TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. D OSA - TOM. 794

MEDICA - ODONTOLOGICA

QUANTITATIVE PROTEOMICS IN THE CHARACTERIZATION OF T HELPER LYMPHOCYTE DIFFERENTIATION

by

Jan-Jonas Filén

TURUN YLIOPISTO
Turku 2008

Turku Centre for Biotechnology, University of Turku and Åbo Akademi University;
Department of Medical Biochemistry and Molecular Biology, University of Turku;
The National Graduate School in Informational and Structural Biology

Supervised by

Professor Riitta Lahesmaa, M.D., Ph.D.
Turku Centre for Biotechnology
University of Turku and Åbo Akademi University
Turku, Finland

and

Docent Tuula Nyman, Ph.D.
Institute of Biotechnology
University of Helsinki
Helsinki, Finland

Reviewed by

Professor Seppo Auriola, Ph.D.
Department of Pharmaceutical Chemistry
University of Kuopio
Kuopio, Finland

and

Docent Panu Kovanen, M.D., Ph.D.
Haartman Institute
Department of Pathology
University of Helsinki
Helsinki, Finalnd

Opponent

Professor Ruedi Aebersold, Ph.D.
Institute of Molecular Systems Biology
Eidgenössische Technische Hochschule Zürich
And Faculty of Science, University of Zurich,
Zürich, Switzerland

ISBN 978-951-29- 3492-8 (PRINT)
ISBN 978-951-29- 3493-5 (PDF)
ISSN 0355-9483
Painosalama Oy – Turku, Finland 2008

To Sanna and Anton

Jan-Jonas Filén

Quantitative proteomics in the characterization of T helper lymphocyte differentiation

Turku Centre for Biotechnology, University of Turku and Åbo Akademi University; Department of Medical Biochemistry and Molecular Biology, University of Turku; The National Graduate School in Informational and Structural Biology

ABSTRACT

The term proteome is used to define the complete set of proteins expressed in cells or tissues of an organism at a certain timepoint. Respectively, proteomics is used to describe the methods, which are used to study such proteomes. These methods include chromatographic and electrophoretic techniques for protein or peptide fractionation, mass spectrometry for their identification, and use of computational methods to assist the complicated data analysis.

A primary aim in this Ph.D. thesis was to set-up, optimize, and develop proteomics methods for analysing proteins extracted from T helper (Th) lymphocytes. First, high-throughput LC-MS/MS and ICAT labeling methods were set-up and optimized for analysing the microsomal fraction proteins extracted from Th lymphocytes. Later, iTRAQ method was optimized to study cytokine regulated protein expression in the nuclei of Th lymphocytes. High-throughput LC-MS/MS analyses, like ICAT and iTRAQ, produce large quantities of data and robust software and data analysis pipelines are needed. Therefore, different software programs used for analysing such data were evaluated. Moreover, a pre-filtering algorithm was developed to classify good-quality and bad-quality spectra prior to the database searches.

Th-lymphocytes can differentiate into Th1 or Th2 cells based on surrounding antigens, co-stimulatory molecules, and cytokines. Both subsets have individual cytokine secretion profiles and specific functions. Th1 cells participate in the cellular immunity against intracellular pathogens, while Th2 cells have important role in the humoral immunity against extracellular parasites. An abnormal response of Th1 and Th2 cells and imbalance between the subsets are charasteristic of several diseases. Th1 specific reactions and cytokines have been detected in autoimmune diseases, while Th2 specific response and cytokine profile is common in allergy and asthma.

In this Ph.D. thesis mass spectrometry-based proteomics was used to study the effects of Th1 and Th2 promoting cytokines IL-12 and IL-4 on the proteome of Th lymphocytes. Characterization of microsomal fraction proteome extracted from IL-12 treated lymphobasts and IL-4 stimulated cord blood CD4+ cells resulted in finding of cytokine regulated proteins. Galectin-1 and CD7 were down-regulated in IL-12 treated cells, while IL-4 stimulation decreased the expression of STAT1, MXA, GIMAP1, and GIMAP4. Interestingly, the transcription of both *GIMAP* genes was up-regulated in Th1 polarized cells and down-regulated in Th2 promoting conditions.

Keywords: Proteomics, mass spectrometry, stable isotope labeling, Th1/Th2 differentiation, cytokine

Jan-Jonas Filén

Kvantitatiivinen proteomiikka T-auttajasolujen erilaistumisen tutkimisessa

Turun Biotekniikan keskus, Turun Yliopisto ja Åbo Akademi; Lääketieteellinen biokemia ja molekyylibiologia, Turun Yliopisto; Kansallinen bioinformatiikan ja rakennebiologian tutkijakoulu

TIIVISTELMÄ

Proteomilla tarkoitetaan organismin solujen tai kudosten tietyllä ajanhetkellä ilmentämiä proteiineja. Proteomiikka puolestaan käsittää menetelmät, joiden avulla tutkitaan proteomeja. Näihin menetelmiin kuuluvat kromatografiset ja elektroforeettiset tekniikat proteiinien ja peptidien fraktiointiin, massaspektrometria niiden tunnistamiseen sekä tietojenkäsittely menetelmät avustamaan tietojen analysointia.

Väitöskirjan osatöiden tarkoituksena oli pystyttää, optimoida ja kehittää proteomiikan tutkimusmenetelmiä T-auttajasoluista (Th) eristettyjen proteiinien tutkimiseen. Aluksi pystytettiin ja optimoitiin LC-MS/MS- ja ICAT-menetelmät. Th-solujen mikrosomaalisten fraktioiden analysoimiseksi. Myöhemmin iTRAQ-menetelmä optimoitiin Th-solujen tuman proteomin tutkimiseksi. ICAT- ja iTRAQ-menetelmien kaltaiset LC-MS/MS-menetelmät tuottavat paljon tietoa, minkä analysoimiseksi tarvitaan tehokkaita tietokoneohjelmia ja tietojen analysointijärjestelmiä. Tämän vuoksi väitöskirjatyössä arvioitiin massaspektrometrillä mitattujen spektrien analysointiin soveltuvia tietokoneohjelmia. Lisäksi kehitettiin algoritmi, jonka avulla voidaan erotella laadultaan hyvät ja huonot spektrit toisistaan ennen tietokantahakuja.

Th-solut voivat erilaistua Th1- tai Th2-soluiksi ympäröivien antigeenien, ko-stimuloivien molekyylien ja sytokiinien vaikutuksesta. Molemmilla alatyypeillä on yksilölliset sytokiinien tuottoprofiilit ja spesifiset tehtävät. Th1-solut osallistuvat soluvälitteiseen immuniteettiin solunsisäisiä taudinaiheuttajia vastaan, kun taas Th2-soluilla on tärkeä rooli vasta-ainevälitteisessä immuniteetissä solunulkoisia loisia vastaan. Th1- ja Th2-solujen epänormaali vaste ja solupopulaatioiden välinen epätasapaino voivat johtaa sairauksiin. Th1-soluille ominaisia reaktioita ja sytokiineja on havaittu autoimmuunisairauksissa, kun taas Th2-soluille tyypillinen vaste ja sytokiiniprofiili ovat ominaisia allergiassa ja astmassa.

Tässä väitöskirjatyössä tutkittiin massaspektrometriaan perustuvan proteomiikan avulla Th1 ja Th2 erilaistavien sytokiinien IL-12 ja IL-4 vaikutusta Th-solujen proteomiin. IL-12:lla stimuloiduista perifeerisen veren Th-soluista ja IL-4:llä stimuloiduista napaveren Th-soluista eristettiin mikrosomaaliset fraktiot, joiden proteomeissa havaittiin sytokiinien säätelemiä proteiineja. Galektiini-1:n ja CD7:n ekspressio väheni IL-12:n vaikutuksesta ja IL-4 puolestaan vähensi STAT1:n, MXA:n, GIMAP1:n ja GIMAP4:n ilmenemistä. Molempien *GIMAP* geenien transkription havaittiin vähenevän IL-4:n vaikutuksesta ja lisääntyvän IL-12:n vaikutuksesta.

Avainsanat: Proteomiikka, massaspektrometria, stabiilit isotooppileimat, Th1/Th2 erilaistuminen, sytokiini

CONTENTS

ABSTRACT **4**
TIIVISTELMÄ **5**
CONTENTS **6**
ABBREVIATIONS **8**
LIST OF ORIGINAL PUBLICATIONS **10**
1 INTRODUCTION **11**
2 REVIEW OF THE LITERATURE **12**
2.1 Proteins, proteome, and proteomics 12
2.2 Methods used in proteomics 13
2.3 Sample preparation 14
2.4 Sample fractionation 15
2.4.1 Subcellular fractionation 15
2.4.2 Purification of protein complexes 16
2.4.3 Electrophoresis-based fractionation 17
2.4.4 Chromatography-based fractionation 17
2.5 Biological mass spectrometry 18
2.5.1 MS instrumentation 18
2.5.2 Identification of proteins with mass spectrometry 19
2.6 Quantification & quantitative labeling 19
2.6.1 Gel-based quantification 20
2.6.1.1 Metabolic labeling with radioactive labels 20
2.6.1.2 Protein staining in the gels with dyes 21
2.6.2 MS-based quantification – isotope-coded tags 23
2.6.2.1 Chemical labeling with stable isotopes – ICAT and iTRAQ 24
2.6.2.2 Metabolic labeling with stable isotopes: SILAC 27
2.6.2.3 Enzymatic labeling with stable isotopes: Proteolytic ^{18}O labeling 28
2.7 Proteomics data analysis 29
2.7.1 2-DE image analysis software 29
2.7.2 Software tools for protein identification 29
2.7.3 Protein quantification algorithms 31
2.7.4 Data analysis pipeline 31
2.8 Lymphocytes present in the immune system 32
2.8.1 Th cells 33
2.8.2 Cytokines promoting Th1 and Th2 differentiation 33
2.8.2.1 IL-12 and Th1 differentiation 34
2.8.2.2 IL-4 and Th2 differentiation 35
2.9 Transcriptome and proteome analyses create novel insights about T helper cell signalling and differentiation 36

3 AIMS OF THE STUDY **38**

4 MATERIALS AND METHODS **39**
4.1 Standard protein mixture 39
4.2 Human Th cell cultures 39

4.2.1 CD8+ depleted peripheral blood mononuclear cells 39
4.2.2 CD4+ cells extracted from PBMC and CB 40
4.3 Flow cytometric analyses 40
4.4 Cell fractionation 40
4.4.1 Microsomal fraction 40
4.4.2 Nuclear extract 41
4.5 ICAT labeling of the samples 41
4.6 iTRAQ labeling of the samples 41
4.7 Chromatographic fractionation of the peptides 41
4.8 nanoLC-ESI-MS/MS analysis 42
4.9 Data processing tools 42
4.10 Western blot analyses 43
4.11 Quantitative real-time RT-PCR 44

5 RESULTS AND DISCUSSION 45
5.1 Characterization of microsomal fraction proteome in Th1 and Th2 polarized lymphocytes by using isotope-coded affinity tags and tandem mass spectrometry (I-II) 45
5.2 Characterization of microsomal fraction proteome in human lymphoblasts reveals the down-regulation of galectin-1 by interleukin-12 (I) 46
5.2.1 Characterization of microsomal fraction proteome in human lymphobasts polarized towards Th1 phenotype by interleukin-12 46
5.2.2 Galectin-1 and CD7 expression was decreased in human lymphoblasts treated with interleukin-12 46
5.3 Characterization of microsomal fraction proteome in human Th lymphocytes polarized towards Th2 phenotype by interleukin-4 (II) 47
5.3.1 Characterization of microsomal fraction proteome in naïve human Th lymphocytes polarized towards Th2 phenotype by interleukin-4 47
5.3.2 GIMAP1, GIMAP4, STAT1, and MxA were down-regulated in interleukin-4 treated Th lymphocytes 48
5.4 Determination of expression changes in nuclear proteins extracted from Th2 polarized cord blood CD4+ cells (unpublished results) 49
5.5 A comparative evaluation of software packages to analyse MS/MS data of ICAT-labeled complex protein mixture (III)...51
5.6 Quality classification of tandem mass spectrometry data (IV)...........................53

6 SUMMARY 55

ACKNOWLEDGEMENTS 57

REFERENCES 59

ORIGINAL PUBLICATIONS 75

ABBREVIATIONS

2-DE	2-dimensional electrophoresis
c-maf	musculoaponeurotic fibrosarcoma -protein
DiGE	difference gel electrophoresis
DTT	dithiothreitol
ER	endoplasmic reticulum
ESI	electrospray ionization
FITC	fluorescein isothiocyanate
FoxP	forkhead box protein
FTICR	fourier transform ion cyclotron resonance
γ c	cytokine receptor common γ–chain
GATA	gata-binding protein
GIMAP	GTPase of the immunity associated protein
GM-CSF	granulocyte-macrophage colony-stimulating factor
ICAT	isotope-coded affinity tag
IEF	isoelectric focusing
IFN	interferon
IL	interleukin
IL-#R	interleukin-# receptor
IPG	immobilized pH gradient
IT	ion trap
iTRAQ	isotope tagged relative and absolute quantitation
JAK	Janus kinase
MALDI	matrix assisted laser desorption ionization
MS	mass spectrometry
MS/MS	tandem mass spectrometry
NK cell	natural killer cell
PAGE	polyacrylamide gel electrophoresis
PBMC	peripheral blood mononuclear cells
PHA	phytohemagglutinin
PMF	peptide mass fingerprinting
PFF	peptide fragment fingerprinting
QQQ	triple quadrupole
RP	reversed-phase chromatography
R-PE	R-phycoerytherin
RT-PCR	reverse transcriptase-polymerase chain reaction
SCX	strong cation exchange
SDS	sodium dodecyl sulfate

SEC	size exclusion chromatography
SILAC	stable isotope labeling with amino acids in cell culture
STAT	signal transducer and activator
TAP	tandem affinity purification
T-bet	t-box expressed in T cells
TBP	tributylphosphine
Tc cell	cytotoxic T cell
TCEP	tris(2-carboxyethyl)phosphine
TCR	T cell receptor
TEAB	triethyl ammonium bicarbonate
TGF	transforming growth factor
Th cell	T helper cell
TNF	tumor necrosis factor
TOF	time-of-flight
Tr cell	regulatory T cell
TYK	tyrosine kinase

LIST OF ORIGINAL PUBLICATIONS

I Filén, J.-J.*, Nyman, T.A.*, Korhonen, J., Goodlett, D. R., and Lahesmaa, R. (2005) Characterization of microsomal fraction proteome in human lymphoblasts reveals the down-regulation of galectin-1 by interleukin-12. *Proteomics*, **5**, 4719-4732.
* Authors with equal contribution

II Filén, J.-J.*, Filén, S.*, Moulder, R., West, A., Kouvonen, P., Kantola, S., Björkman, M., Katajamaa, M., Rasool, O., Nyman, T.A., and Lahesmaa, R. Determination of interleukin-4 regulated proteins in the microsomal fraction of $CD4^+$ cells by using isotope-coded affinity tags. *(Manuscript)*
* Authors with equal contribution

III Moulder, R., Filén, J.-J., Salmi, J., Katajamaa, M., Nevalainen, O.S., Oresic, M., Aittokallio, T., Lahesmaa, R., and Nyman, T. A.. (2005) A comparative evaluation of software for the analysis of liquid chromatography-tandem mass spectrometry data from isotope coded affinity tag experiments. *Proteomics,* **5**, 2748-2760.

IV Salmi, J., Moulder, R., Filén, J.-J., Nevalainen, O. S., Nyman, T. A., Lahesmaa, R., and Aittokallio, T. (2006) Quality classification of tandem mass spectrometry data. *Bioinformatics*, Vol 22. no. **4**, 400-406

This thesis also includes unpublished data.

The original publications have been reprinted with the permissions of the copyright holders.

1 INTRODUCTION

The term proteome can be defined as a complete set of proteins expressed by a cell (Wilkins et al., 1996), while proteomics refers to the large-scale analysis of proteins (Pandey and Mann, 2000). Proteomics includes many different methods to analyse proteins of a proteome, for example, chromatographic and electrophoretic techniques for protein or peptide fractionation, labeling reagents for their quantification, mass spectrometers to analyse them, and software programs to assist in the data analysis (Pandey and Mann, 2000; Righetti et al., 2003; Stults and Arnott, 2005; Zhu et al., 2003). The development of LC-MS/MS instruments and related applications has improved the large-scale analyses of complex protein mixtures. The capacity of the analyses to detect proteins has improved and automation of the methods has increased the throughput of the analyses. Isotope-coded labeling reagents used in combination with LC-MS/MS have improved the quantitative analyses of complex protein mixtures. These stable isotopes can be metabolically, chemically, or enzymatically incorporated to the proteins (Corthals and Rose, 2007; Goshe and Smith, 2003; Ong and Mann, 2005). Despite recent instrumental and methodological developments, the cellular proteome is still far too complicated to be completely analysed. Therefore, the experimental design should be planned carefully based on the biological question.

There are two main subsets of T helper (Th) cells, which both differentiate from naïve Th cells based upon the surrounding antigens, costimulatory molecules and cytokines (Agnello et al., 2003; Constant and Bottomly, 1997; Lee et al., 2006; Mosmann et al., 1986; Mosmann and Coffman, 1989b). Cytokines IL-12 and IL-4 promote Th1 and Th2 differentiation, respectively (Hsieh et al., 1992; Hsieh et al., 1993; Kobayashi et al., 1989; Manetti et al., 1993; Seder et al., 1992; Swain et al., 1990). Both subsets have distinct cytokine secretion profiles and functions. Th1 cells secrete interferon (IFN)-γ , interleukin (IL)-2, and lymphotoxin-β, which promotes cellular immunity against intracellular pathogens (Glimcher and Murphy, 2000; Mosmann and Coffman, 1989a; Paul and Seder, 1994). Th2 cells participate in humoral immunity by secreting IL-4, IL-5, IL-6, IL-9, IL-10 and IL-13 (Glimcher and Murphy, 2000; Mosmann and Coffman, 1989a; Paul and Seder, 1994). Imbalance between Th1 and Th2 subsets, and their selective activation is associated with pathophysiological conditions like allergy, asthma, and certain autoimmune diseases (Romagnani, 1994; Romagnani, 1996).

In this Ph. D. thesis high-throughput LC-MS/MS, stable isotope labeling, and data analysis methods were set-up to elucidate the effects of Th1 and Th2 promoting cytokines IL-12 and IL-4 on the proteome of Th lymphocytes.

2 REVIEW OF THE LITERATURE

2.1 Proteins, proteome, and proteomics

Proteins are organic molecules, which are composed of amino acids linked to each other by peptide bonds. Proteins were discovered in the 19th century by Jöns Jakob Berzelius, who named them πρωτειοξ in Greek meaning "we're number one". As their name indicates these molecules have a crucial role for all living organisms. Proteins can act as enzymes, accelerating many important chemical reactions, for example, in cellular metabolism. Proteins also form the backbone of the cells and different cellular structures. Furthermore, they participate in cell motility, signalling, mitosis, etc. The importance of the proteins is highlighted by their involvement in most of the cellular processes and functions.

The central dogma of molecular biology enunciated by Francis Crick in 1958 (Crick, 1970), describes the protein synthesis in the cells according to the genetic code. Genetic information is stored in DNA, except in some viruses as RNA. A gene represents a DNA sequence, which codes the corresponding protein. In protein synthesis, DNA is first transcribed to a messenger-RNA transcript, which is further translated to a protein by transfer-RNA in ribosomes.

The terms proteome and proteomics were coined, only as late as 1994, by Marc Wilkins, even if proteins have been studied since their discovery and, in particular, after their importance for all living organisms was recognized. The term proteome was abbreviated from "PROTEin complement of the genOME" meaning the complete set of the proteins expressed by the cell (Wilkins et al., 1996). Proteomics, for one, could be described as simply as "the large-scale analysis of proteins" (Pandey and Mann, 2000) or more thoroughly: "Proteomics includes not only the identification and quantification of the proteins, but also the determination of their localization, modifications, interactions, activities, and, ultimately, their function." (Fields, 2001)

Proteins are the active agents in the cells, tissues, and organisms and have an essential role in determining the phenotype of an organism. Even if all the organisms have one unique genome, the proteome of the organism can vary and create different phenotypes for the organism. As a classical example a caterpillar and a butterfly are often shown representing a common genome but different proteomes. Similarly, the phenotypes of many diseases are consequences of abnormalities in the proteome, even if they have genetic background. Human genome consists of ~ 2.85 billion nucleotides divided into 24 chromosomes. It is estimated to contain 20000-25000 genes coding for proteins (International Human Genome Sequencing Consortium, 2004; Lander et al., 2001; Venter et al., 2001). However, mRNA transcripts can be alternatively spliced and translated proteins can be cleaved or chemically modified in the cell. At least 200 different post-translational modifications of the proteins have been documented so far, creating a complex milieu of proteins in the cell (Krishna and Wold, 1993). In addition, functionality of the proteins often demands specific protein-protein interactions forming functional protein complexes (Dziembowski and Seraphin, 2004). Figure 1 schematically illustrates the complexity of the cellular proteome.

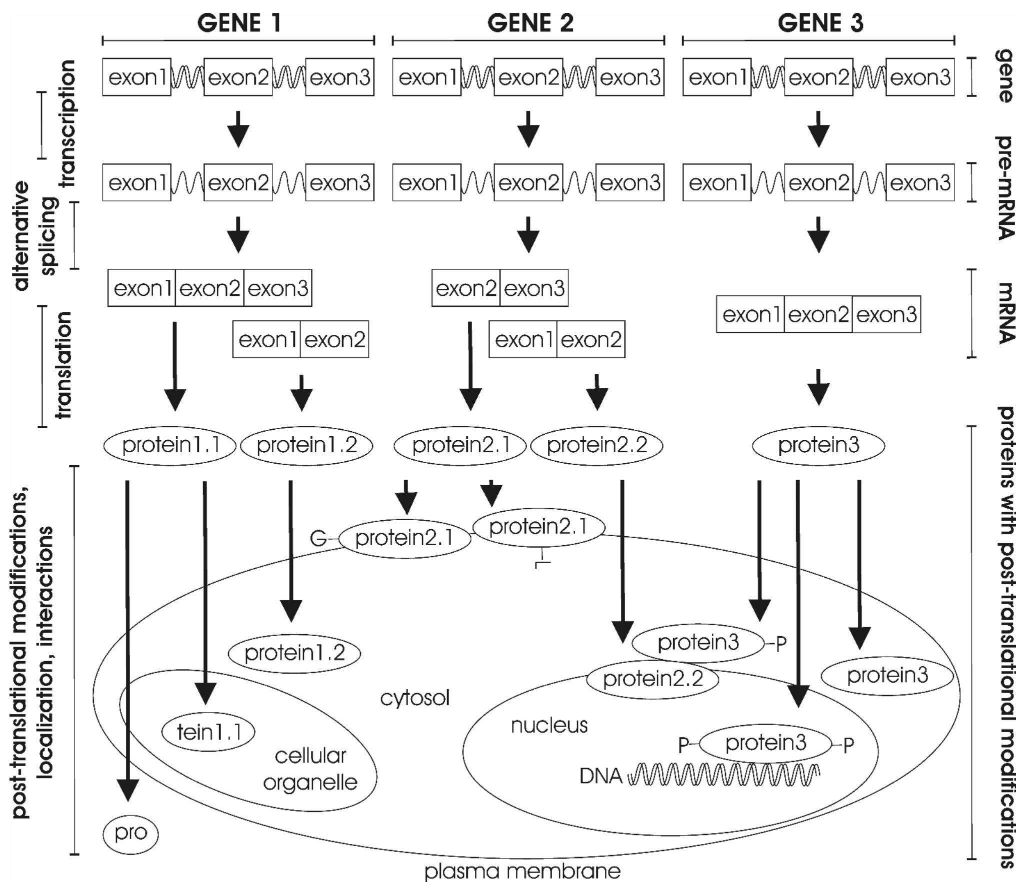


Figure 1: Proteome is much more complex than the corresponding genome. Figure 1 illustrates the proteome created from the expression of three genes. First the genes are transcribed to mRNA, which can be alternatively spliced (in the figure genes 1 and 2 produce two splice variants and gene 1 produces one). Resulting mRNA is translated then to proteins, which can be cleaved and/or modified with different kinds of post-tranlational modifications (in the figure protein 1.1 is cleaved, protein 2.1 is modified either with glycosylation or lipid, protein 3 is phosphorylated). Localization of the proteins and their interactions with other proteins and molecules plays an important role for their functionality. Post-translational modifications can affect to the protein localization and interactions, and *vice versa* (in the figure protein 1.1 is cleaved in the cellular organelle resulting to cleaved protein pro, which is secreted from the cell; protein 1.2 is a cytosolic protein; glycosylated and/or lipidoylated protein 2.1 localizes to the plasma membrane, while protein 2.2 localizes to the nuclear membrane; protein 3 is a cytosolic protein, which can be phosphorylated and binds to protein 2.2; diphosphorylated form of protein 3 localizes to nucleus and binds to DNA). Part of the figure adapted from Peng and Gygi (2001).

2.2 Methods used in proteomics

There is a wide variety of methods, reagents, instrumentation and data analysis tools available to design a proteomics experiment (reviewed for example in the following articles Pandey and Mann, 2000; Righetti et al., 2003; Stults and Arnott, 2005; Zhu et al., 2003). The proteomics experiment consists of four basic stages, namely 1) sample

preparation, 2) sample fractionation, 3) mass spectrometry analysis, and 4) data analysis. In quantitative proteomics proteins are usually labeled during sample preparation and quantified during sample fractionation (2-DE based proteomics) or MS analysis (MS based proteomics). Figure 2 summarizes these stages usually present in proteomics experiments. All the methods used in different stages are interlocked very closely with each other, for example sample preparation goes hand in hand with sample fractionation and a suitable MS analysis method should be selected based on the sample fractionation.

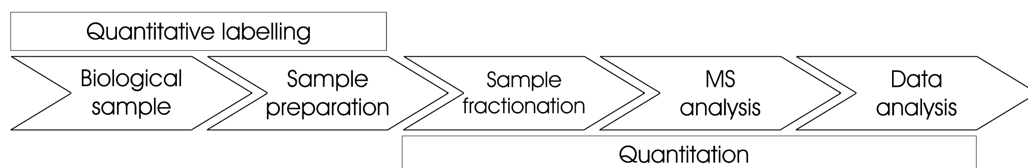


Figure 2: The experimental design in proteomics experiment commonly consists of four stages (sample preparation, sample fractionation, MS analysis, and data analysis). In quantitative proteomics abundances of the proteins are measured during sample fractionation (2-DE based proteomics) or MS analysis (MS based proteomics).

2.3 Sample preparation

The sample material used in proteomics studies varies a lot. Proteins can be extracted from many different sources such as cultured cells, tissues, body fluids etc. A successful proteomics experiment would benefit from having the genome sequence available for the studied organism. Genetic code can be translated in databases to protein sequences, which facilitates the identification of the proteins during data analysis. Without genetic information proteins can be identified either by *de novo* sequencing (Horn et al., 2000; Standing, 2003) or by comparing to the genomes of other organisms (Liska and Shevchenko, 2003) but it will make the data analysis much more complicated and does not favour the high-throughput analysis.

The classical phrase “junk in, junk out” best defines the importance of sample preparation in a proteomics experiment. Contaminated or otherwise poor quality samples just can not be processed into high quality data and results. So, all samples should be handled accurately with carefully chosen sample preparation methods. Special attention should be paid to prevent contamination of samples during sample preparation. Many chemicals, for example, those used in cell biology should be removed carefully from the samples before the MS analysis, because their presence can interfere with the analysis. Nonetheless, the most common contaminant keratin, which is present in all of us, emphasizes the importance of a careful sample handling.

Sample preparation starts usually by culturing and lysing cells, homogenizing tissues, or preparing body fluids to extract the proteins. The extracted proteins are denatured and chemically modified to solubilize them and to prevent their aggregation and protease activity within the sample. Proteins can be denatured either by heat or chemicals such as urea or different detergents (reviewed by Gorg et al., 2004 and

Molloy, 2000). However, detergents can create a background signal in MS that may mask peptide derived signals. Therefore, the amount of detergents should be kept minimal, or alternatively they should be removed or diluted, before protein digestion and MS analysis. Denatured proteins are further reduced and alkylated to prevent their aggregation via formation of crosslinks between cysteinyl thiols. The most commonly used reducing agents are beta-mercaptoethanol, dithiothreitol (DTT), tributylphosphine (TBP) and tris(2-carboxyethyl)phosphine (TCEP), while iodoacetamide, acrylamide derivatives and vinylpyridines are also used as common alkylating reagents (Bai et al., 2005; Herbert et al., 2001; Righetti, 2006; Sebastiano et al., 2003). The final step in the sample preparation is the digestion of the proteins to peptides. A lot of different chemicals and enzymes with different specifities are available for the digestion (see e.g. au.expasy.org/tools/peptidecutter/peptidecutter_enzymes.html). Trypsin is the most commonly used enzyme for protein digestion. It cleaves peptides very specifically after the basic amino acid residues lysine and arginine (Olsen et al., 2004a; Wilkinson, 1986). After trypsin digestion, the resulting peptides remain charged which improves their ionization in the MS analysis. Trypsin is also efficient and cost-effective to use.

2.4 Sample fractionation

Even if proteomics instrumentation has significantly improved during the last decade (Patterson and Aebersold, 2003), the whole cellular proteome is still far too complicated to measure. Therefore, sample fractionation is a crucial step in a proteomics experiment that can increase the number of protein identifications and allow detection of low-abundance proteins. The samples can be fractionated by using different prefractionation methods as well as by electrophoresis- and chromatography-based fractionation. There are many ways to prefractionate the sample during the sample preparation and by that way reduce the complexicity of the sample. Instead of using whole cell lysates subcellular organelles can be isolated and characterized (Pasquali et al., 1999). If protein complexes are studied, these can be enriched by using cloned epitope or affinity purification tags (Fritze and Anderson, 2000; Terpe, 2003). Some high-abundance or artificial proteins can be removed from the samples by specific antibodies, for example serum albumin from plasma samples (Steel et al., 2003). Even if there are many ways to prefractionate the sample, it is usually necessary to continue the fractionation with chromatography or electrophoresis or their combination before the MS analysis.

2.4.1 Subcellular fractionation

The localization of proteins in the cell is important for their functionality. Therefore studying subcellular compartments provides information about the localization of proteins and reduces sample complexity. There are many different methods available for isolation and purification of the organelles (reviewed by Pasquali et al., 1999). These include classical density-gradient centrifugation and affinity purification using antibodies detecting transmembrane proteins in different organelles. Different electrophoresis techniques such as free flow -, high resolution density gradient -, and

immune free flow electrophoresis can be used as well. Organelles can also be sorted with flow cytometry-based methods. During last few years a number of studies have characterized the proteomes of distinct cellular organelles (Andersen and Mann, 2006; Righetti et al., 2005; Yates et al., 2005). The number of protein identifications in these studies varies from tens to thousands of identified proteins depending on the studied organelle and the methods used. Table 1 shows some selected sub-proteome studies and numbers of identified proteins in them.

Table 1: Cellular organelles have been characterized in many subproteome studies. The table summarizes applied fractionation methods and number of protein identifications in the selected studies. The number of protein identifications in the parentheses indicates all the detected spots from the 2-DE gels. However, only some of the protein spots are prepared for the MS analysis.

Abbreviations in the table: ab; antibody based affinity chromatography; de, dimensional electrophoresis; dg, density gradient centrifugation; hilic, hydrophilic interaction chromatography; rp, reversed-phase chromatography; sol, solubility; scx, strong cation-exchange chromatography.

Cells:	Organelle:	Fractionation:	Protein identifications:	Reference:
human liver	nucleus	dg; 2-de	15 (1497)	Jung et al., 2000
hela cells	nucleolus	dg; 1-de; 2-de	213	Scherl et al., 2002
hela cells	nucleolus	dg; 1-de; rp; silac	489	Andersen et al., 2005
rat liver	mitochondrion	dg; 2-de	196 (1596)	Lopez et al., 2000b
human heart	mitochondrion	dg; 1-de	615	Taylor et al., 2003
rat liver	golgi	dg; rp; scx; hilic	421	Wu et al., 2004
rat liver	lysosome	dg; sol; 1-de; rp	215	Bagshaw et al., 2005
yeast	peroxisome	dg; ab; scx; avidin; rp; icat	70	Marelli et al., 2004
mouse liver	endoplasmic reticulum	dg; 2-de	141 (>2000)	Knoblach et al., 2003

2.4.2 Purification of protein complexes

In addition to the correct localization, the function of the proteins often requires interactions with other proteins. The most commonly used methods to purify protein complexes are cloned epitope and purification tags. These tags are usually short polypeptides or small proteins, which are fused to the target proteins by DNA cloning. Proteins of interest and interacting or complexed proteins are then purified by an antibody recognizing the tag by using immunoprecipitation or immunoaffinity chromatography. A number of different tags have been developed and succesfully used in characterization of functional protein complexes (Terpe, 2003). The TAP tag, for example, has been used in more than one hundred different studies (Dziembowski and Seraphin, 2004; Rigaut et al., 1999). Immunoprecipitation and DNA affinity purification are other commonly used methods to enrich protein complexes, which are then characterized by MS-based methods. Such a strategy was used, for example, to characterize protein complexes in yeast cells (Ranish et al., 2003).

2.4.3 Electrophoresis-based fractionation

Gel electrophoresis based techniques are used often in the proteomics studies to separate proteins. Both one- and two-dimensional gels are used depending on the complexity of the sample. In expression proteomics studies of complex protein mixtures, for example, two-dimensional gel electrophoresis (2-DE) is the most popular. 2-DE was originally developed independently by three different research groups (Klose, 1975; O'Farrell, 1975; Scheele, 1975). With 2-DE proteins are fractionated according to their isoelectric point and their molecular weight by isoelectric focusing (IEF) and SDS-PAGE, respectively. Proteins are quantified from the gels by different staining and labeling methods. After this quantitative analysis, the protein spots of interest are cut from the gel, and prepared for the MS analysis for protein identification.

2-DE has many benefits that explain its popularity in the proteomics studies. First, it visualizes the protein map of the studied proteome unlike chromatography-based fractionation methods. This enables easier identification of protein isoforms and their post-translational modifications. These protein maps can also be compared to other existing maps using databases. For example, the SWISS-2DPAGE database contains more than 1000 protein maps (Hoogland et al., 2004). The resolution of 2-DE gels is nowadays sufficient to routinely detect ~2000 proteins. Using larger gels or multiple narrow overlapping IPGs (immobilized pH gradients) allows identification of up to 5000-10000 proteins (Gorg et al., 2004; Wittmann-Liebold et al., 2006). The development of IPGs has removed former limitations of carrier ampholyte-based 2-DE gels such as technical reproducibility and detection of low-abundance, acidic, and basic proteins (Bjellqvist et al., 1982; Gorg et al., 2004). Methodological progress in the 2-DE sample preparation has also advanced the analysis of membrane proteins, which have been traditionally extremely difficult group of proteins to study with 2-DE (Molloy, 2000).

Capillary electrophoresis has also been successfully used to fractionate protein samples. As the name implies proteins are separated by electrophoresis using a capillary instead of a gel. Capillary electrophoresis separations are most often based on proteins charge-to-mass ratio, size, interactions with ligands, or hydrophobicity/ hydrophilicity (Huang et al., 2006).

2.4.4 Chromatography-based fractionation

Chromatography-based techniques are another powerful way to fractionate protein samples. In chromatography-based methods proteins are usually digested to peptides prior to their separation with multidimensional chromatography and automated MS/MS analysis. In multidimensional chromatography different chromatographic fractionations are performed online. Many different kinds of chromatographic stationary phase materials can be used for multidimensional chromatography, like reversed-phase (RP), size exclusion (SEC), strong cation exchange (SCX), anion exchange, affinity, and hydroxyapatite columns (Righetti et al., 2005). In the first proteomics applications of multidimensional chromatography either SEC or SCX were combined online with RP to fractionate peptides (Link et al., 1999; Opiteck et al., 1997). Since these initial experiments many other combinations of chromatography

separations have been introduced and used with HPLC instruments connected online to a MS/MS instrument.

The increasing popularity of the chromatography-based fractionation is a consequence of its capability to overcome some of the limitations encountered with 2-DE-based methods. Chromatography-based techniques have shown to be a powerful approach to study certain proteins that are difficult to study with 2-DE including membrane and low-abundance proteins (Blonder et al., 2004; Han et al., 2001; Peng and Gygi, 2001; Stockwin et al., 2006; Washburn et al., 2001; Wu et al., 2003). The reason for this is simple: In 2-DE proteins are solubilized and fractionated, while in multidimensional chromatography proteins are first digested to peptides and then fractionated. Multidimensional chromatography also allows higher throughput than 2-DE due to the better automation capability. Resolution of multidimensional chromatography is as excellent as with 2-DE enabling the identification of more than 1000 proteins. However, the detection of protein isoforms and post-translational modifications is more complicated than with 2-DE. Recently, the use of stable isotope labeling has increased the use of the chromatography-based fractionation in quantitative proteomics experiments.

Both the electrophoresis- and chromatography-based fractionation methods have advantages and disadvantages and the fractionation method should be selected carefully for each sample and study. These methods can be often used in parallel and by that way obtain complementary information (Kim et al., 2006; Kubota et al., 2003; Schmidt et al., 2004; Schmidt et al., 2006; Tian et al., 2004).

2.5 Biological mass spectrometry

Mass spectrometry (MS) is an analytical method used to measure mass-to-charge (m/z) ratio of molecules. This is achieved by ionisizing the molecules, separating them according to their mass-to-charge ratio, and finally detecting their abundances. The first mass spectrometers were developed in the beginning of the 20th century. However, the development of soft ionization techniques in late 1980's revolutionized the use of MS in life sciences and enabled the analysis of biomolecules including proteins (Fenn et al., 1989; Karas and Hillenkamp, 1988; Tanaka et al., 1988). The importance of electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) techniques were acknowledged in 2002, when their developers were awarded with the Nobel prize in chemistry. Nowadays more powerful MS instruments are available and MS is used routinely in most well-equipped laboratories. In proteomics applications mass spectrometers are used in three major areas; (1) to identify proteins, (2) to detect their covalent modifications including post-translational modifications, and (3) to characterize and control the quality of recombinant proteins (Mann et al., 2001).

2.5.1 MS instrumentation

Different types of mass spectrometers, compatible for peptide and protein analysis, have emerged since the development of soft ionization techniques. In principle, all MS instruments consist of an ion source, a mass analyser, and a detector. These

components can be combined to different configurations to create mass spectrometers with diverse features in terms of speed, resolution, sensitivity, accuracy, mass range, and price (Aebersold and Mann, 2003; Domon and Aebersold, 2006).

In the MS analysis molecules are ionized to gas phase. The ESI is used to ionize the molecules present in liquid phase, while with MALDI molecules are crystallized with matrix and then ionized with laser pulses. Then the ions enter the mass analyser, in which they are separated according to their m/z. The most commonly used mass analysers are quadrupole (Q), time-of-flight (TOF), ion trap (IT), and Fourier transform ion cyclotron resonance (FTICR). From the mass analyser the ions enter the detector, which measures the abundances of each ion with specific m/z. Todays mass spectrometers can detect routinely femtomole quantities of peptides, but even zeptomole sensitivity can be achieved (Belov et al., 2000). At the same time the most accurate FTICR instruments can reach the resolution > 1000000 (meaning the separation of m/z 1000.001 from 1000.000) (Wysocki et al., 2005).

2.5.2 Identification of proteins with mass spectrometry

Peptide mass fingerprinting (PMF) by MS (Henzel et al., 1993; James et al., 1993; Mann et al., 1993; Pappin et al., 1993; Yates et al., 1993) and peptide sequencing by MS/MS (Hunt et al., 1986; Johnson and Biemann, 1987) are the most commonly used MS-based methods to identify proteins. In both methods proteins are digested to peptides, which are then analysed by MS. In the PMF analysis the studied sample should contain peptides from only one or at most a few different proteins. Therefore, PMF is an optimal way to identify proteins derived e.g. from 2-DE gels. In the PMF analysis peptide masses of the digested protein are simply measured by MS. Different search algorithms are then used to match the detected peptide masses against the theoretical peptide masses of different proteins that are present in the protein databases. The PMF is a simple and robust method to identify purified proteins, however MS/MS is needed to analyse complex protein mixtures. In the MS/MS instruments two mass analysers are joined together. In the first analyser a peptide with specific m/z ratio is selected from the complex mixture of peptides. The selected peptide is then dissociated to multiple fragment ions by collision with inert gas. The resulting fragment ion masses are then measured in the second mass analyser. The detected fragment ion masses are used to determine amino acid sequence of the peptide and to identify the corresponding protein. Proteins can be identified from MS/MS spectra by two main approaches, namely *de novo* sequencing and peptide fragment fingerprinting (PFF) (Hernandez et al., 2004). In both approaches peptide sequences are used for protein identification but the major difference is that PFF uses protein and/or DNA sequences present in the databases for peptide sequencing, while *de novo* sequencing is performed without the database information. Powerful computer software algorithms assist the interpretation of the MS/MS spectra in both approaches.

2.6 Quantification & quantitative labeling

Quantification means measuring a property existing in a range of magnitudes, for example, weight, volume, and concentration are typical quantities measured in

chemistry. Quantification can be divided to absolute and relative quantification. Within quantitative proteomics the latter one is much more commonly used. In relative quantification protein abundances are measured between two or more different biological conditions and compared with each other. All the quantitative methodologies described in this thesis represent relative quantification. However, it should be noted that absolute quantities can also be measured in proteomics, for example, by adding known concentrations of internal standard peptides to the samples and comparing their quantitities to the studied peptides present in the sample (Gerber et al., 2003).

Measuring peptide abundances within complex biological samples with a mass spectrometer is not a simple task. Different peptides have different chemical and physical properties, which have an impact on their preparation and fractionation, and finally on their ionization and volatilization in the MS analysis. Therefore, quantities of different peptides can not be compared with each other. In principle, a peptide with low quantity but with good ionization properties might show in the MS spectrum higher intensity than a peptide with higher quantity but worse ionization properties. Therefore, only quantities of identical peptides can be compared with each other. Relative quantification methods used in proteomics studies can be divided to two main categories, namely gel- and MS-based quantifications.

2.6.1 Gel-based quantification

In gel-based quantification proteins are detected and quantified from the gel used for sample fractionation. Due to the popularity of 2-DE in proteomics, a great number of different staining and labeling methods have been developed to detect and quantify proteins from gels (Miller et al., 2006; Patton, 2002). Radioactive labeling and non-radioactive staining with dyes are the main techniques used in gel-based quantification.

2.6.1.1 Metabolic labeling with radioactive labels

For metabolic labeling, cells are cultured either in normal medium or in medium containing amino acids labeled with radioactive isotopes, for example ^{35}S-labeled methionine (Kettman et al., 1986; Westbrook et al., 2001). During cell culture these radioactive amino acids are incorporated into cellular proteins during their synthesis. These radioactively labeled proteins are detected with autoradiography from a 2-DE gel after sample preparation and fractionation. Protein spots of interest can be then cut from the gel and identified either with PMF by MS or with peptide sequencing by MS/MS. Figure 3a summarizes the steps used in the metabolic labeling of proteins with ^{35}S-labeled methionine. Metabolic labeling with radioactively labeled amino acids offers a great sensitivity to detect even low-abundance proteins. Unfortunately metabolic labeling has some serious disadvantages such as the hazardous nature and high-cost of radioactively labeled amino acids. Also, metabolically incorporated radiolabels affect cells by inducing DNA fragmentation, elevating p53 tumor suppressor protein levels, altering cellular morphology, and by causing cell cycle arrest or apoptosis (Patton, 2002). Non-radioactive quantification methods have been developed to avoid these disadvantages.

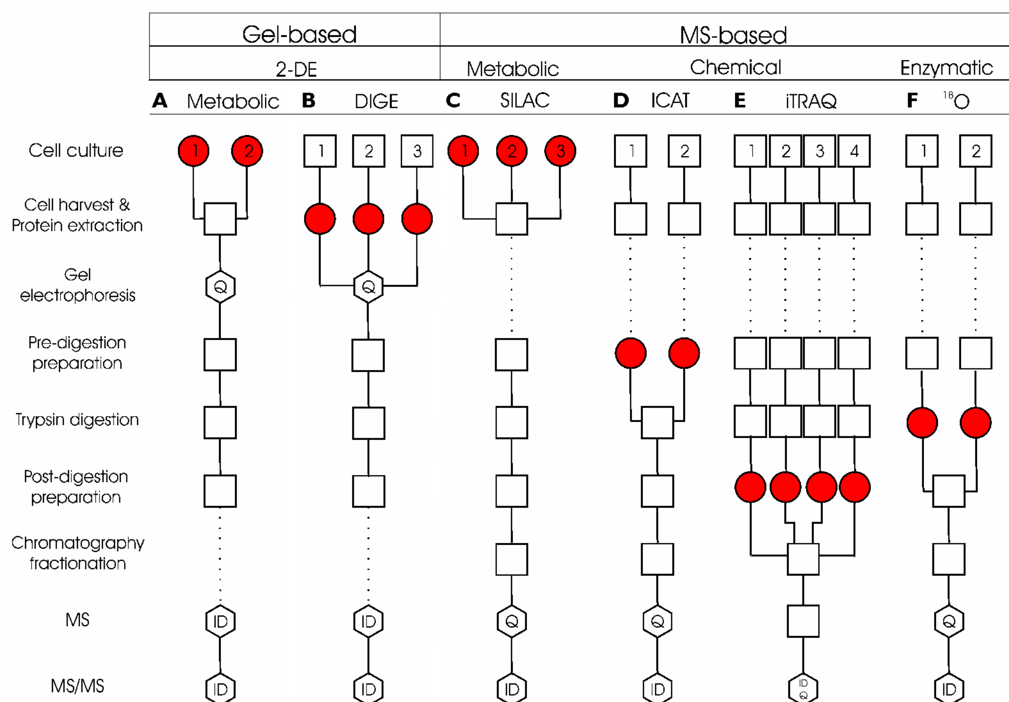


Figure 3: A flow-chart presentation of different stages present in the quantitative proteomics experiments: A) Metabolic labeling, B) DIGE, C) SILAC, D) ICAT, E) iTRAQ, and F) enzymatic ^{18}O incorporation. Methods A and B present gel-based methods, while C-F present MS-based methods. Stable isotopes are incorporated in the MS-based methods metabolically (C), chemically (D, E), or enzymatically (F). Different symbols and solid lines in the figure indicate the stages performed in the method, while the stages marked with dashed lines are not necessary in the method. Numbers inside the symbols describe the number of samples that the method is capable to analyse. Red circles describe the labeling stage, and hexagons with id and/or q indicate the identification and quantification stages, respectively.

Abbreviations: DIGE, difference gel electrophoresis; ICAT, isotope-coded affinity tag; iTRAQ, isotope tagged relative and absolute quantitation; SILAC, stable isotope labeling with amino acids in cell culture

2.6.1.2 Protein staining in the gels with dyes

A major alternative for radioactive labeling within gel-based quantification is the use of different dyes to stain the proteins in the gels. Proteins can be stained either with colorimetric (Coomassie, silver staining) or fluorescence (SYPRO, Cyanine based dyes) detectable dyes either before or after electrophoresis. Table 2 summarizes these staining techniques. Patton (2002) and Miller et al. (2006) have carefully reviewed different protein staining techniques available.

Table 2: SDS-PAGE separated proteins can be stained from the gels for their quantification. Table summarizes the features characteristic for Coomassie, silver, SYPRO ruby, and CyDye stainings.

Dye:	Coomassie:	Silver:	SYPRO Ruby:	CyDyes:
Detection:	colorimetric	colorimetric	fluorescence	fluorescence
Staining:	post-electrophoresis	post-electrophoresis	post-electrophoresis	pre-electrophoresis
Detection limit (ng):	8-10	~ 1	~ 1	~ 2
Linearity range (fold):	10-30	10	1000	1000
# of samples / gel:	1	1	1	2-3
MS compatibility:	++	(+)	+++	++
Comments:		not an endpoint stain, overstaining		minimal protein labeling

Both Coomassie (Fazekas de St Groth, S. et al., 1963) and silver staining methods (Switzer et al., 1979) were developed several decades before the current era of proteomics, yet they are still standard methods to detect proteins in gels. In both methods the proteins are separated by PAGE gels, which are then stained with the dye. Coomassie stain is a disulfonated triphenylmethane textile dye, which binds to the amino groups of proteins in acidic solutions. The benefits of Coomassie staining are its easiness of use, low cost, and compatibility with downstream methods such as MS. Coomassie staining has been optimized a lot to reduce the background staining of the gels as well as to improve the sensitivity and linear range of detection (Candiano et al., 2004; Diezel et al., 1972; Neuhoff et al., 1985). However, the sensitivity and linear range of Coomassie are still the major limitations of the method, in addition to difficulty to control the staining reproducibility between the gels (Patton, 2002). Silver staining is usually performed by using silver nitrate and formaldehyde developer in alkaline carbonate buffer although many alternative protocols exists (Rabilloud, 1990). In silver staining, the gels are saturated with silver, which binds more tightly proteins than the PAGE gel. Gel bound silver is removed by washing steps, while silver bound to the proteins is reduced to visible metallic form. Silver staining is relatively sensitive method enabling the detection of protein quantities of 1 ng but even 0.1 ng levels have been detected (Heukeshoven and Dernick, 1988). The narrow linear range is the major disadvantage of silver staining, and thus the quantification results are not very reliable (Patton, 2002). Moreover silver staining reaction needs to be stopped at an arbitrary time point to prevent over staining. This causes problems with the reproducibility of the stained gels (James et al., 1993). Silver staining also interferes later with the MS analysis by reducing the number of identified peptides and the sequence coverage of the proteins (Scheler et al., 1998). However, careful optimization of the staining protocols and use of robotic staining devices (Sinha et al., 2001) has improved the quantitative results obtained with the silver staining (Giometti et al., 1991).

The development of fluorescent dyes has improved dramatically quantification of proteins from the 2-DE gels (Patton and Beechem, 2002; Patton, 2002). The dyes have increased the dynamic range of quantification as well as the reproducibility. Proteins are stained with fluorescent dyes either before or after electrophoresis, and they are detected after the dye is excited by UV or laser light. SYPRO Ruby is a post-

electrophoresis staining method, while CyDyes are used to label the proteins before electrophoresis. SYPRO Ruby is a ruthenium metal chelate, which binds to basic amino acid residues (Berggren et al., 1999). It is as sensitive as silver staining, but it offers a significantly broader linear dynamic range. It is also easy to use thus improving the reproducibility of the quantitative results. SYPRO Ruby staining is also compatible with MS further stressing its superiority in comparison to conventional colorimetric methods like Coomassie and silver staining (Lopez et al., 2000a). Another option to detect proteins in gels with fluorescent dyes is to use CyDyes, which are used in a method referred to as difference gel electrophoresis (DiGE) (Tonge et al., 2001; Unlu et al., 1997). These succinimidyl esters of the cyanine dyes Cy2, Cy3, and Cy5 covalently link to lysine residues in proteins. Proteins are labeled with CyDyes before 2-DE and three different samples can be labeled, each with a different dye. After labeling the samples are combined and electrophoresed simultaneously on a single 2-DE gel. Finally the stained proteins are detected from the gel with a fluorescence scanner by exciting the dyes with their charasteristic wavelengths of light. DiGE protocol is schematically summarized in Figure 3b. DiGE allows the detection of quantitative differences in protein expression between the samples from one gel, which at the same time both improves the reproducibility of the results and reduces the work needed. Sensitivity of CyDyes is between silver and Coomassie staining, but the dynamic range is much wider. Disadvantages concerning CyDyes mainly relate to necessary minimal labeling of proteins (1-3%) required to keep them soluble during the electrophoresis. This causes a small change in the molecular masses of the labeled proteins resulting in slightly different migration of unlabeled and labeled proteins in SDS-PAGE (Tonge et al., 2001). This is a problem when the detected protein spot, representing the labeled minority of the protein, is cut off from the gel for MS analysis. Therefore, CyDye labeled gels are often stained with another stain, for example, SYPRO Ruby before cutting the spots (Gharbi et al., 2002). Another drawback of the methodology is the high cost of the dyes and the equipment to detect them.

2.6.2 MS-based quantification – isotope-coded tags

In MS-based proteomics relative quantification is often achieved using protein or peptide labeling with stable isotopes. One of the samples is labeled with a light (= natural) isotope, while the other sample is labeled with the corresponding heavy isotope. Thus peptides, labeled either with light or heavy isotope, produce peak pairs in MS analysis with the mass difference of the used isotopes. Relative abundances of the peptides can be measured by comparing areas of these peaks. Alternatively, peptides can be labeled with isobaric tags, in which case the labeled peptides have equal masses. However, in MS/MS analysis these tags produce distinguishing fragment ions, which can be used for the quantification.

Stable isotopes can be incorporated to the samples metabolically, chemically or enzymatically (Fig. 3) (Corthals and Rose, 2007; Goshe and Smith, 2003; Ong and Mann, 2005). In metabolic labeling stable isotopes are incorporated to the proteins during cell culture. Enzymatic incorporation, in turn, is performed during the trypsin digestion of the proteins. In chemical labeling either proteins or the digested peptides are labeled with isotope-coded reagents. Isotope-coded tags are often used to analyse

complex protein mixtures. Therefore, careful sample fractionation is essential for comprehensive identification of different proteins.

Recently, MS-based proteomics applications without protein labeling has been developed. In such applications peptides derived from different protein samples are analysed in multiple LC-MS analyses, so that each analysis represent one of the samples. The LC-MS data sets derived from different samples are aligned by powerful software programs. Then, the areas of selected peptide peaks can be compared to determine the quantities of the peptide present in different samples.

2.6.2.1 Chemical labeling with stable isotopes – ICAT and iTRAQ

Chemical labeling is a commonly used way to incorporate stable isotopes to proteins or peptides. Different isotope-coded chemicals are available with different binding specifities, stable isotopes, and purification tags (Leitner and Lindner, 2004; Leitner and Lindner, 2006). Isotope-coded reagents react chemically with proteins or peptides, usually either with the amino group present in N-terminus of peptides and in lysine residues or with the sulfhydryl groups present in cysteines. There are also reagents with chemical reactivity with carboxyl groups, tryptophan residues, and phosphorylated serine and threonine residues. The stable isotopes used in these reagents are most often hydrogen/deuterium ($^{1}H/^{2}H$), carbon ($^{12}C/^{13}C$), and nitrogen ($^{14}N/^{15}N$). The reagents may also include a purification tag to specifically isolate the labeled peptides, as for example in ICAT reagents, where a biotin tag enables avidin chromatography purification. Table 3 represents isotope-coded reagents available for protein and peptide labeling (reviewed recently by Corthals and Rose, 2007; Leitner and Lindner, 2004; Leitner and Lindner, 2006; Ong and Mann, 2005).

Table 3: Stable isotope labeling reagents used for chemical labeling of proteins and peptides. Table summarizes the specificity of each reagent and the stable isotopes present in the reagent. (* metal ions in chelate).

Target:	Reagent or method:	Isotope:	References:
Amine	iTRAQ (Isotope tagged relative and absolute quantitation)	C, N, O	Ross et al., 2004
Amine	Sulfopropionic acid N-hydroxysuccinimide ester	C	Keough et al., 2003
Amine	N-acetoxysuccinimide	H	Chakraborty and Regnier, 2002; Ji et al., 2000
Amine	Nicotinoyloxysuccinimide (Nic-NHS)	D	Munchbach et al., 2000; Schmidt et al., 2005
Amine	Phenyl isocyanate	H	Mason and Liebler, 2003
Amine	Isotope coded n-terminal sulfonation (4-sulphophenylisothiocyanate)	C	Lee et al., 2004; Samyn et al., 2004
Amine	Anhydrides: Succinic, acetic, propionic	H	Che and Fricker, 2002; Nam et al., 2005; Noga et al., 2005; Wang and Regnier, 2001; Yu et al., 2004; Zappacosta and Annan, 2004; Zhang et al., 2002b
Amine	Formaldehyde	H	Hsu et al., 2003
Amine	N-isotag (aminobutyric acid)	H	Smolka et al., 2005
Amine	Sulfo-NHS-SS-biotin and methyliodide	H, C	Hoang et al., 2003
Amine	Trimethylammonium butyrate	H	Riggs et al., 2005
Amine, Lys	2MEGA (N-terminal dimethylation (2ME) after lysine guanidination (GA))	H, C	Ji et al., 2005
Carboxyl	Methyl/ethyl esterification	H	Goodlett et al., 2001; Syka et al., 2004
Cys	Acrylamide	H	Cahill et al., 2003; Gehanne et al., 2002; Sechi, 2002
Cys	ALICE (Acid-labile isotope-coded extractant)	H	Qiu et al., 2002
Cys	ICAT cleavable	C	Hansen et al., 2003; Li et al., 2003b; Oda et al., 2003
Cys	ECAT (Element coded affinity tags)	*	Whetstone et al., 2004
Cys	HysTag	H	Olsen et al., 2004b
Cys	ICAT	H	Gygi et al., 1999
Cys	Iodoacetanilide	C	Niwayama et al., 2003
Cys	N-t-butyliodoacetamide	H	Pasquarello et al., 2004
Cys	ICAT solid phase	H	Zhou et al., 2002
Cys	Solid phase mass tagging (iodoacetmide)	C	Shi et al., 2004
Cys	Vinylpyridine	H	Sebastiano et al., 2003
Cys	N-alkylmaleimide	H	Niwayama et al., 2001
Cys	N-ethyliodoacetamide (ICROC)	H	Shen et al., 2003
Cys	IBTP [(4-iodobutyl)triphenylphosphonium]	H	Marley et al., 2005
Cys	VICAT (visible isotope-coded affinity tag)	C, N	Lu et al., 2004b
Lys	2-methoxy-4,5-dihydro-1H-imidazole	H	Peters et al., 2001
Lys	Guanidination (O-methylisourea)	C, N	Brancia et al., 2004
Lys	MCAT (mass coded abundance tagging) [Guanidination (o-methylisourea)]	-	Cagney and Emili, 2002
Lys	QUEST (Quantitation using enhanced sequence tags) [Guanidination (o-methylisourea)]	-	Beardsley and Reilly, 2003
O-GlcNac	BEMAD (Beta eliminationand Michael addition with dithiothreitol)	H	Wells et al., 2002
P	Phiat Phosphoprotein isotope-coded affinity tag)	H	Goshe et al., 2001
P	Phist (Phosphoprotein isotope-coded solid-phase tag)	C, N	Qian et al., 2003
Trp	2-nitrobenzenesulfonyl chloride	c	Kuyama et al., 2003
Variable	Tandem mass tag	H	Thompson et al., 2003

The development of ísotope-coded affinity tag (ICAT) labeling (Gygi et al., 1999) was a breakthrough in MS-based quantitative proteomics that enabled the high-throughput identification and quantification of the proteins within complex protein mixtures. The ICAT reagent consists of three functional parts, namely reactive group, isotope-coded linker, and biotin tag (Fig. 4). The reactive group is iodoacetamide, which covalently links the reagent to sulfhydrul groups of cysteines. The biotin tag enables the purification of the ICAT labeled peptides with avidin affinity chromatography. The linker chain between iodoacetamide and biotin is coded with eight deuterium atoms in heavy reagent, while light reagent has eight hydrogen atoms. In cleavable ICAT reagent nine carbon isotopes ($^{12}C/^{13}C$) were used instead of the hydrogen isotopes. With the ICAT labeling protocol, proteins from two samples are labeled, one with light and other with heavy reagent, and then mixed and digested with trypsin. The resulting peptides are first fractionated, for example with SCX chromatography, to several fractions, which are then purified by avidin chromatography. Affinity purified ICAT labeled peptides are separated by RP chromatography in nanoLC prior to MS/MS analysis. In the MS analysis the peptides are quantified by comparing the areas of light and heavy labeled peptide peaks and the peptides are identified by MS/MS. Figure 3d schematically describes the ICAT labeling protocol.

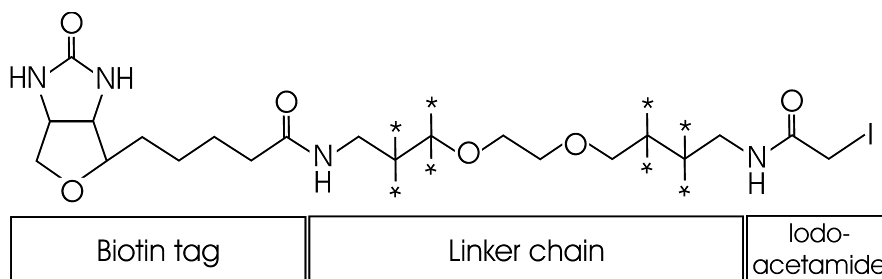


Figure 4: Isotope-coded affinity tag reagent includes cysteine reactive iodoacetamide, a linker chain coded with light or heavy isotopes (indicated with *), and a biotin tag.

ICAT labeling has been used in many proteomics studies to obtain novel insights into cellular mechanisms (Han et al., 2001; Karsan et al., 2005; Li et al., 2003a; Martin et al., 2004; Meehan and Sadar, 2004; Shiio et al., 2002), protein complexes (Ranish et al., 2003; Zhou et al., 2004), and cellular organelles (Dunkley et al., 2004; Marelli et al., 2004). It has also been successfully used to study membrane proteins, which have been difficult to study by 2-DE-based methods (Dunkley et al., 2004; Han et al., 2001; Marelli et al., 2004; Tam et al., 2004). One important benefit of the ICAT methodology is the potential for automated high-throughput analysis. However, such analyses create enormous amounts of data, which is difficult to handle and analyse. Multiple powerful protein identification and quantification algorithms exist, but it is still relatively slow to validate all results obtained from the ICAT analyses. The detection of post-translational modifications with the ICAT is also a more complex than with 2-DE based methods, because the ICAT reagent reacts only with cysteine residues.

Following the success of the ICAT labeling Applied Biosystems developed a new generation of isotope coded reagents, namely isotope tagged relative and absolute

quantitation (iTRAQ) (Ross et al., 2004). There are four isotopically different forms of the iTRAQ reagents (nowadays eight) that are composed of a reactive, a balance, and a reporter group (Fig. 5). The reactive group is N-oxysuccinimide, which reacts with the amino groups of lysine residues and N-termini of peptides. The reporter groups are isotopically labeled N-methylpiperazine derivatives, which have molecular masses of 114, 115, 116 and 117 in the different reagents. These reporter ions are detected in MS/MS analysis and used for peptide quantification. The balance group equalizes the masses of the different reagents, so that the total mass of each reagent is 145. In the iTRAQ labeling protocol proteins are first digested with trypsin to peptides, which are then labeled with the different iTRAQ reagents. All the labeled samples are combined and fractionated, for example, with SCX chromatography and the resulting fractions are analysed by LC-MS/MS. Proteins are identified based on the peptide sequences obtained from the MS/MS analysis and quantified on the grounds of the reporter ion intensities that are fragmented in the MS/MS analysis. Figure 3e summarizes the iTRAQ labeling protocol.

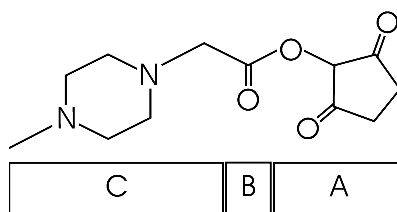


Figure 5: Isotope tagged relative and absolute quantitation reagent consist of amino group reactive N-oxysuccinimide (A), a balance group (B), and a reporter group (C).

iTRAQ reagents represent an advanced labeling method, which allows simultaneous analysis of four samples. In the second generation iTRAQ kit even eight samples can be analysed at the same time. Due to the isobaric nature of the tag all the peptides labeled with the different tags have the same molecular mass meaning four times smaller sample consumption, when compared to traditional tags like ICAT. The iTRAQ reagent is incorporated to all the peptides present in the sample, what increases the confidence of the protein identifications due to higher sequence coverage. However, as iTRAQ does not reduce the sample complexity like ICAT, fractionation is of great importance in iTRAQ experiments. The bottleneck in the iTRAQ experiments, as in the ICAT, is the data handling, storage, and analysis, as very large amounts of data are generated, even from a single sample. Zieske (2006) and Gan et al. (2007) have recently reviewed the iTRAQ labeling applications.

2.6.2.2 Metabolic labeling with stable isotopes: SILAC

Stable isotopes can be incorporated to proteins also metabolically during the cell culture. Cells are grown both in normal medium and in medium enriched with stable isotope labeled compounds such as ^{15}N-labeled salts, ^{13}C-labeled glucose, or amino acids labeled with different stable isotopes (Ong and Mann, 2005). In metabolic labeling cells incorporate the stable isotopes to the synthesized proteins and by that

way the whole proteome is converted to isotopically labeled one. Cells cultured in both conditions are combined already during harvesting, and therefore all sample preparation and fractionation steps are identical for both samples. Simultaneous sample handling improves the accuracy of the quantification. Otherwise, sample preparation and fractionation, as well as identification and quantification of the proteins, are almost similar to the ICAT labeling method. Figure 3a illustrates the principle of metabolic labeling with stable isotopes.

In the initial reports where stable isotopes were used for metabolic labeling, ^{15}N and ^{13}C enriched medium was used (Langen et al., 1998; Oda et al., 1999). The major drawback in this methodology was that the mass change of the peptides could not be predicted, because both backbone and side-chain nitrogen atoms of the peptides were labeled. This makes the quantification of the peptides difficult in complex samples. The development of stable isotope labeling with amino acids in cell culture (SILAC) has circumvented the problem encountered with ^{15}N- and ^{13}C-labeling (Ong et al., 2002). In SILAC, cells are cultured either in a normal medium or in a medium where certain amino acids are labeled with stable isotopes. Therefore, either non-labeled or labeled amino acids are incorporated to the cellular proteins during protein synthesis that creates a predictable mass difference between the labeled peptides. Most commonly used stable isotope labeled amino acids are ^{2}H-leucine, ^{13}C-lysine and $^{13}C/^{15}N$-arginine (Ong et al., 2002; Ong et al., 2003). ^{2}H-leucine is an economical choice and labels ~70% of the amino acids, while all the tryptic peptides are labeled by using ^{13}C-lysine and $^{13}C/^{15}N$-arginine. Recently three different forms of isotopically labeled arginine were used to compare three samples simultaneously (Blagoev et al., 2004). The major benefit of the SILAC and other metabolic labeling techniques is the simultaneous sample preparation and fractionation and their major limitation is that they can be applied only for cell culture samples, so analysing for example proteins from body fluids or tissues is not possible. The interpretation and handling of the data is also quite difficult, as it is with all the MS-based quantification methods.

2.6.2.3 Enzymatic labeling with stable isotopes: Proteolytic ^{18}O labeling

Peptides can be labeled with stable isotopes also during their enzymatic digestion. Most commonly ^{18}O is incorporated to the peptides during trypsin digestion, although other enzymes like Lys-C can be used as well (Desiderio and Kai, 1983; Miyagi and Rao, 2007). Proteins extracted from two different cell states are digested with trypsin in a buffer including either regular H_2O or $H_2{}^{18}O$ (Yao et al., 2001). ^{18}O from the water is transferred to the C-termini of the peptides during the cleaving reaction. These peptides are fractionated and analysed by MS and MS/MS basically in similar way as ICAT or SILAC labeled peptides (Fig. 3f). With enzymatic labeling all peptides present in the sample are labeled that creates a highly complex peptide mixture. Therefore, careful sample fractionation is needed. The incorporation efficiency of ^{18}O during trypsin digestion is another problem related to proteolytic ^{18}O labeling. Some peptides can be even less than 50% labeled with ^{18}O after trypsin digestion, which complicates the quantification (Hicks et al., 2005).

2.7 Proteomics data analysis

Today's proteomics research facilities with MS instruments capable of high-throughput analyses create enormously data. One automated LC-MS/MS run generates thousands of spectra and multiple runs are done every day. As manual interpretation of such data amount is not possible powerful software algorithms and data processing pipelines are necessary. Even though a lot of effort has been invested to developing such software tools, data analysis is still often the bottleneck in proteomics studies. The main tasks of the software programs are identification and quantification of proteins from data involving thousands of MS and MS/MS spectra. In addition, tools for data storing, visualization, validation, and format changing are needed. Recent review articles by Palagi et al. (2006) and Lisacek et al. (2006) summarize the most common bioinformatics tools used to process and analyse proteomics data.

2.7.1 2-DE image analysis software

Protein maps on 2-DE gels can easily have > 1000 protein spots and multiple gels are commonly analysed in a single 2-DE experiment. Accordingly, many software programs have been developed to analyse gel images (recently reviewed by Palagi et al., 2006 and Raman et al., 2002). The basic functions of these software packages include visualization of the gels, spot detection, matching protein maps of different gels, and quantification of the protein spots (Raman et al., 2002).

2.7.2 Software tools for protein identification

Peptide mass fingerprinting by MS and sequencing by MS/MS are currently the mostly used MS-based approaches to identify proteins in proteomics studies. Different software tools have been developed to assist the protein identification by both of the approaches. These software programs compare the experimental spectra with theoretical spectra derived from DNA or protein sequence databases such as UniProt (www.ebi.uniprot.org) and NCBI (www.ncbi.nlm.nih.gov). If there is no genome sequence information available for the studied organism, proteins can be identified by *de novo* sequencing (Hernandez et al., 2006; Standing, 2003).

In PMF peptide masses of a specifically digested protein are simply measured by MS. These observed peptide masses are then compared to theoretical peptide masses of *in silico* digested proteins. Different kind of of PMF tools such as Mascot (Perkins et al., 1999), MS-Fit (Clauser et al., 1999), Profound (Zhang et al., 2002b), PeptideSearch (Mann and Wilm, 1994), PeptIdent (Wilkins and Williams, 1997), Aldente (Tuloup et al., 2003) and PepFrag (Fenyo et al., 1998) assist to evaluate the similarity between experimental and theoretical spectra. All these programs give a score, which describes how well the experimental and theoretical spectra match with each other. Many factors are taken into consideration, when the matching spectra and identified proteins are scored. These include the accuracy of matching peptide peaks, intensities of the peaks, missing peptide peaks, modified amino acids, missed or non-specific cleavages during protein digestion, errors in database sequences, calibration of the instrument, and peaks

originating either from other peptides and contaminant molecules, or background noise (Palagi et al., 2006). In some algorithms the scoring is based simply on the number of matching masses, while the others use more sophisticated algorithms like MOWSE, which is based on the distribution of peptide masses in the databases (Pappin et al., 1993). Due to the differences in search algorithms the results of these PMF may differ. Chamrad et al. (2004b), for example, compared Profound, Mascot, and MS-Fit and noticed that MS-Fit was able to identify less of proteins from their dataset than the two other algorithms. Therefore, it is beneficial to validate the results obtained by one algorithm with another one or alternatively use statistical methods to estimate the significance of identifications.

Sequencing by MS/MS can be done either by *de novo* sequencing or peptide fragment fingerprinting. The latter one is usually used for identification of proteins within complex mixtures, when there is genome sequence information available for the studied organism. In automated MS/MS analyses, the mass spectrometer selects peaks from the MS spectrum for fragmentation and MS/MS analysis. Such analyses easily create hundreds of spectra/hour that illustrates the requirement of efficient data analysis programs. These high-throughput proteomics analyses have facilitated the development of many software packages for analysing MS/MS data such as Mascot (Perkins et al., 1999), SEQUEST (Eng et al., 1994), ProbID (Zhang et al., 2002a), Phenyx (Colinge et al., 2003), X!Tandem (Craig and Beavis, 2004), GutenTag (Tabb et al., 2003), NoDupe (Tabb et al., 2003), Sonar (Field et al., 2002), ProID&ProICAT (ABI), and SpectrumMill (Agilent). All these tools compare experimentally collected MS/MS spectra with theoretical MS/MS spectra of peptides derived from either DNA or protein database sequences and provide a score for the matching spectra. Each software tool applies different specialized algorithms and scoring functions to take into account factors influencing the MS/MS identification of proteins. These are, for example, the presence of contaminants and co-eluting peptides in the analytes, missed and non-specific proteolytic cleavages, modified amino acids, non-annotated mutations and splice variants of proteins, accuracy of the peptides precursor masses as well as fragment masses and their intensities, and sequencing errors present in the databases (Hernandez et al., 2006; Palagi et al., 2006). The performance of some software packages and implemented algorithms has been evaluated and compared. Chamrad et al. (2004b) compared the performance of Mascot and SEQUEST, while Moulder et al. (2005) evaluated the programs SEQUEST, SpectrumMill, and ProICAT. Later Kapp et al. (2005) compared SEQUEST, Mascot, SpectrumMill, X!Tandem, and Sonar. In all these comparisons the different software programs produced overlapping protein identifications, especially with high-abundance proteins, but also unique protein identifications for each individual software program.

There are many MS instruments and software packages available for analysing different proteomics samples. This abundance of methods and tools easily creates variability to the results. Comparing software programs clearly demonstrated that the different software tools produce somewhat different results, even if the search parameters are adjusted as similarly as possible (Chamrad et al., 2004b; Kapp et al., 2005; Moulder et al., 2005). This variability is further increased when different users have different parameters in the data processing (Chamrad et al., 2004b; Omenn et al., 2005). This highlights the importance of validating the protein identifications in order

to increase the confidence of the correct identifications and to prevent false positive identifications. There are many options available to validate the results in addition to manual validation (Nesvizhskii and Aebersold, 2004). Multiple identification algorithms can be used for identification of the proteins. The rate of false positive identifications can be determined by reversed or randomized database searches (Peng et al., 2003). In such searches the amino acid sequences present in the used databases are either reversed or mixed randomly. In addition, software programs are available, which help the validation process like PeptideProphet (Keller et al., 2002), ProteinProphet (Nesvizhskii et al., 2003), and DTASelect (Tabb et al., 2002). Recently, Kim et al. (2007) published a method, which assesses reproducibility of relative quantification strategies used MS-based proteomics. Journals specialized in proteomics have also published guidelines for publishing proteomics data (Carr et al., 2004; Wilkins et al., 2006).

2.7.3 Protein quantification algorithms

In MS-based quantification software programs are needed to calculate the relative quantities of isotopically tagged peptides and corresponding protein abundances. Such programs basically calculate the peptide abundances by integrating areas of the tagged peptides from the ion chromatogram or MS spectrum, determine relative abundances of the peptides by matching corresponding peptides with each others and, finally, combine peptide level information into protein abundances. Many different stable isotope labeling methods exist and often the software programs used for quantification are specialized for certain types of tags. XPRESS (Han et al., 2001), ASAPRatio (Li et al., 2003c), and ProICAT (Applied Biosystems) are specialized in quantifying ICAT labeled peptides, ZoomQuant (Halligan et al., 2005) measures ^{18}O-labeled peptides, and ProQuant (Applied Biosystems) is used to quantify either ICAT or iTRAQ labeled peptides. MSQuant (msquant.sourceforge.net), RelEx (MacCoss et al., 2003), Pepquan in Bioworks (Thermo Electron), Protein Pilot (Applied Biosystems), and Spectrum Mill (Agilent) are able to quantitate different types of stable isotope labels, including SILAC. Later versions of XPRESS are also able to quantify SILAC labeled peptides.

2.7.4 Data analysis pipeline

High-throughput analyses in proteomics also demand automated robust data handling and processing. Therefore so-called data analysis pipelines are often built to analyse the proteomics data in high-throughput fashion. In these pipelines a number of features exist to process and handle the data starting from the raw data and ultimately finishing by integrating the data to data generated by other approaches like transcriptomics and metabolomics. Figure 6 shows schematically some basic functions, which could be performed for the data in such data analysis pipeline. Institute for Systems Biology has been one of the pioneers in developing tools for efficient data analysis pipelines. In their Sashimi project (sashimi.sourceforge.net) they have a variety of data converters to change the raw data format to appropriate format to analyse it with different software programs like SEQUEST and Mascot. Data conversion is necessary, because instruments and software programs manufactured by different companies are quite

often incompatible with each other. The converted data can then be pre-filtered prior to analysis with search program like SEQUEST. PeptideProphet and ProteinProphet can be used to validate peptide and protein identifications generated by SEQUEST, while XPRESS and ASAPRatio are available to quantify the peptides and corresponding proteins. All these results can be then summarized with INTERACT (Han et al., 2001). Finally the proteomics results can be stored and integrated to data from other cellular levels with SBEAMS (Marzolf et al., 2006).

Figure 6: High-throughput proteomics analyses create a lot of data that demand effective automated data analysis pipelines. Figure summarizes common tools used for data analysis in such pipelines.

2.8 Lymphocytes present in the immune system

By exposure to our surrounding environment alone we are constantly exposed to potentially harmful microbes and toxins. During evolution our immune system has therefore evolved to protect us from these hazards. The immune system can be divided into the innate and adaptive immune systems. The innate immune system consists of phagocytes, natural killer (NK) cells, antimicrobial peptides, complement, and cytokines. The innate immune system plays an important role in rapid immune response, when microbes and toxins are encountered. The main function of the adaptive immune system is to generate specific antibodies against antigens. In addition to antibodies the adaptive immune system includes cells such as B- and T-lymphocytes, and different cytokines and chemokines. Cytokines are secreted peptides and proteins that mediate signals of immune system. Cytokines include molecules like interferons (IFN), interleukins (IL), and tumor necrosis factors (TNF). Innate and adaptive immunity consist of these specialized cells and secreted molecules, which function cooperatively to provide an optimal host defence.

Leucocytes mediate cellular functions in the immune system. All leucocytes are derived from pluripotent stem cells, which can differentiate either to lymphoid or myeloid cells. Lymphoid cells can be divided to T, B, and NK cells, which all have divergent functions within the immune system (reviewed by Alam and Gorska, 2003). T cells maturate in the thymus either to T helper cells (Th) or cytotoxic T cells (Tc). Th cells play an important role in immune system by activating and directing other cells involved in the immune response. These functions are mediated both by direct cell-to-cell interactions and by secreting cytokines. Tc cells recognize antigens presented by MHC (major histocompatibility) class I molecules and mediate their cytotoxic functions either by perforin or Fas ligand mechanisms. Tc cells are also able to secrete different cytokines e.g. IFN-γ, IL-4, and IL-5. B cells maturate in the bone-marrow and have a crucial role in adaptive immunity. Their main function is antibody production

and secretion. B cells also participate in antigen presentation to T cells. NK cells have a role both in innate and adaptive immune system. Their main task is to mediate cytotoxicity against cells, which lack or have a reduced amount of MHC class I molecules on their surface, for example, cells infected with certain pathogens and tumour cells. However, NK cells are able also to recognize and destroy cells ´marked´ with antibodies. NK cells produce also cytokines e.g. IL-4, IFN-γ, and TNF-α.

2.8.1 Th cells

Th cells can be divided into two main subsets, namely Th1 and Th2 cells, which both differentiate from precursor Th cells (Mosmann et al., 1986; Mosmann and Coffman, 1989b). Th1 or Th2 polarization of the cells is dependent on many factors such as antigen presentation and co-stimulatory molecules, genetic background of the cells, induction of key transcription factors, and cytokine environment (Agnello et al., 2003; Constant and Bottomly, 1997; Lee et al., 2006). Th1 and Th2 cells display distinct cytokine secretion profiles and different functions within the immune system. Th1 cells secrete cytokines like IFN-γ, IL-2, and lymphotoxin-β, while Th2 cells produce IL-4, IL-5, IL-6, IL-9, IL-10, and IL-13 (Glimcher and Murphy, 2000; Mosmann and Coffman, 1989a; Paul and Seder, 1994). Figure 7 summarizes cytokine secretion profiles charasteristic for Th1 and Th2 subsets. Th1 cells promote cellular immunity to destroy intracellular pathogens. Th2 cells participate in humoral immunity against extracellular pathogens. Imbalance and abnormal activation of the Th1 and Th2 cell populations is associated with certain pathophysiological conditions (Romagnani, 1994; Romagnani, 1996; Romagnani, 2004). Th1 cytokines have been detected in autoimmune diseases, while Th2 specific reactions have been observed in diseases like allergy and asthma. However, the pathophysiologies of such diseases are more complicated including many other factors in addition to an abnormal Th lymphocyte response. Differentiation of Th cells is also much more complex process than just a dichotomy between Th1 and Th2 cells. Recently, new Th cell subtypes have been characterized like FoxP3^{+} T regulatory (Tr) cells, IL-10 producing Tr1 cells, transforming growth factor (TGF)-β-secreting follicular Th cells, and IL-17A-producing Th17 cells (Reinhardt et al., 2006). Th17 subset and regulatory T cells have been reported to play a role in diseases, such as autoimmunity and allergy (reviewed by Reiner, 2007; Azfali et al., 2007; Romagnani, 2004).

2.8.2 Cytokines promoting Th1 and Th2 differentiation

Antigen encounter triggers the cells of the immune system to secrete cytokines. The created cytokine milieu plays an important role in the immune response by directing other cellular responses. The cytokine environment also has a great effect on Th cell differentiation. IL-12 and IL-4 are considered the classical cytokines promoting Th differentiation. IL-12 polarizes the cells to Th1 direction and prevents Th2 differentiation (Hsieh et al., 1993; Kobayashi et al., 1989; Manetti et al., 1993), while IL-4 acts on the contrary by promoting the cells towards Th2 phenotype and inhibiting formation of Th1 cells (Hsieh et al., 1992; Seder et al., 1992; Swain et al., 1990). Both

of the cytokines have been used traditionally for *in vitro* differentiation of Th1 and Th2 cells from Th precursor cells.

In addition to IL-12 several other cytokines and signalling cascades have been shown to affect Th1 differentiation including IL-18, IL-23, IL-27, IFN-α, and IFN-γ (Agnello et al., 2003). STAT1, STAT4, and T-bet are central transcription factors involved in the Th1 differentiation (Berenson et al., 2004; Trinchieri, 2003). Figure 7 summarizes key cytokines and transcription factors involved in Th1 differentiation.

Th2 differentiation also involves many cytokines and signalling molecules. IL-4 is the hallmark Th2 directing cytokine, while cytokines IL-2, IL-6, IL-10, IL-13, and IL-21 have been observed to have minor roles in the differentiation (Agnello et al., 2003; Constant and Bottomly, 1997; Mowen and Glimcher, 2004). GATA3, STAT6, and c-Maf are the key transcription factors in the Th2 cell development (Mowen and Glimcher, 2004). Cytokines and transcription factors playing a crucial role in Th2 differentiation are summarized in figure 7.

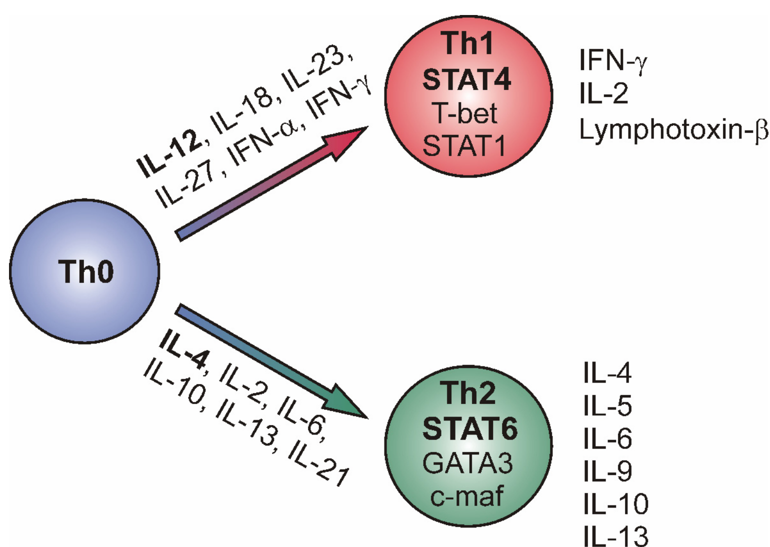


Figure 7: Key cytokines and transcription factors involved in Th1 and Th2 differentiation. Figure also presents cytokines secreted by Th1 and Th2 subsets.

2.8.2.1 IL-12 and Th1 differentiation

IL-12 is a pro-inflammatory cytokine, which plays an important role in the cellular immunity against intracellular pathogens (reviewed by Gately et al., 1998; Trinchieri, 2003). It induces activated T and NK cells to proliferate, enhances their cytolytic properties, and triggers them to produce Th1 specific cytokines like IFN-γ (Chan et al., 1991; Chouaib et al., 1994; Gately et al., 1991; Gately et al., 1992; Gately et al., 1994; Kobayashi et al., 1989; Perussia et al., 1992; Robertson et al., 1992; Robertson et al., 1992; Stern et al., 1990). Most importantly, IL-12 is the key cytokine inducing Th1 differentiation. This has been demonstrated in several studies, for example, by

induction of Th precursor cells towards Th1 phenotype with IL-12, inhibition of Th1 response with IL-12 neutralizing antibodies, and the study of animals genetically deficient for IL-12 signalling, such as IL-12 (Magram et al., 1996), IL-12R (Wu et al., 2000), and STAT4 (Kaplan et al. 1996a) knock-out mice (reviewed by Agnello et al., 2003; Gately et al., 1998; Trinchieri, 2003). All these knock-out mice have impaired Th1 lymhocyte development and reduced IFN-γ production.

IL-12 belongs to the group of long chain type I cytokines (Boulay et al., 2003). This heterodimeric protein consists of p35 and p40 subunits, which do not have sequence homology to each other (Gubler et al., 1991; Kobayashi et al., 1989; Stern et al., 1990). Although the subunit p40 does not have any cytokine homologues, it does belong to the hemopoietin receptor family and has homology to the extracellular domains of the cytokine receptors (Gearing and Cosman, 1991). The p35 subunit shares homology with cytokines IL-6, granulocyte colony stimulating factor, and chicken myelomonocytic growth factor (Merberg et al., 1992). Separate genes located in human chromosomes 3p12-q13.2 and 5q31-33 encode the p35 and p40 subunits, respectively (Sieburth et al., 1992). Neither of the subunits is biologically active alone, instead a heterodimeric form is necessary to biological response.

Intracellular pathogens like bacteria and viruses trigger the production of IL-12 in phagocytes (macrophages, monocytes, and neutrophils) and in dendritic cells (Gazzinelli et al., 1994; Hsieh et al., 1993; Macatonia et al., 1995). IL-12 signalling is mediated through the IL-12 receptor (IL-12R), which is present primarily on the surfaces of activated T cells and NK cells, but is also present on B cells and dendritic cells (Airoldi et al., 2000; Desai et al., 1992; Grohmann et al., 1998; Presky et al., 1996). Functional IL-12R, like its ligand, is composed of two subunits, namely IL12Rβ1 and IL12Rβ2 (Chua et al., 1994; Chua et al., 1995; Presky et al., 1996). TCR, CD28, IFN-γ, IFN-α, TNF, and IL-12 signalling enhances the expression of the β2 subunit, which enables the formation of functional IL-12R and IL-12 signalling (Rogge et al., 1997; Szabo et al., 1997). When IL-12 binds to its receptor, Janus kinase family members JAK2 and TYK2 bind to the intracellular domain of IL-12R, where they are phosphorylated (Bacon et al., 1995a; Zou et al., 1997). JAK2 and TYK2, in turn, activate STAT1, STAT3, STAT4, and STAT5. However, tyrosine phosphorylated STAT4 is the major mediator of IL-12 response (Bacon et al., 1995b; Jacobson et al., 1995). The phosphorylated STAT4 dimer localizes to the nucleus, where it binds to specific DNA sequences, and promotes IL-12 responsive genes (reviewed by Agnello et al., 2003; Gately et al., 1998; Trinchieri, 2003).

2.8.2.2 IL-4 and Th2 differentiation

Cytokine IL-4 participates powerfully in humoral immune response against extracellular parasites by directing activated Th cells towards the Th2 subset and by stimulating B cells (reviewed by Nelms et al., 1999, and Mowen and Glimcher, 2004). Th2 cell differentiation leads to an increased production of IL-4 and other Th2 specific cytokines, thus strengthening the effects of IL-4 signalling. At the same time IL-4 strongly inhibits the development of IFN-γ producing Th1 cells. In addition to its major role in Th2 differentiation, IL-4 also directs B cells in the humoral immune response.

IL-4 switches the isotype of the immunoglobulins produced in B cells towards IgE and IgG1 or IgE and IgG4 in mouse and human, respectively (Coffman et al., 1986; Gascan et al., 1991; Vitetta et al., 1985). IL-4 also acts as a mitogen for B cells and increases their MHC class II expression on the cell surface (Nelms et al., 1999). The impaired Th2 differentiation capacity and decreased IgE production of mice that are genetically deficient for IL-4 signalling, such as IL-4 (Kopf et al., 1993; Kuhn et al., 1991), IL-4R (Noben-Trauth et al., 1997), and STAT6 (Kaplan et al., 1996b; Shimoda et al., 1996; Takeda et al., 1996) knock-outs, demonstrates the importance of IL-4 for Th2 cell development.

IL-4 is a short chain type I cytokine that belongs to the IL-4 family of genes (Boulay et al., 2003). This gene family is clustered within 600 Kb at the chromosomal region 5q31 in human and encodes cytokines IL-4, IL-13, IL-3, IL-5, and GM-CSF. In addition to TCR activated Th2 cells, basophils, and mast cells, NK cells, γ/δ T cells, and eosinophils have also been reported to produce IL-4 (Nelms et al., 1999; Seder and Paul, 1994). Cellular response for IL-4 is mediated through the IL-4 receptor (IL-4R), which is expressed on the surfaces of the cells originating from hematopoietic, endothelial, epithelial, muscle, fibroblast, hepatocyte, and brain tissues (Ohara and Paul, 1987). IL-4R is composed of two subunits, namely IL-4Rα and common γ chain (γ c). IL-4 binds to the IL-4Rα and the receptor-ligand complex thus formed is then recognized by the γ c (Galizzi et al., 1990; Letzelter et al., 1998; Russell et al., 1993). Ligand induced heterodimerization of γ c to IL-4Rα is required for physiological IL-4 signalling (Russell et al., 1993). Signalling via IL-4R phosphorylates JAK1 and JAK3, which will bind to IL-4Rα and γ c, respectively (Leonard and O'Shea, 1998; Nelms et al., 1999; Reichel et al., 1997; Russell et al., 1994; Witthuhn et al., 1994). These JAKs induce phosphorylation of IL-4Rα leading to binding, phosphorylation, and dimerisation of STAT6 that then locates to the nucleus and activates IL-4 responsive genes (reviewed by Andrews et al., 2002; Hebenstreit et al., 2006; Mowen and Glimcher, 2004; Nelms et al., 1999).

2.9 Transcriptome and proteome analyses create novel insights about T helper cell signalling and differentiation

Th cells and their differentiation play an important role in immune responses and pathological conditions. It is therefore important to find factors involved in these processes. The development of transcriptomics and proteomics techniques has enabled large-scale screening of gene and protein expression that has been widely used to study both Th cells and other lymphocytes (reviewed by Cristea et al., 2004; D'Ambrosio et al., 2005; Di Bartolo and Acuto, 2004; Granucci et al., 2001; Thadikkaran et al., 2005). In microarray based transcriptome studies the gene expression of whole genome can be measured in one measurement. Such studies have revealed already several important genes involved in Th cell signalling and differentiation (Bosco et al., 2006; Chen et al., 2003; Chtanova et al., 2001; Chtanova et al., 2004; Chtanova et al., 2005; Lu et al., 2004a; Lund et al., 2003; Lund et al., 2005; Lund et al., 2004; Lund et al., 2007; Rogge et al., 2000). Whilst in transcriptome studies, the gene expression is measured at mRNA level, proteomics studies can offer novel information about the changes in

protein expression and about post-translational modifications. Proteome databases have been established both from primary T cells and T cell lines like Jurkat cells (reviewed by Thadikkaran et al., 2005). At the moment there is not capacity to measure whole cellular proteome. Complexity of the studied proteome can be reduced, for example, by isolating certain cellular organelles. Such an approach also offers additional information about the protein localization. For example, mitochondria (Rezaul et al., 2005; Thiede et al., 2002), plasma membrane/lipid raft (Bini et al., 2003; Loyet et al., 2005; Von Haller et al., 2003), and nuclei (Hwang et al., 2006) of T cells have been characterized. Proteomics applications have also shed light on many biological and physiological functions and responses encountered in T cells, for example, their activation (Bini et al., 2003; Risso et al., 2005; Stentz and Kitabchi, 2004; Von Haller et al., 2003), cytokine signalling and Th differentiation (Filen et al., 2005; Loyet et al., 2005; Nyman et al., 2000; Rautajoki et al., 2004; Rautajoki et al., 2007; Rosengren et al., 2005a; Rosengren et al., 2005b), apoptosis (Hwang et al., 2006; Thiede et al., 2001; Thiede et al., 2002; Thiede et al., 2006), post-translational modifications (Brill et al., 2004; Cao et al., 2006), and response for oxidative stress (Fratelli et al., 2002). Studies by Ideker et al. (2001) and Washburn et al. (2003), and others, have highlighted the importance of integrating biological data from transcriptome and proteome levels to gain a more extensive view on biological processes and phenomena in different biological systems.

3 AIMS OF THE STUDY

The overall aim of this Ph. D. study was to set-up, optimize, develop, and apply proteomics methods which could be used to elucidate the effects of cytokines IL-12 and IL-4 on the proteome of Th lymphocytes. These methods included protein sample preparation, fractionation, high-throughput LC-MS/MS, data analysis, and, in particular, stable isotope labeling methods e.g. ICAT and iTRAQ. These methodologies are compatible for analysing also membrane proteins, which have been difficult to study by 2-dimensional electrophoresis based proteomics. These methods enabled characterization of cellular membranes in Th lymphocytes. The specific aims of the research projects in this Ph. D. thesis were:

- To characterize the microsomal fraction proteome human lymphoblasts cultured with cytokine IL-12 (I).
- To characterize the microsomal fraction and nuclear proteome in human naïve CD4+ lymphocytes cultured with cytokine IL-4 (II, unpublished results).
- To screen IL-4 and IL-12 regulated proteins in human Th lymphocytes (I, II, unpublished results).
- To set up proteomics methods, in particular, ICAT and iTRAQ labeling method to achieve the aims above (I, unpublished results).
- To evaluate the MS/MS data analysis algorithms and build a data analysis pipeline to process MS/MS data created in high-throughput LC-MS/MS analyses (III).
- To create an algorithm to pre-filter good quality MS/MS spectra prior to database searching (IV).

4 MATERIALS AND METHODS

4.1 Standard protein mixture

Two standard protein mixtures were prepared with different protein concentrations. The proteins present in the mixtures and their relative concentrations are shown in Table 4. All the proteins were purchased from Sigma-Aldrich (St. Louis, MO, USA).

Table 4: Proteins and their relative concentrations present in the standard protein mixtures.

Protein:	C_1/C_2:
Alpha-lactalbumin	0.2
Beta-galactosidase	5.0
Beta-lactoglobulin	1.5
Cytochrome C	10
GAPDH	0.1
Glycogen phosphorylase	0.7
Lactoferrin	1.0
Lysozyme C	2.0
Ribonuclease A	0.5

4.2 Human Th cell cultures

Th lymphocytes were isolated either from leukocyte-rich buffy coats (Red Cross Finland Blood Service, Helsinki) or from cord blood samples collected from healthy neonates at Turku University hospital.

4.2.1 CD8+ depleted peripheral blood mononuclear cells

PBMC were isolated from buffy coats by Fiqoll-Paque density gradient centrifugation (Amersham Pharmacia, Uppsala, Sweden). Erythrocytes were lysed with 0.83% Ammonium chloride, after which adherent cells and CD8+ cells were removed by incubating the cells first on a plastic cell culture plate and then with anti-CD8 magnetic beads (Dynal, Oslo, Norway). Aliquots of the remaining cell populations were analysed by flow cytometry. Cells were cultured at cell density 2 x 10^6 /ml in RPMI medium (Sigma-Aldrich) with 10% fetal calf serum (Gibco BRL, Invitrogen Life Technologies, Carlsbad, CA, USA), 2 mM L-glutamine (Sigma-Aldrich), streptomycin (Sigma-Aldrich), and penicillin (Sigma-Aldrich). Cells were activated two days with 500 ng/ml of phytohemagglutinin (PHA) (Murex Diagnostics, Chatillon, France). Cytokine IL-2 (17 ng/ml; R&D Systems, Minneapolis, MN, USA) was then added to induce the proliferation of the cells and IL-12 (2.5 ng/ml; R&D Systems) to polarize the cells towards Th1 phenotype. The cells were then cultured for two days and harvested.

4.2.2 CD4+ cells extracted from PBMC and CB

Mononuclear cells were isolated either from cord blood or buffy coat samples by Fiqoll-Paque and anti-CD4 magnetic beads (Dynal). Purified CD4+ cells were cultured in Yssel's medium (Yssel et al., 1984) supplemented with 1% human AB-serum (Red Cross Finland Blood Service). Cells were activated with 500 ng/ml of platebound anti-CD3 (Immunotech, Marseille, France) and 500 ng/ml of soluble anti-CD28 (Immunotech). Alternatively the cells were activated with 100 ng/ml PHA (Murex Diagnostics, Chatillon, France) and irradiated CD32/B7 transfected mouse L fibroblasts. Cells were polarized towards Th1 or Th2 phenotype by adding either 2.5 ng/ml of IL-12 (R&D Systems) or 10 ng/ml of IL-4 (R&D Systems). No cytokines were used for neutral Th0 condition. Cells were cultured 2 h, 6 h, 12 h, 1 d, 2 d, or 7 d and harvested. In order to enhance proliferation of the cells, 17 ng/ml IL-2 (R&D Systems) was added to the 7 d cultures on the second culture day.

4.3 Flow cytometric analyses

Cell populations of CD8+ depleted peripheral blood mononuclear cells were analysed by FACS Scan flow cytometer (BD Biosciences, San Jose, CA, USA) with CD4, CD8, CD14, CD16, and CD19 antibodies conjugated with R-phycoerytherin (R-PE) (Caltag, Burlingame, CA, USA). FITC conjugated CD69 antibody (BD Biosciences) was used to measure the activation of the cells. CD7 expression was measured with anti-CD7 (BD Biosciences) as primary antibody and antimouse IgG (H+L) (Caltag) as a secondary antibody. Mouse IgG_1-PE (Caltag) and mouse IgG_2-FITC (Caltag) were used for isotype controls.

4.4 Cell fractionation

Microsomal fractions and nuclear extracts were enriched from the harvested cells. During the cell fractionations all the steps were done at +4°C, unless otherwise indicated. Total cell lysates were prepared for Western blot analyses.

4.4.1 Microsomal fraction

Washed cells were incubated 10 min in hypotonic buffer (10 mM Tris-HCl, pH 7.6, 0.5 mM $MgCl_2$) supplemented with Complete Mini protease inhibitor cocktail (Roche, Mannheim, Germany) and then lysed with a dounce homogenizer. After homogenization NaCl was added to the hypotonic buffer at 0.15 M concentration. Cell lysate was centrifuged first at low speed (800 g, 5 min) to collect post-nuclear supernatant, which was then centrifuged (100000 g, 1 h) to collect the microsomal fraction.

4.4.2 Nuclear extract

Enrichment of nuclear fraction including DNA-binding proteins was slightly modified from Andrews and Faller (1991). Washed cells were lysed with 10 min incubation in buffer containing 0.2 % NP-40, 10 mM HEPES pH 7.9, 0.5 mM DTT, 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM NaF, 1 mM Na_3VO_4 and Complete Mini protease inhibitors (Roche). Cell lysates were then centrifuged quickly and pellets were suspended to buffer consisting of 20mM HEPES pH 7.9, 420mM NaCl, 25% glycerol, 0.5 mM DTT, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 1 mM NaF, 1 mM Na_3VO_4, and Complete Mini protease inhibitors (Roche). Samples were incubated 20 min and then centrifuged (13000 rpm, 15 min) to collect the supernatant containing nuclear proteins.

4.5 ICAT labeling of the samples

Microsomal fraction proteins and standard protein mixture were prepared and labeled with ICAT or cleavable ICAT reagents according to the previously described methods (Gygi et al., 1999; Han et al., 2001) and the protocol provided with the ICAT reagents (Applied Biosystems, Foster City, CA, USA). In summary, proteins were solubilized into denaturing buffer (6 M urea, 0.05 % SDS, 5 mM EDTA, and 50 mM Tris, pH 8.5) and reduced with tris(2-carboxyethyl)phosphine (TCEP) for 30 minutes. Then the proteins were labeled either with light or heavy ICAT reagent (2 mM). Labeling reaction was stopped by addition of 10 mM DTT. Proteins labeled with light and heavy ICAT reagents were mixed, diluted, and digested with trypsin (sequencing grade modified trypsin, porcine; Promega, Madison, WI, USA).

4.6 iTRAQ labeling of the samples

iTRAQ labeling for nuclear proteins was performed according to the iTRAQ labeling protocol (Applied Biosystems). Equal amounts of nuclear proteins (90-100 µg) from four cell states were determined with the Bradford assay. Acetone precipitated proteins were dissolved with 0.1 % SDS in 40 µl of triethyl ammonium bicarbonate (TEAB)-buffer provided with the iTRAQ reagents. Proteins were reduced with TCEP (4.5 mM), and their cysteines were blocked with methyl methanethiosulfonate (MMTS; 8.7 mM). Proteins were then digested to peptides with trypsin (10 µg), and proteins of each sample were labeled with different iTRAQ reagents. Labeled peptides were combined and evaporated before the sample fractionation.

4.7 Chromatographic fractionation of the peptides

Both ICAT and iTRAQ labeled peptides were fractionated with strong cation-exchange (SCX) and reversed-phase (RP) chromatography. ICAT labeled peptides were additionally purified with avidin affinity chromatography. SCX fractionation was carried out with BioCAD HPLC (Perseptive Biosystems, Freiburg, Germany) by using a 4.6 mm x 200 mm Poly-sulfoethyl A column (Poly LC, Columbia, MD, USA). Peptides were eluted with KCl gradient in buffer containing 5 mM KH_2PO_4 and 25 %

acetonitrile (pH 3.0). SCX fractions including the ICAT labeled peptides were further purified with the avidin affinity chromatography according to the manufacturer's instructions (Applied Biosystems). Finally, the fractions including either iTRAQ labeled or ICAT labeled peptides were evaporated and dissolved with 1 % HCOOH. The samples were desalted with μ-tips containing RP-material (either Oligo R3, Perseptive Biosystems, Framingham, MA, USA or Empore C18, 3M, St. Paul, MN, USA) prior to nanoLC fractionation and MS/MS analysis.

4.8 nanoLC-ESI-MS/MS analysis

The LC-MS/MS analyses of peptides were achieved by reversed-phase (RP) liquid chromatographic (LC) separation with a nanoLC coupled on-line to a quadrupole time-of-flight mass spectrometer. The LC-MS configuration consisted of nanoLC system, including Famos, Switchos II, and Ultimate (LC Packings, Amsterdam, Netherlands), coupled online to a QSTAR Pulsar mass spectrometer (Applied Biosystems/MDS Sciex). PepMap C18 μ-pre-columns (0.3 mm x 5 mm) (LC Packings) were used for rapid sample loading and desalting and were coupled to either PepMap C18 (75 μm x 15 cm) (LC Packings) or in-house packed Magic C18 (Michrom Bioresources Inc., Auburn, CA, USA), analytical separation columns. The peptides were eluted from the column to the mass spectrometer with an increasing acetonitrile gradient. The mass spectrometer was set to collect 1 s of MS scan followed by 3 s MS/MS scans of the two most intense ions present in the MS scan.

4.9 Data processing tools

Table 5 summarizes all the software programs, databases, and parameters used in LC-MS/MS data processing and analysis.

Table 5: Software programs and databases used in the data analysis of tandem mass spectrometry data.

Software:	Manufacturer:	Description:	Studies:
MS instrument control:			
Analyst QS	ABI	• Controlling MS instrument • Describing IDA parameters • Visualization of raw MS and MS/MS data • Data in *.wiff format	I-V
Data format conversion			
MzStar	Sashimi	• Converting files in *.wiff format to mzXML format	IV, V
mzXML2Other	Sashimi	• Converting files in mzXML format to .dta format	IV, V
Out2Summary	Sashimi	• Converting SEQUEST results files in *.out format to html format	IV
Database search and quantification[1]			
ProICAT	ABI	• Identification and quantification of ICAT or cICAT labeled proteins • Specified modifications: ICAT, cICAT	I, II, IV, V
ProID	ABI	• Identification of non-labeled proteins • Specified modifications: Carbamidomethyl	II

Spectrum Mill	Agilent	• Identification and quantification of non-labeled and stable isotope labeled proteins • Specified modification: ICAT • Trypsin digestion	IV
SEQUEST and XPRESS	Thermo	• Identification and quantification of non-labeled and stable isotope labeled proteins • Compatibility with many useful tools present in Sashimi • Specified modification: ICAT • In XPRESS default mass tolerance 1.0	IV, V
ProQuant	ABI	• Identification and quantification of iTRAQ labeled proteins • Specified modifications: MMTS, iTRAQ	III
Mascot	Matrix Science	• Identification of non-labeled and stable isotope labeled proteins • Available in internet • Specified modifications: MMTS, iTRAQ, ICAT, cICAT	III, V
Databases			
SwissProt		• Human • All organisms • Reversed database	I, II
NCBI		• Human • All organisms • Reversed database	IV, V
Subsetdb	Sashimi	• Creating reversed databases	II
Interrogator	ABI	• Database used with ProICAT and ProQuant software. • Creates a database from any *.fasta format database	I-V
Data validation and annotation			
Peptide Prophet		Keller et al., 2002	II
Protein Prophet		Nesvizhskii et al., 2003	II
PIGOK		Jacob and Cramer, 2006	III
Classification			
WEKA		Witten and Frank, 2000	V
Data sort, comparison, and statistical process			
Interact		Sashimi	II
Excel, Access		Microsoft	I-V
Kensington Discovery edition		Inforsense	I-V

1) In database searches the MS and MS/MS tolerance was set to 0.3 and 0.2, respectively.

Sashimi: free open source software tools at sashimi.sourceforge.net/index.html (Institute for Systems Biology)
ABI: Applied Biosystems, Foster City, CA, USA
Agilent: Agilent Technologies, Palo Alto, CA, USA
Thermo: Thermo Finnigan, San Jose, CA, USA
Inforsense: Inforsense Limited, London, UK

4.10 Western blot analyses

SDS-PAGE loading buffer was added to the cell lysate and microsomal fraction samples and boiled for 5 min prior to loading to the SDS-PAGE gels (10 %, 12 %, and 15 % gels were used). After the electrophoresis proteins were transferred to a nitrocellulose membrane (Amersham Pharmacia; Schleicher & Schuell GmbH, Dassel, Germany). Table 6 summarizes all the antibodies used in the Western blotting. The proteins were finally detected with ECL.

Table 6: Antibodies used in the Western blot analyses.

Antibody:	Organism:	Host:	Dilution:	Manufacturer:
β-actin	human	mouse	1:10000	Sigma-Aldrich
Galectin-1	human	mouse	1:200	Novocastra
GATA3	human	mouse	1:100	Santa Cruz Biotechnology
GIMAP4	human	rabbit	1:5000	Cambot et al., 2002
IgG (HRP conjugated)	mouse	goat	1:5000/1:10000	BD
IgG (HRP conjugated)	rabbit	goat	1:5000/1:10000	Cell Signalling
MxA	human	rabbit	1:2000	Julkunen
Phospho-STAT4	human	rabbit	1:1000	Zymed
Phospho-STAT6	human	rabbit	1:1000	Cell Signalling
STAT1	human	rabbit	1:1000	Cell Signalling

Novocastra Laboratories (Newcastle upon Tyne, UK); Santa Cruz Biotechnology (Santa Cruz, CA, USA); Zymed Laboratories (San Francisco, CA, USA)

4.11 Quantitative real-time RT-PCR

Rneasy Mini kit (Qiagen, Valencia, CA) was used to extract total RNA from the cells. RNA (1 µg) was treated with Dnase I (Invitrogen, Carlsbad, CA) and used for cDNA synthesis by SuperScript II (Invitrogen). Table 7 shows the primers and the probes used for detection of *GIMAP4* and *GIMAP1* mRNA. Housekeeping gene *EEF1A* was used for normalization. ABI Prism 7700 Sequence Detector (Applied Biosystems) was used in the quantitative real-time RT-PCR analyses.

Table 7: Primers and probes present in the table were used in quantitative real-time RT-PCR.

Gene name:	Probe:
GIMAP1	Universal ProbeLibrary probe #24
GIMAP4 (long and short)	5'-FAM-acaaggcaacaggcgctgagca-TAMRA-3'
GIMAP4 (splice variants)	5'-FAM-ttcctgctccggttttacccactaacac-TAMRA-3'
EEF1A	5'-FAM-agcgccggctatgcccctg-TAMRA-3'
Gene Name:	**Primers:**
GIMAP1	5'-cgtggacactccggacat-3'
GIMAP1	5'-tcctcacagccaggatctgt-3'
GIMAP4 (long and short)	5'-tgaccgctactgtgcgttaaa-3'
GIMAP4 (long and short)	5'-tggatcaggcccagcaa-3'
GIMAP4 (splice variants)	5'-gagagggcattcagtgctcc-3'
GIMAP4 (splice variants)	5'-caggggccagttatgggc-3'
GIMAP4 (splice variants)	5'-gcagtgccagaatgaaacactt-3'
EEF1A	5'-ctgaaccatccaggccaaat-3'
EEF1A	5'-gccgtgtggcaatccaat-3'

5 RESULTS AND DISCUSSION

5.1 Characterization of microsomal fraction proteome in Th1 and Th2 polarized lymphocytes by using isotope-coded affinity tags and tandem mass spectrometry (I-II)

The different roles of Th1 and Th2 lymphocytes in immune response also imply differences in their gene expression. Microarray studies, measuring the gene expression at mRNA level, have already brought out several differentially expressed genes between Th1 and Th2 cells (Bosco et al., 2006; Chen et al., 2003; Chtanova et al., 2001; Chtanova et al., 2004; Chtanova et al., 2005; Lu et al., 2004a; Lund et al., 2003; Lund et al., 2005; Lund et al., 2004; Lund et al., 2007; Rogge et al., 2000). Gene expression of Th cells has also been studied at protein level by using gel- and MS-based proteomics. Our group has earlier established a proteome database of primary Th cells and further studied the effects of IL-4, IL-12, and IFN-α on the differentiation by 2-DE based approach (Nyman et al., 2000; Rautajoki et al., 2004; Rautajoki et al., 2007; Rosengren et al., 2005a; Rosengren et al., 2005b). Due to the limitations to study membrane proteins by 2-DE, ICAT was selected to study them in Th cells polarized towards Th1 or Th2 by IL-12 or IL-4, respectively. Figure 8 summarizes the experimental design of the cell cultures used in the studies.

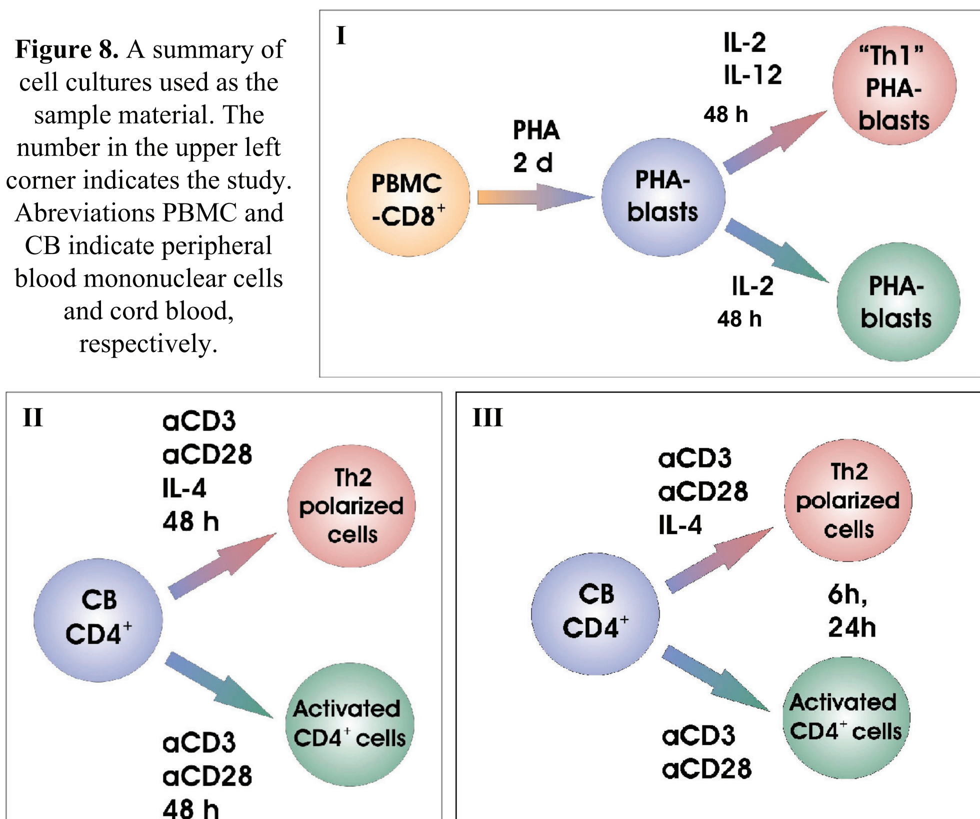


Figure 8. A summary of cell cultures used as the sample material. The number in the upper left corner indicates the study. Abreviations PBMC and CB indicate peripheral blood mononuclear cells and cord blood, respectively.

5.2 Characterization of microsomal fraction proteome in human lymphoblasts reveals the down-regulation of galectin-1 by interleukin-12 (I)

ICAT labeling and analysis protocol was set up during this project to characterize the membrane proteome of Th1 polarized lymphoblasts and by that way to screen novel proteins involved in IL-12 induced Th1 differentiation. Although this project revealed the complexity and some of the problems with the data analysis of complex proteomic samples, it gave an excellent data set, representing complex biological sample, for evaluating and optimizing data analysis tools needed in these studies.

5.2.1 Characterization of microsomal fraction proteome in human lymphobasts polarized towards Th1 phenotype by interleukin-12

The analysis of microsomal fraction proteins labeled with the ICAT reagents resulted in a total of ~ 21000 MS/MS spectra. In this study the data was analysed with software tool ProICAT from Applied Biosystems. The data was filtered with ProICAT stringent filter, which requires peptides to satisfy good-quality scores and high confidences. Altogether 1281 and 965 peptide sequences were identified from the two replicate samples that resulted in identification of 499 and 359 proteins. To prevent false positive identifications, proteins having at least two matching peptides (manually validated, > 5 amino acids), were accepted to high-confidence protein identifications. A total of 120 and 77 proteins fulfilled these criteria (I, Table 2). The identified proteins were categorized based on their cellular localization (I, Fig. 1). With this classification plasma membrane originating proteins were found to represent 25 % of the total. These included several Th cell surface receptors i.e., CD2, CD5, CD6, CD7, CD8, CD44, CD45, and MHC I, which were also identified from plasma membrane preparations by Loyet et al., 2005. Ribosomal and ER proteins constituted 22 % of the proteins, while almost 24 % of the proteins were from mitochondria, golgi, lysosomes, or nuclei. The rest of the proteins were cytoskeletal, cytoplasmic, or non-classified (proteins with unknown localization).

5.2.2 Galectin-1 and CD7 expression was decreased in human lymphoblasts treated with interleukin-12

Altogether 12 and 29 proteins were at least 1.5-fold up- or down-regulated, respectively. Only three of the proteins, namely Galectin-1, malate dehydrogenase, and annexin A-11, had significant expression changes in both the experiments, while the rest were either not identified or did not have significant change (> 1.5-fold) in the other experiment.

Galectin-1 and CD7 were selected from the differentially regulated proteins for further validation due to their immunoregulatory properties. The ICAT analysis showed that both of the proteins had smaller expression in the cell membranes, when the cells were treated with IL-12 (I, Fig. 2). Down-regulation of Galectin-1 was corroborated with Western blotting both in microsomal fraction of PBMC and in CD4+ Th cells isolated from cord blood (I, Fig. 3). CD7 expression on Th cell surface was measured with flow

cytometry, showing a slight down-regulation of CD7 both in IL-12 treated PBMC and cord blood CD4+ cells (I, Fig. 4).

Galectin-1 has been shown to participate in the regulation of cell growth, apoptosis, cell adhesion, and cytokine production (Allione et al., 1998; Baum et al., 2003; Blaser et al., 1998; Gabius, 2001; Liu et al., 2002; Perillo et al., 1995; Rabinovich et al., 1999a; Rabinovich et al., 1999b; Rabinovich et al., 2000; Rabinovich et al., 2002; Santucci et al., 2000; Santucci et al., 2003; Vespa et al., 1999; Zuniga et al., 2001). Gal-1 treatment has also been shown to prevent and suppress the symptoms of experimental autoimmune disorders like autoimmune encephalomyelitis, collagen – induced arthritis, concanavalin A induced hepatitis, experimental colitis, and graft versus host disease (Baum et al., 2003; Offner et al., 1990; Rabinovich et al., 1999b; Santucci et al., 2000; Santucci et al., 2003). In such autoimmune disorders immune response favours the production of Th1 type cytokines, however Gal-1 treatment reduced their production (Baum et al., 2003; Rabinovich et al., 1999b; Santucci et al., 2000; Santucci et al., 2003). Interestingly, Gal-1 is linked to CD7 via apoptosis, because Gal-1 acts as a ligand for CD7 inducing apoptosis (Pace et al., 2000; Perillo et al., 1995; Rabinovich et al., 2000; Vespa et al., 1999). Transmembrane glycoprotein CD7 has also many other immune response related functions, like co-stimulatory activity in T-cell receptor signalling and regulation of integrin adhesiveness (Chan et al., 1997; Stillwell and Bierer, 2001). Our results indicate that Th1 promoting IL-12 signalling reduces the expression of both Gal-1 and CD7. This effect might play a role in a selection of a specific Th cell population and cytokine environment creation during Th cell differentiation.

5.3 Characterization of microsomal fraction proteome in human Th lymphocytes polarized towards Th2 phenotype by interleukin-4 (II)

Characterization of the membrane proteome of Th2 polarized Th lymphocytes and screening of novel proteins involved in the Th2 differentiation was a direct continuation for the corresponding Th1 study. This time the Th lymphocytes were extracted from cord blood. Based on the experiences with complexity of the data analysis, encountered with the Th1 experiment and those obtained from the evaluation of the data analysis tools (III), the data analysis pipeline was automated and improved.

5.3.1 Characterization of microsomal fraction proteome in naïve human Th lymphocytes polarized towards Th2 phenotype by interleukin-4

In this study microsomal fraction proteins were extracted from Th lymphocytes of cord blood, which were activated by TCR stimulation and cultured either in the presence of Th2 promoting cytokine IL-4 or without. High-throughput nanoLC-MS/MS combined with cICAT labeling was used for the characterization of the membrane proteome and to screen for novel proteins involved in IL-4 signalling and Th2 differentiation. The MS/MS data was analysed with the paired algorithms combined with the reversed database searches to identify proteins as confidently and comprehensively as possible. Software programs ProICAT/ProID and SEQUEST/XPRESS were used to analyse the

data. Details of the data analysis protocol and the used parameters are presented in the experimental procedures and in Figure 1 of report II. ProICAT/ProID analysis resulted in 2636 identified peptides, while SEQUEST/XPRESS was able to identify 2007 peptides. These peptides matched to 474 and 432 distinct proteins with at least two matching peptides. 349 of these protein identifications were common to both the algorithms. The ProICAT and XPRESS algorithms were used for relative quantification of the protein abundances between IL-4 treated and non-treated Th lymphocytes that resulted in 304 protein quantifications. Supplementary Table 1 in report II presents protein identifications and quantifications. The rate of false positive identifications was estimated with the reversed database searches by using the relationship described by Peng et al. (2003). False positive rates < 0.1 %, < 1.4 %, and < 8.8 % were determined for proteins identified with ProICAT, SEQUEST, and ProID, respectively. The number of peptide and protein identifications and false positive rates are summarized in Table 1 of report II.

5.3.2 GIMAP1, GIMAP4, STAT1, and MxA were down-regulated in interleukin-4 treated Th lymphocytes

Altogether 304 proteins were quantified with ProICAT and XPRESS algorithms. Filtering of proteins that had at least 1.4-fold difference between IL-4 treated and non-treated cells resulted in eight proteins detected by both of the algorithms, 13 with ProICAT only, and 25 with SEQUEST/XPRESS only. Table 2 and Figure 2 in report II summarize these proteins and the correlation of quantifications between the two different algorithms. Our results indicate that protein identification and quantification with two different algorithms increases the confidence of the results obtained with both algorithms. However, at the same time the number of proteins identified and quantified with a single algorithm increases the scope of the detected proteins. Even if these proteins detected with a single algorithm include more false positive identifications and quantification errors than the proteins detected with two different algorithms, they may be biologically important and offer novel biological information.

In this study we selected four differentially regulated proteins, namely STAT1, MxA, GIMAP1, and GIMAP4, for further analysis based on their immunological and biological properties. Cellular expression levels of STAT1, MxA, and GIMAP4 and their abundances in the microsomal fraction were validated with Western blotting. The expression change of GIMAP1 was detected only with MS (II, Fig. 4), because at the time of this study GIMAP1 antibody wasn't available. The Western blot results show that IL-4 clearly decreased both the cellular and microsomal amount of the studied proteins. Figure 3 in report II represents Western blot results from one of the samples.

GIMAP1 and GIMAP4 belong to a protein family of GTPases of the immune-associated proteins. This family includes seven proteins, which all contain a common AIG1 domain with a GTP binding motif (Krucken et al., 2004). The expression of the GIMAP family members is highest in tissues of the immune system (Cambot et al., 2002; Dion et al., 2005; Krucken et al., 2004; Nitta et al., 2006; Poirier et al., 1999; Schnell et al., 2006; Stamm et al., 2002). GIMAP proteins have been shown to have a significant role and important functions in the immune system. They have been

reported to participate in thymocyte development (Nitta et al., 2006), T cell apoptosis (Carter et al., 2007; Schnell et al., 2006), anti-apoptotic effects (Krucken et al., 2005), T cell survival (Hernandez-Hoyos et al., 1999; Pandarpurkar et al., 2003), autoimmunity (Hellquist et al., 2007; MacMurray et al., 2002) and leukemia (Zenz et al., 2004).

The expression of *GIMAP1* and *GIMAP4* genes was studied in human cord blood CD4+ cells polarized towards Th1 or Th2 lineage. Quantitative real-time RT-PCR analyses showed that IL-12 up-regulated and IL-4 down-regulated both *GIMAP1* and *GIMAP4* (II, Fig. 5). Interestingly, the gene expression profiles of the genes were remarkably similar. Two alternatively spliced forms of *GIMAP4* were found, when the gene was cloned. In addition to the previously described short isoform (e.g. NM_018326) (Cambot et al., 2002), there is a 42 amino acids longer isoform described as a full-length EST in GenBank (e.g. CF594134). Both of the splice variants were up-regulated in IL-12 treated cells and down-regulated in IL-4 treated cells (II, Fig. 6). Even if the long isoform showed more difference in the expression than the short one, the latter is expressed ~44-fold more in Thp cells.

The expression of *GIMAP1* and *GIMAP4* genes and the corresponding proteins suggest their importance for Th differentiation. The role of GIMAP4 in Th differentiation is further supported by the phenotype of inbred Brown Norway rats. These rats carry a natural hypomorphic variant of *GIMAP4* that leads to Th2 type immune response and elevated IgE levels (Carter et al., 2007; Hylkema et al., 2000; Prouvost-Danon et al., 1981). Moreover, these rats are susceptible to Th2 mediated autoimmunity (Fournie et al., 2001).

5.4 Determination of expression changes in nuclear proteins extracted from Th2 polarized cord blood CD4+ cells (unpublished results)

Transcription factors play an essential role in the Th cell differentiation. Therefore we focused to detect changes in the nuclear proteome of cord blood CD4+ during the early stages of Th2 polarization. LC-MS/MS and four-plex iTRAQ labeling were used to screen the changes in the nuclear proteome of cells cultured under activating conditions with or without IL-4 for 6 or 24 hours.

To increase the statistical weight and confidence of the results, altogether three biological and technical replicates were analysed. In addition, all the biological samples represented cell pools, which were extracted from at least six different individuals. In the LC-MS/MS twenty SCX purified fractions were analysed from each sample resulting in 180 LC-MS/MS runs for the complete experiment. The data set was processed with software programs ProQuant and Mascot against SwissProt database. Reversed database was used to estimate the false positive rate. The data analysis of the biological and technical replicates established a list of 700 proteins confidently identified in all three biological replicates representing a false positive rate of 1%. PIGOK software was used to annotate the dataset to ascertain information of the protein functions and localizations. The data analysis by ProQuant also included the quantification of the peptides. All the quantifications of unique peptides were

normalized, weighted, and averaged according to ProQuant and ProGroup user manual and Gan et al., 2007. Weighed averages were also calculated for all the proteins quantified in different samples. Z-scores describing the confidence of the quantitative results were determined for protein quantifications that were based on at least five measurements. Although the magnitude of all the detected changes in the protein expressions were small (< 2-fold), they were statistically significant for many proteins (z > 1.96). We detected, for example, interferon-regulated proteins, which were down-regulated in the nuclei of Th2 polarized cells (Table 8a). Interferons are known to polarize Th cells towards Th1 phenotype and it is known that IL-4 inhibits production of interferon-γ. We also detected proteins, which are known to be up-regulated under Th2 promoting conditions (Table 8b). In addition to these Th1/Th2 'markers' we were also able to find other differentially regulated proteins, which were novel in this context. A number of differentially regulated proteins were selected as candidates for validation by Western blotting based on the observed results and protein functions. These studies are in progress.

In a technical respect, we were able to identify more proteins with the iTRAQ method than with the earlier employed ICAT method. However, in this study the sample material was different and more analyses were performed. Therefore, the direct comparison of this study with studies I and II is difficult. iTRAQ method also improved the confidence of the protein identifications, because all the peptides are labeled and more peptides are present in the samples for each protein. Therefore, the detection of post-translational modifications is also easier than with the ICAT labeling method.

Table 8: Interferon-regulated proteins, which are up-regulated in Th1 specific conditions, were down-regulated in the nuclear fraction of IL-4 treated cells (A). SATB1 and STAT6, which are associated to Th2 differentiation were both up-regulated in the nuclear fractions of IL-4 treated cells (B).

A)

Protein_Name:	**Calcs.**	**6h**	**+/-**	**Z**	**24h**	**+/-**	**Z**
Interferon-induced guanylate-binding protein 2	11	1,01	0,07	0,14	0,82	0,06	2,93
Gamma-interferon-inducible protein Ifi-16	145	1,00	0,02	0,14	0,85	0,02	7,78
Interferon regulatory factor 8	16	1,13	0,06	2,15	0,80	0,06	2,97
Interferon-stimulated gene 20 kDa protein	6	1,03	0,06	0,41	0,81	0,08	2,23
Interferon-induced GTP-binding protein Mx1	12	0,99	0,06	0,17	0,38	0,06	6,86
Signal transducer and activator of transcription 1	21	0,96	0,05	0,67	0,85	0,06	2,52
T-box transcription factor TBX21	5	0,98	0,10	0,22	0,64	0,04	7,78
Tripartite motif protein 22	12	0,86	0,05	2,37	0,82	0,07	2,36

B)

Protein_Name:	**Calcs.**	**6h**	**+/-**	**Z**	**24h**	**+/-**	**Z**
DNA-binding protein SATB1	47	1,10	0,03	3,46	1,21	0,05	5,13
Signal transducer and activator of transcription 6	11	1,42	0,10	5,06	1,16	0,08	2,24

5.5 A comparative evaluation of software packages to analyse MS/MS data of ICAT-labeled complex protein mixture (III)

Analysing MS/MS data of complex protein mixtures demands powerful and robust software programs to allow for high-throughput analyses. Since the manual validation of all the MS/MS spectra of the fragmented peptides is a highly labour intensive process, the performance of three different MS/MS data processing programs that could be used in the high-throughput MS/MS data analysis pipeline were tested. The programs ProICAT (Applied Biosystems), SEQUEST (Eng et al., 1994) [now included as a part of Bioworks package (Thermo Electron)], and Spectrum Mill (Agilent) were compared in terms of protein identification and quantification. As a test material we used a standard protein mixture and a complex biological sample (I), which both were labeled with ICAT reagents. In the standard protein mixture altogether nine proteins were present at a range of relative concentration ratios from 0.1 to 10, while microsomal fractions prepared from IL-12 induced human Th lymphocytes represented the complex biological sample (I).

When the standard protein mixture was analysed with the different programs, they all identified eight of the nine proteins present in the mixture. None of the programs were able to identify ICAT-labeled peptides of Cytochrome C. The number of identified peptides/protein varied between the different programs as illustrated in a Venn diagram (III, Fig. 2). Altogether 60 peptides were identified with at least some of the programs (72 cysteine containing peptides with molecular mass 500-3500 would be theoretically present in the mixture). All the programs were able to identify 21 common peptides, while 25 peptides were identified with two software programs, and the rest 14 with a single algorithm. ProICAT identified most of the peptides, altogether 51, while SEQUEST and Spectrum Mill were able to identify 43 and 33 peptides, respectively. These results indicate a reasonable concordance of qualitative results between the different software algorithms, however it is notable that the programs produce complementary peptide identifications with each other. Quantification properties of the software programs were also evaluated with the protein mixture with known quantities of proteins. Quantitative results generated by the different programs showed no significant differences, when compared to the known quantities of the proteins (III, Table 3). When the sum of errors (comparison between software value and known value) were calculated SEQUEST showed the smallest value, while Spectrum Mill showed the largest one. One important observation in the analysis of the standard protein mixture was that the parameters used to filter significant peptide hits have dramatic effect also on quantitative results. When the results generated by ProICAT were filtered more stringently, the standard deviation of lactoferrin changed from 0.33 to 0.10. With Spectrum Mill similar kind of change in standard deviation from 0.47 to 0.09 was observed.

A complex protein mixture extracted from a biological sample presents a much more demanding task in respect of data analysis. When the MS/MS data from such a complex biological sample was analysed, similar differences were observed with the peptide and protein identifications as with the standard protein mixture. The different programs were able to identify altogether 720 peptides. 175 of these were identified with all the three algorithms, while 227 peptides were identified with two programs,

and the rest 318 peptides were identified with a single algorithm. ProICAT was able to identify the highest number of peptides and proteins, 544 and 328, respectively, while SEQUEST produced 484 peptides and 318 proteins, and Spectrum Mill 269 and 146. Figure 5 in report III summarizes all the peptide and protein identifications generated from different software programs. Quantitative results of the common peptide and protein identifications showed no statistically significant differences between the software programs (III, Table 3).

In similar kind of studies Chamrad et al. have compared the performance of software programs SEQUEST and Mascot (Chamrad et al., 2004a), and Kapp et al. have evaluated SEQUEST, Mascot, Spectrum Mill, Sonar, and X ! Tandem (Kapp et al., 2005). In the study by Chamrad et al. MALDI-PSD was used in the analysis of 2-DE separated mouse brain proteins. Mascot identified 16 proteins and SEQUEST produced 33 identifications out of which 15 were common to both programs. Kapp et al. compared the performance of altogether five software programs to analyse chromatographically fractionated blood samples from HUPO Plasma Proteome Project. Interestingly, among the tested software tools were Mascot, SEQUEST, and Spectrum Mill, thus enabling the comparison with our results and the results of Chamrad et al. In comparison between SEQUEST and Mascot Kapp et al. found 463 common peptides with 29 for Mascot alone and 63 for just SEQUEST. In the case of SEQUEST and Spectrum Mill, there were 402 peptides identified by both, while SEQUEST identified 124 peptides alone and Spectrum Mill 74. Figure 9 illustrates the performance of SEQUEST, Mascot, and Spectrum Mill in the different studies.

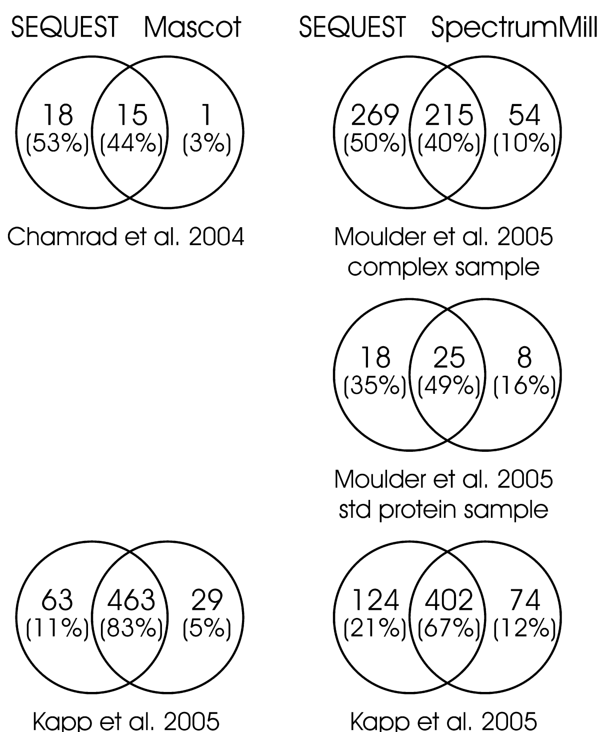


Figure. 9: Comparison of results between Chamrad et al. 2004, Moulder et al. 2005, and Kapp et al. 2005.

5.6 Quality classification of tandem mass spectrometry data (IV)

After the evaluation of MS/MS data analyzing software packages, the project was continued with creating a new data algorithm to help in MS/MS data analysis process. A pre-filtering software tool was created in order to classify MS/MS spectra to two categories: 1) Good-quality spectra containing valuable information for peptide identification and 2) bad-quality spectra originating not from peptides or containing insufficient information for interpretation. Pre-filtering is performed prior to data analysis with database search programs, thus allowing the removal of bad-quality spectra before the database searching. In this way both the database search time is reduced and the opportunities for false positive identifications decreased.

The classification of the spectra was achieved by monitoring certain spectral features and using decision-tree supervised classification techniques to predict the quality of the spectra. Previously, Bern et al. (2007) had described a set of seven spectral features (B1-B7) that were used to classify and prefilter MS/MS data. In this study additional features were tried to find to improve the classification results. Nine features (F1-F9) were employed that measure both the overall (F1-F3) and specific (F4-F9) attributes of the spectrum. The latter features were developed in concordance with the guidelines for manual interpretation of MS/MS spectra (Kinter and Sherman, 2000). Table 9 summarizes all the used spectral features, which are described more carefully in the reports by Bern et al. 2004 and Salmi et al. 2005 (IV). After the features were computed for each spectrum present in the MS/MS data, supervised classification methods were used to classify the good- and bad-quality spectra.

Table 9: Following spectral features were monitored to assist in the classification of good- and bad-quality spectra. Column "Freq." indicates frequency of the features selected for classification by Random Forest.

Feat.:	Description:	Freq.:
F1	The average intensity of the peaks	4
F2	The standard deviation of the peaks	1
F3	The total intensity of the exceptionally high peaks	10
F4	The presence of immonium ions	1
F5	The total intensity of peaks resulting from the ICAT reagent	2
F6	The total intensity of y1-ion peak of tryptic peptide	7
F7	The total intensity of precursor ion	4
F8	The total intensity of ions y_{n-2}, b_2, and b_{n-1}	2
F9	A score based on the mass-ladder, a preliminary peptide sequence, built using rank-normalized peaks	1
B1	The total intensity	3
B2	The intensity balance	3
B3	The number of peaks	3
B4	The total intensity of peaks with water losses	1
B5	The total intensity of peaks with isotopes	1
B6	The total intensity of peak pairs with mass gap similar to a known mass gap of an amino acid	0
B7	The total intensity of pairs of peaks which have the sum of masses similar to the parent ion mass	5

MS/MS data originating from a standard protein mixture (III) and from a complex biological protein sample (I) were classified both by features F1-F9 and B1-B7. The resulting classifications were then compared against results obtained by manual classification. ROC-curves describe the performance of the two sets of the features to classify the MS/MS spectra present in the different data sets (solid line represents classification with features F1-F9 and dotted line with features B1-B7) (IV, Fig. 2). In Figure 2 the vertical axis represents the true positive rate (TPR) and the horizontal axis the false positive rate (FPR). In the optimal classification the TPR=1, the FPR=0, and the area under curve (AUC) should be 1, meaning that the line should be as close the upper left corner as possible. The averages of the calculated AUC values and their standard deviations were 0.84±0.04, 0.77±0.03, and 0.84±0.04 for the features F1-F9, B1-B7, and the both features, respectively (IV, Table 2). FPR values were also calculated when the TPR value was fixed to 0.9 meaning a 10% loss of good-quality spectra (IV, Table 2). The FPR averages were 0.39±0.10 for F1-F9, 0.50±0.06 for B1-B7, and 0.38±0.11 for combined features. These results indicate that features F1-F9 classify the MS/MS spectra in this study material slightly better than features B1-B7 resulting in improved pre-filtering. However, it should be noted that the analysed MS/MS data (different types of MS-instruments, ICAT labeling) and different data classification and validation methods might have effect on the results.

The importance of the different features for the classification was also studied. Therefore the features, which were selected most often for the classification by the Random Forest, were determined. Table 9 summarizes the frequency of the selected features. Feature F3 monitoring the total intensity of high peaks in the spectrum was the only feature, which was selected in every test material. However, features F6, F7, and B7 monitoring the presence of y1-ion peak of tryptic peptides and the precursor ion peaks were frequently selected in the test materials. Nevertheless, as all the features were selected at least in one of the test materials, all the features (B1-B7 and F1-F9) were included in the classification in order to achieve the best possible results.

6 SUMMARY

In this Ph.D. thesis, proteomics applications were used to study the early regulation of Th lymphocyte differentiation. The aim of the studies was to elucidate the effects of IL-4 and IL-12 on the proteome of Th lymphocytes. In particular, the goal was to screen and identify proteins with novel roles in respect to Th1 and Th2 differentiation. To achieve the aims, high-throughput LC-MS/MS, stable isotope labeling, and data analysis tools were set-up. Even though proteomics methods are nowadays technically quite straightforward, careful optimization of the methods is needed for each research question. Therefore, sample preparation, fractionation, and analysis methods were optimized for ICAT and iTRAQ labeling in these studies (I, III, IV, unpublished data).

Th lymphocytes were extracted from either leucocyte-rich buffy coat samples or cord blood samples. The cells were activated and cultured with either Th1 promoting IL-12 or Th2 promoting IL-4. Microsomal fraction and nuclear proteins were isolated and studied with stable isotope labeling, microsomal fractions with ICAT (I, II) and nuclear proteins with iTRAQ (unpublished results). The increasing number of identified and quantified proteins that were determined during the studies reflected technological development. Even if the sample material, the number of analysed samples, labeling methods, and data analysis programs and parameters differed between these experiments, the trend with the improved results was clear. Differentially regulated proteins were also successfully identified in these proteomics applications.

In the characterization of microsomal fraction proteins extracted from IL-12 treated lymphoblasts Galectin-1 and CD7 were found to be down-regulated (I). Interestingly, Galectin-1 is a ligand for CD7 and this ligand-receptor pair induces apoptosis. Gal-1 has also been demonstrated to affect on cytokine production. In particular the treatment of experimental autoimmune diseases with Gal-1 has been shown to reduce the production of Th1 type cytokines. These observations suggest that Gal-1 expression influences the cytokine production and vice versa cytokine signalling has an effect on Gal-1 expression. Proteome profiling of microsomal fractions from IL-4 treated cord blood CD4+ cells resulted in finding of STAT1, MXA and GIMAP family members 1 and 4, which all were down-regulated (II). GIMAP1 and GIMAP4 belong to the same protein family of GTPases of the immune-associated proteins. Quantitative real-time RT-PCR analyses showed that IL-12 up-regulated and IL-4 down-regulated both *GIMAP1* and *GIMAP4*, while Western blotting showed the up-regulation of GIMAP4 by IL-12 and down-regulation by IL-4. These results imply that GIMAP family members have a role in Th cell differentiation. Biological findings, related to the characterization of nuclear proteins extracted from IL-4 treated cells, are currently the focus of functional studies.

Stable isotope labeling methods, like ICAT and iTRAQ, combined with high-throughput LC-MS/MS quickly create a lot of data. Therefore, powerful software programs are needed for data processing, handling, and storing. The need to improve and automate the data analysis pipeline, after the first study, resulted in the evaluation of database search programs ProICAT, SEQUEST, and SpectrumMill (III). This comparative study and similar studies by Kapp et al., 2005 and Chamrad et al., 2004b

have shown that different data analysis algorithms have variability in the protein identifications, especially when analyzing data from complex protein samples. However, the correlation of quantitative results was excellent between all the evaluated software tools. This software evaluation was continued by creating a pre-filtering algorithm, which is able to distinguish the good- and bad-quality spectra (IV). By this means, bad-quality data is removed prior to database searches thus reducing the time spent for data processing and potentially reducing the number of false positive protein identifications. The data classification was based on a set of spectral features, which were monitored from the MS/MS spectra and their quality was predicted with a decision-tree supervised classification techniques. The performance of the algorithm to classify the spectra was compared with a pre-filtering too described earlier Bern et al., 2004. Our results showed that the new pre-filtering algorithm classified good- and bad-quality data better than the earlier one. Furthermore these results indicated its suitability for incorporation in our data analysis pipeline in the processing of MS/MS data originating from ICAT labeled samples.

Altogether, these studies have revealed new aspects of the proteome of Th cells polarized towards Th1 or Th2 subset with IL-12 or IL-4, respectively. Furthermore, differentially regulated proteins, novel in this context, were identified. In addition to these biological discoveries, the establishment of these methods in our research laboratories, and associated national collaborations, has brought methodological progress for proteomics research in Finland.

ACKNOWLEDGEMENTS

I started my Ph.D. studies almost seven years ago at Turku Centre for Biotechnology. During these years I have got a lot of support and encouragement from many people to achieve the goal that will be celebrated soon. I wish to express my gratitude for all those people.

During my Ph.D. studies, I have been privileged to get scientific guidance from two excellent mentors: Professor Riitta Lahesmaa and Docent Tuula Nyman. I am grateful for all their support and pieces of advice. Professor Riitta Lahesmaa is acknowledged for giving me a chance to learn and do immunological research in her Molecular Immunology group, ATLAS, with modern instruments and techniques in excellent working facilities. Especially, I am grateful for her support regarding to scientific training in different courses and symposia as well as in research visits. Docent Tuula Nyman is acknowledged for guiding me in proteomics research. I have been able to learn so much about proteins and proteomics from you. I am extremely grateful for all that information.

Professor David Goodlett and Professor Seppo Auriola are acknowledged for following the progress of my research and studies as members of the supervisory committee. I appreciate very much all the excellent comments and advices you have provided to me.

I wish to express my gratitude to the National Graduate School in Informational and Structural Biology (ISB) and its Director, Professor Mark Johnsson. I am grateful for the financial support to do my Ph.D. studies and to participate on different symposia abroad. Professor Jukka Finne and Medical Biochemistry and Molecular Biology department of Medical Faculty are acknowledged for offering a student position to perform the theoretical studies.

Professor Seppo Auriola and Docent Panu Kovanen are acknowledged for carefully reviewing my Ph.D. thesis. I am thankful for all the excellent comments and suggestions. Dr. Robert Moulder is greatly acknowledged for all the comments and thorough language revision of the thesis.

I wish to thank all the collaborators and co-authors who worked in the projects of this thesis: Professor Olli Nevalainen, Professor David Goodlett, Dr. Robert Moulder, Dr. Omid Rasool, Dr. Anne West, Dr. Jussi Salmi, Dr. Tero Aittokallio, Dr. Matej Oresic, Sanna Filén, Petri Kouvonen, Suvi Kantola, Juha Korhonen, Mikko Katajamaa, and Mari Björkman. I appreciate very much your important contribution for these projects.

Turku Centre for Biotechnology has been a great environment to do research with modern instruments in excellent facilities. Especially, I wish to express my gratitude to the skilful personnel of the Centre. During my Ph.D. studies I have spent a lot of time in the proteomics lab and got a lot of help for my studies from the Proteomics and Mass Spectrometry Unit. I wish to greatly thank the members of the unit: Dr. Garry Corthals, Dr. Anne Rokka, Petri Kouvonen, Raija Andersen, Jani Korhonen, and Jonna Lindholm-Ventola. The secretarial and technical staff of CBT has always had time to patiently help me in many things. I appreciate a lot all your efforts. Especially, I wish

to thank Sirkku Grönroos, Aila Jasmanvaara, and Eva Hirvensalo from the office, it-support Mårten Hedman and Petri Vahakoski, technical maintenance Pasi Viljakainen, Juha Stranden, and Rolf Sara, Virpi Korpiranta and Hannele Vuori from instrument maintenance, and coordinators Dr. Minttu Jaakkola, Outi Irjala, Heidi Sid, and Dr. Heli Salminen-Mankonen.

I want to thank all the current and former members of ATLAS group. It has been a pleasure to do research with you. Especially, I want to thank Helena Ahlfors, Dr. Laura Elo, Sanna Filén, Dr. Tiina Henttinen, Dr. Waltteri Hosia, Suvi Kantola, Juha Korhonen, Minna Kyläniemi, Tapio Lönnberg, Dr. Maritta Löytömäki, Dr. Robert Moulder, Dr. Juha Mykkänen, Tuomas Nikula, Pekka Ojala, Dr. Eeva Rainio, Dr. Omid Rasool, Dr. Kirsi Rautajoki, Dr. Arsi Rosengren, Johanna Tavanainen, and Soile Tuomela for all your support and help. Special thanks go to the skilful technicians of ATLAS group, who have helped me a lot in the experiments: Marjo Hakkarainen, Sarita Heinonen, Marjo Linja, Outi Melin, Paula Suominen, and Marju Niskala.

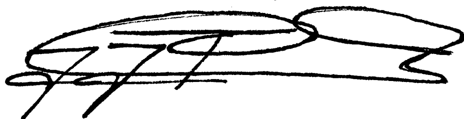
My warmest thanks are given to all my friends. I appreciate all the memorable moments spent with you.

I am extremely thankful for my and Sanna's families for the endless support and motivation you have given to do studies. I wish to warmly thank Bengt, Marja, Tomas, Marika, Reijo, Marjatta, and Tanja. You are all very special to me.

Finally, I express my gratitude to the most important persons of my life, my wife Sanna and my son Anton by dedicating this thesis to them. You both bring so much happiness and light to me every day. I am extremely grateful for every day I can share with you. My beloved Sanna, words can't describe my gratitude for all the love, joy, and happiness you have given to me. Thank you very much for all the support and help you have given to reach these final words of my thesis.

This work was financially supported by The Academy of Finland, The National Technology Agency of Finland (TEKES), Sigrid Juselius Foundation, Turku University Hospital Research Fund, National Graduate School in Informational and Structural Biology (ISB), Finnish Cultural Foundation, Väinö and Laina Kivi Foundation, Ida Montin Foundation, Tampere Tuberculosis Foundation, and Turku University Foundation.

Kirkkonummi, January 20th, 2008



REFERENCES

Aebersold, R. and M. Mann. 2003. Mass spectrometry-based proteomics. Nature 422:198-207.

Agnello, D., C.S. Lankford, J. Bream, A. Morinobu, M. Gadina, J.J. O'Shea and D.M. Frucht. 2003. Cytokines and transcription factors that regulate T helper cell differentiation: New players and new insights. J. Clin. Immunol. 23:147-161.

Airoldi, I., G. Gri, J.D. Marshall, A. Corcione, P. Facchetti, R. Guglielmino, G. Trinchieri and V. Pistoia. 2000. Expression and function of IL-12 and IL-18 receptors on human tonsillar B cells. J. Immunol. 165:6880-6888.

Alam, R. and M. Gorska. 2003. 3. Lymphocytes. J. Allergy Clin. Immunol. 111:s476-s485.

Allione, A., V. Wells, G. Forni, L. Mallucci and F. Novelli. 1998. Beta-galactoside-binding protein (beta GBP) alters the cell cycle, up-regulates expression of the alpha- and beta-chains of the IFN-gamma receptor, and triggers IFN-gamma-mediated apoptosis of activated human T lymphocytes. J. Immunol. 161:2114-2119.

Andersen, J.S. and M. Mann. 2006. Organellar proteomics: Turning inventories into insights. EMBO Rep. 7:874-879.

Andersen, J.S., Y.W. Lam, A.K. Leung, S.E. Ong, C.E. Lyon, A.I. Lamond and M. Mann. 2005. Nucleolar proteome dynamics. Nature 433:77-83.

Andrews, N.C. and D.V. Faller. 1991. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. 19:2499.

Andrews, R.P., M.B. Ericksen, C.M. Cunningham, M.O. Daines and G.K. Hershey. 2002. Analysis of the life cycle of stat6. continuous cycling of STAT6 is required for IL-4 signaling. J. Biol. Chem. 277:36563-36569.

Azfali, B., G. Lombardi, R.I. Lechler and G.M. Lord. 2007. The role of T helper 17 (Th17) and regulatory T cells (Treg) in human organ transplantationand autoimmune disease. Clin. Exp. Immunol. 148:32-46.

Bacon, C.M., D.W. McVicar, J.R. Ortaldo, R.C. Rees, J.J. O'Shea and J.A. Johnston. 1995a. Interleukin 12 (IL-12) induces tyrosine phosphorylation of JAK2 and TYK2: Differential use of janus family tyrosine kinases by IL-2 and IL-12. J. Exp. Med. 181:399-404.

Bacon, C.M., E.F. Petricoin 3rd, J.R. Ortaldo, R.C. Rees, A.C. Larner, J.A. Johnston and J.J. O'Shea. 1995b. Interleukin 12 induces tyrosine phosphorylation and activation of STAT4 in human lymphocytes. Proc. Natl. Acad. Sci. U. S. A. 92:7307-7311.

Bagshaw, R.D., D.J. Mahuran and J.W. Callahan. 2005. A proteomic analysis of lysosomal integral membrane proteins reveals the diverse composition of the organelle. Mol. Cell. Proteomics 4:133-143.

Bai, F., S. Liu and F.A. Witzmann. 2005. A "de-streaking" method for two-dimensional electrophoresis using the reducing agent tris(2-carboxyethyl)-phosphine hydrochloride and alkylating agent vinylpyridine. Proteomics 5:2043-2047.

Baum, L.G., D.P. Blackall, S. Arias-Magallano, D. Nanigian, S.Y. Uh, J.M. Browne, D. Hoffmann, C.E. Emmanouilides, M.C. Territo and G.C. Baldwin. 2003. Amelioration of graft versus host disease by galectin-1. Clin. Immunol. 109:295-307.

Beardsley, R.L. and J.P. Reilly. 2003. Quantitation using enhanced signal tags: A technique for comparative proteomics. J. Proteome Res. 2:15-21.

Belov, M.E., M.V. Gorshkov, H.R. Udseth, G.A. Anderson and R.D. Smith. 2000. Zeptomole-sensitivity electrospray ionization--fourier transform ion cyclotron resonance mass spectrometry of proteins. Anal. Chem. 72:2271-2279.

Berenson, L.S., N. Ota and K.M. Murphy. 2004. Issues in T-helper 1 development--resolved and unresolved. Immunol. Rev. 202:157-174.

Berggren, K., T.H. Steinberg, W.M. Lauber, J.A. Carroll, M.F. Lopez, E. Chernokalskaya, L. Zieske, Z. Diwu, R.P. Haugland and W.F. Patton. 1999. A luminescent ruthenium complex for ultrasensitive detection of proteins immobilized on membrane supports. Anal. Biochem. 276:129-143.

Bern, M., D. Goldberg, W.H. McDonald and J.R. Yates 3rd. 2004. Automatic quality assessment of peptide tandem mass spectra. Bioinformatics 20 Suppl 1:i49-54.

Bini, L., S. Pacini, S. Liberatori, S. Valensin, M. Pellegrini, R. Raggiaschi, V. Pallini and C.T. Baldari. 2003. Extensive temporally regulated reorganization of the lipid raft proteome following T-cell antigen receptor triggering. Biochem. J. 369:301-309.

Bjellqvist, B., K. Ek, P.G. Righetti, E. Gianazza, A. Gorg, R. Westermeier and W. Postel. 1982. Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications. J. Biochem. Biophys. Methods 6:317-339.

Blagoev, B., S.E. Ong, I. Kratchmarova and M. Mann. 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat. Biotechnol. 22:1139-1145.

Blaser, C., M. Kaufmann, C. Muller, C. Zimmermann, V. Wells, L. Mallucci and H. Pircher. 1998. Beta-galactoside-binding protein secreted by activated T cells inhibits antigen-induced proliferation of T cells. Eur. J. Immunol. 28:2311-2319.

Blonder, J., A. Terunuma, T.P. Conrads, K.C. Chan, C. Yee, D.A. Lucas, C.F. Schaefer, L.R. Yu, H.J. Issaq, T.D. Veenstra and J.C. Vogel. 2004. A proteomic characterization of the plasma membrane of human epidermis by high-throughput mass spectrometry. J. Invest. Dermatol. 123:691-699.

Bosco, A., K.L. McKenna, C.J. Devitt, M.J. Firth, P.D. Sly and P.G. Holt. 2006. Identification of novel Th2-associated genes in T memory responses to allergens. J. Immunol. 176:4766-4777.

Boulay, J.L., J.J. O'Shea and W.E. Paul. 2003. Molecular phylogeny within type I cytokines and their cognate receptors. Immunity 19:159-163.

Brancia, F.L., H. Montgomery, K. Tanaka and S. Kumashiro. 2004. Guanidino labeling derivatization strategy for global characterization of peptide mixtures by liquid chromatography matrix-assisted laser desorption/ionization mass spectrometry. Anal. Chem. 76:2748-2755.

Brill, L.M., A.R. Salomon, S.B. Ficarro, M. Mukherji, M. Stettler-Gill and E.C. Peters. 2004. Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry. Anal. Chem. 76:2763-2772.

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

Cahill, M.A., W. Wozny, G. Schwall, K. Schroer, K. Holzer, S. Poznanovic, C. Hunzinger, J.A. Vogt, W. Stegmann, H. Matthies and A. Schrattenholz. 2003. Analysis of relative isotopologue abundances for quantitative profiling of complex protein mixtures labelled with the acrylamide/D3-acrylamide alkylation tag system. Rapid Commun. Mass Spectrom. 17:1283-1290.

Cambot, M., S. Aresta, B. Kahn-Perles, J. de Gunzburg and P.H. Romeo. 2002. Human immune associated nucleotide 1: A member of a new guanosine triphosphatase family expressed in resting T and B cells. Blood 99:3293-3301.

Candiano, G., M. Bruschi, L. Musante, L. Santucci, G.M. Ghiggeri, B. Carnemolla, P. Orecchia, L. Zardi and P.G. Righetti. 2004. Blue silver: A very sensitive colloidal coomassie G-250 staining for proteome analysis. Electrophoresis 25:1327-1333.

Cao, L., K. Yu and A.R. Salomon. 2006. Phosphoproteomic analysis of lymphocyte signaling. Adv. Exp. Med. Biol. 584:277-288.

Carr, S., R. Aebersold, M. Baldwin, A. Burlingame, K. Clauser, A. Nesvizhskii and Working Group on Publication Guidelines for Peptide and Protein Identification Data. 2004. The need for guidelines in publication of peptide and protein identification data: Working group on publication guidelines for peptide and protein identification data. Mol. Cell. Proteomics 3:531-533.

Carter, C., C. Dion, S. Schnell, W.J. Coadwell, M. Graham, L. Hepburn, G. Morgan, A. Hutchings, J.C. Pascall, H. Jacobs, J.R. Miller and G.W. Butcher. 2007. A natural hypomorphic variant of the apoptosis regulator Gimap4/IAN1. J. Immunol. 179:1784-1795.

Chakraborty, A. and F.E. Regnier. 2002. Global internal standard technology for comparative proteomics. J. Chromatogr. A 949:173-184.

Chamrad, D.C., G. Korting, K. Stuhler, H.E. Meyer, J. Klose and M. Bluggel. 2004a. Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data. Proteomics 4:619-628.

Chamrad, D.C., G. Korting, K. Stuhler, H.E. Meyer, J. Klose and M. Bluggel. 2004b. Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data. Proteomics 4:619-628.

Chan, A.S., J.L. Mobley, G.B. Fields and Y. Shimizu. 1997. CD7-mediated regulation of integrin adhesiveness on human T cells involves tyrosine phosphorylation-dependent activation of phosphatidylinositol 3-kinase. J. Immunol. 159:934-942.

Chan, S.H., B. Perussia, J.W. Gupta, M. Kobayashi, M. Pospisil, H.A. Young, S.F. Wolf, D. Young, S.C. Clark and G. Trinchieri. 1991. Induction of interferon gamma production by natural killer cell stimulatory factor: Characterization of the responder cells and synergy with other inducers. J. Exp. Med. 173:869-879.

Che, F.Y. and L.D. Fricker. 2002. Quantitation of neuropeptides in cpe(fat)/Cpe(fat) mice using differential isotopic tags and mass spectrometry. Anal. Chem. 74:3190-3198.

Chen, Z., R. Lund, T. Aittokallio, M. Kosonen, O. Nevalainen and R. Lahesmaa. 2003. Identification of novel IL-4/Stat6-regulated genes in T lymphocytes. J. Immunol. 171:3627-3635.

Chouaib, S., J. Chehimi, L. Bani, N. Genetet, T. Tursz, F. Gay, G. Trinchieri and F. Mami-Chouaib. 1994. Interleukin 12 induces the differentiation of major histocompatibility complex class I-primed cytotoxic T-lymphocyte precursors into allospecific cytotoxic effectors. Proc. Natl. Acad. Sci. U. S. A. 91:12659-12663.

Chtanova, T., R.A. Kemp, A.P. Sutherland, F. Ronchese and C.R. Mackay. 2001. Gene microarrays reveal extensive differential gene expression in both CD4(+) and CD8(+) type 1 and type 2 T cells. J. Immunol. 167:3057-3063.

Chtanova, T., S.G. Tangye, R. Newton, N. Frank, M.R. Hodge, M.S. Rolph and C.R. Mackay. 2004. T follicular helper cells express a distinctive transcriptional profile, reflecting their role as non-Th1/Th2 effector cells that provide help for B cells. J. Immunol. 173:68-78.

Chtanova, T., R. Newton, S.M. Liu, L. Weininger, T.R. Young, D.G. Silva, F. Bertoni, A. Rinaldi, S. Chappaz, F. Sallusto, M.S. Rolph and C.R. Mackay. 2005. Identification of T cell-restricted genes, and signatures for different T cell responses, using a comprehensive collection of microarray datasets. J. Immunol. 175:7837-7847.

Chua, A.O., V.L. Wilkinson, D.H. Presky and U. Gubler. 1995. Cloning and characterization of a mouse IL-12 receptor-beta component. J. Immunol. 155:4286-4294.

Chua, A.O., R. Chizzonite, B.B. Desai, T.P. Truitt, P. Nunes, L.J. Minetti, R.R. Warrier, D.H. Presky, J.F. Levine and M.K. Gately. 1994. Expression cloning of a human IL-12 receptor component. A new member of the cytokine receptor superfamily with strong homology to gp130. J. Immunol. 153:128-136.

Clauser, K.R., P. Baker and A.L. Burlingame. 1999. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal. Chem. 71:2871-2882.

Coffman, R.L., J. Ohara, M.W. Bond, J. Carty, A. Zlotnik and W.E. Paul. 1986. B cell stimulatory factor-1 enhances the IgE response of lipopolysaccharide-activated B cells. J. Immunol. 136:4538-4541.

Colinge, J., A. Masselot, M. Giron, T. Dessingy and J. Magnin. 2003. OLAV: Towards high-throughput tandem mass spectrometry data identification. Proteomics 3:1454-1463.

Constant, S.L. and K. Bottomly. 1997. Induction of Th1 and Th2 CD4+ T cell responses: The alternative approaches. Annu. Rev. Immunol. 15:297-322.

Corthals, G.L. and K. Rose. 2007. Quantitation in proteomics. *In* Wilkins (ed.) Proteome research: New frontiers in functional genomics. Springer-Verlag, .

Craig, R. and R.C. Beavis. 2004. TANDEM: Matching proteins with tandem mass spectra. Bioinformatics 20:1466-1467.

Crick, F. 1970. Central dogma of molecular biology. Nature 227:561-563.

Cristea, I.M., S.J. Gaskell and A.D. Whetton. 2004. Proteomics techniques and their application to hematology. Blood 103:3624-3634.

D'Ambrosio, C., L. Gatta and S. Bonini. 2005. The future of microarray technology: Networking the genome search. Allergy 60:1219-1226.

Desai, B.B., P.M. Quinn, A.G. Wolitzky, P.K. Mongini, R. Chizzonite and M.K. Gately. 1992. IL-12 receptor. II. distribution and regulation of receptor expression. J. Immunol. 148:3125-3132.

Desiderio, D.M. and M. Kai. 1983. Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. Biomed. Mass Spectrom. 10:471-479.

Di Bartolo, V. and O. Acuto. 2004. Large-scale screening for genes involved in T-cell signaling: Do we know all the players now? Trends Immunol. 25:399-402.

Diezel, W., G. Kopperschlager and E. Hofmann. 1972. An improved procedure for protein staining in polyacrylamide gels with a new type of coomassie brilliant blue. Anal. Biochem. 48:617-620.

Dion, C., C. Carter, L. Hepburn, W.J. Coadwell, G. Morgan, M. Graham, N. Pugh, G. Anderson, G.W. Butcher and J.R. Miller. 2005. Expression of the ian family of putative GTPases during T cell development and description of an ian with three sets of GTP/GDP-binding motifs. Int. Immunol. 17:1257-1268.

Domon, B. and R. Aebersold. 2006. Mass spectrometry and protein analysis. Science 312:212-217.

Dunkley, T.P., R. Watson, J.L. Griffin, P. Dupree and K.S. Lilley. 2004. Localization of organelle proteins by isotope tagging (LOPIT). Mol. Cell. Proteomics 3:1128-1134.

Dziembowski, A. and B. Seraphin. 2004. Recent developments in the analysis of protein complexes. FEBS Lett. 556:1-6.

Eng, J., A. McCormack and J.R. Yates. 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.

Fazekas de St Groth, S., R.G. Webster and A. Datyner. 1963. Two new staining procedures for quantitative estimation of proteins on electrophoresis strips. Biochim. Biophys. Acta 71:377-391.

Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong and C.M. Whitehouse. 1989. Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64-71.

Fenyo, D., J. Qin and B.T. Chait. 1998. Protein identification using mass spectrometric information. Electrophoresis 19:998-1005.

Field, H.I., D. Fenyo and R.C. Beavis. 2002. RADARS, a bioinformatics solution that automates proteome mass spectral analysis, optimises protein identification, and archives data in a relational database. Proteomics 2:36-47.

Fields, S. 2001. Proteomics. proteomics in genomeland. Science 291:1221-1224.

Filen, J.J., T.A. Nyman, J. Korhonen, D.R. Goodlett and R. Lahesmaa. 2005. Characterization of microsomal fraction proteome in human lymphoblasts reveals the down-regulation of galectin-1 by interleukin-12. Proteomics 5:4719-4732.

Fournie, G.J., B. Cautain, E. Xystrakis, J. Damoiseaux, M. Mas, D. Lagrange, I. Bernard, J.F. Subra, L. Pelletier, P. Druet and A. Saoudi. 2001. Cellular and genetic factors involved in the difference between brown norway and lewis rats to develop respectively type-2 and type-1 immune-mediated diseases. Immunol. Rev. 184:145-160.

Fratelli, M., H. Demol, M. Puype, S. Casagrande, I. Eberini, M. Salmona, V. Bonetto, M. Mengozzi, F. Duffieux, E. Miclet, A. Bachi, J. Vandekerckhove, E. Gianazza and P. Ghezzi. 2002. Identification by redox proteomics of glutathionylated proteins in oxidatively stressed human T lymphocytes. Proc. Natl. Acad. Sci. U. S. A. 99:3505-3510.

Fritze, C.E. and T.R. Anderson. 2000. Epitope tagging: General method for tracking recombinant proteins. Methods Enzymol. 327:3-16.

Gabius, H.J. 2001. Probing the cons and pros of lectin-induced immunomodulation: Case studies for the mistletoe lectin and galectin-1. Biochimie 83:659-666.

Galizzi, J.P., C.E. Zuber, N. Harada, D.M. Gorman, O. Djossou, R. Kastelein, J. Banchereau, M. Howard and A. Miyajima. 1990. Molecular cloning of a cDNA encoding the human interleukin 4 receptor. Int. Immunol. 2:669-675.

Gan, C.S., P.K. Chong, T.K. Pham and P.C. Wright. 2007. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). J. Proteome Res. 6:821-827.

Gascan, H., J.F. Gauchat, M.G. Roncarolo, H. Yssel, H. Spits and J.E. de Vries. 1991. Human B cell clones can be induced to proliferate and to switch to IgE and IgG4 synthesis by interleukin 4 and a signal provided by activated CD4+ T cell clones. J. Exp. Med. 173:747-750.

Gately, M.K., A.G. Wolitzky, P.M. Quinn and R. Chizzonite. 1992. Regulation of human cytolytic lymphocyte responses by interleukin-12. Cell. Immunol. 143:127-142.

Gately, M.K., L.M. Renzetti, J. Magram, A.S. Stern, L. Adorini, U. Gubler and D.H. Presky. 1998. The interleukin-12/interleukin-12-receptor system: Role in normal and pathologic immune responses. Annu. Rev. Immunol. 16:495-521.

Gately, M.K., R.R. Warrier, S. Honasoge, D.M. Carvajal, D.A. Faherty, S.E. Connaughton, T.D. Anderson, U. Sarmiento, B.R. Hubbard and M. Murphy. 1994. Administration of recombinant IL-12 to normal mice enhances cytolytic lymphocyte activity and induces production of IFN-gamma in vivo. Int. Immunol. 6:157-167.

Gately, M.K., B.B. Desai, A.G. Wolitzky, P.M. Quinn, C.M. Dwyer, F.J. Podlaski, P.C. Familletti, F. Sinigaglia, R. Chizonnite and U. Gubler. 1991. Regulation of human lymphocyte proliferation by a heterodimeric cytokine, IL-12 (cytotoxic lymphocyte maturation factor). J. Immunol. 147:874-882.

Gazzinelli, R.T., M. Wysocka, S. Hayashi, E.Y. Denkers, S. Hieny, P. Caspar, G. Trinchieri and A. Sher. 1994. Parasite-induced IL-12 stimulates early IFN-gamma synthesis and resistance during acute infection with toxoplasma gondii. J. Immunol. 153:2533-2543.

Gearing, D.P. and D. Cosman. 1991. Homology of the p40 subunit of natural killer cell stimulatory factor (NKSF) with the extracellular domain of the interleukin-6 receptor. Cell 66:9-10.

Gehanne, S., D. Cecconi, L. Carboni, P.G. Righetti, E. Domenici and M. Hamdan. 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.

Gerber, S.A., J. Rush, O. Stemman, M.W. Kirschner and S.P. Gygi. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc. Natl. Acad. Sci. U. S. A. 100:6940-6945.

Gharbi, S., P. Gaffney, A. Yang, M.J. Zvelebil, R. Cramer, M.D. Waterfield and J.F. Timms. 2002. Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. Mol. Cell. Proteomics 1:91-98.

Giometti, C.S., M.A. Gemmell, S.L. Tollaksen and J. Taylor. 1991. Quantitation of human leukocyte proteins after silver staining: A study with two-dimensional electrophoresis. Electrophoresis 12:536-543.

Glimcher, L.H. and K.M. Murphy. 2000. Lineage commitment in the immune system: The T helper lymphocyte grows up. Genes Dev. 14:1693-1711.

Goodlett, D.R., A. Keller, J.D. Watts, R. Newitt, E.C. Yi, S. Purvine, J.K. Eng, P. von Haller, R. Aebersold and E. Kolker. 2001. Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. Rapid Commun. Mass Spectrom. 15:1214-1221.

Gorg, A., W. Weiss and M.J. Dunn. 2004. Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665-3685.

Goshe, M.B. and R.D. Smith. 2003. Stable isotope-coded proteomic mass spectrometry. Curr. Opin. Biotechnol. 14:101-109.

Goshe, M.B., T.P. Conrads, E.A. Panisko, N.H. Angell, T.D. Veenstra and R.D. Smith. 2001. Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. Anal. Chem. 73:2578-2586.

Granucci, F., P.R. Castagnoli, L. Rogge and F. Sinigaglia. 2001. Gene expression profiling in immune cells using microarray. Int. Arch. Allergy Immunol. 126:257-266.

Grohmann, U., M.L. Belladonna, R. Bianchi, C. Orabona, E. Ayroldi, M.C. Fioretti and P. Puccetti. 1998. IL-12 acts directly on DC to promote nuclear localization of NF-kappaB and primes DC for IL-12 production. Immunity 9:315-323.

Gubler, U., A.O. Chua, D.S. Schoenhaut, C.M. Dwyer, W. McComas, R. Motyka, N. Nabavi, A.G. Wolitzky, P.M. Quinn and P.C. Familletti. 1991. Coexpression of two distinct genes is required to generate secreted bioactive cytotoxic lymphocyte maturation factor. Proc. Natl. Acad. Sci. U. S. A. 88:4143-4147.

Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb and R. Aebersold. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat. Biotechnol. 17:994-999.

Halligan, B.D., R.Y. Slyper, S.N. Twigger, W. Hicks, M. Olivier and A.S. Greene. 2005. ZoomQuant: An application for the quantitation of stable isotope labeled peptides. J. Am. Soc. Mass Spectrom. 16:302-306.

Han, D.K., J. Eng, H. Zhou and R. Aebersold. 2001. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat. Biotechnol. 19:946-951.

Hansen, K.C., G. Schmitt-Ulms, R.J. Chalkley, J. Hirsch, M.A. Baldwin and A.L. Burlingame. 2003. Mass spectrometric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and multidimensional chromatography. Mol. Cell. Proteomics 2:299-314.

Hebenstreit, D., G. Wirnsberger, J. Horejs-Hoeck and A. Duschl. 2006. Signaling mechanisms, interaction partners, and target genes of STAT6. Cytokine Growth Factor Rev. 17:173-188.

Hellquist, A., M. Zucchelli, K. Kivinen, U. Saarialho-Kere, S. Koskenmies, E. Widen, H. Julkunen, A. Wong, M.L. Karjalainen-Lindsberg, T. Skoog, J. Vendelin, D.S. Cunninghame-Graham, T.J. Vyse, J. Kere and C.M. Lindgren. 2007. The human GIMAP5 gene has a common polyadenylation polymorphism increasing risk to systemic lupus erythematosus. J. Med. Genet. 44:314-321.

Henzel, W.J., T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley and C. Watanabe. 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc. Natl. Acad. Sci. U. S. A. 90:5011-5015.

Herbert, B., M. Galvani, M. Hamdan, E. Olivieri, J. MacCarthy, S. Pedersen and P.G. Righetti. 2001. Reduction and alkylation of proteins in preparation of two-dimensional map analysis: Why, when, and how? Electrophoresis 22:2046-2057.

Hernandez, P., M. Muller and R.D. Appel. 2006. Automated protein identification by tandem mass spectrometry: Issues and strategies. Mass Spectrom. Rev. 25:235-254.

Hernandez-Hoyos, G., S. Joseph, N.G. Miller and G.W. Butcher. 1999. The lymphopenia mutation of the BB rat causes inappropriate apoptosis of mature thymocytes. Eur. J. Immunol. 29:1832-1841.

Heukeshoven, J. and R. Dernick. 1988. Improved silver staining procedure for fast staining in PhastSystem development unit. I. staining of sodium dodecyl sulfate gels. Electrophoresis 9:28-32.

Hicks, W.A., B.D. Halligan, R.Y. Slyper, S.N. Twigger, A.S. Greene and M. Olivier. 2005. Simultaneous quantification and identification using 18O labeling with an ion trap mass spectrometer and the analysis software application "ZoomQuant". J. Am. Soc. Mass Spectrom. 16:916-925.

Hoang, V.M., T.P. Conrads, T.D. Veenstra, J. Blonder, A. Terunuma, J.C. Vogel and R.J. Fisher. 2003. Quantitative proteomics employing primary amine affinity tags. J. Biomol. Tech. 14:216-223.

Hoogland, C., K. Mostaguir, J.C. Sanchez, D.F. Hochstrasser and R.D. Appel. 2004. SWISS-2DPAGE, ten years later. Proteomics 4:2352-2356.

Horn, D.M., R.A. Zubarev and F.W. McLafferty. 2000. Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. Proc. Natl. Acad. Sci. U. S. A. 97:10313-10317.

Hsieh, C.S., A.B. Heimberger, J.S. Gold, A. O'Garra and K.M. Murphy. 1992. Differential regulation of T helper phenotype development by interleukins 4 and 10 in an alpha beta T-cell-receptor transgenic system. Proc. Natl. Acad. Sci. U. S. A. 89:6065-6069.

Hsieh, C.S., S.E. Macatonia, C.S. Tripp, S.F. Wolf, A. O'Garra and K.M. Murphy. 1993. Development of TH1 CD4+ T cells through IL-12 produced by listeria-induced macrophages. Science 260:547-549.

Hsu, J.L., S.Y. Huang, N.H. Chow and S.H. Chen. 2003. Stable-isotope dimethyl labeling for quantitative proteomics. Anal. Chem. 75:6843-6852.

Huang, Y.F., C.C. Huang, C.C. Hu and H.T. Chang. 2006. Capillary electrophoresis-based separation techniques for the analysis of proteins. Electrophoresis 27:3503-3522.

Hunt, D.F., J.R. Yates 3rd, J. Shabanowitz, S. Winston and C.R. Hauer. 1986. Protein sequencing by tandem mass spectrometry. Proc. Natl. Acad. Sci. U. S. A. 83:6233-6237.

Hwang, S.I., D.H. Lundgren, V. Mayya, K. Rezaul, A.E. Cowan, J.K. Eng and D.K. Han. 2006. Systematic characterization of nuclear proteome during apoptosis: A quantitative proteomic study by differential extraction and stable isotope labeling. Mol. Cell. Proteomics 5:1131-1145.

Hylkema, M.N., M. van der Deen, J.M. Pater, J. Kampinga, P. Nieuwenhuis and H. Groen. 2000. Single expression of CD45RC and RT6 in correlation with T-helper 1 and T-helper 2 cytokine patterns in the rat. Cell. Immunol. 199:89-96.

Ideker, T., V. Thorsson, J.A. Ranish, R. Christmas, J. Buhler, J.K. Eng, R. Bumgarner, D.R. Goodlett, R. Aebersold and L. Hood. 2001. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292:929-934.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. Nature 431:931-945.

Jacob, R.J. and R. Cramer. 2006. PIGOK: Linking protein identity to gene ontology and function. J. Proteome Res. 5:3429-3432.

Jacobson, N.G., S.J. Szabo, R.M. Weber-Nordt, Z. Zhong, R.D. Schreiber, J.E. Darnell Jr and K.M. Murphy. 1995. Interleukin 12 signaling in T helper type 1 (Th1) cells involves tyrosine phosphorylation of signal transducer and activator of transcription (stat)3 and Stat4. J. Exp. Med. 181:1755-1762.

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

Ji, C., N. Guo and L. Li. 2005. Differential dimethyl labeling of N-termini of peptides after guanidination for proteome analysis. J. Proteome Res. 4:2099-2108.

Ji, J., A. Chakraborty, M. Geng, X. Zhang, A. Amini, M. Bina and F. Regnier. 2000. Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. J. Chromatogr. B Biomed. Sci. Appl. 745:197-210.

Johnson, R.S. and K. Biemann. 1987. The primary structure of thioredoxin from chromatium vinosum determined by high-performance tandem mass spectrometry. Biochemistry 26:1209-1214.

Jung, E., C. Hoogland, D. Chiappe, J.C. Sanchez and D.F. Hochstrasser. 2000. The establishment of a human liver nuclei two-dimensional electrophoresis reference map. Electrophoresis 21:3483-3487.

Kaplan, M.H., Y.L. Sun, T. Hoey, M.J. Grusby. 1996a. Impaired IL-12 responses and enhanced development of Th2 cells in Stat4-deficient mice. Nature 382:174-177.

Kaplan, M.H., U. Schindler, S.T. Smiley and M.J. Grusby. 1996b. Stat6 is required for mediating responses to IL-4 and for development of Th2 cells. Immunity 4:313-319.

Kapp, E.A., F. Schutz, L.M. Connolly, J.A. Chakel, J.E. Meza, C.A. Miller, D. Fenyo, J.K. Eng, J.N. Adkins, G.S. Omenn and R.J. Simpson. 2005. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: Sensitivity and specificity analysis. Proteomics 5:3475-3490.

Karas, M. and F. Hillenkamp. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal. Chem. 60:2299-2301.

Karsan, A., I. Pollet, L.R. Yu, K.C. Chan, T.P. Conrads, D.A. Lucas, R. Andersen and T. Veenstra. 2005. Quantitative proteomic analysis of sokotrasterol sulfate-stimulated primary human endothelial cells. Mol. Cell. Proteomics 4:191-204.

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

Keough, T., R.S. Youngquist and M.P. Lacey. 2003. Sulfonic acid derivatives for peptide sequencing by MALDI MS. Anal. Chem. 75:156A-165A.

Kettman, J.R., L. Kuhn, P. Young and I. Lefkovits. 1986. Parameters of the labeling of mitogen-activated murine lymphocytes by [35S]methionine for two-dimensional gel electrophoresis. I. effect of culture conditions. J. Immunol. Methods 88:53-64.

Kim, Y.H., K. Cho, S.H. Yun, J.Y. Kim, K.H. Kwon, J.S. Yoo and S.I. Kim. 2006. Analysis of aromatic catabolic pathways in pseudomonas putida KT 2440 using a combined proteomic approach: 2-DE/MS and cleavable isotope-coded affinity tag analysis. Proteomics 6:1301-1318.

Kim, Y.J., P. Zhan, B. Feild, S.M. Ruben, and T. He. 2007. Reproducibility assessment of relative quantitation strategies for LC-MS based proteomics. Anal. Chem. 79:5651-5658.

Kinter, M. and N.E. Sherman. 2000. Protein sequencing and identification using tandem mass spectrometry. Wiley, New York.

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

Knoblach, B., B.O. Keller, J. Groenendyk, S. Aldred, J. Zheng, B.D. Lemire, L. Li and M. Michalak. 2003. ERp19 and ERp46, new members of the thioredoxin family of endoplasmic reticulum proteins. Mol. Cell. Proteomics 2:1104-1119.

Kobayashi, M., L. Fitz, M. Ryan, R.M. Hewick, S.C. Clark, S. Chan, R. Loudon, F. Sherman, B. Perussia and G. Trinchieri. 1989. Identification and purification of natural killer cell stimulatory factor (NKSF), a cytokine with multiple biologic effects on human lymphocytes. J. Exp. Med. 170:827-845.

Kopf, M., G. Le Gros, M. Bachmann, M.C. Lamers, H. Bluethmann and G. Kohler. 1993. Disruption of the murine IL-4 gene blocks Th2 cytokine responses. Nature 362:245-248.

Krishna, R.G. and F. Wold. 1993. Post-translational modification of proteins. Adv. Enzymol. Relat. Areas 67:265-298.

Krucken, J., M. Epe, W.P. Benten, N. Falkenroth and F. Wunderlich. 2005. Malaria-suppressible expression of the anti-apoptotic triple GTPase mGIMAP8. J. Cell. Biochem. 96:339-348.

Krucken, J., R.M. Schroetel, I.U. Muller, N. Saidani, P. Marinovski, W.P. Benten, O. Stamm and F. Wunderlich. 2004. Comparative analysis of the human gimap gene cluster encoding a novel GTPase family. Gene 341:291-304.

Kubota, K., K. Wakabayashi and T. Matsuoka. 2003. Proteome analysis of secreted proteins during osteoclast differentiation using two different methods: Two-dimensional electrophoresis and isotope-coded affinity tags analysis with two-dimensional chromatography. Proteomics 3:616-626.

Kuhn, R., K. Rajewsky and W. Muller. 1991. Generation and analysis of interleukin-4 deficient mice. Science 254:707-710.

Kuyama, H., M. Watanabe, C. Toda, E. Ando, K. Tanaka and O. Nishimura. 2003. An approach to quantitative proteome analysis by labeling tryptophan residues. Rapid Commun. Mass Spectrom. 17:1642-1650.

Lander, E.S., L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M.

Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrinos, M.J. Morgan, P. de Jong, J.J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y.J. Chen and International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Langen, H., M. Fountoulakis, S. Evers, B. Wipf and P. Berndt. 1998. 15N and 13C labeling of cells for identification and quantification of proteins on 2D gels. *In* 15N and 13C labeling of cells for identification and quantification of proteins on 2D gels. Siena, Italy. 1998.

Lee, G.R., S.T. Kim, C.G. Spilianakis, P.E. Fields and R.A. Flavell. 2006. T helper cell differentiation: Regulation by cis elements and epigenetics. Immunity 24:369-379.

Lee, Y.H., H. Han, S.B. Chang and S.W. Lee. 2004. Isotope-coded N-terminal sulfonation of peptides allows quantitative proteomic analysis with increased de novo peptide sequencing capability. Rapid Commun. Mass Spectrom. 18:3019-3027.

Leitner, A. and W. Lindner. 2006. Chemistry meets proteomics: The use of chemical tagging reactions for MS-based proteomics. Proteomics 6:5418-5434.

Leitner, A. and W. Lindner. 2004. Current chemical tagging strategies for proteome analysis by mass spectrometry. J. Chromatogr. B. Analyt Technol. Biomed. Life. Sci. 813:1-26.

Leonard, W.J. and J.J. O'Shea. 1998. Jaks and STATs: Biological implications. Annu. Rev. Immunol. 16:293-322.

Letzelter, F., Y. Wang and W. Sebald. 1998. The interleukin-4 site-2 epitope determining binding of the common receptor gamma chain. Eur. J. Biochem. 257:11-20.

Li, J., H. Steen and S.P. Gygi. 2003a. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: The yeast salinity stress response. Mol. Cell. Proteomics 2:1198-1204.

Li, J., H. Steen and S.P. Gygi. 2003b. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: The yeast salinity stress response. Mol. Cell. Proteomics 2:1198-1204.

Li, X.J., H. Zhang, J.A. Ranish and R. Aebersold. 2003c. Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. Anal. Chem. 75:6648-6657.

Link, A.J., J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik and J.R. Yates 3rd. 1999. Direct analysis of protein complexes using mass spectrometry. Nat. Biotechnol. 17:676-682.

Lisacek, F., S. Cohen-Boulakia and R.D. Appel. 2006. Proteome informatics II: Bioinformatics for comparative proteomics. Proteomics.

Liska, A.J. and A. Shevchenko. 2003. Expanding the organismal scope of proteomics: Cross-species protein identification by mass spectrometry and its implications. Proteomics 3:19-28.

Liu, F.T., R.J. Patterson and J.L. Wang. 2002. Intracellular functions of galectins. Biochim. Biophys. Acta 1572:263-273.

Lopez, M.F., K. Berggren, E. Chernokalskaya, A. Lazarev, M. Robinson and W.F. Patton. 2000a. A comparison of silver stain and SYPRO ruby protein gel stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. Electrophoresis 21:3673-3683.

Lopez, M.F., B.S. Kristal, E. Chernokalskaya, A. Lazarev, A.I. Shestopalov, A. Bogdanova and M. Robinson. 2000b. High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. Electrophoresis 21:3427-3440.

Loyet, K.M., W. Ouyang, D.L. Eaton and J.T. Stults. 2005. Proteomic profiling of surface proteins on Th1 and Th2 cells. J. Proteome Res. 4:400-409.

Lu, B., P. Zagouras, J.E. Fischer, J. Lu, B. Li and R.A. Flavell. 2004a. Kinetic analysis of genomewide gene expression reveals molecule circuitries that control T cell activation and Th1/2 differentiation. Proc. Natl. Acad. Sci. U. S. A. 101:3023-3028.

Lu, Y., P. Bottari, F. Turecek, R. Aebersold and M.H. Gelb. 2004b. Absolute quantification of specific proteins in complex mixtures using visible isotope-coded affinity tags. Anal. Chem. 76:4104-4111.

Lund, R., T. Aittokallio, O. Nevalainen and R. Lahesmaa. 2003. Identification of novel genes regulated by IL-12, IL-4, or TGF-beta during the early polarization of CD4+ lymphocytes. J. Immunol. 171:5328-5336.

Lund, R., H. Ahlfors, E. Kainonen, A.M. Lahesmaa, C. Dixon and R. Lahesmaa. 2005. Identification of genes involved in the initiation of human Th1 or Th2 cell commitment. Eur. J. Immunol. 35:3307-3319.

Lund, R.J., Z. Chen, J. Scheinin and R. Lahesmaa. 2004. Early target genes of IL-12 and STAT4 signaling in th cells. J. Immunol. 172:6775-6782.

Lund, R.J., M. Loytomaki, T. Naumanen, C. Dixon, Z. Chen, H. Ahlfors, S. Tuomela, J. Tahvanainen, J. Scheinin, T. Henttinen, O. Rasool and R. Lahesmaa. 2007. Genome-wide identification of novel genes involved in early Th1 and Th2 cell differentiation. J. Immunol. 178:3648-3660.

Macatonia, S.E., N.A. Hosken, M. Litton, P. Vieira, C.S. Hsieh, J.A. Culpepper, M. Wysocka, G. Trinchieri, K.M. Murphy and A. O'Garra. 1995. Dendritic cells produce IL-12 and direct the development of Th1 cells from naive CD4+ T cells. J. Immunol. 154:5071-5079.

MacCoss, M.J., C.C. Wu, H. Liu, R. Sadygov and J.R. Yates 3rd. 2003. A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. Anal. Chem. 75:6912-6921.

MacMurray, A.J., D.H. Moralejo, A.E. Kwitek, E.A. Rutledge, B. Van Yserloo, P. Gohlke, S.J. Speros, B. Snyder, J. Schaefer, S. Bieg, J. Jiang, R.A. Ettinger, J. Fuller, T.L. Daniels, A. Pettersson, K. Orlebeke, B. Birren, H.J. Jacob, E.S. Lander and A. Lernmark. 2002. Lymphopenia in the BB rat model of type 1 diabetes is due to a mutation in a novel immune-associated nucleotide (ian)-related gene. Genome Res. 12:1029-1039.

Magram, J., S.E. Connaughton, R.R. Warrier, D.M. Carvajal, C.Y. Wu, J. Ferrante, C. Stewart, U. Sarmiento, D.A. Faherty, M.K. Gately. 1996. IL-12-deficient mice are defective in IFN gamma production and type 1 cytokine responses. Immunity 4:471-481.

Manetti, R., P. Parronchi, M.G. Giudizi, M.P. Piccinni, E. Maggi, G. Trinchieri and S. Romagnani. 1993. Natural killer cell stimulatory factor (interleukin 12 [IL-12]) induces T helper type 1 (Th1)-specific immune responses and inhibits the development of IL-4-producing th cells. J. Exp. Med. 177:1199-1204.

Mann, M. and M. Wilm. 1994. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal. Chem. 66:4390-4399.

Mann, M., R.C. Hendrickson and A. Pandey. 2001. Analysis of proteins and proteomes by mass spectrometry. Annu. Rev. Biochem. 70:437-473.

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

Marelli, M., J.J. Smith, S. Jung, E. Yi, A.I. Nesvizhskii, R.H. Christmas, R.A. Saleem, Y.Y. Tam, A. Fagarasanu, D.R. Goodlett, R. Aebersold, R.A. Rachubinski and J.D. Aitchison. 2004. Quantitative mass spectrometry reveals a role for the GTPase Rho1p in actin organization on the peroxisome membrane. J. Cell Biol. 167:1099-1112.

Marley, K., D.T. Mooney, G. Clark-Scannell, T.T. Tong, J. Watson, T.M. Hagen, J.F. Stevens and C.S. Maier. 2005. Mass tagging approach for mitochondrial thiol proteins. J. Proteome Res. 4:1403-1412.

Martin, D.B., D.R. Gifford, M.E. Wright, A. Keller, E. Yi, D.R. Goodlett, R. Aebersold and P.S. Nelson. 2004. Quantitative proteomic analysis of proteins released by neoplastic prostate epithelium. Cancer Res. 64:347-355.

Marzolf, B., E.W. Deutsch, P. Moss, D. Campbell, M.H. Johnson and T. Galitski. 2006. SBEAMS-microarray: Database software supporting genomic expression analyses for systems biology. BMC Bioinformatics 7:286.

Mason, D.E. and D.C. Liebler. 2003. Quantitative analysis of modified proteins by LC-MS/MS of peptides labeled with phenyl isocyanate. J. Proteome Res. 2:265-272.

Meehan, K.L. and M.D. Sadar. 2004. Quantitative profiling of LNCaP prostate cancer cells using isotope-coded affinity tags and mass spectrometry. Proteomics 4:1116-1134.

Merberg, D.M., S.F. Wolf and S.C. Clark. 1992. Sequence similarity between NKSF and the IL-6/G-CSF family. Immunol. Today 13:77-78.

Miller, I., J. Crawford and E. Gianazza. 2006. Protein stains for proteomic applications: Which, when, why? Proteomics 6:5385-5408.

Miyagi, M. and K.C. Rao. 2007. Proteolytic 18O-labeling strategies for quantitative proteomics. Mass Spectrom. Rev. 26:121-136.

Molloy, M.P. 2000. Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. Anal. Biochem. 280:1-10.

Mosmann, T.R. and R.L. Coffman. 1989a. TH1 and TH2 cells: Different patterns of lymphokine secretion lead to different functional properties. Annu. Rev. Immunol. 7:145-173.

Mosmann, T.R. and R.L. Coffman. 1989b. Heterogeneity of cytokine secretion patterns and functions of helper T cells. Adv. Immunol. 46:111-147.

Mosmann, T.R., H. Cherwinski, M.W. Bond, M.A. Giedlin and R.L. Coffman. 1986. Two types of murine helper T cell clone. I. definition according to profiles of lymphokine activities and secreted proteins. J. Immunol. 136:2348-2357.

Moulder, R., J.J. Filen, J. Salmi, M. Katajamaa, O.S. Nevalainen, M. Oresic, T. Aittokallio, R. Lahesmaa and T.A. Nyman. 2005. A comparative evaluation of software for the analysis of liquid chromatography-tandem mass spectrometry data from isotope coded affinity tag experiments. Proteomics 5:2748-2760.

Mowen, K.A. and L.H. Glimcher. 2004. Signaling pathways in Th2 development. Immunol. Rev. 202:203-222.

Munchbach, M., M. Quadroni, G. Miotto and P. James. 2000. Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. Anal. Chem. 72:4047-4057.

Nam, H.W., R. Simpson and Y.S. Kim. 2005. N-terminal isotope tagging with propionic anhydride: Proteomic analysis of myogenic differentiation of C2C12 cells. J. Chromatogr. B. Analyt Technol. Biomed. Life. Sci. 826:91-107.

Nelms, K., A.D. Keegan, J. Zamorano, J.J. Ryan and W.E. Paul. 1999. The IL-4 receptor: Signaling mechanisms and biologic functions. Annu. Rev. Immunol. 17:701-738.

Nesvizhskii, A.I. and R. Aebersold. 2004. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. Drug Discov. Today 9:173-181.

Nesvizhskii, A.I., A. Keller, E. Kolker and R. Aebersold. 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75:4646-4658.

Neuhoff, V., R. Stamm and H. Eibl. 1985. Clear background and highly sensitive protein staining with coomassie blue dyes in polyacrylamide gels: A systematic analysis. Electrophoresis 6:427-448.

Nitta, T., M. Nasreen, T. Seike, A. Goji, I. Ohigashi, T. Miyazaki, T. Ohta, M. Kanno and Y. Takahama. 2006. IAN family critically regulates survival and development of T lymphocytes. PLoS Biol. 4:e103.

Niwayama, S., S. Kurono and H. Matsumoto. 2003. Synthesis of 13C-labeled iodoacetanilide and application to quantitative peptide analysis by isotope differential mass spectrometry. Bioorg. Med. Chem. Lett. 13:2913-2916.

Niwayama, S., S. Kurono and H. Matsumoto. 2001. Synthesis of d-labeled N-alkylmaleimides and application to quantitative peptide analysis by isotope differential mass spectrometry. Bioorg. Med. Chem. Lett. 11:2257-2261.

Noben-Trauth, N., L.D. Shultz, F. Brombacher, J.F. Urban Jr, H. Gu and W.E. Paul. 1997. An interleukin 4 (IL-4)-independent pathway for CD4+ T cell IL-4 production is revealed in IL-4 receptor-deficient mice. Proc. Natl. Acad. Sci. U. S. A. 94:10838-10843.

Noga, M.J., J.J. Lewandowski, P. Suder and J. Silberring. 2005. An enhanced method for peptides sequencing by N-terminal derivatization and MS. Proteomics 5:4367-4375.

Nyman, T.A., S. Matikainen, T. Sareneva, I. Julkunen and N. Kalkkinen. 2000. Proteome analysis reveals ubiquitin-conjugating enzymes to be a new family of interferon-alpha-regulated genes. Eur. J. Biochem. 267:4011-4019.

Oda, Y., K. Huang, F.R. Cross, D. Cowburn and B.T. Chait. 1999. Accurate quantitation of protein expression and site-specific phosphorylation. Proc. Natl. Acad. Sci. U. S. A. 96:6591-6596.

Oda, Y., T. Owa, T. Sato, B. Boucher, S. Daniels, H. Yamanaka, Y. Shinohara, A. Yokoi, J. Kuromitsu and T. Nagasu. 2003. Quantitative chemical proteomics for identifying candidate drug targets. Anal. Chem. 75:2159-2165.

O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250:4007-4021.

Offner, H., B. Celnik, T.S. Bringman, D. Casentini-Borocz, G.E. Nedwin and A.A. Vandenbark. 1990. Recombinant human beta-galactoside binding lectin suppresses clinical and histological signs of experimental autoimmune encephalomyelitis. J. Neuroimmunol. 28:177-184.

Ohara, J. and W.E. Paul. 1987. Receptors for B-cell stimulatory factor-1 expressed on cells of haematopoietic lineage. Nature 325:537-540.

Olsen, J.V., S.E. Ong and M. Mann. 2004a. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol. Cell. Proteomics 3:608-614.

Olsen, J.V., J.R. Andersen, P.A. Nielsen, M.L. Nielsen, D. Figeys, M. Mann and J.R. Wisniewski. 2004b. HysTag--a novel proteomic quantification tool applied to differential display analysis of membrane proteins from distinct areas of mouse brain. Mol. Cell. Proteomics 3:82-92.

Omenn, G.S., D.J. States, M. Adamski, T.W. Blackwell, R. Menon, H. Hermjakob, R. Apweiler, B.B. Haab, R.J. Simpson, J.S. Eddes, E.A. Kapp, R.L. Moritz, D.W. Chan, A.J. Rai, A. Admon, R. Aebersold, J. Eng, W.S. Hancock, S.A. Hefta, H. Meyer, Y.K. Paik, J.S. Yoo, P. Ping, J. Pounds, J. Adkins, X. Qian, R. Wang, V. Wasinger, C.Y. Wu, X. Zhao, R. Zeng, A. Archakov, A. Tsugita, I. Beer, A. Pandey, M. Pisano, P. Andrews, H. Tammen, D.W. Speicher and S.M. Hanash. 2005.

Overview of the HUPO plasma proteome project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. Proteomics 5:3226-3245.

Ong, S.E. and M. Mann. 2005. Mass spectrometry-based proteomics turns quantitative. Nat. Chem. Biol. 1:252-262.

Ong, S.E., L.J. Foster and M. Mann. 2003. Mass spectrometric-based approaches in quantitative proteomics. Methods 29:124-130.

Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey and M. Mann. 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.

Opiteck, G.J., K.C. Lewis, J.W. Jorgenson and R.J. Anderegg. 1997. Comprehensive on-line LC/LC/MS of proteins. Anal. Chem. 69:1518-1524.

Pace, K.E., H.P. Hahn, M. Pang, J.T. Nguyen and L.G. Baum. 2000. CD7 delivers a pro-apoptotic signal during galectin-1-induced T cell death. J. Immunol. 165:2331-2334.

Palagi, P.M., P. Hernandez, D. Walther and R.D. Appel. 2006. Proteome informatics I: Bioinformatics tools for processing experimental data. Proteomics.

Pandarpurkar, M., L. Wilson-Fritch, S. Corvera, H. Markholst, L. Hornum, D.L. Greiner, J.P. Mordes, A.A. Rossini and R. Bortell. 2003. Ian4 is required for mitochondrial integrity and T cell survival. Proc. Natl. Acad. Sci. U. S. A. 100:10382-10387.

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

Pappin, D.J., P. Hojrup and A.J. Bleasby. 1993. Rapid identification of proteins by peptide-mass fingerprinting. Curr. Biol. 3:327-332.

Pasquali, C., I. Fialka and L.A. Huber. 1999. Subcellular fractionation, electromigration analysis and mapping of organelles. J. Chromatogr. B Biomed. Sci. Appl. 722:89-102.

Pasquarello, C., J.C. Sanchez, D.F. Hochstrasser and G.L. Corthals. 2004. N-t-butyliodoacetamide and iodoacetanilide: Two new cysteine alkylating reagents for relative quantitation of proteins. Rapid Commun. Mass Spectrom. 18:117-127.

Patterson, S.D. and R.H. Aebersold. 2003. Proteomics: The first decade and beyond. Nat. Genet. 33 Suppl:311-323.

Patton, W.F. 2002. Detection technologies in proteome analysis. J. Chromatogr. B. Analyt Technol. Biomed. Life. Sci. 771:3-31.

Patton, W.F. and J.M. Beechem. 2002. Rainbow's end: The quest for multiplexed fluorescence quantitative analysis in proteomics. Curr. Opin. Chem. Biol. 6:63-69.

Paul, W.E. and R.A. Seder. 1994. Lymphocyte responses and cytokines. Cell 76:241-251.

Peng, J. and S.P. Gygi. 2001. Proteomics: The move to mixtures. J. Mass Spectrom. 36:1083-1091.

Peng, J., J.E. Elias, C.C. Thoreen, L.J. Licklider and S.P. Gygi. 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. J. Proteome Res. 2:43-50.

Perillo, N.L., K.E. Pace, J.J. Seilhamer and L.G. Baum. 1995. Apoptosis of T cells mediated by galectin-1. Nature 378:736-739.

Perkins, D.N., D.J. Pappin, D.M. Creasy and J.S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567.

Perussia, B., S.H. Chan, A. D'Andrea, K. Tsuji, D. Santoli, M. Pospisil, D. Young, S.F. Wolf and G. Trinchieri. 1992. Natural killer (NK) cell stimulatory factor or IL-12 has differential effects on the proliferation of TCR-alpha beta+, TCR-gamma delta+ T lymphocytes, and NK cells. J. Immunol. 149:3495-3502.

Peters, E.C., D.M. Horn, D.C. Tully and A. Brock. 2001. A novel multifunctional labeling reagent for enhanced protein characterization with mass spectrometry. Rapid Commun. Mass Spectrom. 15:2387-2392.

Poirier, G.M., G. Anderson, A. Huvar, P.C. Wagaman, J. Shuttleworth, E. Jenkinson, M.R. Jackson, P.A. Peterson and M.G. Erlander. 1999. Immune-associated nucleotide-1 (IAN-1) is a thymic selection marker and defines a novel gene family conserved in plants. J. Immunol. 163:4960-4969.

Presky, D.H., H. Yang, L.J. Minetti, A.O. Chua, N. Nabavi, C.Y. Wu, M.K. Gately and U. Gubler. 1996. A functional interleukin 12 receptor complex is composed of two beta-type cytokine receptor subunits. Proc. Natl. Acad. Sci. U. S. A. 93:14002-14007.

Prouvost-Danon, A., A. Abadie, C. Sapin, H. Bazin and P. Druet. 1981. Induction of IgE synthesis and potentiation of anti-ovalbumin IgE antibody response by HgCl2 in the rat. J. Immunol. 126:699-792.

Qian, W.J., M.B. Goshe, D.G. Camp 2nd, L.R. Yu, K. Tang and R.D. Smith. 2003. Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures. Anal. Chem. 75:5441-5450.

Qiu, Y., E.A. Sousa, R.M. Hewick and J.H. Wang. 2002. Acid-labile isotope-coded extractants: A class of reagents for quantitative mass spectrometric analysis of complex protein mixtures. Anal. Chem. 74:4969-4979.

Rabilloud, T. 1990. Mechanisms of protein silver staining in polyacrylamide gels: A 10-year synthesis. Electrophoresis 11:785-794.

Rabinovich, G.A., N. Rubinstein and M.A. Toscano. 2002. Role of galectins in inflammatory and immunomodulatory processes. Biochim. Biophys. Acta 1572:274-284.

Rabinovich, G.A., C.R. Alonso, C.E. Sotomayor, S. Durand, J.L. Bocco and C.M. Riera. 2000. Molecular mechanisms implicated in galectin-1-induced apoptosis: Activation of the AP-1 transcription factor and downregulation of bcl-2. Cell Death Differ. 7:747-753.

Rabinovich, G.A., A. Ariel, R. Hershkoviz, J. Hirabayashi, K.I. Kasai and O. Lider. 1999a. Specific inhibition of T-cell adhesion to extracellular matrix and proinflammatory cytokine secretion by human recombinant galectin-1. Immunology 97:100-106.

Rabinovich, G.A., G. Daly, H. Dreja, H. Tailor, C.M. Riera, J. Hirabayashi and Y. Chernajovsky. 1999b. Recombinant galectin-1 and its genetic delivery suppress collagen-induced arthritis via T cell apoptosis. J. Exp. Med. 190:385-398.

Raman, B., A. Cheung and M.R. Marten. 2002. Quantitative comparison and evaluation of two commercially available, two-dimensional electrophoresis image analysis software packages, Z3 and melanie. Electrophoresis 23:2194-2202.

Ranish, J.A., E.C. Yi, D.M. Leslie, S.O. Purvine, D.R. Goodlett, J. Eng and R. Aebersold. 2003. The study of macromolecular complexes by quantitative proteomics. Nat. Genet. 33:349-355.

Rautajoki, K., T.A. Nyman and R. Lahesmaa. 2004. Proteome characterization of human T helper 1 and 2 cells. Proteomics 4:84-92.

Rautajoki, K.J., E.M. Marttila, T.A. Nyman and R. Lahesmaa. 2007. Interleukin-4 inhibits caspase-3 by regulating several proteins in the fas pathway during initial stages of human T helper 2 cell differentiation. Mol. Cell. Proteomics 6:238-251.

Reichel, M., B.H. Nelson, P.D. Greenberg and P.B. Rothman. 1997. The IL-4 receptor alpha-chain cytoplasmic domain is sufficient for activation of JAK-1 and STAT6 and the induction of IL-4-specific gene expression. J. Immunol. 158:5860-5867.

Reiner, S.L. 2007. Development in motion: Helper T cells at work. Cell 129:33-36.

Reinhardt, R.L., S.J. Kang, H.E. Liang and R.M. Locksley. 2006. T helper cell effector fates--who, how and where? Curr. Opin. Immunol. 18:271-277.

Rezaul, K., L. Wu, V. Mayya, S.I. Hwang and D. Han. 2005. A systematic characterization of mitochondrial proteome from human T leukemia cells. Mol. Cell. Proteomics 4:169-181.

Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann and B. Seraphin. 1999. A generic protein purification method for protein complex characterization and proteome exploration. Nat. Biotechnol. 17:1030-1032.

Riggs, L., E.H. Seeley and F.E. Regnier. 2005. Quantification of phosphoproteins with global internal standard technology. J. Chromatogr. B. Analyt Technol. Biomed. Life. Sci. 817:89-96.

Righetti, P.G. 2006. Real and imaginary artefacts in proteome analysis via two-dimensional maps. J. Chromatogr. B. Analyt Technol. Biomed. Life. Sci. 841:14-22.

Righetti, P.G., A. Castagna, P. Antonioli and E. Boschetti. 2005. Prefractionation techniques in proteome analysis: The mining tools of the third millennium. Electrophoresis 26:297-319.

Righetti, P.G., A. Castagna, F. Antonucci, C. Piubelli, D. Cecconi, N. Campostrini, G. Zanusso and S. Monaco. 2003. The proteome: Anno domini 2002. Clin. Chem. Lab. Med. 41:425-438.

Risso, A., G. Tell, C. Vascotto, A. Costessi, S. Arena, A. Scaloni and M.E. Cosulich. 2005. Activation of human T lymphocytes under conditions similar to those that occur during exposure to microgravity: A proteomics study. Proteomics 5:1827-1837.

Robertson, M.J., R.J. Soiffer, S.F. Wolf, T.J. Manley, C. Donahue, D. Young, S.H. Herrmann and J. Ritz. 1992. Response of human natural killer (NK) cells to NK cell stimulatory factor (NKSF): Cytolytic activity and proliferation of NK cells are differentially regulated by NKSF. J. Exp. Med. 175:779-788.

Rogge, L., L. Barberis-Maino, M. Biffi, N. Passini, D.H. Presky, U. Gubler and F. Sinigaglia. 1997. Selective expression of an interleukin-12 receptor component by human T helper 1 cells. J. Exp. Med. 185:825-831.

Rogge, L., E. Bianchi, M. Biffi, E. Bono, S.Y. Chang, H. Alexander, C. Santini, G. Ferrari, L. Sinigaglia, M. Seiler, M. Neeb, J. Mous, F. Sinigaglia and U. Certa. 2000. Transcript imaging of the development of human T helper cells using oligonucleotide arrays. Nat. Genet. 25:96-101.

Romagnani, S. 2004. Immunologic influences on allergy and the Th1/Th2 balance. J. Allergy Clin. Immunol. 113:395-400.

Romagnani, S. 1996. Th1 and Th2 in human diseases. Clin. Immunol. Immunopathol. 80:225-235.

Romagnani, S. 1994. Lymphokine production by human T cells in disease states. Annu. Rev. Immunol. 12:227-257.

Rosengren, A.T., T.A. Nyman and R. Lahesmaa. 2005a. Proteome profiling of interleukin-12 treated human T helper cells. Proteomics 5:3137-3141.

Rosengren, A.T., T.A. Nyman, S. Syyrakki, S. Matikainen and R. Lahesmaa. 2005b. Proteomic and transcriptomic characterization of interferon-alpha-induced human primary T helper cells. Proteomics 5:371-379.

Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson and D.J. Pappin. 2004. Multiplexed protein quantitation in saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol. Cell. Proteomics 3:1154-1169.

Russell, S.M., J.A. Johnston, M. Noguchi, M. Kawamura, C.M. Bacon, M. Friedmann, M. Berg, D.W. McVicar, B.A. Witthuhn and O. Silvennoinen. 1994. Interaction of IL-2R beta and gamma c chains with Jak1 and Jak3: Implications for XSCID and XCID. Science 266:1042-1045.

Russell, S.M., A.D. Keegan, N. Harada, Y. Nakamura, M. Noguchi, P. Leland, M.C. Friedmann, A. Miyajima, R.K. Puri and W.E. Paul. 1993. Interleukin-2 receptor gamma chain: A functional component of the interleukin-4 receptor. Science 262:1880-1883.

Salmi, J., R. Moulder, J.-J. Filén, O.S. Nevalainen, T.A. Nyman, R. Lahesmaa, and T. Aittokallio (2006) Quality classification of tandem mass spectrometry data. Bioinformatics, 22, 400-406

Samyn, B., G. Debyser, K. Sergeant, B. Devreese and J. Van Beeumen. 2004. A case study of de novo sequence analysis of N-sulfonated peptides by MALDI TOF/TOF mass spectrometry. J. Am. Soc. Mass Spectrom. 15:1838-1852.

Santucci, L., S. Fiorucci, F. Cammilleri, G. Servillo, B. Federici and A. Morelli. 2000. Galectin-1 exerts immunomodulatory and protective effects on concanavalin A-induced hepatitis in mice. Hepatology 31:399-406.

Santucci, L., S. Fiorucci, N. Rubinstein, A. Mencarelli, B. Palazzetti, B. Federici, G.A. Rabinovich and A. Morelli. 2003. Galectin-1 suppresses experimental colitis in mice. Gastroenterology 124:1381-1394.

Scheele, G.A. 1975. Two-dimensional gel analysis of soluble proteins. charaterization of guinea pig exocrine pancreatic proteins. J. Biol. Chem. 250:5375-5385.

Scheler, C., S. Lamer, Z. Pan, X.P. Li, J. Salnikow and P. Jungblut. 1998. Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS). Electrophoresis 19:918-927.

Scherl, A., Y. Coute, C. Deon, A. Calle, K. Kindbeiter, J.C. Sanchez, A. Greco, D. Hochstrasser and J.J. Diaz. 2002. Functional proteomic analysis of human nucleolus. Mol. Biol. Cell 13:4100-4109.

Schmidt, A., J. Kellermann and F. Lottspeich. 2005. A novel strategy for quantitative proteomics using isotope-coded protein labels. Proteomics 5:4-15.

Schmidt, F., B. Dahlmann, K. Janek, A. Kloss, M. Wacker, R. Ackermann, B. Thiede and P.R. Jungblut. 2006. Comprehensive quantitative proteome analysis of 20S proteasome subtypes from rat liver by isotope coded affinity tag and 2-D gel-based approaches. Proteomics 6:4622-4632.

Schmidt, F., S. Donahoe, K. Hagens, J. Mattow, U.E. Schaible, S.H. Kaufmann, R. Aebersold and P.R. Jungblut. 2004. Complementary analysis of the mycobacterium tuberculosis proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. Mol. Cell. Proteomics 3:24-42.

Schnell, S., C. Demolliere, P. van den Berk and H. Jacobs. 2006. Gimap4 accelerates T-cell death. Blood 108:591-599.

Sebastiano, R., A. Citterio, M. Lapadula and P.G. Righetti. 2003. A new deuterated alkylating agent for quantitative proteomics. Rapid Commun. Mass Spectrom. 17:2380-2386.

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.

Seder, R.A. and W.E. Paul. 1994. Acquisition of lymphokine-producing phenotype by CD4+ T cells. Annu. Rev. Immunol. 12:635-673.

Seder, R.A., W.E. Paul, M.M. Davis and Fazekas de St Groth,B. 1992. The presence of interleukin 4 during in vitro priming determines the lymphokine-producing potential of CD4+ T cells from T cell receptor transgenic mice. J. Exp. Med. 176:1091-1098.

Shen, M., L. Guo, A. Wallace, J. Fitzner, J. Eisenman, E. Jacobson and R.S. Johnson. 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.

Shi, Y., R. Xiang, J.K. Crawford, C.M. Colangelo, C. Horvath and J.A. Wilkins. 2004. A simple solid phase mass tagging approach for quantitative proteomics. J. Proteome Res. 3:104-111.

Shiio, Y., S. Donohoe, E.C. Yi, D.R. Goodlett, R. Aebersold and R.N. Eisenman. 2002. Quantitative proteomic analysis of myc oncoprotein function. EMBO J. 21:5088-5096.

Shimoda, K., J. van Deursen, M.Y. Sangster, S.R. Sarawar, R.T. Carson, R.A. Tripp, C. Chu, F.W. Quelle, T. Nosaka, D.A. Vignali, P.C. Doherty, G. Grosveld, W.E. Paul and J.N. Ihle. 1996. Lack of IL-4-induced Th2 response and IgE class switching in mice with disrupted Stat6 gene. Nature 380:630-633.

Sieburth, D., E.W. Jabs, J.A. Warrington, X. Li, J. Lasota, S. LaForgia, K. Kelleher, K. Huebner, J.J. Wasmuth and S.F. Wolf. 1992. Assignment of genes encoding a unique cytokine (IL12) composed of two unrelated subunits to chromosomes 3 and 5. Genomics 14:59-62.

Sinha, P., J. Poland, M. Schnolzer and T. Rabilloud. 2001. A new silver staining apparatus and procedure for matrix-assisted laser desorption/ionization-time of flight analysis of proteins after two-dimensional electrophoresis. Proteomics 1:835-840.

Smolka, M.B., C.P. Albuquerque, S.H. Chen, K.H. Schmidt, X.X. Wei, R.D. Kolodner and H. Zhou. 2005. Dynamic changes in protein-protein interaction and protein phosphorylation probed with amine-reactive isotope tag. Mol. Cell. Proteomics 4:1358-1369.

Stamm, O., J. Krucken, H.P. Schmitt-Wrede, W.P. Benten and F. Wunderlich. 2002. Human ortholog to mouse gene imap38 encoding an ER-localizable G-protein belongs to a gene family clustered on chromosome 7q32-36. Gene 282:159-167.

Standing, K.G. 2003. Peptide and protein de novo sequencing by mass spectrometry. Curr. Opin. Struct. Biol. 13:595-601.

Steel, L.F., M.G. Trotter, P.B. Nakajima, T.S. Mattu, G. Gonye and T. Block. 2003. Efficient and specific removal of albumin from human serum samples. Mol. Cell. Proteomics 2:262-270.

Stentz, F.B. and A.E. Kitabchi. 2004. Transcriptome and proteome expression in activated human CD4 and CD8 T-lymphocytes. Biochem. Biophys. Res. Commun. 324:692-696.

Stern, A.S., F.J. Podlaski, J.D. Hulmes, Y.C. Pan, P.M. Quinn, A.G. Wolitzky, P.C. Familletti, D.L. Stremlo, T. Truitt and R. Chizzonite. 1990. Purification to homogeneity and partial characterization of cytotoxic lymphocyte maturation factor from human B-lymphoblastoid cells. Proc. Natl. Acad. Sci. U. S. A. 87:6808-6812.

Stillwell, R. and B.E. Bierer. 2001. T cell signal transduction and the role of CD7 in costimulation. Immunol. Res. 24:31-52.

Stockwin, L.H., J. Blonder, M.A. Bumke, D.A. Lucas, K.C. Chan, T.P. Conrads, H.J. Issaq, T.D. Veenstra, D.L. Newton and S.M. Rybak. 2006. Proteomic analysis of plasma membrane from hypoxia-adapted malignant melanoma. J. Proteome Res. 5:2996-3007.

Stults, J.T. and D. Arnott. 2005. Proteomics. Methods Enzymol. 402:245-289.

Swain, S.L., A.D. Weinberg, M. English and G. Huston. 1990. IL-4 directs the development of Th2-like helper effectors. J. Immunol. 145:3796-3806.

Switzer, R.C.,3rd, C.R. Merril and S. Shifrin. 1979. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. Anal. Biochem. 98:231-237.

Syka, J.E., J.A. Marto, D.L. Bai, S. Horning, M.W. Senko, J.C. Schwartz, B. Ueberheide, B. Garcia, S. Busby, T. Muratore, J. Shabanowitz and D.F. Hunt. 2004. Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. J. Proteome Res. 3:621-626.

Szabo, S.J., A.S. Dighe, U. Gubler and K.M. Murphy. 1997. Regulation of the interleukin (IL)-12R beta 2 subunit expression in developing T helper 1 (Th1) and Th2 cells. J. Exp. Med. 185:817-824.

Tabb, D.L., A. Saraf and J.R. Yates 3rd. 2003. GutenTag: High-throughput sequence tagging via an empirically derived fragmentation model. Anal. Chem. 75:6415-6421.

Tabb, D.L., W.H. McDonald and J.R. Yates 3rd. 2002. DTASelect and contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. 1:21-26.

Tabb, D.L., M.J. MacCoss, C.C. Wu, S.D. Anderson and J.R. Yates 3rd. 2003. Similarity among tandem mass spectra from proteomic experiments: Detection, significance, and utility. Anal. Chem. 75:2470-2477.

Tahvanainen, J., M. Pykalainen, T. Kallonen, H. Lahteenmaki, O. Rasool and R. Lahesmaa. 2006. Enrichment of nucleofected primary human CD4+ T cells: A novel and efficient method for studying gene function and role in human primary T helper cell differentiation. J. Immunol. Methods 310:30-39.

Takeda, K., T. Tanaka, W. Shi, M. Matsumoto, M. Minami, S. Kashiwamura, K. Nakanishi, N. Yoshida, T. Kishimoto and S. Akira. 1996. Essential role of Stat6 in IL-4 signalling. Nature 380:627-630.

Tam, E.M., C.J. Morrison, Y.I. Wu, M.S. Stack and C.M. Overall. 2004. Membrane protease proteomics: Isotope-coded affinity tag MS identification of undescribed MT1-matrix metalloproteinase substrates. Proc. Natl. Acad. Sci. U. S. A. 101:6917-6922.

Tanaka, K., H. Waki, Y. Ido, S. Akita, Y. Yoshida and T. Yoshida. 1988. Protein and polymer analyses up to mlz 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 2:151-153.

Taylor, S.W., E. Fahy, B. Zhang, G.M. Glenn, D.E. Warnock, S. Wiley, A.N. Murphy, S.P. Gaucher, R.A. Capaldi, B.W. Gibson and S.S. Ghosh. 2003. Characterization of the human heart mitochondrial proteome. Nat. Biotechnol. 21:281-286.

Terpe, K. 2003. Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. 60:523-533.

Thadikkaran, L., M.A. Siegenthaler, D. Crettaz, P.A. Queloz, P. Schneider and J.D. Tissot. 2005. Recent advances in blood-related proteomics. Proteomics 5:3019-3034.

Thiede, B., A. Kretschmer and T. Rudel. 2006. Quantitative proteome analysis of CD95 (Fas/Apo-1)-induced apoptosis by stable isotope labeling with amino acids in cell culture, 2-DE and MALDI-MS. Proteomics 6:614-622.

Thiede, B., F. Siejak, C. Dimmler and T. Rudel. 2002. Prediction of translocation and cleavage of heterogeneous ribonuclear proteins and rho guanine nucleotide dissociation inhibitor 2 during apoptosis by subcellular proteome analysis. Proteomics 2:996-1006.

Thiede, B., C. Dimmler, F. Siejak and T. Rudel. 2001. Predominant identification of RNA-binding proteins in fas-induced apoptosis by proteome analysis. J. Biol. Chem. 276:26044-26050.

Thompson, A., J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A.K. Mohammed and C. Hamon. 2003. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. 75:1895-1904.

Tian, Q., S.B. Stepaniants, M. Mao, L. Weng, M.C. Feetham, M.J. Doyle, E.C. Yi, H. Dai, V. Thorsson, J. Eng, D. Goodlett, J.P. Berger, B. Gunter, P.S. Linseley, R.B. Stoughton, R. Aebersold, S.J. Collins, W.A. Hanlon and L.E. Hood. 2004. Integrated genomic and proteomic analyses of gene expression in mammalian cells. Mol. Cell. Proteomics 3:960-969.

Tonge, R., J. Shaw, B. Middleton, R. Rowlinson, S. Rayner, J. Young, F. Pognan, E. Hawkins, I. Currie and M. Davison. 2001. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 1:377-396.

Trinchieri, G. 2003. Interleukin-12 and the regulation of innate resistance and adaptive immunity. Nat. Rev. Immunol. 3:133-146.

Tuloup, M., C. Hernandez, I. Coro, C. Hoogland, P.-. Binz and R.D. Appel. 2003. Aldente and BioGraph: An improved peptide mass fingerprinting protein identification environment. p. 174-176. *In* Aldente and BioGraph: An improved peptide mass fingerprinting protein identification environment. Swiss proteomics society 2003 congress: Understanding biological systems through proteomics, Basel, Switzerland. 2-4 Dec. 2003 2003. Ed. FontisMedia, .

Unlu, M., M.E. Morgan and J.S. Minden. 1997. Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. Electrophoresis 18:2071-2077.

Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M.L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y.H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N.N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J.F. Abril, R. Guigo, M.J. Campbell, K.V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y.H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh and X. Zhu. 2001. The sequence of the human genome. Science 291:1304-1351.

Vespa, G.N., L.A. Lewis, K.R. Kozak, M. Moran, J.T. Nguyen, L.G. Baum and M.C. Miceli. 1999. Galectin-1 specifically modulates TCR signals to enhance TCR apoptosis but inhibit IL-2 production and proliferation. J. Immunol. 162:799-806.

Vitetta, E.S., J. Ohara, C.D. Myers, J.E. Layton, P.H. Krammer and W.E. Paul. 1985. Serological, biochemical, and functional identity of B cell-stimulatory factor 1 and B cell differentiation factor for IgG1. J. Exp. Med. 162:1726-1731.

Von Haller, P.D., E. Yi, S. Donohoe, K. Vaughn, A. Keller, A.I. Nesvizhskii, J. Eng, X.J. Li, D.R. Goodlett, R. Aebersold and J.D. Watts. 2003. The application of new software tools to quantitative protein profiling via isotope-coded affinity tag (ICAT) and tandem mass spectrometry: I. statistically annotated datasets for peptide sequences and proteins identified via the application of ICAT and tandem mass spectrometry to proteins copurifying with T cell lipid rafts. Mol. Cell. Proteomics 2:426-427.

Wang, S. and F.E. Regnier. 2001. Proteomics based on selecting and quantifying cysteine containing peptides by covalent chromatography. J. Chromatogr. A 924:345-357.

Washburn, M.P., D. Wolters and J.R. Yates 3rd. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. 19:242-247.

Washburn, M.P., A. Koller, G. Oshiro, R.R. Ulaszek, D. Plouffe, C. Deciu, E. Winzeler and J.R. Yates 3rd. 2003. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. 100:3107-3112.

Wells, L., K. Vosseller, R.N. Cole, J.M. Cronshaw, M.J. Matunis and G.W. Hart. 2002. Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. Mol. Cell. Proteomics 1:791-804.

Westbrook, J.A., J.X. Yan, R. Wait and M.J. Dunn. 2001. A combined radiolabelling and silver staining technique for improved visualisation, localisation, and identification of proteins separated by two-dimensional gel electrophoresis. Proteomics 1:370-376.

Whetstone, P.A., N.G. Butlin, T.M. Corneillie and C.F. Meares. 2004. Element-coded affinity tags for peptides and proteins. Bioconjug. Chem. 15:3-6.

Wilkins, M.R. and K.L. Williams. 1997. Cross-species protein identification using amino acid composition, peptide mass fingerprinting, isoelectric point and molecular mass: A theoretical evaluation. J. Theor. Biol. 186:7-15.

Wilkins, M.R., J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser and K.L. Williams. 1996. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. Biotechnol. Genet. Eng. Rev. 13:19-50.

Wilkins, M.R., R.D. Appel, J.E. Van Eyk, M.C. Chung, A. Gorg, M. Hecker, L.A. Huber, H. Langen, A.J. Link, Y.K. Paik, S.D. Patterson, S.R. Pennington, T. Rabilloud, R.J. Simpson, W. Weiss and M.J. Dunn. 2006. Guidelines for the next 10 years of proteomics. Proteomics 6:4-8.

Wilkinson, J.M. 1986. Fragmentation of polypeptides by enzymic methods. *In* A. Darbre (ed.) Practical protein chemistry: A handbook. John Wiley and Sons, New York.

Witten, I.H. and E. Frank. 2000. Data mining: Practical machine learning tools and techniques with java implementations. Morgan Kaufmann, San Francisco, USA.

Witthuhn, B.A., O. Silvennoinen, O. Miura, K.S. Lai, C. Cwik, E.T. Liu and J.N. Ihle. 1994. Involvement of the jak-3 janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells. Nature 370:153-157.

Wittmann-Liebold, B., H.R. Graack and T. Pohl. 2006. Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry. Proteomics 6:4688-4703.

Wu, C., X. Wang, M. Gadina, J. J. O'Shea, D. H. Presky, and J. Magram. 2000. IL-12 receptor beta 2 (IL-12R beta 2)-deficient mice are defective in IL-12-mediated signaling despite the presence of high affinity IL-12 binding sites. J. Immunol. 165:6221-6228.

Wu, C.C., M.J. MacCoss, K.E. Howell and J.R. Yates 3rd. 2003. A method for the comprehensive proteomic analysis of membrane proteins. Nat. Biotechnol. 21:532-538.

Wu, C.C., M.J. MacCoss, G. Mardones, C. Finnigan, S. Mogelsvang, J.R. Yates 3rd and K.E. Howell. 2004. Organellar proteomics reveals golgi arginine dimethylation. Mol. Biol. Cell 15:2907-2919.

Wysocki, V.H., K.A. Resing, Q. Zhang and G. Cheng. 2005. Mass spectrometry of peptides and proteins. Methods 35:211-222.

Yao, X., A. Freas, J. Ramirez, P.A. Demirev and C. Fenselau. 2001. Proteolytic 18O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. Anal. Chem. 73:2836-2842.

Yates, J.R.,3rd, A. Gilchrist, K.E. Howell and J.J. Bergeron. 2005. Proteomics of organelles and large cellular structures. Nat. Rev. Mol. Cell Biol. 6:702-714.

Yates, J.R.,3rd, S. Speicher, P.R. Griffin and T. Hunkapiller. 1993. Peptide mass maps: A highly informative approach to protein identification. Anal. Biochem. 214:397-408.

Yssel, H., J.E. De Vries, M. Koken, W. Van Blitterswijk and H. Spits. 1984. Serum-free medium for generation and propagation of functional human cytotoxic and helper T cell clones. J. Immunol. Methods 72:219-227.

Yu, Y., J. Cui, X. Wang, Y. Liu and P. Yang. 2004. Studies on peptide acetylation for stable-isotope labeling after 1-D PAGE separation in quantitative proteomics. Proteomics 4:3112-3120.

Zappacosta, F. and R.S. Annan. 2004. N-terminal isotope tagging strategy for quantitative proteomics: Results-driven analysis of protein abundance changes. Anal. Chem. 76:6618-6627.

Zenz, T., A. Roessner, A. Thomas, S. Frohling, H. Dohner, B. Calabretta and L. Daheron. 2004. hIan5: The human ortholog to the rat Ian4/Iddm1/lyp is a new member of the ian family that is overexpressed in B-cell lymphoid malignancies. Genes Immun. 5:109-116.

Zhang, N., R. Aebersold and B. Schwikowski. 2002a. ProbID: A probabilistic algorithm to identify peptides through sequence database searching using tandem mass spectral data. Proteomics 2:1406-1412.

Zhang, X., Q.K. Jin, S.A. Carr and R.S. Annan. 2002b. N-terminal peptide labeling strategy for incorporation of isotopic tags: A method for the determination of site-specific absolute phosphorylation stoichiometry. Rapid Commun. Mass Spectrom. 16:2325-2332.

Zhou, H., J.A. Ranish, J.D. Watts and R. Aebersold. 2002. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. Nat. Biotechnol. 20:512-515.

Zhou, Y., G. Gu, D.R. Goodlett, T. Zhang, C. Pan, T.J. Montine, K.S. Montine, R.H. Aebersold and J. Zhang. 2004. Analysis of alpha-synuclein-associated proteins by quantitative proteomics. J. Biol. Chem. 279:39155-39164.

Zhu, H., M. Bilgin and M. Snyder. 2003. Proteomics. Annu. Rev. Biochem. 72:783-812.

Zieske, L.R. 2006. A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J. Exp. Bot. 57:1501-1508.

Zou, J., D.H. Presky, C.Y. Wu and U. Gubler. 1997. Differential associations between the cytoplasmic regions of the interleukin-12 receptor subunits beta1 and beta2 and JAK kinases. J. Biol. Chem. 272:6073-6077.

Zuniga, E., G.A. Rabinovich, M.M. Iglesias and A. Gruppi. 2001. Regulated expression of galectin-1 during B-cell activation and implications for T-cell apoptosis. J. Leukoc. Biol. 70:73-79.

ORIGINAL PUBLICATIONS