

Review article

The potential of proteomics and peptidomics for allergy and asthma research

Progress in the field of proteomics, the branch of biology that studies the full set of proteins derived from a given genome, is moving fast. Two-dimensional gel electrophoresis (2DG) separation of complex protein mixtures and the subsequent analysis of isolated protein spots by mass spectrometry allow fast and accurate identification of proteins. The comparison of spots from different samples separated on customized 2D gels allows the detection of punctual differences in their mobility and facilitates tracing back differences in protein expression, presence of isoforms, splice variants and posttranslational modifications by mass spectrometry. In spite of significant analytical challenges owing to the high complexity of the proteome and the challenge deriving from the necessity to process huge amounts of raw data generated by mass spectrometric profiling, proteomics has evolved to an indispensable tool in life sciences. A restricted window of the proteome that consists of peptides and small proteins not easily manageable by conventional gel electrophoresis prompted the development of separation methods based on liquid chromatography. This new research field termed peptidomics already contributed, together with proteomics to enlarge our knowledge about biological processes and supported by sophisticated bioinformatics tools, to the discovery of new diagnostic and therapeutic targets. The technological capabilities of biophysical separation, mass spectrometry and bioinformatics form the basis of discovery programs that aim at mining the proteome starting from microgram amounts of protein extracts derived from body fluids and tissues. Proteomics and peptidomics have a great potential to speed up allergy and asthma research, where disease- and tissue-specific samples are easy to obtain.

R. Cramer

Swiss Institute of Allergy and Asthma Research (SIAF), Davos, Switzerland

Key words: allergy; asthma; mass spectrometry; peptidomics; protein separation; proteomics.

Prof Reto Cramer
Head Molecular Allergology
Swiss Institute of Allergy and Asthma Research (SIAF)
Obere Strasse 22
CH-7270 Davos
Switzerland

Accepted for publication 3 February 2005

The human genome project (1–3) initiated a dramatic change of the research strategy applied in life sciences from the punctual investigation of defined projects to global approaches. The ‘omics’ projects aim at studying the full characteristics of a given biological sample in a single set of experiments. Despite the impact of genomics in molecular biology (4) and medicine (5), the information encoded by genomes does not represent more than a rough ground plan for an organism. This information might restrict but does not describe the dynamics of the life processes occurring in cells and organisms. Healthy as well as disease status are governed by complex processes, which start from the activation of a (limited) number of genes located in the genome either on the same chromosome or in different chromosome segments. Gene activation, translation, transcription and maturation of mRNAs, as well as posttrans-

lational modifications give rise to a multitude of mature proteins and peptides (6) with structures, which are not directly deducible from genome sequences. Although proteins and peptides are products of gene expression, there are much more different forms of functional proteins and peptides than genes encoding them. As a consequence of the tightly regulated steady state between synthesis, posttranslational modifications and degradation every cell, tissue and organ of an organism contains a complex pool of proteins, protein fragments and peptides adequately modified to fulfill their biological functions. This result in the equilibrium of different functional forms of a given protein in varying concentrations embedded into extremely complex interaction networks involving a multitude of other different proteins or peptides (Fig. 1). Because of the dynamic nature of gene expression, protein synthesis and posttranslational modifications, the identification and quantification of proteins alone is not sufficient to understand functional interactions. In this context it is important to realize that changes as small as the addition of a single phosphate, cleavage of a leader peptide, amidation or

Abbreviations: 2DG, two-dimensional gel electrophoresis; ESI, electrospray ionization; HPLC, high pressure liquid chromatography; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; pI, isoelectric point; TOF, time of flight.

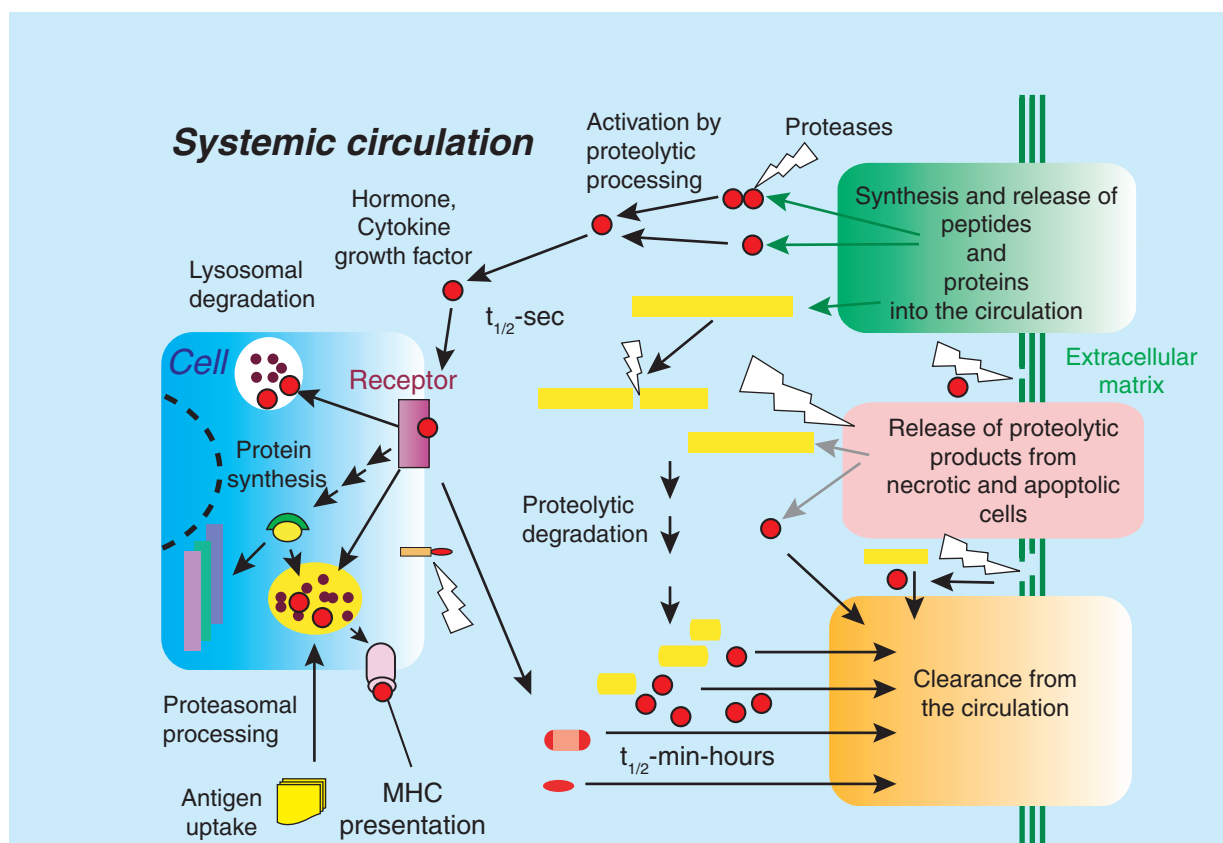


Figure 1. The complexity of proteome and peptidome. Proteins (■) and peptides (●) undergo several modifications after synthesis. The actual state of proteome and peptidome at the time of sampling represents a frozen snapshot of the equilibrium between synthesis, posttranslational modifications, degradation and clearance of each protein and peptide interacting in complex networks (modified from Ref. 123).

oxidation processes can drastically alter the biological function of a protein (7). Since the whole protein and peptide status of a biological sample at a given time point should be detected in a proteomic approach, sophisticated methods for sample preparation as well as for separation and exact molecular characterization of each single protein and peptide are required. However, each biological sample analyzed depicts only a static snapshot of the ‘frozen’ proteome and reflects the state of the system at the moment at which the sample is taken. Information about dynamic processes can only be reconstructed by analyzing arrays of single snapshots taken at different times by complex correlation-associated networks (8, 9). Mass spectrometry (MS) is currently making significant contributions to the understanding of the composition and structure of proteins and, in turn, to the elucidation of structure–function relationships. This sensitive technology allows the exact determination of molecular masses of proteins and peptides, direct verification of gene-derived sequences, determination of the number and localization of disulphide bonds and rapid structural elucidation of posttranslational modifications (10, 11). As sequence data of genes and proteins are rapidly accumulating in databases, it becomes

easier to generate hypotheses about the biological function and about the network of interactions of newly discovered proteins and peptides. Hierarchical cluster analysis (8, 12), relevance networks (13, 14) and correlation-associated networks (9, 15) based on statistical analysis of biomolecules with a common co-transcription, co-expression, co-secretion or co-clearance pathway facilitate the generation of robust predictions about putative protein interactions. The combination of comprehensive sequencing data derived from large proteome projects and sophisticated bioinformatics tools hold great promises for the fast discovery of diagnostic and therapeutic targets as well as for the understanding of the pathophysiological mechanisms involved in any disease. Because allergies and specially asthma, represent genetic disorders with a heterogeneous phenotype attributed to the interaction among many genes and environmental factors, proteomic and peptidomic experimental set-ups bear a great potential to provide us with a more detailed understanding of the molecular mechanisms involved in the development of these diseases. Traditionally, the application of emerging technologies in allergy and asthma research lag behind technological developments and this review aims at

motivating scientists to think about chances and pitfalls of proteomic and peptidomic applications to this field of research.

Conceptual background

The human genome deals 'only' with about 30 000 genes (1, 2) while the human proteome has to deal with several hundred thousands of proteins and peptides deriving from alternative splicing, posttranslational modifications and proteolytic processes (16–18). The word 'proteome' was first proposed in 1994 to describe the whole content of proteins present in a cell, tissue or body fluid at a given time (19). In analogy the term 'peptidome', coined in 2001, aims at describing the whole content of peptides present in a biological sample (20). Both, proteomic and peptidomic technologies, targeting different windows of the 'proteome', represent global strategies, where all proteins or peptides present in a biological sample derived from a cell, tissue or body liquid are simultaneously visualized and identified (20–23). These approaches do not depend on any experimental hypothesis but can, based on correlation-associated network analyses of proteomic or peptidomic profiles, generate hypotheses about relations between structurally and biologically related proteins (9). The spectrum of predicted functional relationships covering posttranslational modifications (24, 25), proteolytic processes (9), co-secretion of proteins and peptides (9) and protein–protein interactions (26, 27) can be easily verified by sequencing of the corresponding peak sized from the proteomic or peptidomic profile of the analyzed sample. This process is greatly facilitated by the availability of the whole sequence of the human genome, which allows the identification of a protein with minimal sequence information (28, 29). In contrast to the relatively high stability of the genome, the proteome is highly dynamic and reflects the current metabolic status of a given biological system. Although genomic analysis can provide important information about the susceptibility to acquire a certain disease, it is equally important to elucidate metabolic changes at proteome and peptidome level which might determine severity and prognosis of a disease in order to discover new diagnostic and therapeutic targets.

Technological background

In contrast to genomic information which is primarily stored in the four base linear sequence of easily isolable, biochemically homogeneous and amplifiable DNA, at least 21 amino acids determine the primary structure of a protein. The inherent physical properties of individual proteins including solubility, hydrophobicity, size and isoelectric point (pI) differ markedly. As a consequence thereof, the analysis of complex protein mixtures is more

laborious and less customizable than the analysis of DNA. To identify and characterize proteins and peptides from extremely complex mixtures of thousands of compounds with a wide dynamic concentration range the use of technologies with a high resolving power and extreme sensitivity is mandatory.

Protein separation techniques

The current standard method for the separation of complex proteomic samples is still 2D polyacrylamide gel electrophoresis (PAGE), a method developed more than two decades ago (30), which allows high resolution of single components (31, 32). The first dimension (isoelectric focusing) is carried out using gels with a fixed pH gradient to separate proteins according to their pI. Subsequently, the proteins are separated according to their relative masses in a second dimension by SDS-PAGE. After separation, protein spots are visualized, extracted from the gels using different methods (33, 34) and subjected to mass spectrometric analysis to determine their identity (35). However, sample preparation is still the most critical step in two-dimensional gel electrophoresis (2DG) and should be optimized for each type of sample. To obtain gene expression profiles, proteins have to be extracted from complex biological samples such as tissue biopsies, serum plasma or cerebrospinal fluid. One should be aware of the fact that possible protein degradation or modifications related to the extraction procedure will directly influence the proteome profile of the sample. Moreover, the presence of abundant proteins, like albumin in serum, which interfere with 2D-PAGE separation, will mask low abundance proteins (36, 37). Analyses of such complex samples remain challenging in view of the large number of proteins expected to be present and the wide dynamic range of protein concentrations known to span at least 11–12 orders of magnitude (38). These problems have strongly supported technological development aiming to improve the quality of sample preparations through enrichment of modified proteins (39, 40), prefractionation of body fluids by column chromatography (41, 42) and selective depletion of abundant proteins (43, 44). The molecular mass (apparent mass) in a proteomic experiment typically ranges from 10 to 200 kDa and the pI from pH 4 to pH 10. Staining the 2D gel already gives a visual image, which allows one to estimate the resolution quality, the number of spots and the protein quantity in individual spots (Fig. 2). The reproducibility of the first and second dimension in gel electrophoresis has improved dramatically allowing a good resolution and relatively high throughput (45–47). Quantification of stained spots is mainly achieved by spot integration after staining. Identification is performed either by recovering the proteins from the gel and conventional Edman sequence determination or by digestion of each protein 'in-gel', recovery of the fragments and precise

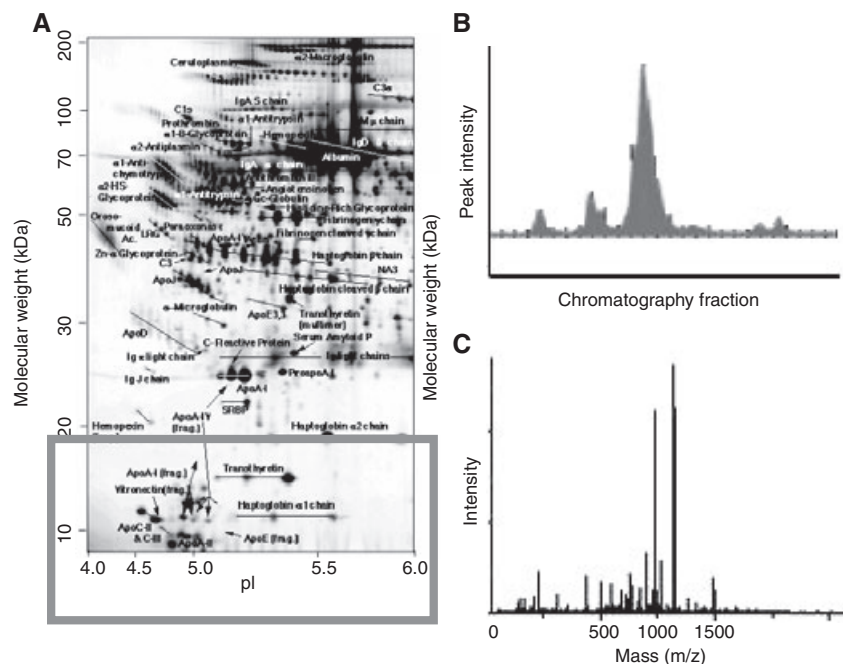


Figure 2. The 2D gel visualizing proteins in the ranges from about 10 to >200 kDa with different pI (A). In typical proteomic experiments isolated protein spots are subsequently analyzed by MS. The window marked in red corresponds to the molecular weight range typically analyzed in peptidomics experiments (1–20 kDa). Small protein and peptide samples are first separated into 96 fractions by HPLC (B) and each fraction is individually analyzed by MS (C) (modified from Ref. 124).

analysis of the fragment masses by MS (48, 49). Although still suffering from some practical limitations (50), 2DG-based proteome analysis technology has been shown to be a fantastic tool for the identification of novel biomarkers (51–53) and therapeutic targets (54–57), for the study of differential protein expression in healthy and disease status in different pathophysiological conditions (58, 59) up to the industrial scale analysis of human blood plasma (38).

Analysis of peptides from complex mixtures differs from commonly employed proteomics strategies. In contrast to large proteins, peptides can be separated with high resolution by column chromatography, preferably using high quality reverse phase material and analytical ion exchangers (60) and as a consequence thereof the complexity of peptide mixtures can be reduced (61). A definitive advantage of chromatographic separation methods over 2DG is that the analytes always remain in the liquid phase, minimizing loss of substances. Moreover, column chromatography is a scalable process covering a range of sub-microgram amounts in capillary chromatography to grams processed by preparative chromatography (62). However, body fluids such as plasma or serum contain high amounts of protein (*ca.* 70 mg/ml) compared to relatively low amounts of peptides ranging from 50 to 200 µg/ml (63). Obviously, the high protein content of these fluids interferes with chromatographic separations of the peptides in terms of resolution and loading capacity. Therefore a prior separation of peptides from larger proteins (> 25–30 kDa) is a

necessary step for efficient subsequent chromatographic fractionation and MS analysis of peptides (20). The fractionation of protein-depleted samples by high-pressure reversed phase liquid chromatography (RP-HPLC) into 96 fractions prior to mass spectrometry reduces the complexity of the sample and increases the resolution significantly (64, 65). The molecular mass window covered by peptidomics includes peptide hormones, neuropeptides, cytokines and many other circulating peptides representing messenger molecules involved in the control of thousands of biological and regulatory processes (66) not covered by classical proteomic approaches. In the field of allergy and asthma where cytokines and chemokines (67, 68), small (peptide) mediators released from effector cells (69) and MHC classes I and II presented peptides (70, 71) play a pivotal role in the pathogenesis, peptidomics bear a great potential for both, analysis of the concentration of known peptides in biological samples and for the discovery of new, so far unknown small molecules involved in the onset, maintenance and progression of the diseases. Advantages and disadvantages of the most commonly used separation techniques applied in proteomics and peptidomics experiments are summarized in Table 1.

Mass spectrometry

The biochemical entity of protein spots isolated from gels or of peaks derived from chromatographic fractions is usually assessed by mass spectrometric ‘fingerprinting’

Table 1. Commonly used protein and peptide separation techniques in proteomics and peptidomics

Technique	Proteins	Peptides	Advantages	Disadvantages	Ref.
2D-SDS gels	Applicable	Not applicable	Good resolution of high MW proteins	Coverage 30–50% of the proteome, gel to gel variation	31, 32
2D-DIGE†	Applicable to differently labeled samples	Not applicable	Increased reproducibility, differential expression	Gel to gel variations, loss of low concentration proteins	36
HPLC	Applicable to 'in gel' spot digests	Applicable after protein depletion	High resolution, samples directly analysable by MS	Bad resolution of high MW proteins	61
MuPit‡	Applicable for sample fractionation	Applicable	High resolution, reduction of sample complexity	Preferential detection of abundant proteins/peptides	126

†Two-dimensional difference gel electrophoresis.

‡Multi-dimensional protein identification technology.

(72). The MS allows a sensitive and precise detection of the total peptide content even of complex mixtures (73). Basically two techniques, electrospray ionization (ESI) (74) and matrix-assisted laser desorption/ionization (MALDI) are used (75). In the ESI mode, peptides and proteins are multiple protonated and then gently transferred from solution to gas phase yielding multiple charged ions (76). The MALDI technique transfers molecules straight from solid to gas phase with the help of laser irradiation and the so-called matrix, which ionizes the peptides and proteins and transfers sufficient energy to the substrate molecules (77) yielding predominantly singly charged ions. These techniques analyze the mass-to-charge ratio (m/z) of peptide fragments obtained by proteolytic digestion of isolated protein spots or of peptide peaks separated by HPLC or other chromatographic methods. The real mass analyzer is usually time of flight (TOF) if coupled to MALDI or ion trap and quadrupled if ESI is used. In either case, peptides are separated according to their specific mass to charge quotient (m/z) plotted against peak intensities. The raw mass spectra have to be processed because of noise, peak broadening, instrument distortion, saturation effects, miscalibration or presence of contaminants (Fig. 3). The

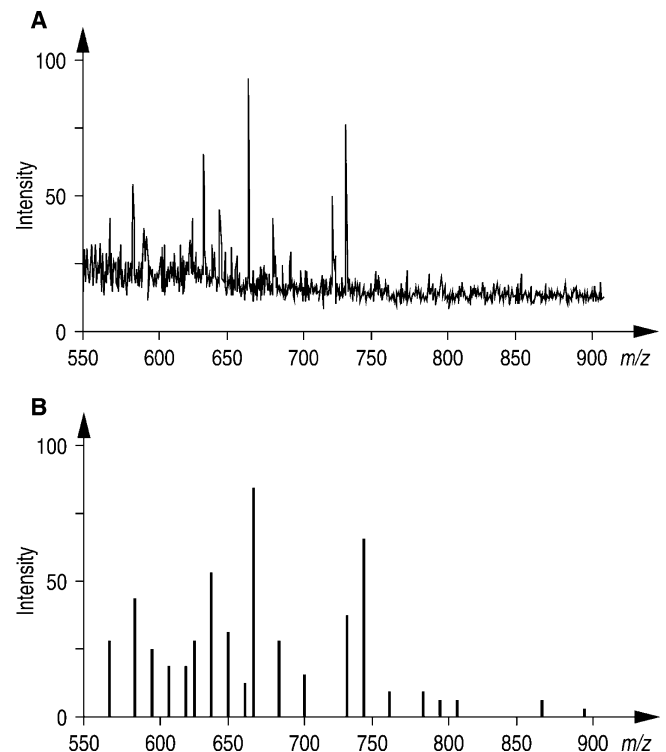


Figure 3. Representation of a raw mass spectrum and of a processed peptide mass fingerprint. The raw output of an MS spectrum (A) has to be processed to correct noise, peak broadening due to the detection system, instrument distortion and saturation effects to generate discrete peptide mass fingerprints (B) (modified from Ref. 125).

Table 2. Workflow of typical proteomics and peptidomics experiments

Step	Proteomics	Peptidomics
1	2D-SDS sample separation	Sample depletion from proteins
2	Spot identification by gel staining	Fractionation by HPLSC
3	'in gel' digestion of spots	MS/MS spectra of each fraction
4	MS/MS peptide fingerprinting	Candidate list identification
5	Sequence determination	Sequence determination
6	Database search	Database search
7	Protein identification	Peptide identification

introduction of tandem MS (MS/MS), i.e., fragmentation of peptides that are selectively isolated in the mass spectrometer in order to read out the (partial) amino acid sequence of the peptide, further improved the analysis procedures. Today, a vast variety of (hybrid) mass spectrometers incorporating various and sometimes multiple, ionization sources and analyzers in one instrument are available. Biomarker profiling from body fluids by MALDI has been complemented and accelerated by the SELDI (surface-enhanced laser desorption/ionization) technology, a combination of miniaturized chromatographic prefractionation on a strip and subsequent MALDI-TOF analysis of the sub-fractions (78). However, one should be aware of the fact that while proteomics has an enormous discovery potential, this also brings the risk of simply getting lost in too many data. This demands for powerful methods for data analysis and therefore bioinformatics is an indispensable complement to any global proteomic or peptidomic approach (79). Typical workflows for proteomics and peptidomics experiments are outlined in Table 2.

The bioinformatics challenge

The modern biochemical profiling technologies including DNA microarrays, proteomics, peptidomics and metabolomics provide the research community with enormous amounts of raw data. However, our ability to generate data now outstrips our ability to analyze it (80). The process from proteomics profiling by 2DG or peptidomics profiling to the identification of (differentially expressed) proteins or peptides by comparison of mass spectrometric results with reference databases includes many components of data processing (81). Data mining represents a modern computer-based statistical approach to provide a system-level understanding by organization of genomic, functional genomic and proteomic data deriving from genetic and environmental parameter variations and is indispensable to build a descriptive and mechanistic model of biological phenomena. The final goal of the bioinformatics challenge is to build a mathematical framework with predictive abilities. Simply knowing the existence of genes and proteins does not tell us much about the biological processes in which they are involved (82). The analysis

of MS data can be a hard task due to the complexity of the information to decipher, the accuracy and dynamic range of quantitative analysis, the availability of appropriate software and the overwhelming size of data files. However, progress in this field is fast and methods for the systematic modeling of proteomics experimental data (83), to describe proteomic analysis of posttranslational modifications (84) and for the analysis of correlation-associated networks (9, 15), are emerging. It is exciting to anticipate a time when data from thousands of gene expression experiments and from proteomic profiling will be available for meta-analysis (85), which has the potential to balance out artefacts from many individual studies, thus leading to more robust results. However, this will require that data adhere to some type of uniform structure and format that would ideally be independent of the particular technology used to generate it (86) and at present we are quite far from such an ideal situation.

Emerging applications of proteomics and peptidomics in allergy and asthma research

Obviously, the first and simplest applications of proteomic approaches in allergy and asthma research consist in the fast identification of new allergenic molecules. It is well known that many allergenic sources like fungi, moulds or latex are complex and contain large numbers of IgE-binding molecules (87, 88). They can easily be separated on 1D or 2D SDS gels and visualized by IgE immunoblotting (89). Proteomics offer a straightforward strategy for the identification of new IgE-binding proteins through 2DG, site-specific fragmentation, MS analysis of the fragments and database search. Use and improvement of the available allergen databases (90) will strongly contribute to the fast identification of new and cross-reactive allergenic structures (91). However, the available computational tools for the study of allergens (92) will need to be improved and further developed to allow a full exploitation of the large knowledge about allergen structures cumulated during the past decade. The 2DG approach has been successfully applied to identify new allergens from shrimp, Japanese cedar and latex (93–95) without the need to isolate the individual proteins in relevant amounts, which is required by classical methods of sequence determinations. However, the greatest potential of the technology derives from the ability to determine posttranslational modifications, which are not detectable by cDNA cloning and difficult to analyze by classical chemical methods. Kolarich and Altmann (96) showed that an electrophoretically separated band corresponding to 4 µg of glycoprotein is generally sufficient to allow detection of the major N-glycan species by MS. Thus mass spectrometry represents a fast and sensitive tool for the characterization of natural allergens identified in Western blot analyses (97) and for comprehensive

structural comparisons with the corresponding recombinant allergens (98, 99). The rapid progress in structure elucidation facilitated by proteomic methods will significantly contribute to the elucidation of the repertoire of allergenic molecules and its micro-heterogeneity deriving from posttranslational modifications and from the presence of allergen isoforms in natural extracts (100, 101).

However, more significant progresses can be expected from proteome wide protein expression profiling (102), which can be applied to any clinical sample. In multifactorial diseases, such as allergy and asthma, it is more likely that disease correlations will be more closely associated with a set of biomarkers or a protein profile rather than with one single marker. Basically, samples derived from cell cultures of biopsy material (103, 104), from fractionated or nonfractionated biopsies (105) and bronchioalveolar lavage are available to study lung diseases (106). Although these areas of research are still in the infancy, first reports on mouse models of asthma using proteomic approaches are emerging (107, 108). It can be expected that such complex experiments will reveal arrays of up- and down-regulated proteins in comparison to healthy and disease status. The proteins found may not all play a role in the immunological mechanisms and/or pathophysiology of allergy or asthma. This is mainly due to the sensitivity of the system, problems in identifying false-negative and false-positive signals (80) and the relatively low throughput of 2D proteomic experiments, which do not yet allow comparison of high numbers of replicates for a robust statistical analysis of single observations. A more promising window for the analysis of allergy and asthma is located in the world of small peptides, which are known to be involved in the pathogenesis of these diseases. Progress in peptidomics is, although young, very impressive (109). Peptidomic biomarker discovery in body fluids (110–112), cell cultures (113, 114) and animal models (115, 116) demonstrate the power of the technology, which is particularly suitable for investigations in the field of immunology (117–119). First reports of quantitative profiling of plasma peptides in asthmatic mice (120), of identification of disease-associated humoral antigens (121) and of activated platelet-derived growth factor pathways in human asthmatic pulmonary-derived mesenchymal cells (122) indicate the potential of peptidomics for allergy and asthma research.

Outlook

At present both proteomic and peptidomic technologies are progressing fast making predictions about future developments a hazardous task. Progress in understanding the pathophysiology of allergy and asthma will depend on the discovery of interaction networks between already known molecules and on the discovery of new peptides and proteins interacting with these networks. In turn, the discovery of new disease-related members of these networks will depend on technology development. Progress and robustness of the results will clearly depend on the ability to develop bioinformatics tools able to deal with huge amounts of data, on the ability to eliminate false-positive and false-negative signals in MS profiles and on the development of high-throughput proteomics and peptidomics to be able to compare large sample collections. Every progress in one of these separate research fields will synergetically influence the development of the other complementary fields and therefore I would expect an exponential growth in technology improvements culminating in the fast discovery of large numbers of diagnostic and therapeutic targets. The way from target discovery to target validation is long and disseminated with pitfalls and the way from the discovery of new proteins or peptides to the elucidation of their roles in the complex networks influencing the pathogenesis of the diseases will not be shorter. However, molecules involved in the pathogenesis of allergy and asthma such as cytokines, chemokines, their receptors, major histocompatibility complex-binding peptides and transcription factors are accessible for proteomic and peptidomic investigations and the power of these technologies will likely lead to the definition of correlation-associated networks, to fill in gaps by the discovery of new network members and finally, to the establishment of more effective prophylaxis and therapy for these diseases.

Acknowledgements

I am grateful to the Swiss National Science Foundation, the OPO Stiftung Zürich and to the EMDO Stiftung Zürich for continuous financial support. I thank Prof. HD Zucht, BioVision AG, Hannover, Germany and Dr. M. Kussmann, Nestlé Research Center, Lausanne, Switzerland for stimulating discussion. I am indebted to Prof. K. Blaser for his continuous support and to Dr. S. Flückiger, Dr. C. Rhyner and Dr. M. Weichel for careful comments on the manuscript.

References

1. Venter JC, Adams MD, Myers EW, Smith HO, Miklos GLG, Wides L et al. The sequence of the human genome. *Science* 2001;**291**:1304–1351.
2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J et al. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.
3. Schmutz J, Wheeler J, Grimwood J, Dickson M, Yang J, Caoile C et al. Quality assessment of the human genome sequence. *Nature* 2004;**429**:365–368.

4. Hood L, Galas D. The digital code of DNA. *Nature* 2003;**421**:444–448.
5. Taylor JG, Choi EH, Foster CB, Chanock SJ. Using genetic variation to study human disease. *Trends Mol Med* 2001;**7**:507–512.
6. Rehfeld JF, Goetze JP. The posttranslational phase of gene expression: new possibilities in molecular diagnosis. *Curr Mol Med* 2003;**3**:25–38.
7. Huber SC, Hardin SC. Numerous posttranslational modifications provide opportunities for the intricate regulation of metabolic enzymes at multiple levels. *Curr Opin Plant Biol* 2004;**7**:318–322.
8. Eisen MB, Spellman PT, Brown PD, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;**95**:14863–14868.
9. Lamerz J, Selle H, Scapozza L, Cramer R, Schulz-Knappe P, Mohring T et al. Correlation-associated peptide networks of human cerebrospinal fluid. *Proteomics* (in press).
10. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. *Nat Biotechnol* 2003;**21**:255–261.
11. Standing KG. Peptide and protein de novo sequencing by mass spectrometry. *Curr Opin Struct Biol* 2003;**13**:595–601.
12. Chiang DY, Brown PO, Eisen MB. Visualizing associations between genome sequences and gene expression data using genome-mean expression profiles. *Bioinformatics* 2001;**17**:S49–S55.
13. Butte AJ, Tamayo P, Slonim D, Golub TR, Kohane JS. Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. *Proc Natl Acad Sci USA* 2000;**97**:12182–12186.
14. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL. The large scale organization of metabolic networks. *Nature* 2000;**407**:651–654.
15. Lamerz J, Cramer R, Scapozza L, Mohring T, Selle H, Zucht HD. Peptide sequence predictions from thousands of peptide signals based on correlation-associated networks. *Combin Chem High Throughput Screen* 2005 (in press).
16. Roberts GC, Smith CW. Alternative splicing: combinatorial output from the genome. *Curr Opin Chem Biol* 2002;**6**:375–383.
17. Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 2004;**8**:33–41.
18. Lopez-Otin C, Overall CM. Protease degradomics: a new challenge for proteomics. *Nat Rev Mol Cell Biol* 2002;**3**:509–519.
19. Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF et al. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 1996;**13**:19–50.
20. Schulz-Knappe P, Zucht HD, Heine G, Jürgens M, Hess R, Schrader M. Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. *Combin Chem High Throughput Screen* 2001;**4**:207–217.
21. Baggerman G, Verleyen P, Clynen E, Huybechts J, de Louf A, Schoofs L. Peptidomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 2004;**803**:3–16.
22. Naistat DM, Leblanc R. Proteomics. *J Environ Pathol Toxicol Oncol* 2004;**23**:161–178.
23. Domingues FS, Lengauer T. Protein function from sequence and structure data. *Appl Bioinform* 2003;**2**:3–12.
24. Meri S, Baumann M. Proteomics: posttranslational modifications, immune responses and current analytical tools. *Biomol Eng* 2001;**18**:213–220.
25. Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 2004;**8**:33–41.
26. Hermjakob H, Montecchi-Palazzi L, Bader G, Wojcik J, Salwinski L, Ceol A et al. The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. *Nat Biotechnol* 2004;**22**:177–183.
27. Cho S, Park SG, Lee du H, Park BC. Protein–protein interaction networks: from interactions to networks. *J Biochem Mol Biol* 2004;**37**:45–52.
28. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem* 2001;**70**:437–473.
29. Domingues FS, Lengauer T. Protein function from sequence and structure data. *Appl Bioinform* 2003;**2**:3–12.
30. O'farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975;**250**:4007–4021.
31. Rabilloud T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* 2002;**3**:3–10.
32. Fey SL, Larsen PM. 2D or not 2D. Two-dimensional gel electrophoresis. *Curr Opin Chem Biol* 2001;**9**:26–33.
33. Joo WA, Lee DY, Kim CW. Development of an effective sample preparation method for the proteome analysis of body fluids using 2-D gel electrophoresis. *Biosci Biotechnol Biochem* 2003;**67**:1574–1577.
34. Manabe T. Analysis of complex protein-polypeptide systems for proteomic studies. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003;**787**:29–41.
35. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;**422**:198–207.
36. Hoving S, Voshol H, van Oostrum J. Towards high performance two-dimensional gel electrophoresis using ultrazoom gels. *Electrophoresis* 2000;**21**:2617–2621.
37. Zuo X, Speicher DW. Comprehensive analysis of complex proteomes using microscale solution isoelectrofocusing prior to narrow pH range two-dimensional electrophoresis. *Proteomics* 2002;**2**:58–68.
38. Rose K, Bougueler L, Baussant T, Bohm G, Botti P, Colinge J et al. Industrial-scale proteomics: from liters of plasma to chemically synthesized proteins. *Proteomics* 2004;**4**:2125–2150.
39. Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol* 2001;**19**:379–382.
40. Wilkins JA, Xiang R, Horvath C. Selective enrichment of low-abundance peptides in complex mixtures by elution-modified displacement chromatography and their identification by electrospray ionization mass spectrometry. *Anal Chem* 2002;**74**:3933–3941.
41. Bodnar WM, Blackburn RK, Krise JM, Moseley MA. Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J Am Mass Spectrom* 2003;**14**:971–979.
42. Shefcheck K, Yao X, Fenselau C. Fractionation of cytosolic proteins on an immobilized heparin column. *Anal Chem* 2003;**75**:1691–1698.
43. Govorukhina NT, Keizer-Gunnink A, van der Zee AG, DE Jong S, de Bruijn HW, Bischoff R. Sample preparation of human serum for the analysis of tumor markers. Comparison of different approaches for albumin and gamma-globulin depletion. *J Chromatogr A* 2003;**1009**:171–178.
44. Hinerfeld D, Innamorati D, Pirro J, Tam SW. Serum/Plasma depletion with chicken immunoglobulin Y antibodies for proteomic analysis from multiple Mammalian species. *J Biomol Tech* 2004;**15**:184–190.

45. Quadroni M, James P. Proteomics and automation. *Electrophoresis* 1999;**20**:664–677.
46. Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R et al. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 2000;**21**:1037–1053.
47. Lin Z, Crockett DK, Lim MS, Elenitoba-Johanson KS. High-throughput analysis of protein/peptide complexes by immunoprecipitation and automated LC-MS/MS. *J Biol Technol* 2003;**14**:149–155.
48. Jensen ON, Wilm M, Shevchenko A, Mann M. Peptide sequencing of 2DE gel-isolated proteins by nanoelectrospray tandem mass spectrometry. *Methods Mol Biol* 1999;**112**:571–588.
49. Borchers C, Peter JF, Hall MC, Kunkel TA, Tomer KB. Identification of in-gel digested proteins by complementary peptide mass fingerprinting and tandem mass spectrometry data obtained on an electrospray ionization quadrupole time-of-flight mass spectrometer. *Anal Chem* 2000;**72**:1163–1168.
50. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci USA* 2000;**97**:9390–9395.
51. Colburu WA, Lee JW. Biomarkers, validation and pharmacokinetic-pharmacodynamic modelling. *Clin Pharmacokinet* 2003;**42**:997–1022.
52. Kantur AB, Wang W, Lin H, Govindarajan H, Anderle M, Perrone A et al. Biomarker discovery by comprehensive phenotyping for autoimmune diseases. *Clin Immunol* 2004;**111**:186–195.
53. Chanin TD, Merrick DT, Franklin WA, Hirsch FR. Recent developments in biomarkers for the early detection of lung cancer: perspectives based on publications 2003 to present. *Curr Opin Pulm Med* 2004;**10**:242–247.
54. Seliger B, Kellner R. Design of proteome-based studies in combination with serology for the identification of biomarkers and novel targets. *Proteomics* 2002;**2**:1641–1651.
55. Petricoin E, Wulfkuhle J, Spina V, Liotta LA. Clinical proteomics: Revolutionizing disease detection and patient tailoring therapy. *Proteome Res* 2004;**3**:209–217.
56. Oh P, Li Y, Yu J, Burr E, Krasinska KM, Carver LA et al. Subtractive proteomic mapping of the endothelial surface in lung and solid tumors for tissue-specific therapy. *Nature* 2004;**429**:629–635.
57. Verma M, Kagan J, Sidransky D, Srivastava S. Proteomic analysis of cancer-cell mitochondria. *Nat Rev Cancer* 2003;**3**:789–795.
58. Petricoin EF, Liotta LA. Clinical applications of proteomics. *J Nutr* 2003;**133**:2476S–2484S.
59. Mootha VK, Lepage P, Miller K, Bunkenborg J, Reich M, Hjerrild M et al. Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. *Proc Natl Acad Sci USA* 2003;**100**:605–610.
60. Ramstrom M, Bergquist J. Miniaturized proteomics and peptidomics using capillary liquid separation and high resolution mass spectrometry. *FEBS Lett* 2004;**567**:92–95.
61. Issaq HJ, Conrads TP, Ianini GM, Veenstra TD. Methods for fractionation, separation and profiling of proteins and peptides. *Electrophoresis* 2002;**23**:3048–3061.
62. Stickel JJ, Fotopoulos A. Pressure-flow relationships for packed beds of compressible chromatography media at laboratory and production scale. *Biotechnol Prog* 2001;**17**:744–751.
63. Schulz-Knappe P, Schrader M, Standker L, Richter R, Hess R, Jürgens M et al. Peptide bank generated by large-scale preparation of circulating human peptides. *J Chromatogr A* 1997;**776**:125–132.
64. Skold K, Svensson M, Kaplan A, Björkstén L, Åström J, Andren PE. A neuroproteomic approach to targeting neuropeptides in the brain. *Proteomics* 2002;**2**:447–454.
65. Heine G, Zucht HD, Scuhmann MU, Burger K, Jürgens M, Zumkeiler M et al. High-resolution peptide mapping of cerebrospinal fluid: a novel concept for diagnosis and research in central nervous system diseases. *J Chromatogr B Analyt Technol Biomed Life Sci* 2002;**782**:353–361.
66. Schrader M, Schulz-Knappe P. Peptidomics technologies for human body fluids. *Trends Biotechnol* 2001;**19**(Suppl. 10):S55–S60.
67. Romagnani S. Cytokines and chemottractants in allergic inflammation. *Mol Immunol* 2002;**38**:881–885.
68. Lloyd CM, Rankin SM. Chemokines in allergic airway disease. *Curr Opin Pharmacol* 2003;**3**:443–448.
69. Krishnaswamy G, Kelly J, Johnson D, Youngberg G, Stone W, Hung SK et al. The human mast cell: functions in physiology and disease. *Front Biosci* 2001;**6**:109–127.
70. Margalit H, Altuvia Y. Insights from MHC-bound peptides. *Novartis Found Symp* 2003;**254**:77–90.
71. Purcell AW, German JJ. Immunoproteomics: Mass spectrometry-based methods to study the targets of the immune response. *Mol Cell Proteomics* 2004;**3**:193–208.
72. Larsen MR, Roepstorff P. Mass spectrometric identification of proteins and characterization of their post-translational modifications in proteome analysis. *Fresenius J Anal Chem* 2000;**366**:677–690.
73. Yates JR III. Mass spectrometry and the age of the proteome. *J Mass Spectrom* 1998;**33**:1–19.
74. Fenn JB. Electrospray wings for molecular elephants (Nobel lecture). *Angew Chem Int Ed Engl* 2003;**42**:3871–3894.
75. Tanaka K. The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew Chem Int Ed Engl* 2003;**42**:3860–3870.
76. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989;**246**:64–71.
77. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10 000 daltons. *Anal Chem* 1988;**60**:2299–2301.
78. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002;**359**:572–577.
79. Wiemer JC, Prokudin A. Bioinformatics in proteomics: application, terminology, and pitfalls. *Pathol Res Pract* 2004;**200**:173–178.
80. Patterson SD. Data analysis – the Achilles heel of proteomics. *Nat Biotechnol* 2003;**21**:221–222.
81. Stanislaus R, Jiang LH, Swartz M, Arthur J, Almeida JS. An XML standard for the dissemination of annotated 2D gel electrophoresis data complemented with mass spectrometry. *BMC Bioinform* 2004;**5**:9.
82. Ng SK, Tan SH. Discovering protein-protein interactions. *J Bioinform Comput Biol* 2004;**1**:711–741.
83. Taylor CF, Paton NW, Grawood KL, Kirby PD, Stead DA, Yin Z et al. A systematic approach to modelling, capturing, and disseminating proteomics experimental data. *Nat Biotechnol* 2003;**21**:247–254.
84. Mann M, Jensen ON. Proteomic analysis of post-translational modification. *Nat Biotechnol* 2003;**21**:255–261.
85. Hedges LV, Olkin I. Statistical methods for meta-analysis. New York: Academic Press, 1985.

86. Ermoeva O, Rastogi M, Pruitt KD, Schuler GD, Bittner ML, Chen Y et al. Data management and analysis for gene expression arrays. *Nat Genet* 1998;**20**:19–23.
87. Cramer R, Blaser K. Allergy and immunity to fungal infections and colonization. *Eur Respir J* 2002;**19**:151–157.
88. Niggemann B, Breiteneder H. Latex allergy in children. *Int Arch Allergy Immunol* 2000;**121**:98–107.
89. Poulsen LK. *In-vitro* diagnosis: serum-based methods used for risk assessment of allergenic food. *Curr Opin Allergy Clin Immunol* 2001;**1**:249–254.
90. Brusic V, Millot M, Petrovsky N, Gendel SM, Gignozac O, Stelman SJ. Allergen databases. *Allergy* 2003;**58**:1093–1100.
91. Ferreira F, Hawranek T, Gruber P, Wopfner N, Mari A. Allergic cross-reactivity: from gene to the clinic. *Allergy* 2004;**59**:243–267.
92. Brusic V, Petrovsky N, Gendel SM, Millot M, Gignozac O, Stelman SJ. Computational tools for the study of allergens. *Allergy* 2004;**58**:1083–1092.
93. Yu CJ, Lin YF, Chiang BL, Show LP. Proteomics and immunological analysis of a novel shrimp allergen, Pen m 2. *J Immunol* 2003;**170**:445–453.
94. Fujimura T, Shigeta S, Kawamoto S, Aki T, Masubuchi M, Hayashi T et al. Two-dimensional IgE-binding spectrum of Japanese cedar (*Cryptomeria japonica*) pollen allergens. *Int Arch Allergy Immunol* 2004;**133**:125–135.
95. Haishima Y, Tsuchiya T, Tomitaka-Yagami A, Kano H, Matsunaga K. Proteomic analysis of putative latex allergens. *Int Arch Allergy Immunol* 2004;**135**:3–11.
96. Kolarich D, Altmann F. N-Glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen Ole e 1. *Anal Biochem* 2000;**285**:64–75.
97. Raftery MJ, Saldanha RG, Geczy CL, Kumar RK. Mass spectrometric analysis of electrophoretically separated allergens and proteases in grass pollen diffusates. *Respir Res* 2003;**4**:10.
98. Grönlund H, Bergman T, Sandström K, Alvelius G, Reininger R, Verdino P et al. Formation of disulfide bonds and homodimers of the major cat allergen Fel d 1 equivalent to the natural allergen by expression in *Escherichia coli*. *J Biol Chem* 2003;**278**:40144–40151.
99. Westphal S, Kolarich D, Foetisch K, Lauer I, Altmann F, Conti A et al. Molecular characterization and allergenic activity of Lyc e 2 (beta-fructofuranosidase), a glycosylated allergen of tomato. *Eur J Biochem* 2003;**270**:1327–1337.
100. Blume C, Lindner B, Pecker WM, Petersen A. Microheterogeneity of the major grass group 6 allergen Phl p 6: analysis by mass spectrometry. *Proteomics* 2004;**4**:1366–1371.
101. Rautiainen J, Auriola S, Kontinen A, Virtanen T, Rytönen-Nissinen M, Zeiler T et al. Two new variants of the lipocalin allergen Bos d2. *J Chromatogr B Biomed Sci Appl* 2001;**763**:91–98.
102. Boggs SF. Protein profiling in respiratory disease: techniques and impact. *Expert Rev Proteomics* 2004;**1**:29–36.
103. Westergren-Thorsson G, Malmström J, Marko-Varga G. Proteomics – the protein expression technology to study connective tissue biology. *J Pharm Biomed Anal* 2001;**24**:815–824.
104. Malmström J, Tufvesson E, Lofdahl CG, Hansson L, Marko-Varga G, Westergren-Thorsson G. Activation of platelet-derived growth factor pathway in human asthmatic pulmonary-derived mesenchymal cells. *Electrophoresis* 2003;**24**:276–285.
105. Fehninger TE, Sato-Folatre JG, Malmström J, Berglund M, Lindberg C, Brange C et al. Exploring the context of the lung proteome within the airway mucosa following allergen challenge. *J Proteome Res* 2004;**3**:307–320.
106. Sepper R, Prikk K. Proteomics: is it an approach to understand the progression of chronic lung disorders? *J Proteome Res* 2004;**3**:277–281.
107. Houtman R, Krijgsvelde J, Kool M, Romijn EP, Redegeld FA, Nijkamp FP et al. Lung proteome alterations in a mouse model for nonallergic asthma. *Proteomics* 2003;**3**:2008–2018.
108. Wheelock AM, Zhang L, Tran MU, Morin D, Penn S, Buckpitt AR et al. Isolation of rodent airway epithelial cell proteins facilitates *in vivo* proteomics studies of lung toxicity. *Am J Physiol Lung Cell Mol Physiol* 2004;**286**:399–410.
109. Cramer R, (editor). Peptidomics: a new dimension in functional genomics. Hot topics Issue, *Combin Chem High Throughput Screen*. 2005 (in press).
110. Svensson M, Skold K, Svenningsson P, Andren PE. Peptidomic-based discovery of neuropeptides. *Proteome Res* 2003;**2**:213–219.
111. Schrader M, Schulz-Knappe P. Peptidomics technologies for human body fluids. *Trends Biotechnol* 2001;**19**:S55–S60.
112. Tammen H, Möhring T, Kellmann M, Pich A, Kreipe HH, Hess R. Mass spectrometric phenotyping of Val34-Lue polymorphism of blood coagulation factor XIII by differential peptide display. *Clin Chem* 2004;**50**:545–551.
113. Tammen H, Kreipe H, Hess R, Kellmann M, Lehmann U, Pich A et al. Expression profiling of breast cancer cells by differential peptide display. *Breast Cancer Res Treat* 2003;**79**:83–93.
114. Sasaki K, Sato K, Akiyama Y, Yanagihara K, Oka M, Yamaguchi K. Peptidomics-based approach reveals the secretion of the 29-residue COOH-terminal fragment of the putative tumor suppressor protein DMBT1 from pancreatic adenocarcinoma cell lines. *Cancer Res* 2002;**62**:4894–4898.
115. Pierson J, Norris JL, Aerni HR, Svenningsson P, Caprioli RM, Andren PE. Molecular profiling of experimental Parkinson's disease: direct analysis of peptides and proteins on brain tissue sections by MALDI mass spectrometry. *J Proteome Res* 2004;**3**:289–295.
116. Budde P, Schulte I, Appelt A, Kellmann M, Neitt S, Hess R et al. Peptidomics biomarker discovery in mouse models of obesity and type 2 diabetes. *Combin Chem High Throughput Screen* (in press).
117. de Jong A. Contribution of mass spectrometry to contemporary immunology. *Mass Spectrom Rev* 1998;**17**:311–335.
118. Dowd KM. Contributions of mass spectrometry to structural immunology. *J Mass Spectrom* 2000;**35**:493–503.
119. Halapi E, Hakonarson H. Recent development in genomic and proteomic research for asthma. *Curr Opin Pulm Med* 2004;**10**:22–30.
120. Yeo S, Roh GS, Kim DH, Lee JM, Seo SW, Cho JW et al. Quantitative profiling of plasma peptides in asthmatic mice using liquid chromatography and mass spectrometry. *Proteomics* 2004;**4**:3308–3317.
121. Gires O, Munz M, Schaffrik M, Kieu C, Rauch J, Abelmann M et al. Profile identification of disease-associated humoral antigens using AMIDA, a novel proteomics-based technology. *Cell Mol Life Sci* 2004;**61**:1198–1207.

122. Malmstrom J, Tufvesson E, Lofdahl CG, Hansson L, Marko-Varga G, Westergren-Thorsson G. Activation of platelet-derived growth factor pathway in human asthmatic pulmonary-derived mesenchymal cells. *Electrophoresis* 2003;**24**:276–285.
123. Lamerz J. Peptides and peptide networks in human cerebrospinal fluid. PhD Thesis No 15623, Switzerland: Swiss Federal Institute of Technology, Zürich, 2004.
124. Schulz-Knappe P, Budde P, Jürgens M, Schrader M, Selle H, Tammen H et al. Biomarker- und Wirkstoffsuche von Peptiden und kleinen Proteinen. *Lab- orwelt* 2004;**5**:40–45.
125. Englbrecht CC, Facius A. Bioinformatics challenges in proteomics. *Combin Chem High Throughput Screen* 2005 (in press).
126. Washburn MP, Wolters D, Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001;**25**:242–247.