A joint meeting of the Nutrition Society and the UK Molecular Epidemiology Group was held at Wren Room, Royal Institute of British Architects, Portland Place, London on 2 December 2010

Conference on 'Functional genomic biomarkers, nutrition and disease susceptibility'

Nutriproteomics: technologies and applications for identification and quantification of biomarkers and ingredients

Sandra Sénéchal[1]* and Martin Kussmann[2,3]

[1]*Nestlé Research Center, Functional Genomics, Lausanne, Switzerland*
[2]*Proteomics and Metabonomics Core, Nestlé Institute of Health Science, Lausanne, Switzerland*
[3]*Faculty of Science, Aarhus University, Aarhus, Denmark*

Nutrition refers to the process by which a living organism ingests and digests food and uses the nutrients therein for growth, tissue maintenance and all other functions essential to life. Food components interact with our body at molecular, cellular, organ and system level. Nutrients come in complex mixtures, in which the presence and concentration of single compounds as well as their interactions with other compounds and the food matrix influence their bioavailability and bioefficacy. Traditionally, nutrition research mainly concentrated on supplying nutrients of quality to nourish populations and on preventing specific nutrient deficiencies. More recently, it investigates health-related aspects of individual ingredients or of complete diets, in view of health promotion, performance optimisation, disease prevention and risk assessment. This review focuses on proteins and peptides, their role as nutrients and biomarkers and on the technologies developed for their analysis. In the first part of this review, we provide insights into the way proteins are currently characterised and analysed using classical and emerging proteomic approaches. The scope of the second part is to review major applications of proteomics to nutrition, from characterisation of food proteins and peptides, via investigation of health-related food benefits to understanding disease-related mechanisms.

Nutriproteomics: Biomarkers: Ingredients

Proteins and peptides as ingredients and biomarkers

Proteins form a major class of macronutrients and they participate in most biological processes in the body. Enzymes are those proteins that catalyse virtually every biochemical reaction involved in metabolism (e.g. the digestive enzymes pepsin and trypsin). Proteins also ensure structural and mechanical functions and they are involved in cell signalling and immune response. Therefore, the body needs relatively large amounts of proteins to ensure proper function and to cope with the continuous synthesis and degradation (protein turnover).

Proteins and peptides are composed of amino acids that are arranged in a sequential fashion to form higher structures and functions. The essential amino acids cannot be synthesised by the body and must therefore be taken up with the diet[1] (i.e. for adults: leucine, isoleucine, valine, lysine, threonine, tryptophan, methionine, phenylalanine and histidine). Food proteins differ in their amino acid composition depending on their origin, i.e. animal or plant source. Consequently, a balanced diet covers proteins from complementary sources (e.g. meat, vegetables, cereals, grains and legumes) in order to avoid amino acid deficiencies. The nutritional quality of proteins is described by the amino acid composition, digestibility and absorptive ability[2].

Proteins, by nature, are key actors in all biological processes in the human organism. Proteomics is a powerful

---

tool for the elucidation of such molecular events related to nutrition: it can identify and quantify bioactive proteins and peptides, shed light on their effects at protein/peptide level (biomarkers) and thereby addresses questions of nutritional bioefficacy. In this paper, we will first describe how proteins are analysed in a holistic, high-throughput fashion; and in the second part we will review major applications of proteomics in the nutrition field.

Proteomic technologies

Proteomics builds on a combination of several techniques and has been propelled by technological progresses including genome sequencing, biomolecule separation, MS and bioinformatics. The discipline has evolved as an analogue to genomics and has traditionally aimed at identifying all proteins present in a given sample at a given time. Over the last two decades, proteomics has been developed into an established technology for biomarker discovery[3,4], clinical applications[5], disease profiling and diagnostics[6,7], the study of protein interactions[8] and of the dynamics of signalling pathways[9].

The proteome is highly dynamic and constantly changing in response to environmental stimuli including nutrition. Nutritional proteomics holds great promise to (a) profile and characterise body and dietary proteins, including digestion and absorption of the latter; (b) identify biomarkers of nutritional status and health/disease conditions; and (c) understand functions of nutrients and other dietary factors in growth, reproduction and health[10].

The classical proteomic workflow, sometimes referred to as bottom-up approach, combines techniques of protein separation, protein digestion into peptides, mass spectrometric analysis, protein identification by comparison with protein databases and protein quantification[11]. The development of gentle ionisation strategies such as electrospray ionisation[12] and matrix-assisted laser desorption/ionisation (MALDI)[13] MS, as well as the publication of genome sequences of many organisms including human were breakthrough factors for the successful development and deployment of proteomics.

The sequence of amino acids defines the primary structure of a protein. Secondary, tertiary and quaternary structures concern the arrangement of the amino acid chain in space, resulting from covalent or non-covalent interactions. Nowadays, the primary structures of proteins are routinely analysed by MS. Post-translational modifications (PTM) and protein–ligand interactions can also be elucidated this way.

Protein separation on gels and columns

A major analytical challenge of proteomics is the dynamic range of protein concentrations (e.g. estimated 10^{12} in human blood)[14]. Current MS-based proteomic platforms can deliver a dynamic range of 10^4. This means that the low-abundant proteome has to be addressed by depletion of the most abundant proteins[15] or by selective enrichment of low-abundant proteins[16–18]. After protein depletion and/or enrichment, further separation is performed at protein

and/or peptide level, based on two-dimensional (2D) gels or on liquid chromatography (LC) or on hybrid approaches (Gel-LC). Fig. 1 summarises the classical proteomic workflows.

The gel strategy for protein separation offers the advantage of visualisation of proteins, and, to some extent, of their modifications and, therefore, it preserves the protein context. After protein separation on gel, the gel spots are then excised, digested with trypsin and further processed in LC-tandem MS (MS/MS) or MALDI-MS instruments for protein identification. However, major drawbacks of this method are that (i) hydrophobic proteins such as membrane proteins, very basic proteins and low-abundant proteins are not well represented on the gels and (ii) the procedure is low throughput and difficult to standardise and automate.

The 'gel-free' discovery approach for protein identification is referred to as 'shotgun' analysis: here, the sample or protein mixture is directly digested in solution. The resulting peptide mixture is then separated on an HPLC usually coupled online to a mass spectrometer for peptide mass identification. The development of ultra-HPLC delivers improved peptide separation[19] and column efficiency by means of higher mass sensitivity, analytical resolution and speed. The resolving capability of the HPLC separation can be enhanced with 2D HPLC techniques[20] that combine ion-exchange chromatography followed by reverse-phase HPLC.

Peptides can also be separated using gas-phase fractionation, which is defined as iterative mass spectrometric interrogations of a sample over multiple smaller mass-to-charge (m/z) ranges. By doing so, a higher number of unique peptides compared to the ions selected from the wide mass range scan in standard LC-MS/MS analysis can be addressed. Gas-phase fractionation is described as a means to achieve higher proteome coverage than classical LC-MS/MS analyses of a complex peptide mixture[21–23].

Protein identification by MS

Mass spectrometers identify proteins and peptides by determination of their exact masses and by generating information on their amino acid sequences. Today the main ionisation methods deployed are electrospray ionisation[24] and MALDI[25]. These ion sources are combined with various mass analysers that separate the ions by m/z. The most popular analysers in proteomics are ion traps, triple-quadrupoles, time-of-flight tubes, orbitrap and Fourier-transform ion cyclotron resonance, with their specific advantages: high sensitivity and multiple-stage fragmentation for ion traps; high selectivity for triple-quadrupoles; high sensitivity and speed for time-of-flight. Current top-end proteomic machines are orbitraps[26] and Fourier-transform ion cyclotron resonance instruments[27], which provide very high mass accuracy and resolution compared to the other analysers. MS/MS consists of either the fragmentation of a selected precursor peptide ion to generate specific fragment ions for sequence elucidation (data-dependent acquisition); or uncoupled acquisitions of intact and fragment masses with retrospective reconstitution of
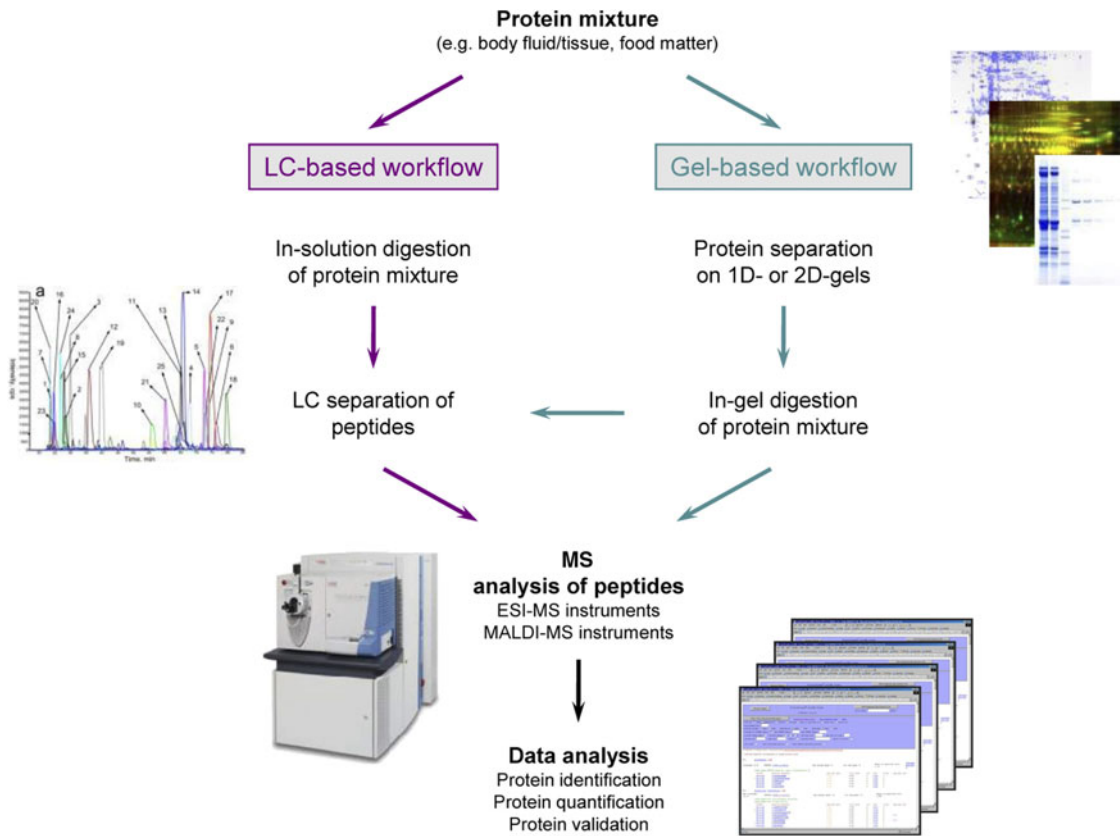
Fig. 1. (Colour online) Classical workflows in MS-based discovery proteomics. Both gel-based and liquid chromatography (LC) based approaches are represented, with the latter increasingly taking over the proteomics business. ESI, electrospray ionisation; MALDI, matrix-assisted laser desorption/ionisation.

the parent–daughter ion context (data-independent acquisition; see Emerging Technologies)[28].

Considering the amount of data generated in a single shotgun run, search algorithms have been developed to process the raw data automatically and to compare the measured MS/MS spectra against theoretical fragment ion spectra generated by *in silico* digestion of protein databases. Comparison of experimental with theoretical spectra results in a list of possible peptide matches, each with an associated score that quantifies the quality of the match. The peptide(s) with the highest score(s) is(are) generally considered for protein identification. Several of these search engines have been commercialised (e.g. Sequest, Mascot or Phenyx), but others are freely available (open source programs, e.g. XTandem or OMSSA). They usually agree on 80% of the identifications[29]. The two most established applications are Sequest[30] and Mascot[31].

Identified peptides and proteins can be further validated by software that applies statistical algorithms to calculate additional scores and probabilities, thereby distinguishing correct from incorrect assignments. PeptideProphet and ProteinProphet from the TransProteomic Pipeline (ISB, Seattle)[32,33] determine respectively at peptide and protein level the probabilities of correctness associated with false discovery rate. These tools provide the researcher with means to assess the quality of the data in a dataset-dependent manner and to control the trade-off between

false positives (specificity) and false negatives (sensitivity)[34]. The second strategy to elucidate the false-positive/false-negative relationship relies on a database search using a target-decoy database[35], which is typically generated by reversing the sequences of target protein database. The search is done against the target and the decoy database and allows the estimation of false positives.

Protein quantification

Once qualitative analysis of a sample is achieved (first discovery mode, protein catalogue), very often quantitative information is required to obtain more insight into proteome differences between conditions or over time.

Relative quantification enables the comparison of two or more biological samples/conditions and the identification of (candidate) biomarkers, i.e. proteins that are more or less abundant (up- or down-regulated) in a certain condition compared to another.

The classical method for relative quantification evolved from 2D gels: 2D difference gel electrophoresis is characterised by differential labelling of proteins with fluorescent dyes prior to separation according to isoelectric point and molecular weight[36]. An internal standard is used to match the protein patterns across gels thereby facilitating gel alignment, spot matching and quantification. Dedicated
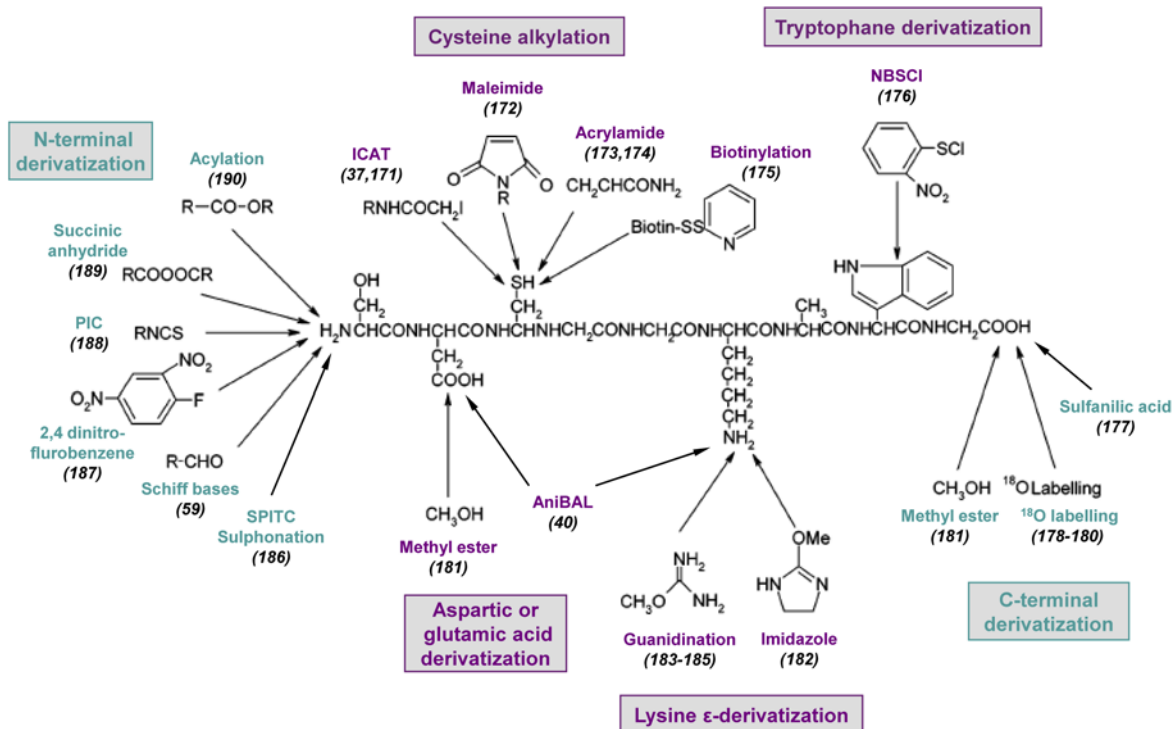
Adapted from Julka et al., 2004 [170]

Fig. 2. (Colour online) Protein quantification based on chemical stable-isotope labelling. Labels targeting N- and C-termini are reported in light print, whereas labels targeting side chains of amino acids within a polypeptide are reported in dark print. Derivatisations occurring at the N-terminus also affect the ε-amino group of lysine. For each type of label, a reference is indicated[37,40,42–63]. Adapted from Julka and Regnier[43], authorisation no. 2597551000090.

software (e.g. Progenesis SameSpots and DeCyder) allow for gel comparison and spot quantification.

Other procedures for relative quantification of proteins rely on metabolic or chemical labelling of proteins by incorporation of stable isotopes (usually ^{13}C or ^{15}N) in the samples to be compared and quantified at MS level. A summary of popular chemical labels used for relative protein quantification is shown on Fig. 2 that depicts a generic peptide with its side chains and modification options. Labelling of proteins and peptides is performed by *in vitro* chemical or enzymatic derivatisation and include for example: isotope-coded affinity tag[37], isotope tags for relative and absolute quantification[38], isotope-coded protein label[39] or aniline and benzoic acid labelling[40]. The differentially labelled samples are mixed and infused into an LC-MS/MS instrument. While these techniques involving labelling are accurate, they need specific sample preparation and involve added costs due to stable-isotope labelled reagents. In contrast to chemical labelling, stable isotope labelling by amino acids in cell culture[41] consists in labelling of proteins already during cell growth and division by incorporation of labelled amino acids.

Recently, methods for label-free relative quantification have been developed. In this case, each sample is separately analysed by LC-MS/MS and resulting data are processed using software designed for LC-MS/MS run alignment, extraction of peptide intensities and peptide counts (e.g. Progenesis LCMS, DeCyder MS, SuperHirn,

SpecArray or MSQuant). Currently, two label-free quantification strategies can be used: (a) measuring and comparing the mass spectrometric signal intensity of peptide precursor ions belonging to a particular protein[64,65]; and (b) counting and comparing the number of fragment spectra identifying peptides of a given protein[66]. In comparison with stable isotopes, label-free protein quantification is simpler to perform; there is technically no limit to the number of sample/conditions to be compared (except the analytical capacity of software/computers/servers); and it can yield an improved proteome coverage with a broader dynamic range[67]. However, label-free proteome comparisons are limited to not too complex proteomes, otherwise the peptide-to-peptide alignment becomes too difficult due to LC elution and m/z overlaps.

Absolute quantification of proteins relies on the addition of an internal standard at a known concentration. The Absolute QUAntification method[68] uses synthetic stable-isotope labelled proteotypic peptides as internal standards that are otherwise identical to the peptides to be quantified. This approach requires preliminary analyses to select the peptides to be used to quantify the protein(s) of interest (proteotypic peptides). Recently, the QconCAT technology has been developed for parallel production of labelled proteotypic peptides which are then used in multiplexed quantification assays[69]. QconCAT consists of an artificial gene, inserted into a vector for expression in *Escherischia coli*, which is designed to express

artificial proteins comprising a concatenation of proteo-typic peptides. This latter technology is highly useful for repetitive, multiplexed analysis, i.e. a protein assay-like situation.

The systematic large-scale approach for absolute protein quantification has been referred to as selected-reaction monitoring or multiple-reaction monitoring[70]. This type of protein quantification is based on the quantification of relevant proteotypic peptides and is exclusively performed with triple-quadrupole mass spectrometers. The masses of the peptides and of their most abundant fragments are defined in the method and the mass spectrometer only scans for these as well as for the corresponding stable-isotope labelled proteotypic peptides. A chromatographic peak proportional to the peptide amount appears only if both parent and fragment masses are present (referred to as one transition). These peaks can then be integrated and the peptide and protein concentrations can be calculated by comparison with the internal peptide standard. This method enables the targeted, multiplexed, high-throughput quanti-fication of low-abundance proteins in highly complex mixtures.

Analysis of post-translational modifications: protein functionality

PTM of amino acids give functionality to the proteins through the attachment of functional groups such as phos-phate, carbohydrates, acetate or lipids. PTM play crucial roles in regulating the biology of the cell since they can change a protein's physical or chemical property, activity, localisation or stability. Some PTM can be added and removed dynamically as a mechanism for reversibly con-trolling protein function and cell signalling. Several pro-teomic techniques have been developed to identify and quantify PTM and allow the study of modifications such as phosphorylation, acetylation, glycosylation or lipid modifications.

Phosphorylation of proteins controls many cellular pro-cesses such as growth, differentiation, metabolism, signal-ling and cell death, and is itself regulated by enzymatic activity (i.e. by kinases and phosphatases). The challenges of phosphoproteomics lie, as for proteomics, in the com-plexity, dynamic range and temporal dynamics of protein isoforms[71]. Several methods have been developed to enrich phosphoproteins and include anti-phosphotyrosine antibodies[72]; immobilised metal affinity chromato-graphy[73]; and chemical modification and strong exchange chromatography[74]. The analysis of phosphorylated pro-teins has been facilitated by the development of new fragmentation techniques such as electron capture dis-sociation[75] and electron transfer dissociation[76,77] that help identify and determine the location of phosphoryla-tions that cannot be as efficiently characterised by standard collision-induced dissociation.

Protein glycosylation is prevalent in proteins that are involved in mechanisms like cell–cell interactions, immune system (e.g. antibodies, MHC) or transport (e.g. transfer-rin). Also, glycoproteins account for a major proportion of milk and human blood proteomes for which it has been estimated that 70 and 50 % of all proteins are glycosylated,

respectively. There are two types of glycoproteins: (a) N-glycosylated proteins with carbohydrates linked to the side chain of the asparagine; and (b) O-glycosylated proteins, in which carbohydrates are bound to the side chain of serine, threonine, hydroxylysine or hydroxyproline. The functions of glycoproteins are still incompletely understood[78]. MS can provide information on molecular mass, composition, sequence and sometimes branching of a glycan chain[79,80]. As glycoprotein forms are often minor constituents com-pared to the non-glycosylated proteins, enrichment meth-ods such as cell surface-capture technology[42] or affinity capture with lectins[81] have been set-up.

Acetylation is a PTM that has been associated with several biological processes, especially gene expression regulation by histones, which pack the chromosomes so that they fit into the cell nucleus[82]. Acetylation usually occurs at lysines or on N-terminal groups of peptides or proteins and is involved in the destabilisation of chromatin and recruitment of effector proteins[83–86]. Especially the mass spectral deciphering of the so-termed and above dis-cussed histone codes, a key epigenetic mechanism, is expected to shed light on the phenomenon of metabolic programming: there is compelling evidence that the human body retains a memory of environmental, such as nutri-tional, impacts and may thereby be 'metabolically (re)wired'. Such events have been associated with epi-genetics, of which histone (de)acetylation is a central mechanism[87]. Recently O-aceylated serine and threonine residues have been identified[88]. Acetylation is stable to peptide fragmentation and can be detected by its charac-teristic mass shift from unmodified form. Trypsin cleavage at acetyllysine residues is usually blocked, so the acety-lated peptides are detected as 'missed cleavage' product when performing database searches[89]. Enrichment of acetylated peptides is difficult and therefore studies have generally characterised protein acetylation on partially purified mixtures (e.g. histones). Immunoaffinity techni-ques have been developed to purify acetylated peptides: acetyllysine sites were mapped by enriching acetylated peptides using resin-coupled antibodies to acetyllysine[90].

Lipoproteins are lipid–protein complexes whose primary function is thought to be transport of cholesterol and other lipids and that include five protein classes: chylomicrons, VLDL, intermediate density lipoproteins, LDL and HDL. Many lines of evidence strongly link them to the immune system and macrophage biology[91–93]. Several approaches have been developed for lipoprotein isolation[94–97]; deli-pidation[98,99]; sample preparation (solubilisation and digestion)[100–103]; and the characterisation of protein component of lipoproteins using MS[104,105].

Emerging proteomic technologies

The full characterisation of a given proteome remains a challenge today. This is due to factors such as the large dynamic range of protein expression, complexity of the mixture in terms of numbers of proteins, as well as lack of methods to amplify proteins. However, proteomics is a rapidly expanding field and new analytical approaches are emerging.

Usually, proteomic profiling is done in a data-dependent acquisition mode in which the most abundant ionised peptides from each MS scan are selected for subsequent MS/MS analysis. A data-independent acquisition method, referred to as Precursor Acquisition Independent From Ion Count, consists of the acquisition of MS/MS spectra at every m/z value regardless of whether a precursor ion is observed or not: precursor ion scans (MS scan) are no longer conducted[28]. This strategy yields better proteome coverage, higher numbers of identified proteins and an extended dynamic range compared to the classical data-dependent method[106].

Imaging MS is a technology enabling the direct examination of the distribution of biomolecules (e.g. proteins, peptides, etc) in cells or tissues[107–109]. Imaging MS is principally used for clinical applications[110,111] and biomarker discovery in diseased tissue[112,113]. In imaging MS, frozen tissue sections are mounted on a target plate, covered with a suitable matrix, dried and inserted into a MALDI-MS for spectra acquisition. The mass spectrometer records the spatial distribution of peptides and proteins by scanning the tissue surface with consecutive laser shots. Alternatively, *in situ* tryptic digestion of a spot on a tissue followed by peptide sequencing of a predicted fragment by MALDI-MS/MS can be done[114]. Specific methods have been developed for the analysis of formalin-fixed paraffin-embedded sections[115]. Often histological staining, either on the same section[116] or on a serial section[117], is used to guide the placement of matrix and provides the capability of focusing on areas having a high content of a cell type of interest.

The trend towards biological analysis at decreasing scale, ultimately down to an individual cell, continues, and MS with sensitivity of detecting a few to single molecules will be necessary. Recently, a prototype for a mass spectrometer with single-molecule sensitivity for single-cell proteomics has been designed[118]. Another method for the analysis of protein complexes in single cells, so-called visual proteomics, has been developed[119] and consists of the combination of quantitative MS with cryo-electron tomography for the detection, counting and localisation of protein complexes.

Proteomics today: a paradigm shift

A recent publication reviews proteome coverage and reports on the detection of protein abundance over seven orders of magnitude with today's high-end platforms[120]. This impressive power is a combined result of highly improved mass spectrometric instrumentation and data acquisition/processing as well as of highly sophisticated fractionation, enrichment and depletion techniques.

However, given the complexity and dynamics of proteomes, proteomics experiences nowadays a paradigm shift. Strategically speaking, the original hypothesis-free discovery workflow is being increasingly complemented or followed up by either hypothesis-driven analysis or even by candidate-based targeted analysis and validation: a recent review puts the discovery, directed and targeted proteomics approaches into perspective[121]. Proteomics has thereby developed from a pure discovery to a screening

and validation tool. The discovery workflow (or shotgun approach) aims at identifying large protein sets in a sample, and can include protein quantification (with or without prior protein labelling). The directed proteomics workflow consists of two successive analyses of the same sample. The first analysis is a survey scan aiming at the definition of a list of target peptides used for a second analysis surveying exclusively the peptides of the target list. This approach allows the quantification of less abundant proteins. Finally, targeted proteomics is a hypothesis-driven approach focusing on the detection and quantification of specific peptides associated with the proteins of interest (selected-reaction monitoring or multiple-reaction monitoring).

The other change of proteomic 'philosophy' roots in the increasing appreciation of peptides as bioactive, health beneficial food components[122]. The analysis of such peptides requires a different analytical approach because these entities vary much more in their chemical nature than classical tryptic peptides generated in shotgun proteomics workflows for protein biomarker identification: multiple-processing parameters and digestive enzymes come into play and these generate not only a large variety in peptide length, sequence and terminal residues but also a number of peptide modifications. Moreover, there is only a single possibility to identify and quantify the native peptide of interest as such molecule is not one of several representatives of a parent protein, as it is typically the case in biomarker research. In view of this food peptidome complexity, it becomes evident that proteomic tools must be further developed and adapted from biomarker to bioactive research[123].

Proteomics and nutrition: major applications

Several groups including ours have contributed to the introduction and adaptation of proteomics to the field of nutrition and health[124–127]. Numerous studies have shown the prominent role of nutrition for maintaining and improving health. In this view, proteomics has been deployed in fields such as characterisation of bioactive proteins and peptides[128]; elucidation of immune-related disorders[129,130]; investigation of metabolism-related disorders[131]; dietary intervention studies for recovery[132,133]; and mechanistic elucidation of nutrient action[134]. The following selected citations cover topics from characterisation of the food matter itself via investigation of health-related food benefits to understanding disease-related mechanisms.

Protein and peptides as food ingredients

Milk. Milk is an essential component for infant nutrition since it represents the major source of feeding for newborns and infants[135]. It is a rich source of functional peptides and proteins beneficial for human health. Extensive characterisation of milk from different species has been reported[135–139] and the composition of milk in terms of major proteins, lipids and carbohydrates has been established.

Technical milk fractionation by successive centrifugation steps yields three major fractions: caseins, whey and milk fat globule membrane (MFGM)[140]. Each of these fractions contains different protein functionalities which have been studied using proteomics.

The whey protein fraction is dominated by a small number of abundant proteins which constitute over 80% of its protein content[141]. In particular, β-lactoglobulin alone constitutes 50% of whey. In consequence, further fractionation was necessary to identify less abundant proteins. A gel-based approach[142] enabled the identification of a large number of minor whey proteins, for example, a cluster of osteopontin peptides suggesting novel bioactivities. Another study, based on the use of electrospray ionisation and MALDI ionisation sources in parallel, allowed enhanced protein identification[143]: a total of thirty-nine bovine milk proteins were identified with a high degree of confidence.

The MFGM is a milk fraction rich in bioactive proteins. A qualitative and a quantitative proteomic profiling of two MFGM enriched milk fractions, whey protein concentrate and buttermilk protein concentrate was reported by our group[144]: using an LC-MS/MS-based shotgun approach, we could reveal the presence of 244 proteins in whey protein concentrate and 133 in buttermilk protein, respectively, and provided an extensive characterisation of the protein content in those two fractions. Then, a label-free profiling approach delivered semi-quantitative comparison of both fractions and yielded protein fingerprints. Finally, we performed absolute quantification by combining stable-isotope dilution and multiple-reaction monitoring in order to precisely quantify seven major MFGM proteins.

PTM of milk proteins (phosphorylation and glycosylation) were investigated by mass spectrometric technologies. Different approaches, from 2D gels to LC separation, were used to determine the phosphorylation pattern of caseins in human, bovine, equine, goat and buffalo milk[144–148]. In particular, a study of buffalo skim milk, whey and MFGM reported phosphorylation data on caseins providing scientific basis to coagulation/cheese making processes used in dairy productions[149]. Glycosylation of milk proteins has also been investigated by MS[150,151]. It is estimated that, in milk, glycoproteins may account for up to 70% of the total protein content, whereas it is about 50% for all human proteins[136]. Indeed, the most abundant proteins in milk including casein, lactoferrin and the Ig are all glycoproteins. The position and extent of glycosylation of these proteins affect their degradation, the resulting released peptides and glycopeptides and the function they provide. Indeed, there is emerging evidence for the involvement of milk glycoproteins in infant protection against pathogen infection[136,152–154]. For example, the glycoproteins in MFGM are considered to operate as specific bacterial and viral ligands preventing the pathogens from binding to the intestinal mucosa of the infants[155–157]. Hydrophilic interaction chromatography was used to enrich glycoproteins from human milk[158] and enabled the identification of thirty-two glycoproteins and sixty-three N-glycosylated sites. Immunocompetent complexes, membrane fat globule enzymes, proteins involved in lipid metabolism and specific receptors figured among these glycoproteins.

Lactoferrin is a major Fe-binding mammalian milk glycoprotein that impacts the defence system of the human host: it can for example prevent microbial growth, by direct interaction with the membrane of Gram-negative bacteria[159]. Lactoferrin and its derived peptides are also known to influence cytokine production in cell cultures experiments mimicking immune and inflammatory processes[160].

The potential benefits of food-derived peptides in terms of reduced risk of CVD have been reviewed[134]: the favourable properties for blood pressure, oxidative stress, homoeostasis, appetite and lipid metabolism have been discussed. Also, the benefits of lactotripeptides on hypertension are well established: the tripeptides VPP (Valine–Proline–Proline) and IPP (Isoleucine–Proline–Proline) form upon fermentation of a milk product by *Lactobacillus helveticus* and *Saccharomyces cerevisiae*. When this fermented milk was fed to rats, the animals' blood pressure was lowered[161].

Probiotics. Probiotics are live micro-organisms which, when administered in adequate amounts, confer a health benefit on the host[162]. Probiotics are commonly consumed as part of fermented foods with specially added active live cultures, such as in yoghurt or dietary supplements. Lactic acid bacteria and bifidobacteria are the most common types of micro-organisms used as probiotics. While their health benefits have been documented in clinical trials, their mechanisms of action are still poorly understood. The benefits of probiotics include stimulation of the mucosal immunity, reduction of mucosal alterations and interaction with mediators of inflammation[163,164]. Numerous proteomic studies aim at characterising the microbial proteomes and at understanding how probiotics interact with the gastrointestinal tract.

Bacteria release a wide range of compounds into their environment in order to communicate and coordinate their activities. Recently, it was shown that the co-culture of two Bifidobacteria strains (*Bifidobacterium longum* and *Bifidobacterium breve*) induced changes in each bacteria's proteome[165]. Indeed, 2D gel analysis followed by LC-MS/MS analysis resulted in the identification of sixteen proteins, whose abundances were drastically changed when bifidobacteria were grown in co-culture compared to mono-culture. Differentially regulated proteins were grouped into ribosomal proteins and proteins involved in carbohydrate metabolism, gene regulation, cell envelope biogenesis as well as transport.

Another study investigated the surface-associated proteins from the probiotic *Lactobacillus plantarum*[166]. Cell surface proteins were migrated on one-dimensional gels and identified using LC-MS/MS. A total of twenty-nine proteins were identified and many of these proteins had previously been described of being capable to bind components of the human intestinal mucosa. In a related investigation, three different *L. plantarum* strains showing different adhesion rates were analysed using proteomics[167]. Several proteins, previously reported to be involved in bacterial adhesion, were found to be more abundant in the cell wall proteome of the most

highly-adhesive strain (elongation factor EF-tu (Elongation Factor Tu), GroEL (60 kDa chaperonin), DnaK (Chaperone protein Dna K) and glyceraldehyde-3-phosphate dehydrogenase). The association of proteomic profiles with particular probiotic properties opens the way for the selection of probiotics with specific, targeted benefits.

Allergens: protein and peptides as food-derived causes of hazard

Food allergies arise from the intake of allergenic food components, which can induce a response from the immune system and lead to clinical symptoms ranging from mild to life threatening[168]. The prevalence of food allergy is rising; indeed 2% of adults and 5–8% of children in industrialised countries are affected[169–171]. Over 180 protein allergens have been identified so far, the major ones occurring in common foods such as cow's milk, egg, peanut, soyabean, wheat, fish and tree nut[172]. Sensitive consumers have to be protected from undesirable allergic reactions and, therefore, proteomic methods have been developed for accurate allergen identification and quantification[130,172,173].

The classical proteomic strategy to identify food allergens consists of separating food proteins on 2D-PAGE, followed by electro-transfer onto a nitrocellulose membrane and subsequent IgE reactive protein detection by IgE immunoblotting using sera from allergic patients. This method was used to study allergens in wheat[174,175], apple[176], maize[177] or sesame seeds[178]. A systematic proteomic analysis of rice (*Oryza sativa*) leaf, root and seed using 2D gels followed by MS/MS allowed for the detection and identification of more than 2500 proteins[179] including several previously characterised allergenic proteins. The 2D difference gel electrophoresis method was also used to study several peanut varieties in order to show their low content of major allergens[180]. Recently, a method based on spectral counting was developed and successfully applied to the analysis of transgenic peanut lines containing reduced levels of certain major allergens[181].

Biomarkers: proteins and peptides as indicators of health and disease

Biomarkers are measurable indicators of different stages in a biological process, ranging from healthy functioning via deviation from such healthy equilibrium to disease onset and development[182]. Proteins and peptides, which are the main effectors in the body, can be used as such biomarkers. In practice, biomarkers are used for diagnostics, for prognostics, and to measure bioefficacy of nutrients in an intervention study.

Intestinal health and disease: Inflammatory bowel disease as an example. Inflammatory bowel diseases (IBD), including ulcerative colitis and Crohn's disease, are chronic, heterogeneous and multi-factorial inflammatory disorders of the gastrointestinal tract[182,183]. Proteomic investigations of the intestinal tissue of patients *v.* controls have the potential to deliver insights into gut dysfunction and may provide disease biomarkers.

A study of protein expression in intestinal epithelial cells led to the characterisation of changes in protein profiles of patients with Crohn's disease or ulcerative colitis compared to controls[184]. 2D-PAGE followed by MALDI-time-of-flight protein identification delivered the identification of nine proteins significantly different in IBD patients (e.g. Rho GDI alpha (Rho GDP-dissociation inhibitor 1), L-lactate dehydrogenase A, etc.). In ulcerative colitis patients, forty differentially expressed proteins were identified among which thirteen were associated with energy metabolism, which is in line with chronic intestinal inflammation being characterised by energy deficiency and alteration of the oxidative metabolism of epithelial cells[185,186].

A related study with 120 serum samples collected from four patients groups (Crohn's disease, ulcerative colitis, inflammatory controls, healthy controls) was performed to identify serum IBD biomarkers[187]: four new serum biomarkers were identified, namely PF4 (platelet factor 4), MRP8 (migration inhibitory factor-related protein 8), FIBA (fibrinogen alpha chain) and Hp alpha 2 (haptaglobin alpha 2). Another study focused on stool analysis: apart from S100A8 and S100A9 already associated with IBD, S100A12 was identified a possible new IBD marker[188].

However, the aetiology of IBD is still poorly characterised. A major symptom of IBD comprises malnutrition since inflammation of the gastrointestinal tract perturbs normal food intake and nutrient absorption[189,190]. The mechanisms involved in malnutrition include decreased food intake, malabsorption, increased nutrient loss, increased energy requirements and drug–nutrient interactions. In consequence, nutrition aspects play important roles in IBD and complement drug treatment. Adequate food intake is important for treatment, remission, remission maintenance, relapse prevention and prevention of IBD.

Conclusions and outlook

By balancing their diet, consumers want to optimise some health aspects without compromising others. Holistic and integrative approaches are therefore primordial. Proteomics is a central platform in nutrigenomics, which attempts to holistically understand how our genome is expressed as a response to diet. From a molecular perspective, nutritional proteomics covers two dimensions: characterisation of food proteins and peptides; biomarker as well as bioactive discovery and quantification. Nutritional proteomic biomarkers must be interconnected with other genomic and genetic markers: nutrigenetics investigates our genetic predisposition and susceptibility towards diet; epigenetics encompasses DNA sequence-unrelated biochemical modifications of both DNA itself and DNA-binding proteins and appears to provide a format for metabolic programming. Proteomics plays a key role here, too, as it can address PTM (e.g. acetylation) of DNA-packaging proteins and thereby help decipher the so-termed histone code.

Nutrition is still an expanding field for proteomics compared to well-established clinical and medical applications. The success of proteomics in nutrition and health will depend on multiple factors. The proteomic technology

per se will benefit from ever improving protein/peptide separation, depletion and enrichment on the one hand and more sensitive and specific mass spectrometers on the other hand. The second area of platform-related improvements is bioinformatics with rapidly improving tools to assess data quality and to convert data into interpretable information. The third room for improvement concerns the analytical strategy: focusing on proteome subsets – be it at the level of cell organelles, protein subclasses, the mass spectral level (targeted proteomics, gas-phase fractionation) – will provide deeper insights into molecular networks.

Apart from this expected progress at platform level, the technology will increasingly benefit from its cross-correlation with gene expression analysis and metabolite profiling. An option of addressing the interrelated timing of gene and protein expression is the investigation of protein turnover at proteomic scale but single-protein resolution, i.e. interpreting protein abundance changes as a result of both protein synthesis and degradation rather than taking proteomic snapshots.

In a nutshell, proteomics in nutrition delivers both bio-markers and bioactives. In this sense, proteomics will continue to drive (nutritional) systems biology, as it not only can identify and quantify the 'molecular robots' that do all the work in biological systems but also can map the networks of their physical interactions, between each other and with DNA, nutrients, drugs and other small molecules.

Acknowledgements

References

1. Young VR (1994) Adult amino acid requirements: The case for a major revision in current recommendations. *J Nutr* **124**, 1517S–1523S.
2. Desai B (2000) Proteins. In *Handbook of Nutrition and Diet*, pp. 1–816 [M Dekker, editor]. New York: CRC Press.
3. Lescuyer P, Hochstrasser D & Rabilloud T (2007) How shall we use the proteomics toolbox for biomarker discovery? *J Proteome Res* **6**, 3371–3376.
4. Schrattenholz A & Groebe K (2007) What does it need to be a biomarker? Relationships between resolution, differential quantification and statistical validation of protein surrogate biomarkers. *Electrophoresis* **28**, 1970–1979.
5. Mischak H, Coon JJ, Novak J *et al.* (2009) Capillary electrophoresis-mass spectrometry as a powerful tool in biomarker discovery and clinical diagnosis: An update of recent developments. *Mass Spectrom Rev* **28**, 703–724.
6. Marko-Varga G, Lindberg H, Lofdahl CG *et al.* (2005) Discovery of biomarker candidates within disease by protein profiling: Principles and concepts. *J Proteome Res* **4**, 1200–1212.
7. Vitzthum F, Behrens F, Anderson NL *et al.* (2005) Proteomics: From basic research to diagnostic application. A review of requirements and needs. *J Proteome Res* **4**, 1086–1097.
8. Gingras AC, Gstaiger M, Raught B *et al.* (2007) Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* **8**, 645–654.
9. Scholten A, van Veen TA, Vos MA *et al.* (2007) Diversity of cAMP-dependent protein kinase isoforms and their anchoring proteins in mouse ventricular tissue. *J Proteome Res* **6**, 1705–1717.
10. Wang J, Li D, Dangott LJ *et al.* (2006) Proteomics and its role in nutrition research. *J Nutr* **136**, 1759–1762.
11. Aebersold R & Mann M (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198–207.
12. Fenn JB, Mann M, Meng CK *et al.* (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71.
13. Karas M & Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **60**, 2299–2301.
14. Jacobs JM, Adkins JN, Qian WJ *et al.* (2005) Utilizing human blood plasma for proteomic biomarker discovery. *J Proteome Res* **4**, 1073–1085.
15. Gong Y, Li X, Yang B *et al.* (2006) Different immunoaffinity fractionation strategies to characterize the human plasma proteome. *J Proteome Res* **5**, 1379–1387.
16. Mechref Y, Madera M & Novotny MV (2008) Glycoprotein enrichment through lectin affinity techniques. *Methods Mol Biol* **424**, 373–396.
17. Thingholm TE, Jensen ON & Larsen MR (2009) Analytical strategies for phosphoproteomics. *Proteomics* **9**, 1451–1468.
18. Wollscheid B, Bausch-Fluck D, Henderson C *et al.* (2009) Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol* **27**, 378–386.
19. Eschelbach JW & Jorgenson JW (2006) Improved protein recovery in reversed-phase liquid chromatography by the use of ultrahigh pressures. *Anal Chem* **78**, 1697–1706.
20. Washburn MP, Wolters D & Yates JR, 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**, 242–247.
21. Blonder J, Rodriguez-Galan MC, Lucas DA *et al.* (2004) Proteomic investigation of natural killer cell microsomes using gas-phase fractionation by mass spectrometry. *Biochim Biophys Acta* **1698**, 87–95.
22. Scherl A, Shaffer SA, Taylor GK *et al.* (2008) Genome-specific gas-phase fractionation strategy for improved shotgun proteomic profiling of proteotypic peptides. *Anal Chem* **80**, 1182–1191.
23. Yi EC, Marelli M, Lee H *et al.* (2002) Approaching complete peroxisome characterization by gas-phase fractionation. *Electrophoresis* **23**, 3205–3216.
24. Fenn JB (2003) Electrospray wings for molecular elephants (Nobel lecture). *Angew Chem Int Ed Engl* **42**, 3871–3894.
25. Tanaka K (2003) The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew Chem Int Ed Engl* **42**, 3860–3870.
26. Makarov A, Denisov E, Kholomeev A *et al.* (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem* **78**, 2113–2120.
27. Nielsen ML, Savitski MM & Zubarev RA (2005) Improving protein identification using complementary fragmentation techniques in Fourier transform mass spectrometry. *Mol Cell Proteomics* **4**, 835–845.
28. Panchaud A, Scherl A, Shaffer SA *et al.* (2009) Precursor acquisition independent from ion count: How to dive deeper into the proteomics ocean. *Anal Chem* **81**, 6481–6488.
29. Deutsch EW, Lam H & Aebersold R (2008) Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. *Physiol Genomics* **33**, 18–25.

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

31. Perkins DN, Pappin DJ, Creasy DM *et al.* (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567.

32. Keller A, Nesvizhskii AI, Kolker E *et al.* (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383–5392.

33. Nesvizhskii AI, Keller A, Kolker E *et al.* (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **75**, 4646–4658.

34. Urfer W, Grzegorczyk M & Jung K (2006) Statistics for proteomics: A review of tools for analyzing experimental data. *Proteomics* **6** Suppl. 2, 48–55.

35. Elias JE & Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4**, 207–214.

36. Unlu M, Morgan ME & Minden JS (1997) Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071–2077.

37. Gygi SP, Rist B, Gerber SA *et al.* (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994–999.

38. Ross AR, Ambrose SJ, Cutler AJ *et al.* (2004) Determination of endogenous and supplied deuterated abscisic acid in plant tissues by high-performance liquid chromatography-electrospray ionization tandem mass spectrometry with multiple reaction monitoring. *Anal Biochem* **329**, 324–333.

39. Schmidt A, Kellermann J & Lottspeich F (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**, 4–15.

40. Panchaud A, Hansson J, Affolter M *et al.* (2008) ANIBAL, stable isotope-based quantitative proteomics by aniline and benzoic acid labeling of amino and carboxylic groups. *Mol Cell Proteomics* **7**, 800–812.

41. Ong SE, Blagoev B, Kratchmarova I *et al.* (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376–386.

42. Zhang H, Li XJ, Martin DB *et al.* (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* **21**, 660–666.

43. Julka S & Regnier F (2004) Quantification in proteomics through stable isotope coding: A review. *J Proteome Res* **3**, 350–363.

44. Oda Y, Owa T, Sato T *et al.* (2003) Quantitative chemical proteomics for identifying candidate drug targets. *Anal Chem* **75**, 2159–2165.

45. Qiu Y, Sousa EA, Hewick RM *et al.* (2002) Acid-labile isotope-coded extractants: A class of reagents for quantitative mass spectrometric analysis of complex protein mixtures. *Anal Chem* **74**, 4969–4979.

46. Sechi S (2002) A method to identify and simultaneously determine the relative quantities of proteins isolated by gel electrophoresis. *Rapid Commun Mass Spectrom* **16**, 1416–1424.

47. Gehanne S, Cecconi D, Carboni L *et al.* (2002) Quantitative analysis of two-dimensional gel-separated proteins using isotopically marked alkylating agents and matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom* **16**, 1692–1698.

48. Shen M, Guo L, Wallace A *et al.* (2003) Isolation and isotope labeling of cysteine- and methionine-containing tryptic peptides: Application to the study of cell surface proteolysis. *Mol Cell Proteomics* **2**, 315–324.

49. Kuyama H, Watanabe M, Toda C *et al.* (2003) An approach to quantitative proteome analysis by labeling tryptophan residues. *Rapid Commun Mass Spectrom* **17**, 1642–1650.

50. Panchaud A, Guillaume E, Affolter M *et al.* (2006) Combining protein identification and quantification: C-terminal isotope-coded tagging using sulfanilic acid. *Rapid Commun Mass Spectrom* **20**, 1585–1594.

51. Yao X, Freas A, Ramirez J *et al.* (2001) Proteolytic 18O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Anal Chem* **73**, 2836–2842.

52. Kosaka T, Takazawa T & Nakamura T (2000) Identification and C-terminal characterization of proteins from two-dimensional polyacrylamide gels by a combination of isotopic labeling and nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* **72**, 1179–1185.

53. Reynolds KJ, Yao X & Fenselau C (2002) Proteolytic 18O labeling for comparative proteomics: Evaluation of endoprotease Glu-C as the catalytic agent. *J Proteome Res* **1**, 27–33.

54. Goodlett DR, Keller A, Watts JD *et al.* (2001) Differential stable isotope labeling of peptides for quantitation and *de novo* sequence derivation. *Rapid Commun Mass Spectrom* **15**, 1214–1221.

55. Peters EC, Horn DM, Tully DC *et al.* (2001) A novel multifunctional labeling reagent for enhanced protein characterization with mass spectrometry. *Rapid Commun Mass Spectrom* **15**, 2387–2392.

56. Beardsley RL, Karty JA & Reilly JP (2000) Enhancing the intensities of lysine-terminated tryptic peptide ions in matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom* **14**, 2147–2153.

57. Karty JA, Ireland MM, Brun YV *et al.* (2002) Defining absolute confidence limits in the identification of Caulobacter proteins by peptide mass mapping. *J Proteome Res* **1**, 325–335.

58. Cagney G & Emili A (2002) De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat Biotechnol* **20**, 163–170.

59. Guillaume E, Panchaud A, Affolter M *et al.* (2006) Differentially isotope-coded N-terminal protein sulphonation: Combining protein identification and quantification. *Proteomics* **6**, 2338–2349.

60. Chen X, Chen YH & Anderson VE (1999) Protein cross-links: Universal isolation and characterization by isotopic derivatization and electrospray ionization mass spectrometry. *Anal Biochem* **273**, 192–203.

61. Mason DE & Liebler DC (2003) Quantitative analysis of modified proteins by LC-MS/MS of peptides labeled with phenyl isocyanate. *J Proteome Res* **2**, 265–272.

62. Zhang R & Regnier FE (2002) Minimizing resolution of isotopically coded peptides in comparative proteomics. *J Proteome Res* **1**, 139–147.

63. Zhang R, Sioma CS, Thompson RA *et al.* (2002) Controlling deuterium isotope effects in comparative proteomics. *Anal Chem* **74**, 3662–3669.

64. Chelius D & Bondarenko PV (2002) Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* **1**, 317–323.

65. Mueller LN, Rinner O, Schmidt A *et al.* (2007) SuperHirn – a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* **7**, 3470–3480.

66. Liu H, Sadygov RG & Yates JR, 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* **76**, 4193–4201.

67. Bantscheff M, Schirle M, Sweetman G *et al.* (2007) Quantitative mass spectrometry in proteomics: A critical review. *Anal Bioanal Chem* **389**, 1017–1031.

68. Gerber SA, Rush J, Stemman O *et al.* (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci USA* **100**, 6940–6945.

69. Rivers J, Simpson DM, Robertson DH *et al.* (2007) Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol Cell Proteomics* **6**, 1416–1427.

70. Lange V, Picotti P, Domon B *et al.* (2008) Selected reaction monitoring for quantitative proteomics: A tutorial. *Mol Syst Biol* **4**, 222.

71. Nita-Lazar A, Saito-Benz H & White FM (2008) Quantitative phosphoproteomics by mass spectrometry: Past, present, and future. *Proteomics* **8**, 4433–4443.

72. Pandey A, Andersen JS & Mann M (2000) Use of mass spectrometry to study signaling pathways. *Sci STKE* **2000**, pl1.

73. Andersson L & Porath J (1986) Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal Biochem* **154**, 250–254.

74. Beausoleil SA, Jedrychowski M, Schwartz D *et al.* (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci USA* **101**, 12130–12135.

75. Zubarev RA, Horn DM, Fridriksson EK *et al.* (2000) Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal Chem* **72**, 563–573.

76. Syka JE, Coon JJ, Schroeder MJ *et al.* (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci USA* **101**, 9528–9533.

77. Wiesner J, Premsler T & Sickmann A (2008) Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. *Proteomics* **8**, 4466–4483.

78. Funakoshi Y & Suzuki T (2009) Glycobiology in the cytosol: The bitter side of a sweet world. *Biochim Biophys Acta* **1790**, 81–94.

79. Dell A & Morris HR (2001) Glycoprotein structure determination by mass spectrometry. *Science* **291**, 2351–2356.

80. Tissot B, North SJ, Ceroni A *et al.* (2009) Glycoproteomics: Past, present and future. *FEBS Lett* **583**, 1728–1735.

81. Yang Z & Hancock WS (2004) Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. *J Chromatogr A* **1053**, 79–88.

82. Allfrey VG, Faulkner R & Mirsky AE (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci USA* **51**, 786–794.

83. Grant PA (2001) A tale of histone modifications. *Genome Biol* **2**, REVIEWS0003.

84. Taverna SD, Li H, Ruthenburg AJ *et al.* (2007) How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**, 1025–1040.

85. Verreault A, Kaufman PD, Kobayashi R *et al.* (1998) Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr Biol* **8**, 96–108.

86. Strahl BD & Allis CD (2000) The language of covalent histone modifications. *Nature* **403**, 41–45.

87. Bonenfant D, Coulot M, Towbin H *et al.* (2006) Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. *Mol Cell Proteomics* **5**, 541–552.

88. Mukherjee S, Hao YH & Orth K (2007) A newly discovered post-translational modification – the acetylation of serine and threonine residues. *Trends Biochem Sci* **32**, 210–216.

89. Witze ES, Old WM, Resing KA *et al.* (2007) Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* **4**, 798–806.

90. Kim SC, Sprung R, Chen Y *et al.* (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell* **23**, 607–618.

91. Barter PJ, Nicholls S, Rye KA *et al.* (2004) Antiinflammatory properties of HDL. *Circ Res* **95**, 764–772.

92. Laberge MA, Moore KJ & Freeman MW (2005) Atherosclerosis and innate immune signaling. *Ann Med* **37**, 130–140.

93. Shiflett AM, Bishop JR, Pahwa A *et al.* (2005) Human high density lipoproteins are platforms for the assembly of multicomponent innate immune complexes. *J Biol Chem* **280**, 32578–32585.

94. Havel RJ, Eder HA & Bragdon JH (1955) The distribution and chemical composition of ultracentrifugally separated lipoproteins in human serum. *J Clin Invest* **34**, 1345–1353.

95. Kostner GM (1983) Apolipoproteins and lipoproteins of human plasma: Significance in health and in disease. *Adv Lipid Res* **20**, 1–43.

96. Kunitake ST, Carilli CT, Lau K *et al.* (1994) Identification of proteins associated with apolipoprotein A-I-containing lipoproteins purified by selected-affinity immunosorption. *Biochemistry* **33**, 1988–1993.

97. Zechner R, Moser R & Kostner GM (1986) Isolation of pure LpB from human serum. *J Lipid Res* **27**, 681–686.

98. Folch J, Lees M & Sloane Stanley GH (1957) A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* **226**, 497–509.

99. Karlsson H, Leanderson P, Tagesson C *et al.* (2005) Lipoproteomics I: Mapping of proteins in low-density lipoprotein using two-dimensional gel electrophoresis and mass spectrometry. *Proteomics* **5**, 551–565.

100. Farwig ZN, Campbell AV & Macfarlane RD (2003) Analysis of high-density lipoprotein apolipoproteins recovered from specific immobilized pH gradient gel pI domains by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal Chem* **75**, 3823–3830.

101. Ogorzalek Loo RR, Yam L, Loo JA *et al.* (2004) Virtual two-dimensional gel electrophoresis of high-density lipoproteins. *Electrophoresis* **25**, 2384–2391.

102. Rezaee F, Casetta B, Levels JH *et al.* (2006) Proteomic analysis of high-density lipoprotein. *Proteomics* **6**, 721–730.

103. Stahlman M, Davidsson P, Kanmert I *et al.* (2008) Proteomics and lipids of lipoproteins isolated at low salt concentrations in D_2O/sucrose or in KBr. *J Lipid Res* **49**, 481–490.

104. Mancone C, Amicone L, Fimia GM *et al.* (2007) Proteomic analysis of human very low-density lipoprotein by two-dimensional gel electrophoresis and MALDI-TOF/TOF. *Proteomics* **7**, 143–154.

105. Vaisar T, Pennathur S, Green PS *et al.* (2007) Shotgun proteomics implicates protease inhibition and complement activation in the anti-inflammatory properties of HDL. *J Clin Invest* **117**, 746–756.

106. Bern M, Finney G, Hoopmann MR *et al.* (2010) Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Anal Chem* **82**, 833–841.

107. Chaurand P & Caprioli RM (2002) Direct profiling and imaging of peptides and proteins from mammalian cells and tissue sections by mass spectrometry. *Electrophoresis* **23**, 3125–3135.

108. Seeley EH & Caprioli RM (2008) Molecular imaging of proteins in tissues by mass spectrometry. *Proc Natl Acad Sci USA* **105**, 18126–18131.

109. Stoeckli M, Chaurand P, Hallahan DE *et al*. (2001) Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. *Nat Med* **7**, 493–496.

110. Franck J, Arafah K, Elayed M *et al*. (2009) MALDI imaging mass spectrometry: State of the art technology in clinical proteomics. *Mol Cell Proteomics* **8**, 2023–2033.

111. Samsi SS, Krishnamurthy AK, Groseclose M *et al*. (2009) Imaging mass spectrometry analysis for follicular lymphoma grading. *Conf Proc IEEE Eng Med Biol Soc* **1**, 6969–6972.

112. Schwamborn K, Krieg RC, Reska M *et al*. (2007) Identifying prostate carcinoma by MALDI-Imaging. *Int J Mol Med* **20**, 155–159.

113. Wong SC, Chan CM, Ma BB *et al*. (2009) Advanced proteomic technologies for cancer biomarker discovery. *Expert Rev Proteomics* **6**, 123–134.

114. Groseclose MR, Andersson M, Hardesty WM *et al*. (2007) Identification of proteins directly from tissue: *In situ* tryptic digestions coupled with imaging mass spectrometry. *J Mass Spectrom* **42**, 254–262.

115. Lemaire R, Desmons A, Tabet JC *et al*. (2007) Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J Proteome Res* **6**, 1295–1305.

116. Chaurand P, Schwartz SA & Capriolo RM (2004) Profiling and imaging proteins in tissue sections by MS. *Anal Chem* **76**, 87A–93A.

117. Chaurand P, Schwartz SA, Reyzer ML *et al*. (2005) Imaging mass spectrometry: Principles and potentials. *Toxicol Pathol* **33**, 92–101.

118. Naik AK, Hanay MS, Hiebert WK *et al*. (2009) Towards single-molecule nanomechanical mass spectrometry. *Nat Nanotechnol* **4**, 445–450.

119. Beck M, Malmstrom JA, Lange V *et al*. (2009) Visual proteomics of the human pathogen Leptospira interrogans. *Nat Methods* **6**, 817–823.

120. Zhang Q, Faca V & Hanash S (2011) Mining the plasma proteome for disease applications across seven logs of protein abundance. *J Proteome Res* **10**, 46–50.

121. Domon B & Aebersold R (2010) Options and considerations when selecting a quantitative proteomics strategy. *Nat Biotechnol* **28**, 710–721.

122. Minkiewicz P, Dziuba J, Darewicz M *et al*. (2008) Food peptidomics. *Food Technol Biotechnol* **46**, 1–10.

123. Kussmann M, Panchaud A & Affolter M (2010) Proteomics in nutrition: Status quo and outlook for biomarkers and bioactives. *J Proteome Res* **9**, 4876–4887.

124. de Roos B & McArdle HJ (2008) Proteomics as a tool for the modelling of biological processes and biomarker development in nutrition research. *Br J Nutr* **99** Suppl. 3, S66–S71.

125. Fuchs D, Winkelmann I, Johnson IT *et al*. (2005) Proteomics in nutrition research: Principles, technologies and applications. *Br J Nutr* **94**, 302–314.

126. Kussmann M & Affolter M (2009) Proteomics at the center of nutrigenomics: Comprehensive molecular understanding of dietary health effects. *Nutrition* **25**, 1085–1093.

127. Schweigert FJ (2007) Nutritional proteomics: methods and concepts for research in nutritional science. *Ann Nutr Metab* **51**, 99–107.

128. Mamone G, Picariello G, Caira S *et al*. (2009) Analysis of food proteins and peptides by mass spectrometry-based techniques. *J Chromatogr A* **1216**, 7130–7142.

129. Alex P, Gucek M & Li X (2009) Applications of proteomics in the study of inflammatory bowel diseases: Current status and future directions with available technologies. *Inflamm Bowel Dis* **15**, 616–629.

130. Kirsch S, Fourdrilis S, Dobson R *et al*. (2009) Quantitative methods for food allergens: A review. *Anal Bioanal Chem* **395**, 57–67.

131. Sundsten T & Ortsater H (2009) Proteomics in diabetes research. *Mol Cell Endocrinol* **297**, 93–103.

132. Fuchs D, Dirscherl B, Schroot JH *et al*. (2007) Proteome analysis suggests that mitochondrial dysfunction in stressed endothelial cells is reversed by a soy extract and isolated isoflavones. *J Proteome Res* **6**, 2132–2142.

133. Tang JE, Moore DR, Kujbida GW *et al*. (2009) Ingestion of whey hydrolysate, casein, or soy protein isolate: Effects on mixed muscle protein synthesis at rest and following resistance exercise in young men. *J Appl Physiol* **107**, 987–992.

134. Erdmann K, Cheung BW & Schroder H (2008) The possible roles of food-derived bioactive peptides in reducing the risk of cardiovascular disease. *J Nutr Biochem* **19**, 643–654.

135. German JB, Freeman SL, Lebrilla CB *et al*. (2008) Human milk oligosaccharides: Evolution, structures and bioselectivity as substrates for intestinal bacteria. *Nestle Nutr Workshop Ser Pediatr Program* **62**, 205–218; discussion 218–222.

136. Casado B, Affolter M & Kussmann M (2009) OMICS-rooted studies of milk proteins, oligosaccharides and lipids. *J Proteomics* **73**, 196–208.

137. Conti A, Giuffrida MG & Cavaletto M (2007) Proteomics of human milk. In *Proteomics of Human Body Fluids*, pp. 437–451 [V Thongboonkerd, editor]. Totowa: Humana Press.

138. Fox PF (2003) Milk proteins: General and historical aspects. In *Advanced Dairy Chemistry: Proteins*, 3rd ed., pp. 1–48 [PF Fox & PLH McSweeney, editors]. New York: Kluwer Academic/Plenum Publishers.

139. Kanwar JR, Kanwar RK, Sun X *et al*. (2009) Molecular and biotechnological advances in milk proteins in relation to human health. *Curr Protein Pept Sci* **10**, 308–338.

140. Mange A, Bellet V, Tuaillon E *et al*. (2008) Comprehensive proteomic analysis of the human milk proteome: Contribution of protein fractionation. *J Chromatogr B Analyt Technol Biomed Life Sci* **876**, 252–256.

141. Tremblay L, Laporte M, Leonil J *et al*. (2003) Quantitation of proteins in milk and milk products. In *Advanced Dairy Chemistry: Proteins*, 3rd ed., pp. 49–138 [PF Fox & PL McSweeney, editors]. New York: Kluwer Academic/Plenum Publishers.

142. Fong YP, Norris CS & Palmano KP (2008) Fractionation of bovine whey proteins and characterisation by proteomic techniques. *Int Dairy J* **18**, 23–46.

143. Molle D, Jardin J, Piot M *et al*. (2009) Comparison of electrospray and matrix-assisted laser desorption ionization on the same hybrid quadrupole time-of-flight tandem mass spectrometer: Application to bidimensional liquid chromatography of proteins from bovine milk fraction. *J Chromatogr A* **1216**, 2424–2432.

144. Affolter M, Grass L, Vanrobaeys F *et al*. (2010) Qualitative and quantitative profiling of the bovine milk fat globule membrane proteome. *J Proteomics* **73**, 1079–1088.

145. Feligini M, Bonizzi I, Buffoni JN *et al*. (2009) Identification and quantification of alphaS1, alphaS2, beta, and kappa-caseins in water buffalo milk by reverse phase-high

performance liquid chromatography and mass spectrometry. *J Agric Food Chem* **57**, 2988–2992.

146. Kjeldsen F, Haselmann KF, Budnik BA *et al.* (2003) Complete characterization of posttranslational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage. *Anal Chem* **75**, 2355–2361.

147. Mateos A, Girardet JM, Molle D *et al.* (2009) Two-dimensional cartography of equine beta-casein variants achieved by isolation of phosphorylation isoforms and control of the deamidation phenomenon. *J Dairy Sci* **92**, 2389–2399.

148. Roncada P, Gaviraghi A, Liberatori S *et al.* (2002) Identification of caseins in goat milk. *Proteomics* **2**, 723–726.

149. D'Ambrosio C, Arena S, Salzano AM *et al.* (2008) A proteomic characterization of water buffalo milk fractions describing PTM of major species and the identification of minor components involved in nutrient delivery and defense against pathogens. *Proteomics* **8**, 3657–3666.

150. Harvey DJ (2005) Proteomic analysis of glycosylation: Structural determination of N- and O-linked glycans by mass spectrometry. *Expert Rev Proteomics* **2**, 87–101.

151. Park Y & Lebrilla CB (2005) Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides. *Mass Spectrom Rev* **24**, 232–264.

152. Hamosh M (1998) Protective function of proteins and lipids in human milk. *Biol Neonate* **74**, 163–176.

153. Newburg DS (1999) Human milk glycoconjugates that inhibit pathogens. *Curr Med Chem* **6**, 117–127.

154. Newburg DS & Walker WA (2007) Protection of the neonate by the innate immune system of developing gut and of human milk. *Pediatr Res* **61**, 2–8.

155. Hamosh M, Peterson JA, Henderson TR *et al.* (1999) Protective function of human milk: The milk fat globule. *Semin Perinatol* **23**, 242–249.

156. Peterson JA, Patton S & Hamosh M (1998) Glycoproteins of the human milk fat globule in the protection of the breast-fed infant against infections. *Biol Neonate* **74**, 143–162.

157. Peterson JA, Scallan CD, Ceriani RL *et al.* (2001) Structural and functional aspects of three major glycoproteins of the human milk fat globule membrane. *Adv Exp Med Biol* **501**, 179–187.

158. Picariello G, Ferranti P, Mamone G *et al.* (2008) Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* **8**, 3833–3847.

159. Farnaud S & Evans RW (2003) Lactoferrin – a multifunctional protein with antimicrobial properties. *Mol Immunol* **40**, 395–405.

160. Crouch SP, Slater KJ & Fletcher J (1992) Regulation of cytokine release from mononuclear cells by the iron-binding protein lactoferrin. *Blood* **80**, 235–240.

161. Masuda O, Nakamura Y & Takano T (1996) Antihypertensive peptides are present in aorta after oral administration of sour milk containing these peptides to spontaneously hypertensive rats. *J Nutr* **126**, 3063–3068.

162. World Health Organization (2001) *Health and Nutritional Properties of Probiotics in Food including Powder Milk with Live Lactic Acid Bacteria. Joint WHO/FAO Expert Consultation*. Geneva: WHO.

163. Ljungh A & Wadstrom T (2006) Lactic acid bacteria as probiotics. *Curr Issues Intest Microbiol* **7**, 73–89.

164. Salminen SJ, Gueimonde M & Isolauri E (2005) Probiotics that modify disease risk. *J Nutr* **135**, 1294–1298.

165. Ruiz L, Sanchez B, de Los Reyes-Gavilan CG *et al.* (2009) Coculture of *Bifidobacterium longum* and *Bifidobacterium breve* alters their protein expression profiles and enzymatic activities. *Int J Food Microbiol* **133**, 148–153.

166. Beck HC, Madsen SM, Glenting J *et al.* (2009) Proteomic analysis of cell surface-associated proteins from probiotic *Lactobacillus plantarum*. *FEMS Microbiol Lett* **297**, 61–66.

167. Izquierdo E, Horvatovich P, Marchioni E *et al.* (2009) 2-DE and MS analysis of key proteins in the adhesion of *Lactobacillus plantarum*, a first step toward early selection of probiotics based on bacterial biomarkers. *Electrophoresis* **30**, 949–956.

168. Bush RK & Hefle SL (1996) Food allergens. *Crit Rev Food Sci Nutr* **36**, Suppl., S119–S163.

169. Burks W, Kulis M & Pons L (2008) Food allergies and hypersensitivity: A review of pharmacotherapy and therapeutic strategies. *Expert Opin Pharmacother* **9**, 1145–1152.

170. Kagan RS (2003) Food allergy: An overview. *Environ Health Perspect* **111**, 223–225.

171. Ortolani C, Ispano M, Scibilia J *et al.* (2001) Introducing chemists to food allergy. *Allergy* **56**, Suppl., **67**, 5–8.

172. Poms RE, Klein CL & Anklam E (2004) Methods for allergen analysis in food: A review. *Food Addit Contam* **21**, 1–31.

173. Monaci L & Visconti A (2009) Mass spectrometry-based proteomics methods for analysis of food allergens. *Trends Anal. Chem.* **28**, 581–591.

174. Akagawa M, Handoyo T, Ishii T *et al.* (2007) Proteomic analysis of wheat flour allergens. *J Agric Food Chem* **55**, 6863–6870.

175. Sotkovsky P, Hubalek M, Hernychova L *et al.* (2008) Proteomic analysis of wheat proteins recognized by IgE antibodies of allergic patients. *Proteomics* **8**, 1677–1691.

176. Guarino C, Arena S, De Simone L *et al.* (2007) Proteomic analysis of the major soluble components in Annurca apple flesh. *Mol Nutr Food Res* **51**, 255–262.

177. Fasoli E, Pastorello EA, Farioli L *et al.* (2009) Searching for allergens in maize kernels via proteomic tools. *J Proteomics* **72**, 501–510.

178. Beyer K, Bardina L, Grishina G *et al.* (2002) Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: Seed storage proteins as common food allergens. *J Allergy Clin Immunol* **110**, 154–159.

179. Koller A, Washburn MP, Lange BM *et al.* (2002) Proteomic survey of metabolic pathways in rice. *Proc Natl Acad Sci USA* **99**, 11969–11974.

180. Schmidt H, Gelhaus C, Latendorf T *et al.* (2009) 2-D DIGE analysis of the proteome of extracts from peanut variants reveals striking differences in major allergen contents. *Proteomics* **9**, 3507–3521.

181. Stevenson SE, Chu Y, Ozias-Akins P *et al.* (2009) Validation of gel-free, label-free quantitative proteomics approaches: Applications for seed allergen profiling. *J Proteomics* **72**, 555–566.

182. De Roos B (2009) Nutrition proteomics and biomarker discovery. *Expert Rev Proteomics* **6**, 349–351.

183. Podolsky DK (2002) Inflammatory bowel disease. *N Engl J Med* **347**, 417–429.

184. Shkoda A, Werner T, Daniel H *et al.* (2007) Differential protein expression profile in the intestinal epithelium from patients with inflammatory bowel disease. *J Proteome Res* **6**, 1114–1125.

185. Fukushima K & Fiocchi C (2004) Paradoxical decrease of mitochondrial DNA deletions in epithelial cells of active ulcerative colitis patients. *Am J Physiol Gastrointest Liver Physiol* **286**, G804–G813.

186. Roediger WE (1980) The colonic epithelium in ulcerative colitis: An energy-deficiency disease? *Lancet* **2**, 712–715.

187. Meuwis MA, Fillet M, Geurts P et al. (2007) Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling. Biochem Pharmacol 73, 1422–1433.

188. Foell D, Wittkowski H, Ren Z et al. (2008) Phagocyte-specific S100 proteins are released from affected mucosa and promote immune responses during inflammatory bowel disease. J Pathol 216, 183–192.

189. Shamir R (2009) Nutritional aspects in inflammatory bowel disease. J Pediatr Gastroenterol Nutr 48, Suppl. 2, S86–S88.

190. Vagianos K, Bector S, McConnell J et al. (2007) Nutrition assessment of patients with inflammatory bowel disease. JPEN J Parenter Enteral Nutr 31, 311–319.

Proceedings of the Nutrition Society