 1), in which only the epithelium is infiltrated by intraepithelial lymphocytes i.e. more than 40 per 100 enterocytes. In destructive (type 3) lesions were modified later into three categories i.e. 3a, 3b and 3c on the basis of villous atrophy. So 3a is mild villous atrophy, 3b is marked villous atrophy while 3c is total villous atrophy. During the diagnosis, other morphological parameters e.g. villous height crypt depth ratio and number of IELs can also be considered[142]. The sampling of the small bowel mucosa should be performed in all patients because the macroscopic changes e.g. mosaic appearance of mucosa, absent or reduced duodenal folds or scalloping of mucosal folds through an endoscopy are not specific for CD[141]. A person with the symptoms of CD should go for duodenal biopsy independent of serologic test results. A duodenal biopsy is mainly recommended for the person if he is positive for celiac autoantibodies[143 144].

2.6 Treatment

Currently, the only treatment is the use of gluten free diet. So the food should not contain wheat, rye, and barley. The food items e.g. potato, rice, soy, amaranth, quinoa, buckwheat, or bean flour

instead of wheat flour, meat, fish, fruits, and vegetables can be used. Gluten is also used in some medications. So the patient should take care in using such medicines. The recovery after using the gluten free diet is very quick moreover in most of the patients, the symptoms of the disease disappear and the effected part of the intestine heals. To stay healthy, the gluten free diet should be taken throughout the life.

In Table 6, the list of allowed foods and the foods to avoid for celiac patient is given. All gluten containing food should be avoided in order to stay healthy.

Table 6: The table showing the list of allowed foods and the foods to avoid for celiac patient

Allowed foods	**Foods to avoid**
Rice	Wheat including einkorn
Corn flax	Emmer
Seeds	Spelt
Sorghum	Kamut
Cassava	Wheat starch
Buck wheat	Wheat bran
Arrow root	Wheat germ
Wild rice	Cracked wheat
Indian rice grass	Hydrolyzed wheat protein
Potatoes	Rye

3 Aims of the research

The aims of the present research describe as follow:

1- To setup a protocol suitable to prepare serum samples derived from celiac disease patients for proteomic analysis.

2- To differentiate between two methods of serum protein depletion in order to have a better and accurate proteomic analysis.

3- To study the circulating proteins in the blood of CD patients during active disease in order to come across the differences between three groups i.e. control, CD patients and the same patients after one year of gluten free diet.

4 Materials and methods

4.1 Materials

4.1.1 Samples and patients

Serum samples were collected from three groups.

1- Celiac disease patients at the time of diagnosis by duodenal biopsy together with the presence of anti-tissue transglutaminase 2 antibodies, as well as Immunoglobulin A (IgA), Immunoglobulin G (IgG) and antiendomysial antibodies.
2- The same patients after one year of gluten free diet with the normal biopsy as well as negative for antibodies.
3- Healthy donors HLA DQ2 or DQ8 positive without celiac disease hence negative for antibodies and normal biopsy.

There were 10 samples in each group and altogether 30 samples were collected. Ethical permission was given to accomplish the present studies. In addition, all patients gave their written inform consent and all records are stored according to the regulations. The samples were stored at -20ºC until analysis.

4.1.2 Instruments and reagents

Agilent Human 14 multiple affinity removal system (Agilent technologies, Wilmington, DE, UK) was used to remove the 14 most abundant proteins in serum. Albumin/IgG removal kit (EMD Biosciences, La Jolla, CA) was used to remove albumin and IgG. Tris base, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), Dimethylformamide (DMF), thiourea, Dithiothrietol (DTT), Sodium Dodecyl Sulfate (SDS), Ammonium pursulfate (APS), Tetramethylethylenediamine (TEMED), glycine, urea, glycerol, agarose and bromophenol blue were purchased from Sigma Aldrich (St. Louis, MO, USA). Acrylamide was purchased from BioRad (Hercules, CA, USA). Cydye dyes and destreak rehydration solution were purchased from GE healthcare (Little Chalfont, Buckinghamshire, UK). Proteoextract was purchased from

calbiochem (Merck KGaA, Darmstadt, Germany). IPGphor and DALTsix machines (GE Healthcare, Little Chalfont, Buckinghamshire, UK) were used for the first and second dimension electrophoresis respectively. Immobiline drystrip 24cm and 3-11 linear pH (GE Healthcare) was used for Isoelectric focusing (IEF) and 2D electrophoresis. The results were analyzed by Decyder 2D software, version 6.5. GE Healthcare. Typhoon 9400 (GE Healthcare) scanner was used to scan fluorescent gels.

4.2 Methods

4.2.1 Sample preparation

4.2.2 Removal of albumin and IgG (method A)

In method A, two major proteins i.e. albumin and IgG were removed from serum. For this purpose, 25 µl samples were used for the analysis. For dilution, 350 µl binding buffer was added in the sample. Then albumin and IgG were removed through the removal column. First the columns were washed by 850 µl of binding buffer. Then the diluted samples were put in the columns followed by 600 µl binding buffer two times. So the collected samples were free of albumin and IgG. Then the samples were centrifuged by using amicon centrifuge tubes (Amicon) at 5000 rpm for 20 minutes at room temperature. Almost 80% of albumin and IgG are removed by this method. After the removal of major proteins, a pool of all samples from one group was prepared. So ten samples from each group were combined and at the end, three vials of three groups were prepared.

4.2.3 Removal of 14 major proteins (method B)

In method B, 14 major serum proteins were removed. For this purpose, 8 µl serum from all the groups was diluted with 192 µl of buffer A to make the final volume of 200 µl. The diluted sample was filtered through 0.22 µm provided spin filter at 3500 rpm/3 min RT. The buffer from cartridge was removed through a transfer pipette. Then the 200 µl of diluted sample was added in the cartridge and centrifuged for one minute at 100 g (1000 rpm) after placing in a 2 µl collection tube.

The flow through fraction was collected in collection tube. Cartridge was placed at room temperature for 5 minutes. 400 µl of buffer A was added at the top of resin bed and centrifuged for 2.5 minutes at 100 g (1000 rpm). The flow through fraction was collected in the same collection tube. The 5 ml plastic syringe was filled to 2.5 ml buffer B and high abundant proteins were eluted in to a new collection tube after passing buffer B slowly through cartridge. The cartridge was washed after passing 5 ml buffer A through cartridge. The cartridge was placed at 4ºC after leaving the small amount of buffer on the top of the frit. So the 14 major proteins include albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, α-2-macroglobulin, α-1-acid glycoprotein, IgM, apolipoprotein AI, apolipoprotein AII, complement C3, and transthyretin were removed by using the kit. More than 90% of these proteins were removed by this method as stated by the user guide manual. After the removal of major proteins, a pool of all samples from one group was prepared. So ten samples from each group were combined and at the end, three vials of three groups were prepared.

4.2.4 Protein precipitation, concentration and resuspension

After the removal of major proteins by using method A or method B, it is mandatory to proceed with the removal of salts as well as further concentration using the comercial kit i.e. protein precepitation kit (calbiochem). To this end, samples were mixed with four volumes equivalent of cold precipitation agent and mixed briefly. Then they are incubated at -20ºC for 60 minutes. The sample tubes were centrifuged at 5000 rpm for 20 minutes at room temperature. The supernatant was discarded and 1ml cold wash solution was added and mixed until detach of pellet. So the pellet was resuspended and transferred to a clean micro centrifuge tube of 1.5 µl. The tubes were centrifuged for 2 minutes at 14000 rpm at room temperature. Then again the supernatant was aspirated and 500 µl cold wash solution was mixed in the pellet. Tubes were centrifuged for 2 minutes at 14000 rpm at room temperature. The wash solution was aspirated completely and pellet was dried for 1 hour by leaving the tubes open on lab bench. Pellet was resuspended by using the

lysis solution (7 M urea, 2 M thiourea, 30 mM tris, 4% CHAPS). The lysis solution (90 µl) was added in the pellet and placed it at room temperature for one hour. The tubes were vortexed after every 15 minutes. Then tubes were centrifuged at 14000 rpm for 5 minutes to remove any impurities. pH of the solution was determined which should be 8.5 for protein labelling (CyDye). Then the protein concentration was determined in order to get 50 µg of protein per group for CyDye labelling.

4.2.5 Sample labelling

Three tubes were prepared for labelling. Each tube consisted of one of three samples. Cy2 dye was used to label control sample, Cy3 dye was used to label untreated sample while Cy5 dye was used to label gluten free sample. Working dye solution (1 µl) was added to the micro centrifuge tube contained 50 µg protein sample. Dye was mixed in the sample and centrifuges briefly and left on ice for 30 minutes in the dark. At the end, 1 µl of 10 mM lysine was added to stop the labelling reaction then mixed and spin briefly and left on ice for 10 minutes in the dark. Three samples were combined in a single tube.

4.2.6 Isoelectic focusing (IEF)

Destreak rehydration solution was equilibrated for 30 minutes at room temperature. IPG buffer (60 µl) was added in 3 ml stock solution. DTT (20 mg) and IPG buffer (20 µl) were added in 1 ml 2X sample buffer stock (8 M urea, 2% CHAPS). Equal amount of 2X sample buffer was added in the sample and destreak rehydration solution was added to make the final volume 450 µl. This 450 µl volume was loaded in the strip tray and then strip was placed over it. Mineral oil was put after removing air bubbles. The IEF was performed according to the following protocol.

In the first step, rehydration was performed at 50 V for 14 hours with step and hold. The second step was performed in three steps as follow.

1^{st} step = 100 V for 0.5 hour gradient

2nd step = 4000 V for 8 hour gradient

3rd step = 8000 V for 6 hours step and hold

After IEF, 2nd dimension was performed immediately in DALTsix.

4.2.7 Second dimension

Second dimension was performed in a 12% polyacrylamide gel using DALTsix (GE Healthcare). SDS running buffer (1X) was used to run the second dimension after the dilution of 10X SDS running buffer stock (250 mM Tris, 1.92 M glycine, 1% SDS, pH 8.3). IPG strips were then equilibrated in SDS equilibration buffer (75 mM Tris, 6 M urea, 29.3% glycerol, 2% SDS, 0.002% bromophenol blue, pH 8.8) supplemented with 1% DTT to maintain the fully reduced state of proteins, followed by 2.5% iodoacetamide to prevent reoxidation of thiol groups during electrophoresis. Agarose sealing solution (0.5% agarose, 10% 10X SDS running buffer, 0.002% bromophenol blue) was put after loading the strip in the gel. Lower and upper chambers were filled with 1X SDS running buffer and the empty slots were filled with cassettes. The second dimension was run at 1 Watt per gel for one hour and 2 Watts per gel for 16 hours.

4.2.8 Gel scanning and analysis

The scanning of the gels was performed through typhoon scanner (GE Healthcare) at 450 PMT and 100 micron pixels. The gels were analyzed by Decyder software (GE Healthcare) to get the statistically significant spots.

4.2.9 Silver staining

To stain the gel, a modified silver staining method was used[145]. Briefly, the process starts under gentle shaking of staining tray with 250 ml fixing solution. The fixation was performed two times for 15 minutes each in fixing solution (40% methanol, 10% acetic acid). The sensitization was then performed in sensitizing solution (30% methanol, 4% sodium thiosulphate, 6.8% sodium acetate) for 30 minutes. After washing three times in distilled water for 5 minutes each, silver staining was

performed in silver solution (10% silver nitrate solution) for 20 minutes. Developing was performed in developing solution (2.5% sodium carbonate, 0.04% formaldehyde) for 3 minutes after two washings in distilled water. After developing the protein spots, the gel was shifted to stop solution (1.46% EDTA-Na2.2$H_2$0) for 10 minutes. The gel was stored in distilled water after three washings until the further analysis.

4.2.10 Spot excision and protein identification

After the silver staining, the gel from method B was used for further analysis as method B produced more number of spots compare to method A. All those spots, which were differentiated after a threshold of 1.5 in Decyder software, were considered statistically significant. These statistically significant spots were excised and the proteins were identified by MALDI TOF TOF. Briefly the spots were destained and then digestion was performed by trypsin. After incubating the samples overnight, the resulting peptides were analyzed by MALDI TOF TOF to get the mass to charge ratio.

4.2.11 Data mining for identified proteins

The mascot database was used to identify the specific proteins while NCBI database was used to find the details of the identified proteins. Briefly, NCBI pubmed and OMIM databases were used to find the association of the protein in CD and different other autoimmune diseases while NCBI gene database was used to find the location of the protein in the genome. The results were also checked by previous articles in pubmed.

.

5 Results

Serum derived from celiac disease, same donors after one year of gluten free diet and healthy donors were prepared for a proteomic analysis. Two different methods were used in order to remove the more abundant proteins. In method A, the depletion of albumin and IgG was performed and a total number of 348 spots ere found after the analysis. While in method B, the depletion of 14 major proteins (albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, α-2-macroglobulin, α-1-acid glycoprotein, IgM, apolipoprotein AI, apolipoprotein AII, complement C3, and transthyretin) was performed. A total number of 689 protein spots were detected after the analysis. Hence, gel from method B was further processed.

5.1 Difference between method A and method B

The gels obtained in method A and method B were compared and 169 spots in gel A were correlated with gel B by Decyder software. Figure 8 presents a comparison between both methods. Extra spots detected by method B are circled.

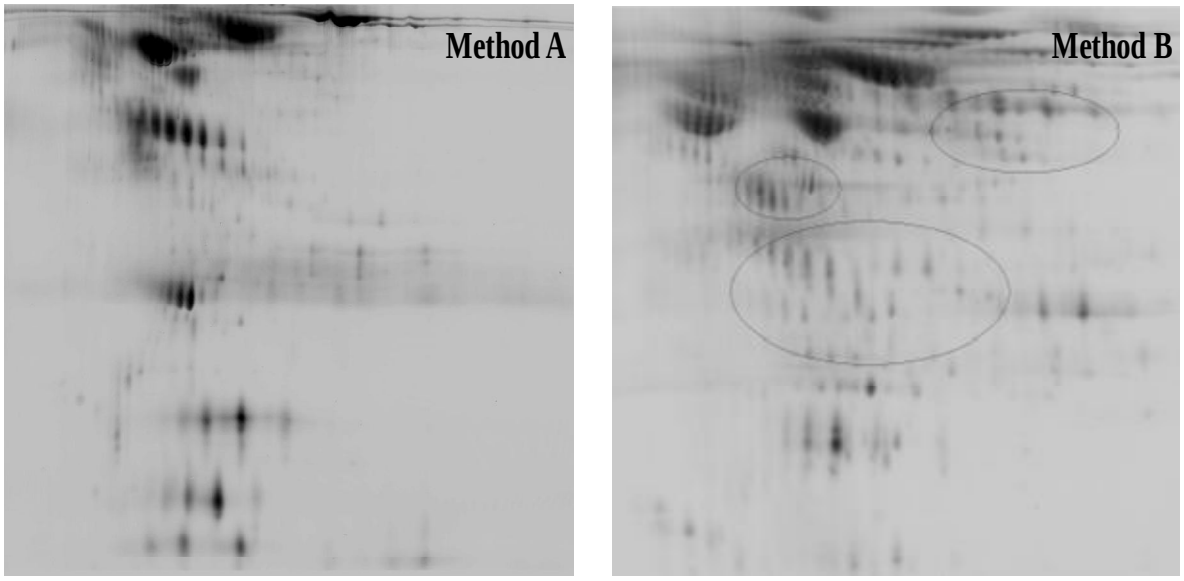

Figure 7: Extra spots are circled which are detected by Decyder software by the use of method B. Due to the removal of more major proteins, the small proteins get florescence and are detected by the software. More protein spots were detected by method B (right) compare to method A (left).

5.2 Proteins identified by MALDITOF TOF

Statistically significant spots were identified by MALDI TOF TOF. All those proteins which were different between CD during active celiac disease and GFD when were similar to control were identified. Total six proteins were identified in this group as are presented in the Table 7.

Table 7: Proteins which are different between the CD patients and the gluten free diet (GFD) groups

Spot number	Protein	Score	Mw	Expression ratio
24	Retinol-binding protein 4	194	23337	-2,26
27	β-2-glycoprotein 1	197	39584	-2,16
28	Vitronectin	129	55069	-2,30
29	Serotransferrin	53	79280	-2,10
33	Kininogen-1	275	72996	2,35
684	Serum amyloid P-component	60	25485	1.6

During mass spectrometry, the digested peptides were identified by using mascot database to identify the protein. In Mascot, the ions score for an MS/MS match is based on the calculated probability i.e. P while The protein score in the result report from an MS/MS search is derived from the ions scores. For a search that contains a small number of queries, the protein score is the sum of the highest ions score for each distinct sequence. The protein score works well for small searches, and provides a logical order to the report.

Mw is protein molecular weight which is measured in dalton (Da). It is a unit that is used for indicating mass on an atomic or molecular scale. One Da is approximately equal to the mass of one proton or one neutron. During the second step of electrophoresis, the proteins in the gel separate on the basis of their molecular weight. Hence, during the separation of proteins on the basis of their

Mw, the proteins with less molecular weight travel longer on the gel than the proteins with high molecular weight.

Expression ratio indicates the comparison of the protein between different states. For example, in Table 7, the proteins are compared between active celiac disease condition, GFD condition and control. The negative values in the expression ratio indicate that the protein is down-regulated in the active celiac disease condition compare to the GFD condition and control. While positive values indicate their up-regulation. Hence, kininogen-1 and serum amyloid P-component proteins were present in high abundance (up-regulated) while all other proteins were present in low abundance (down-regulated) in the active celiac disease condition compare to the GFD condition and control.

6 Discussion

Proteomics is a tool using2-D gel electrophoresis to study the proteins present in biological fluid e.g. serum. We aimed to setup a protocol suitable for the proteomic analysis of serum samples derived from celiac disease (CD) patients and healthy donors. Moreover, we aimed to differentiate between two methods of major protein removal in order to get better proteomic analysis and results. Therefore, high abundant proteins were depleted by two different methods A and B. Method A removed 80% of two major proteins from the samples i.e. albumin and IgG while method B removed more than 90% of 14 major proteins i.e. albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, α-2-macroglobulin, α-1-acid glycoprotein, IgM, apolipoprotein AI, apolipoprotein AII, complement C3, and transthyretin. The remaining proteins were separated by IEF on the basis of their PI and DALTsix on the basis of their molecular weight. The resulting gel was scanned by typhoon scanner and stained by silver staining. The spots detected by using method A were 348 while by method B, 689 spots were detected. So method B produced better results compare to method A, due to the removal of more proteins.

During proteomic analysis, it is necessary to remove the high abundant proteins. The labelling reaction occurs more intensely with the high abundant proteins compare to the low abundant proteins. So due to the presence of high abundant proteins, their florescence will be high and the low abundant proteins would not show their effect on polyacrylamide gel. In celiac disease patients, the abundance of IgG and IgA is very high[146]. Hence, these major proteins should also be removed in order to get better analysis. As the method B removed more proteins than method A, so the numbers of detected spots were also more. The comparison of both methods shows that method B is more efficient than method A. Both methods A and B removed IgG antibodies. Moreover, method B removed more than 90% of IgG compare to method A which removed 80% of this protein.

On the other hand, in the present study, 10 serum samples were used in each group. A pool of all the samples in each group was analysed. We used 10 samples according to the minimum number in

order to get statistical significance between proteins. Thus by using a power analysis, as previously described[147], a pool of the samples was made. Celiac disease is more frequent in females than in males[127], hence the pool made and analysed in both sex at the same ratio while all the patients were having the characteristic features of the disease. Therefore, the values as well as other clinical symptoms similarity among the individuals reduced the variability. However, further studies should be performed in bigger number of samples in order to find differential spots and validation of the proteins.

One interesting fact is that, only 169 protein spots from method A were correlated to the protein spots obtained by method B. This may be due to the sensitivity of the system due to the presence of major proteins as well as in gel variation and biological variations during the experiment. As major proteins disturbed the detection of others and probably these proteins are responsible of the variation in the gel, and the differences might be related to the behavior of the fluorescence in the presence of the high abundant proteins[148].

The protein spots detected by Decyder software can be placed in different categories. In one category of identified proteins, all those proteins can be placed which show different behaviour between the CD group and GFD. Such proteins change their abundance between two conditions i.e. the active disease condition and after the use of gluten free diet. Hence, these proteins can be used to study the pathogenesis of the disease as they change their behaviour during changing of the disease from active phase to recovery phase. The proteins detected by MALDI TOF TOF for this group has shown in Table 7. Total 6 spots were analysed and identified from this group. The identified proteins were retinol-binding protein 4, β-2-glycoprotein 1, vitronectin, serotransferrin, kininogen-1 and serum amyloid P-component. Among these proteins, kininogen-1 and serum amyloid P-component proteins were present in high abundance (up regulated) while all other proteins were present in low abundance (down regulated) in the active celiac disease condition compare to the GFD condition and control.

The location of retinol-binding protein 4 (RBP4) is chromosome 10 at 10q23-q24. The level of serum RBP4, increased in insulin-resistant states. Experiments in mice suggested that elevated RBP4 levels cause insulin resistance[149]. Serum RBP4 level also correlated with the magnitude of insulin resistance in human with obesity and type 2 diabetes while in nonobese, nondiabetic subjects, with a strong family history of type 2 diabetes[150]. The protein level was low during the active celiac disease compare to the GFD condition in the present studies. The protein was studied in context of CD and changes in retinol binding protein values did not correlate with mucosal improvement[151]. Other diseases associated with RBP4 are type 2 diabetes mellitus[152], insulin resistance and the metabolic syndrome[153], subclinical hypothyroidism[154] and obesity and insulin resistance[155].

β-2-glycoprotein (β2GPI) 1 is located on chromosome 17 at 17q23-qter. The protein also known as apolipoprotein H. It binds to and neutralizes negatively charged phospholipid macromolecules, thereby diminishing inappropriate activation of the intrinsic blood coagulation cascade. β-2 GPI has been implicated in a variety of physiologic pathways, including blood coagulation, hemostasis, and the production of antiphospholipid antibodies[156]. The protein level was low during the active celiac disease compare to the GFD condition in the present studies. In context of CD, Antibodies against β-2-glycoprotein-I epitopes (GRTCPKPDDLP) were more prevalent in EMA-positive patients[157]. In another study, there was no statistically significant difference for the prevalence of IgG and IgA β2GPI between patients and controls[158]. The protein is Protective against lupus nephritis and renal damage[159].

Vitronectin is located on chromosome 17 at 17q11. Other names of the protein are serum spreading factor or complement S-protein. The somatomedin B domain of vitronectin binds to plasminogen activator inhibitor-1 (PAI-1), and stabilizes it[160]. The protein is considered to be involved in haemostasis[161] and tumour malignancy[162]. The protein relation was not found in context of CD while in the present studies, the level was low during the active celiac disease compare to the GFD

condition. The protein is a serum marker for early breast cancer detection[163]. Moreover, vitronectin was down regulated in healthy cigarette smoker compare to the healthy non smokers[164].

Serotransferrin is coded by chromosome 3 at 3q29. Transferrin is an iron binding, monomeric glycoprotein that carries ferric iron between the sites of its absorption, storage, and utilization[165]. Transferrin receptors are recently recognized as CD71 and IgA1 receptor14[166]. IgA1 molecules are not pathognomonic for celiac disease[167]. In one studies, it was concluded that CD71-mediated transcytosis of gliadin peptides that may participate in the pathogenesis of CD in genetically predisposed individuals[168]. In the present studies, the protein level was low during the active celiac disease compare to the GFD condition.

Kininogen-1 (KNG1) is located on chromosome 3 at 3q27. The KNG1 protein is essential for blood coagulation[169]. In context of CD, no relation of this protein was found. However, the protein is a potential marker for diagnosis of colorectal cancer[170]. KNG1 gene was associated with essential hypertension and blood pressure traits in the Chinese male population[171]. Administration of an aflatoxin contaminated diet to growing barrows alters hepatic gene expression, and in particular apoptosis genes e.g. Kininogen-1[172]. In the present studies, the protein was up regulated during the active celiac disease compare to the GFD condition.

Serum amyloid P-component (SAP) protein is coded on chromosome 1 at 1q21-q23. It belongs to a superfamily of proteins known as the pentraxins. This protein superfamily is highly conserved throughout nature and is known for its calcium-dependent ligand binding and lectin properties[173]. No relation of this protein was found in context of CD. In the present studies, the protein was present in more abundance in the active celiac disease compare to the GFD group. The protein has a role in different diseases e.g. Autoimmune encephalomyelitis[174], Parkinson's disease[175] and Amyloidosis[176].

However, these proteins should be further analyzed and validated and studies should be performed to confirm the protein profile.

7 Conclusion

We setup a suitable protocol to perform the proteomic analysis by using serum. The serum samples were prepared by two different methods A and B. The difference between both methods was determined. We found more spots by using method B due to the removal of more than 90% of 14 major proteins. So we concluded that the removal of 14 major proteins (method B) is better than the removal of albumin and IgG proteins (method A) in context of celiac disease (CD).

Moreover, we found 6 proteins altered during the active disease in comparison with GFD when were similar to control. Most of them are related with vascular biology and its homeostasis.

8 References

1. Petriction EF, Zoon KC, Kohn EC (2002). Clinical proteomics : Translating benchside promise into beside reality. Nat. Rev. Drug discov. 1: 683-695.

2. Rai AJ, Chan DW. (2004). Cancer proteomics: Serum diagnostics for tumor marker discovery. Ann N Y Acad Sci: Jun;1022: 286-94.

3. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, Mills GB, Simone C, Fishman DA, Kohn EC, Liotta. (2002). Use of proteomic patterns in serum to identify ovarian cancer. Lancet 359: 572-577

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

5. Lee J, Garrett WM, Cooper B (2007). Shotgun proteomic analysis of *Arabidopsis thaliana* leaves. J Sep Sci 30, 2225–2230.

6. Westermeier R, Marouga R (2005). Protein detection methods in proteomics research. Biosci Rep 25, 19–32.

7. Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 21, 1037–1053.

8. Gygi SP, Corthals GL, Zhang Y (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proc. Natl. Acad. Sci. USA 97: 9390-9395.

9. Bae MS, Cho EJ, Choi EY, Park OK (2003). Analysis of the Arabidopsis nuclear proteome and its response to cold stress. Plant J 36, 652–663.

10. Millar AH, Sweetlove LJ, Giege P, Leaver CJ (2001). Analysis of the Arabidopsis mitochondrial proteome. Plant Physiol 127, 1711–1727.

11. Choudhary MK, Basu D, Datta A, Chakraborty N, Chakraborty S (2009). Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective. Mol Cell Proteomics 8, 1579–1598.

12. Peltier JB, Friso G, Kalume DE, Roepstorff P, Nilsson F, Adamska I, Van Wijk KJ (2000). Proteomics of the chloroplast: Systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. Plant Cell 12, 319–341.

13. Sparkman, O. David (2000). Mass spectrometry desk reference. Pittsburgh: Global View Pub. ISBN 0-9660813-2-3

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

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

16. Aebersold R, Mann M (2003). Mass spectrometry based proteomics. Nature 422, 198–207.

17. Burtis CA, Ashwood ER (2001). Tietz fundamentals of clinical chemistry. 5th Ed. W.B. Saunders Co. Philadelphia, 74-90.

18. Görg A, Weiss W, Dunn MJ (2004). Current two-dimensional electrophoresis technology for proteomics. Proteomics 4 (12), 3665–85.

19. Burbaum, Jonathan et al (2005). High resolution functional proteomics by active site peptide profiling. PNAS Mar, 4996-5001.

20. Madin, Kairat et al (2000). A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: Plants apparently contain a suicide system directed at ribosomes. Applied Biological Sciences. Jan;97(2), 559-564.

21. MacBeath, G. (2002), 'Protein microarrays and proteomics. Nat. Genet., Vol. 32 (Suppl.), pp. 526–532.

22. Figeys, D. (2002), ‘Adapting arrays and lab-ona- chip technology for proteomics’, Proteomics, 2, pp. 373–382.

23. Blagoev, B. and Pandey, A. (2001), ‘Microarrays go live —New prospects for proteomics’, Trends Biochem. Sci., 26, pp. 639–641

24. Macbeath G, Schreiber S (2000). Printing proteins as microarrays for high-throughput function determination. Science. Sep 08; 289 (5485), 1760-1764.

25. Ichihara T., Akada J.K., Kamei S., Ohshiro S., Sato D., Fujimoto M., Kuramitsu Y., Nakamura K., (2006) J Proteome ;5(9):2144-51

26. Zhu H., Klemic J.F., Chang S., Bertone P., Casamayor A., Klemic K.G., Smith D., Gerstein M., Reed M.A., Snyder M. (2000) Nat Genet.;26(3):283-9.

27. Biomarkers definitions working group (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin. Pharmacol. Ther. 69, 89-95

28. Reynolds, Mark A et al (2003). Early biomarkers of stroke. Clinical Chemistry 49, 1733-1739.

29. Manoj Kumar and Shiv K Sarin (2010). Biomarkers of diseases in medicine. Current Trends in Science. 15, 403-406.

30. Pennisi E (2000). Genomics comes of age. Science, 290, 2220-2221.

31. Blackstock WP, Weir MP (1999). Proteomics: quantitative and physical mapping of cellular proteins. Trends Biotech 17, 121-127.

32. Pandey A, Mann M (2000). Proteomics to study genes and genomes. Nature 405,837-846.

33. Akhilesh P, Matthias M (2010). Proteomics to study genes and genomes. Nature, 840-841.

34. Nuwaysir L, Stults JT (1993). ESI mass spectrometry of phosphopeptides isolated by on-line immobilized metal affinity chromatography. J. Am. Soc. Mass spectrom. 4, 662–669.

35. Betts JC, Blackstock WP, WardMA, Anderton BH (1997). Identification of hosphorylation sites on neurofilament proteins by nanoelectrospray mass spectrometry. J. Biol. Chem. 272, 12922–12927.

36. Neubauer G, Mann M (1999). Mapping of phosphorylation sites of gel-isolated proteins by nanoelectrospray tandem mass spectrometry: Potentials and limitations. Anal. Chem. 71, 235–242.

37. Zhang X et al (1998). Identification of phosphorylation sites in proteins separated by polyacrylamide gel electrophoresis. Anal. Chem. 70, 2050–2059.

38. Cortez D, Wang Y, Qin J, Elledge SJ (1999). Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. Science 286, 1162–1166.

39. Matthias Mann (1999). Quantitative proteomics? Nature Biotechnology 17, 954-955.

40. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B (2007). Quantitative mass spectrometry in proteomics: A critical review. Anal Bioanal Chem 389 (4), 1017–31.

41. Ong SE, Mann M (2005). Mass spectrometry-based proteomics turns quantitative. Nature Chemical Biology 1 (5), 252–262.

42. Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC (2000). Electrophoresis 21, 1104-1115.

43. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R (2000). Proc. Natl. Acad. Sci. USA 97, 9390–9395.

44. Tanke HJ (2007). Genomics and proteomics: The potential role of oral diagnostics. Ann N Y Acad Sci; 1098, 330-4.

45. Garcia I, Tabak LA (2008). Beyond the "omics": Translating science into improved health. J Am Dent Assoc; 139, 392-5.

46. Wright JT, Hart TC (2002). The genome projects: Implications for dental practice and education. J Dent Educ; 66, 659-71.

47. Covani U, Marconcini S, Giacomelli L, Sivozhelevov V, Barone A, Nicolini C (2008). Bioinformatic prediction of leader genes in human periodontitis. J periodontol; 79, 1974-83.

48. Klose J (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. Humangenetik 26, 231–243.

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

50. Martin Blueggel, Daniel Chamrad and Helmut E. Meyer (2004). Bioinformatics in Proteomics. Current Pharmaceutical Biotechnology, 5, 79-88.

51. Giacomelli L, Marconcini S, Derchi G, Barone A, Covani U (2007). Leader genes: A data mining approach for molecular genomics. AMIA Annu Symp Proc; 963.

52. Lehner B, Crombie C, Tischler J, Fortunato A, Fraser AG (2006). Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. Nat Genet; 38, 896-903.

53. Sivozhelezov V, Giacomelli L, Tripathi S, Nicolini C (2006). Gene expression in the cell cycle of human T lymphocytes: Predicted gene and protein networks. J Cell Biochem; 97, 1137-50.

54. Giacomelli L, Nicolini C (2006). Gene expression in human T lymphocytes: Experimental and bioinformatics analysis. J Cell Biochem; 99, 1326-33.

55. Sivozhelezov V, Braud C, Giacomelli L, et al (2008). Immunosuppressive drug-free operational immune tollerance in human kidney transplant recipient: Non-statistical gene microarray analysis, J Cell Biochem; 103, 1693-706.

56. Kuo WP (2003). Overview of bioinformatics and its application to oral genomics. Adv Dent Res; 17, 89-94.

57. Smilansky Z (2001). Automatic registration for images of two-dimensional protein gels. Electrophoresis, 22(9), 1616-26.

58. Klose J, Kobalz U (1995). Two-dimensional electrophoresis of proteins: An updated protocol and implications for a functional analysis of the genome. Electrophoresis, 16(6), 1034-1059.

59. Zischka H.; Weber G.; Weber P.J.; et al (2003) Improved proteome analysis of Saccharomyces cerevisiae mitochondria by free-flow electrophoresis. Proteomics, 3(6), 906-916

60. Yates JR (1998). Mass spectrometry and the age of the proteome. J. Mass Spectrom, 33(1), 1-19.

61. Klose J, Kobalz U (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis, 16(6), 1034-1059.

62. Courchesne PL, Jones MD, Robinson JH, Spahr CS, McCracken S, Bentley DL, Luethy R, Patterson SD (1998). Optimization of capillary chromatography ion trap-mass spectrometry for identification of gel-separated proteins. Electrophoresis, 19(6), 956-967.

63. Yates JR, McCormack AL, Schieltz D, Carmack E, Link A (1997). Direct analysis of protein mixtures by tandem mass spectrometry. J. Protein Chem., 16(5), 495-497.

64. Quadroni M, James P (1999). Proteomics and automation. Electrophoresis, 20(4-5), 664-677.

65. Moore RE, Young MK, Lee TD, (2002). Qscore: An algorithm for evaluating SEQUEST database search results. J. Am. Soc. Mass Spectrom, 13(4), 378-386.

66. Eddes JS, Kapp EA, Frecklington DF, Connolly LM, Layton MJ, Moritz RL, Simpson RJ (2002). CHOMPER: a bioinformatic tool for rapid validation of tandem mass spectrometry search results associated with high-throughput proteomic strategies. Proteomics, 2(9), 1097-1103.

67. Coombes KR, Fritsche HA, Clarke C, Chen JN, Baggerly KA, Morris JS, Xiao LC, Hung MC, Kuerer HM (2003). Quality control and peak finding for proteomics data collected from nipple aspirate fluid by surface-enhanced laser desorption and ionization. Clin. Chem., 49(10), 1615-1623.

68. Breen EJ, Hopwood FG, Williams KL, Wilkins MR (2000). Automatic poisson peak harvesting for high throughput protein identification. Electrophoresis, 21(11), 2243-2251.

69. Gras R, Muller M, Gasteiger E, Gay S, Binz PA, Bienvenut W, Hoogland C, Sanchez JC, Bairoch A, Hochstrasser DF, Appel RD (1999). Improving protein identification from peptide mass

fingerprinting through a parameterized multi-level scoring algorithm and an optimized peak detection. Electrophoresis, 20(18), 3535-3550.

70. Ferrige AG, Seddon MJ, Jarvis S (1991). Maximum entropy deconvolution in electrospray mass-spectrometry. Rapid Commun. Mass Spectrom., 5(8), 374-377.

71. Zhang Z, Marshall AG (1998). A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. J. Am. Soc. Mass Spectrom., 9(3), 225-233.

72. Zheng H, Ojha PC, McClean S, Black ND, Hughes JG, Shaw C (2003). Heuristic charge assignment for deconvolution of electrospray ionization mass spectra. Rapid Commun. Mass Spectrom., 17(5), 429-436.

73. Gras R, Muller M, Gasteiger E, Gay S, Binz PA, Bienvenut W, Hoogland C, Sanchez JC, Bairoch A, Hochstrasser DF, Appel RD (1999). Improving protein identification from peptide mass fingerprinting through a parameterized multi-level scoring algorithm and an optimized peak detection. Electrophoresis, 20(18), 3535-3550.

74. Lee K, Bae D, Lim D (2002). Evaluation of parameters in peptide mass fingerprinting for protein identification by MALDI-TOF mass spectrometry. Mol. Cells, 13(2), 175-184.

75. Mortz E, Vorm O, Mann M, Roepstorff P (1994). Identification of proteins in polyacrylamide gels by mass spectrometric peptide mapping combined with database search. Biol. Mass Spectrom., 23(5), 249-261.

76. Wool A, Smilansky Z (2002). Precalibration of matrix-assisted laser desorption/ionization-time of flight spectra for peptide mass fingerprinting. Proteomics, 2(9), 1365-1373.

77. Chamrad DC, Koerting G, Gobom J, Thiele H, Klose J, Meyer HE, Blueggel M (2003). Interpretation of mass spectrometry data for high-throughput proteomics. Anal. Bioanal. Chem., 376(7), 1014-1022.

78. Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C (1993). Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc. Natl. Acad. Sci. USA, 90(11), 5011-5015.

79. Pappin DJ, Hojrup P, Bleasby A (1993). Rapid identification of proteins by peptide-mass fingerprinting. Curr. Biol., 3, 327-332.

80. James P, Quadroni M, Carafoli E, Gonnet G (1993). Protein identification by mass profile fingerprinting. Biochem. Biophys. Res. Commun., 195(1), 58-64.

81. Mann M, Hojrup P, Roepstorff P (1993). Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol. Mass Spectrom., 22(6), 338-345.

82. Yates JR, Speicher S, Griffin PR, Hunkapiller T (1993). Peptide mass maps: a highly informative approach to protein identification. Anal. Biochem., 214(2), 397-408.

83. Zhang W, Chait BT (2000). ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. Anal. Chem., 72(11), 2482-2489.

84. Jensen ON, Podtelejnikov AV, Mann M (1997). Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. Anal. Chem., 69(23), 4741-4750.

85. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF (1999). Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol., 112, 531-552.

86. Clauser KR, Baker P, Burlingame AL (1999). Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal. Chem., 71(14), 2871-2882.

87. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis, 20(18), 3551-3567.

88. Eriksson J, Chait BT, Fenyo D (2000). A statistical basis for testing the significance of mass spectrometric protein identification results. Anal. Chem., 72(5), 999-1005.

89. Yates JR, Eng JK, McCormack AL, Schieltz D (1995). Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal. Chem., 67(8), 1426-1436.

90. Mann M, Wilm M (1994). Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal. Chem., 66(24), 4390-4399.

91. Eng JK, McCormack AL, Yates JR (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein databaseJ. Am. Soc. Mass Spec., 5, 976-989.

92. Tabb DL, McDonald WH, Yates JR (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome. Res., 1(1), 21-26.

93. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem., 74(20), 5383-5392.

94. Anderson, D.C.; Li, W.; Payan, D.G.; Noble, W.S. (2003). A new algorithm for the evaluation of shotgun peptide sequencing in proteomics: support vector machine classification of peptide MS/MS spectra and SEQUEST scores. J. Proteome. Res., 2(2), 137-146.

95. Fenyo D, Qin J, Chait BT (1998). Protein identification using mass spectrometric information. Electrophoresis, 19(6), 998-1005.

96. Field HI, Fenyo D, Beavis RC (2002). RADARS, a bioinformatics solution that automates proteome mass spectral analysis, optimises protein identification, and archives data in a relational database. Proteomics, 2(1), 36-47.

97. Johnson RS, Biemann K (1989). Computer program (SEQPEP) to aid in the interpretation of high-energy collision tandem mass spectra of peptides. Biomed. Environ. Mass Spectrom., 18(11), 945-957.

98. Zidarov D, Thibault P, Evans MJ, Bertrand MJ (1990). Determination of the primary structure of peptides using fast atom bombardment mass spectrometry. Biomed. Environ. Mass Spectrom., 19(1), 13-26.

99. Fernandez-de-Cossio J, Gonzalez J, Satomi Y, Shima T, Okumura N, Besada V, Betancourt L, Padron G, Shimonishi Y, Takao T (2000). Automated interpretation of low-energy collision-induced dissociation spectra by SeqMS, a software aid for de novo sequencing by tandem mass spectrometry. Electrophoresis, 21(9), 1694-1699.

100. Taylor JA, Johnson RS (1997). Sequence database searches via de novo peptide sequencing by tandem mass spectrometry. Rapid Commun. Mass Spectrom., 11(9), 1067-1075.

101. Dancik V, Addona TA, Clauser KR, Vath JE, Pevzner PA (1999). De novo peptide sequencing via tandem mass spectrometry. J. Comput. Biol., 6(3-4), 327-342.

102. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. 31, 365-370.

103. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. J. Mol. Biol., 5, 215(3), 403-410.

104. Goddard CJR, Gillett HR (2006). Complications of celiac disease: are all patients are at risk? Postgrad Med J 82 (973), 705–712.

105. Abdulkarim AS, Murray JA (2003). Review article: The diagnosis of coeliac disease. Aliment Pharmacol Ther 17(8), 987–995.

106. Mustalahti K, Collin P, Sievänen H, Salmi J, Mäki M (1999). Osteopenia in patients with clinically silent coeliac disease warrants screening. Lancet; 354: 744-745.

107. Fry, L (1995). Dermatitis herpetiformis. Baillieres Clin Gastroenterol; 9: 371-393.

108. Hadjivassiliou, M, Sanders, DS, Grunewald, RA, Woodroofe, N, Boscolo, S, Aeschlimann, D (2010). Gluten sensitivity: from gut to brain. Lancet Neurol; 9: 318-330.

109. Rubio-Tapia A, Murray JA (2007). The liver in celiac disease. Hepatology; 46: 1650-1658.

110. Chakravarty K, Scott DG (1992). Oligoarthritis--a presenting feature of occult coeliac disease. Br J Rheumatol; 31: 349-350.

111. Ludvigsson JF, Montgomery SM, Ekbom A (2005). Celiac disease and risk of adverse fetal outcome: a population-based cohort study. Gastroenterology; 129: 454-463.

112. Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, Greco L et al (2003). HLA types in celiac disease patients not carrying the DQA1*05-DQB1*02 (DQ2) heterodimer: results from the European Genetics Cluster on Celiac Disease. Hum Immunol; 64: 469-477.

113. Briani C, Samaroo D, Alaedini A (2008). Celiac disease: from gluten to autoimmunity. Autoimmun Rev 7(8), 644–650.

114. Freeman HJ (2008). Adult celiac disease in the elderly. World J Gastroenterol 14(45):6911–6914.

115. Plot L, Amital H (2009). Infectious associations of celiac disease. Autoimmun Rev 8(4), 316–319.

116. Torres MI, López Casado MA, Ríos A (2007). New aspects in celiac disease. World J Gastroenterol 13(8), 1156–1161.

117. American Gastroenterological Association medical position statement: Celiac sprue (2001). Gastroenterology 120 (6), 1522–5.

118. Tursi A, Brandimarte G, Giorgetti G (2003). High prevalence of small intestinal bacterial overgrowth in celiac patients with persistence of gastrointestinal symptoms after gluten withdrawal. Am J Gastroenterol 98 (4), 839–43.

119. Lohi S, Mustalahti K, Kaukinen K, Laurila K, Collin P, Rissanen H et al (2007). Increasing prevalence of coeliac disease over time. Aliment Pharmacol Ther; 26: 1217-1225.

120. Catassi C (2005). The world map of celiac disease. Acta Gastroenterol Latinoam ;35, 37e55.

121. Rubio-Tapia A, Kyle RA, Kaplan EL et al (2009). Increased prevalence and mortality in undiagnosed celiac disease. Gastroenterology :137, 88e93.

122. Lohi S, Mustalahti K, Kaukinen K et al (2007). Increasing prevalence of coeliac disease over time. Aliment Pharmacol Ther ;26, 1217e25.

123. Ascher H, Krantz I, Kristiansson B (1991). Increasing incidence of coeliac disease in Sweden. Arch Dis Child 66(5), 608–611.

124. Catassi C, Fabiani E, Rätsch IM, Coppa GV, Giorgi PL, Pierdomenico R, Alessandrini S, Iwanejko G, Domenici R, Mei E, Miano A, Marani M, Bottaro G, Spina M, Dotti M, Montanelli A, Barbato M, Viola F, Lazzari R, Vallini M, Guariso G, Plebani M, Cataldo F, Traverso G, Ventura A et al (1996). The coeliac iceberg in Italy. A multicentre antigliadin antibodies screening for coeliac disease in school-age subjects. Acta Paediatr Suppl 412, 29–35.

125. Kolho KL, Färkkilä MA, Savilahti E (1998). Undiagnosed coeliac disease is common in Finnish adults. Scand J Gastroenterol 33, 1280–1283.

126. Fasano A, Berti I, Gerarduzzi T, Not T, Colletti R, Drago S, Elitsur Y, Green P, Guandalini S, Hill I, Pietzak M, Ventura A, Thorpe M, Kryszak D, Fornaroli F, Wasserman S, Murray J, Horvath K (2003). Prevalence of celiac disease in at-risk and not-at-risk groups in the United States: a large multicenter study. Archives of Internal Medicine 163 (3), 286–92.

127. Green PHR, Stavropoulos SN, Panagi SG, Goldstein SL, Mcmahon DJ, Absan H, Neugut AI (2001). Characteristics of adult celiac disease in the USA: results of a national survey. Am J Gastroenterol 96(1), 126–131.

128. Hausch F, Shan L, Santiago NA, Gray GM, Khosla C (2002). Intestinal digestive resistance of immunodominant gliadin peptides. Am J Physiol Gastrointest Liver Physiol 2002; 283: G996-G1003.

129. Shan L, Molberg O, Parrot I, Hausch F, Filiz F, Gray GM et al (2002). Structural basis for gluten intolerance in celiac sprue. Science; 297: 2275-2279.

130. Molberg O, McAdam SN, Korner R, Quarsten H, Kristiansen C, Madsen L et al (1998). Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. Nat Med; 4: 713-717.

131. Vader LW, de Ru A, van der Wal Y, Kooy YM, Benckhuijsen W, Mearin ML et al (2002). Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. J Exp Med; 195: 643-649.

132. Van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, Papadopoulos G et al (1998). Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. J Immunol; 161: 1585-1588.

133. Alaedini A, Green PH (2008). Autoantibodies in celiac disease. Autoimmunity; 41: 19-26.

134. Marsh M (1992). Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity (celiac sprue). Gastroenterology 102 (1), 330–54.

135. Valdimarsson T, Franzen L, Grodzinsky E, Skogh T and Ström M (1996). Is small bowel biopsy necessary in adults with suspected celiac disease and IgA anti-endomysium antibodies? 100% positive predictive value for celiac disease in adults. Dig Dis Sci 41, 83-7.

136. Seah PP, Fry L, Rossiter MA, Hoffbrand AV and Holborow EJ (1971). Anti-reticulin antibodies in childhood celiac disease. Lancet 2, 681-2.

137. Chorzelski TP, Sulej J, Tchorzewska H, Jablonska S, Beutner EH, Kumar V (1983). IgA class endomysium antibodies in dermatitis herpetiformis and celiac disease. Ann N Y Acad Sci 420, 325-34.

138. Korponay-Szabó I, Dahlbom I, Laurila K, Koskinen S, Woolley N, Partanen J, Kovács J, Mäki M, Hansson T (2003). Elevation of IgG antibodies against tissue transglutaminase as a diagnostic tool for coeliac disease in selective IgA deficiency. Gut 52 (11), 1567–71.

139. Raivio T, Korponay-Szabo I, Collin P, Laurila K, Huhtala H, Kaartinen T, Partanen J, Maki M, Kaukinen K (2007). Performance of a new rapid whole blood coeliac test in adult patients with low prevalence of endomysial antibodies. Dig. Liver Dis: 39, 1057-1063.

140. Paulley JW (1954). Orservation on the aetiology of idiopathic steatorrehoea, jejunal and lymph node biopsies. BMJ, 1318-21.

141. Green PH, Cellier C (2007). Celiac disease. N Engl J Med 357, 1731-43.

142. Ferguson A and Murray D (1971). Quantitation of intraepithelial lymphocytes in human jejunum. Gut 12, 988-994.

143. Collin P, Reunala T, Rasmussen M, Kyrönpalo S, Pehkonen E, Laippala P, Mäki M (1997). High incidence and prevalence of adult celiac disease. Augmented diagnostic approach. Scandinavian Journal of Gastroenterology. 32(11), 1129-33.

144. Mäki M, Kallonen K, Läahdeaho ML, Visakorpi JK (1988). Changing pattern of childhood coeliac disease in Finland. Acta Paediatr Scand 77, 408-12.

145. Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH, Dunn MJ (2000). A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis; 21(17):3666-72.

146. Shiner M, Ballard J (1972). Antigen-antibody reactions in jejunal mucosa in childhood coeliac disease after gluten challenge. Lancet; 1: 1202-1205

147. Gandhi, Fiona C. McKay, Eve Diefenbach, Ben Crossett, Stephen D. Schibeci, Robert N. Heard, Graeme J. Stewart, David R. Booth, Jonathan W. Arthur (2010). Novel approaches to detect serum biomarkers for clinical response to interferon-β treatment in multiple sclerosis. Volume 5, Issue 5, e10484.

148. Bellei E, Bergamini S, Monari E, Fantoni LI, Cuoghi A, Ozben T, Tomasi A (2011). High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. Amino Acids;40(1):145-56

149. Yang Q, Graham TE, Mody N, Preitner F, Peroni OD, Zabolotny JM, Kotani K, Quadro L, Kahn BB (2005). Serum retinol binding protein 4 contributes to insulin resistance in obesity and type 2 diabetes. Nature.;436(7049):356-62.

150. Graham TE, Yang Q, Blüher M, Hammarstedt A, Ciaraldi TP, Henry RR, Wason CJ, Oberbach A, Jansson PA, Smith U, Kahn BB (2006). Retinol-binding protein 4 and insulin resistance in lean, obese, and diabetic subjects. N Engl J Med.;354(24):2552-63.

151. McMillan SA, Dickey W, Douglas JP, Hughes DF (2001). Transthyretin values correlate with mucosal recovery in patients with coeliac disease taking a gluten free diet. J Clin Pathol.;54(10):783-6.

152. Toruner F, Altinova AE, Akturk M, Kaya M, Arslan E, Bukan N, Kan E, Yetkin I, Arslan M (2011). The relationship between adipocyte fatty acid binding protein-4, retinol binding protein-4 levels and early diabetic nephropathy in patients with type 2 diabetes. Diabetes Res Clin Pract.;91(2):203-7.

153. Choi KM, Yannakoulia M, Park MS, Cho GJ, Kim JH, Lee SH, Hwang TG, Yang SJ, Kim TN, Yoo HJ, Baik SH, Kim SM, Mantzoros CS (2011). Serum adipocyte fatty acid-binding protein, retinol-binding protein 4, and adiponectin concentrations in relation to the development of the metabolic syndrome in Korean boys: a 3-y prospective cohort study. Am J Clin Nutr.;93(1):19-26. Epub 2010 Nov 24.

154. Duntas LH, Biondi B (2011). New insights into subclinical hypothyroidism and cardiovascular risk. Semin Thromb Hemost.;37(1):27-34.

155. Lu HY, Li XF, Mu PW, Jiang W, Zeng LY (2010). Depot-specific expression of retinol-binding protein 4 in human adipose tissue and their relationship with obesity and insulin resistance. Zhonghua Yi Xue Za Zhi.;90(48):3395-8.

156. Mehdi H, Manzi S, Desai P, Chen Q, Nestlerode C, Bontempo F, Strom SC, Zarnegar R, Kamboh MI (2003). A functional polymorphism at the transcriptional initiation site in β-2-glycoprotein I (apolipoprotein H) associated with reduced gene expression and lower plasma levels of β-2-glycoprotein I. Eur J Biochem.;270(2):230-8.

157. Shamir R, Shoenfeld Y, Blank M, Eliakim R, Lahat N, Sobel E, Shinar E, Lerner A (2003). The prevalence of coeliac disease antibodies in patients with the antiphospholipid syndrome. Lupus.;12(5):394-9.

158. Karoui S, Sellami MK, Laatar AB, Zitouni M, Matri S, Laadhar L, Fekih M, Boubaker J, Makni S, Filali A (2007). Prevalence of anticardiolipin and anti-β-2-glycoprotein I antibodies in celiac disease. Dig Dis Sci.;52(4):1096-100.

159. Mehrani T, Petri M (2011). IgM anti-ss2 glycoprotein I is protective against lupus nephritis and renal damage in systemic lupus erythematosus. J Rheumatol.;38(3):450-3.

160. Zhou A, Huntington JA, Pannu NS, Carrell RW, Read RJ (2003). How vitronectin binds PAI-1 to modulate fibrinolysis and cell migration. Nat Struct Biol.;10(7):541-4.

161. Preissner KT, Seiffert D (1998). Role of vitronectin and its receptors in haemostasis and vascular remodeling. Thromb Res.;89(1):1-21.

162. Felding-Habermann B, Cheresh DA (1993). Vitronectin and its receptors. Curr Opin Cell Biol.;5(5):864-8.

163. Kadowaki M, Sangai T, Nagashima T, Sakakibara M, Yoshitomi H, Takano S, Sogawa K, Umemura H, Fushimi K, Nakatani Y, Nomura F, Miyazaki M (2011). Identification of vitronectin as a novel serum marker for early breast cancer detection using a new proteomic approach. J Cancer Res Clin Oncol.

164. Bortner JD, Richie JP, Das A, Liao J, Umstead TM, Stanley A, Stanley BA, Belani CP, El-Bayoumy K (2011). Proteomic Pprofiling of human plasma by iTRAQ reveals down-regulation of ITI-HC3 and VDBP by cigarette smoking. J Proteome Res.;10(3):1151-9.

165. Aisen P, Listowsky I (1980). Iron transport and storage proteins. Annu Rev Biochem.;49:357-93.

166. Moura, I.C., M.N. Centelles, M. Arcos-Fajardo, et al (2001). Identification of the transferrin receptor as a novel immunoglobulin (Ig)A1 receptor and its enhanced expression on mesangial cells in IgA nephropathy. J. Exp. Med. 194: 417–425.

167. Lindfors K, Suzuki H, Novak J, Collin P, Saavalainen P, Koskinen LL, Mäki M, Kaukinen K (2011). Galactosylation of serum IgA1 o-glycans in celiac disease. J Clin Immunol.;31(1):74-9.

168. Heyman M, Menard S (2009). Pathways of gliadin transport in celiac disease. Ann N Y Acad Sci.;1165:274-8.

169. Margolis J (1958). Activation of plasma by contact with glass: evidence for a common reaction which releases plasma kinin and initiates coagulation. J Physiol.;144(1):1-22.

170. Qiu Y, Patwa TH, Xu L, Shedden K, Misek DE, Tuck M, Jin G, Ruffin MT, Turgeon DK, Synal S, Bresalier R, Marcon N, Brenner DE, Lubman DM (2008). Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. J Proteome Res.;7(4):1693-703.

171. Zhao W, Wang Y, Wang L, Lu X, Yang W, Huang J, Chen S, Gu D (2009). Gender-specific association between the kininogen 1 gene variants and essential hypertension in Chinese Han population. J Hypertens.;27(3):484-90.

172. Rustemeyer SM, Lamberson WR, Ledoux DR, Wells K, Austin KJ, Cammack KM (2011). Effects of dietary aflatoxin on the hepatic expression of apoptosis genes in growing barrows. J Anim Sci.;89(4):916-25.

173. Pepys, M. B., Booth, D. R., Hutchinson, W. L., Gallimore, J.R. et al (1997). Amyloid P component. A critical review., Amyloid – Int. J. Exp. Clin. Invest. 1997, 4, 274–295.

174. Jastorff AM, Haegler K, Maccarrone G, Holsboer F, Weber F, Ziemssen T, Turck CW (2009). Regulation of proteins mediating neurodegeneration in experimental autoimmune encephalomyelitis and multiple sclerosis. Proteomics Clin Appl.;3(11):1273-87.

175. Chen HM, Lin CY, Wang V (2011). Amyloid P component as a plasma marker for Parkinson's disease identified by a proteomic approach. Clin Biochem.;44(5-6):377-85.

176. Klein CJ, Vrana JA, Theis JD, Dyck PJ, Dyck PJ, Spinner RJ, Mauermann ML, Bergen HR 3rd, Zeldenrust SR, Dogan A (2011). Mass spectrometric-based proteomic analysis of amyloid neuropathy type in nerve tissue. Arch Neurol.;68(2):195-9.