From Department of Medicine Solna, Respiratory Medicine Unit
Karolinska Institutet, Stockholm, Sweden

Quantitative Proteomic Approaches for the Analysis of Human Lung Samples in Pulmonary Sarcoidosis

Ernesto Silva

Karolinska Institutet

Stockholm 2011

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet. Printed by Larserics Digital Print AB

© Ernesto Silva, 2011
ISBN 978-91-7457-296-4

Till min familj

ABSTRACT

This thesis is focused on the analysis of protein expression profiles from the airways of sarcoidosis patients using quantitative proteomics. Sarcoidosis is a multisystemic inflammatory disorder of unknown etiology characterized by the presence of noncaseating granulomas in the affected organs. Previous findings in proteomics have reported identification of several proteins that certainly reflect the ongoing inflammation response but that did not have any specificity for sarcoidosis disease. The overall aim of these studies was therefore to continue searching for new protein candidates that could help us to determine the possible mechanisms behind sarcoidosis.

We first investigated the total protein profile in the lung lumen from two granulomatous disorders, sarcoidosis (HLA-DRB1*15 positive), characterized to present an unresolved chronic inflammation disease and chronic beryllium disease (CBD) by collecting the bronchoalveolar soluble proteins and applying differential gel electrophoresis (DIGE) coupled to mass spectrometry (MS) *(paper I)*. This led to the identification of fourteen proteins with altered profiles, some of these related to inflammation (β2-microglobulin, annexin II, complement C3, apolipoprotein A1, IgG kappa chain and heat shock protein 70) and the oxidative response (peroxiredoxin 5, hemopexin, α1-antitrypsin and superoxide dismutase), hence reflecting the persistent inflammation state in those granulomatous diseases.

Gel-based 2DE techniques have been claimed to be unreliable in quantitative proteomics. The introduction of differential gel electrophoresis (DIGE) in 2DE approches has greatly improves gel-to-gel variation and the reproducibility; however, other sources of variance have been highlighted. In a 2D experimental-related study, we investigated the different sources of variance that were intrinsic to gel-based proteomics. We measured the technical variance related to background subtraction algorithms [4-8%] and the experimental variance related to the 1st and 2nd dimension of 2DE workflow (~30%). In addition, we reported the improvement of the 4th generation image software program SameSpots^{TM} in terms of reduced levels of variance introduced from background algorithms, higher levels of accurate spot-matching and most importantly an improved objectivity of the analysis *(paper II)*.

To further evaluate the protein changes in sarcoidosis we investigated the protein profiles from purified alveolar macrophages (AM). With the intention to improve both the protein resolution we applied two complementary proteomic approaches. First, all soluble AM proteins were resolved using the DIGE technique *(paper III)*. Second all membrane-associated proteins (MAP) were then identified and quantified using the shotgun proteomic, liquid chromatography couple to mass spectrometry LC-MS/MS approach *(paper IV)*. We found similar results from these parallel studies, including several pathways altered in sarcoidosis (*papers III & IV)*. In addition, by applying multivariate regression analysis we could also identify a robust model with a set of 13 proteins able to discriminate sarcoidosis from the healthy group *(paper IV)*.

Taken together, improvements in gel-based image software and clinical sample pre-treatment allow more accurate and quantitative analysis, revealing deeper insight into the proteome. The biological findings presented in this thesis give new perspectives in understanding AM and their role in sarcoidosis disease as well as the possibility to search for disease biomarkers.

LIST OF PUBLICATIONS

I. Silva E, Bourin S, Sabounchi-Schütt F, Laurin Y, Barker E, Newman L.S, Eriksson H, Eklund A, Grunewald J. **A quantitative proteomic analysis of soluble bronchoalveolar fluid proteins from patients with sarcoidosis and chronic beryllium disease.** Sarcoidosis Vasc Diffuse Lung Dis 2007 Mar 24 (1):24-32

II. Ernesto Silva, Martin O'Gorman, Susanne Becker, Gert Auer, Anders Eklund, Johan Grunewald, Åsa M. Wheelock. **In the eye of the beholder: does the master see the SameSpots as the novice?** J Proteome Res. 2010 Mar 5;9(3):1522-32

III. Ernesto Silva, Serhiy Souchelnytskyi, Kie Kasuga, Anders Eklund, Johan Grunewald, Åsa Wheelock **Proteomics approach on alveolar macrophages from sarcoidosis patients.** Manuscript

IV. Ernesto Silva, Hanna Eriksson, Rui Mamede Branca, Anders Eklund, Per-Johan Jakobsson, Johan Grunewald, Janne Lehtiö and Åsa Wheelock. **Proteomic analysis of membrane-associated proteins in alveolar macrophages from patients with pulmonary sarcoidosis**. Manuscript

CONTENTS

1 INTRODUCTION 1
1.1 The lungs 1
1.2 Innate Immunity 1
1.2.1 Fcγ receptors 1
1.2.2 Complement receptors 2
1.2.3 Toll-like receptors 2
1.2.4 Macrophages 2
1.2.5 Neutrophils 3
1.2.6 Dendritic cells 3
1.3 Adaptive immunity 3
1.3.1 MHC and antigen presentation 3
1.3.2 T cells 4
1.4 Sarcoidosis 5
1.4.1 Epidemiology 5
1.4.2 Etiology 5
1.4.3 Genetics 5
1.4.4 Pathology 6
1.4.5 Diagnosis 7
1.4.6 Löfgren's and non-Löfgren's sarcoidosis patients 7
1.5 Proteomics 8
1.5.1 Two-dimensional gel electrophoresis (2DE) 8
1.5.2 Mass Spectrometry 8
1.5.3 Tandem Mass spectrometry LC-MS/MS 9
1.5.4 Isotope Tags for relative and absolute quantification 10
1.5.5 Multivariate Analysis 10
2 CURRENT INVESTIGATION 11
2.1 Aims of the thesis 11
2.2 Comments on methodologies 12
2.2.1 Patient data and sampling collection 12
2.2.2 Bronchoscopy 12
2.2.3 BAL fluid concentration (paper I) 13
2.2.4 Alveolar macrophage (papers III & IV) 13
2.2.5 Microsomal preparation (papers II, III, IV) 13
2.2.6 Differential Gel Electrophoresis (papers I, II & III) 14
2.2.7 Mass Spectrometry (papers I and III) 15
2.2.8 Peptide labeling iTRAQ (paper IV) 15
2.2.9 LC-MS/MS analyses (paper IV) 15
2.3 Results and discussion 17
2.3.1 Paper I 17
2.3.2 Paper II 18
2.3.3 Paper III 19
2.3.4 Paper IV 22
2.4 General conclusions 26
3 ACKNOWLEDGEMENTS 28
4 REFERENCES 29

LIST OF ABBREVIATIONS

AM	Alveolar macrophages
ACTA1	Actin alpha A1
ACTB	Actin beta B1
ACTR3	Actin related protein-3 homolog
APC	Antigen presenting cells
AP	Adaptor-related protein
BALF	Bronchoalveolar lavage fluid
BHL	Bilateral hilar lymphadenopathy
CBD	Chronic beryllium disease
CID	Collision-induced dissociation energy
DIGE	Differential gel electrophoresis
DC	Dendritic cells
2DE	Two-dimensional gel electrophoresis
ERAD	ER associated protein degradation
EZR	Ezrin
ESI	Electronspray ionization
HMOX1	Heme oxigenase-1
HLA	Human leukocyte antigen
IEF	Isoelectric focusing
IL	Interleukin
iTRAQ	Isotope tags for relative and absolute quantification
LC-MS/MS	Liquid chromatography couple to tandem mass spectrometry
MALDI	Matrix-assisted laser desorption ionization
MAP	Membrane associated proteins
MHC	Major histocompatibility complex
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
mKatG	Mycobaterial catalase-peroxidase
NHS	N-Hydroxysuccinimide
OPLS	Orthogonal projection of latent structures
PAMP	Pathogen-associated molecular patterns
PRR	Pattern recognition receptors
PTM	Post translational modification
RAB11A	Ras related protein Rab-11
ROS	Reactive oxigen species
TLN1	Talin-1
TLR	Toll like receptors
TNF	Tumour necrosis factor
TOF	Time of flight
VCP	Valosin contained protein
VIP	Variable important on projection
VAMP3	Vesicle-associated membrane protein-3

1 INTRODUCTION

1.1 THE LUNGS

The primary function of the respiratory system is to supply the body with oxygen and to exchange and exhale carbon dioxide from the blood system by passive diffusion at the alveolar levels. The respiratory system can be anatomically divided into the proximal conducting airways and the distal respiratory region (i.e. alveolar ducts and alveolar airsacs). The average adult's lungs contain about 600 million alveoli that are surrounded by capillaries. The inhaled oxygen passes into the alveoli and then diffuses through the capillaries into the arterial blood. Meanwhile, the blood from the veins releases its carbon dioxide into the alveoli.

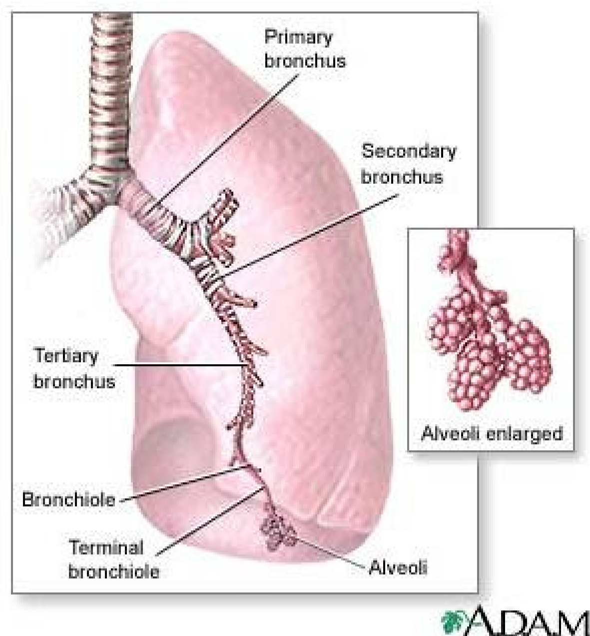

Figure I. Illustration of the anatomic Lung structure

1.2 INNATE IMMUNITY

The innate immune system is our first line of defense against invading organisms. The process of phagocytosis involves a variety of receptors and here I briefly present three of main systems: the Fcγ receptors, the complement receptors, and Toll-like receptors relevant for optimal initiation of innate responses.

1.2.1 Fcγ receptors

Immunoglobulin Fc receptors (FcRs) play a crucial role in immune defense by providing a link between antibody–antigen complexes and cellular phagocyte effector machinery. Antibody opsonization is the process through which a pathogen is marked for ingestion and destruction by AMs or other phagocytes[1]. Binding of immune-complexes to FcRs activates effector cells leading to phagocytosis, endocytosis of IgG-

opsonized particles, the release of inflammatory mediators and antibody-dependent cellular cytotoxicity [2-5]. FcRs have been described for all classes of immunoglobulins[6]; herein I only focus on FcγR since it is relevant in the studies performed in this thesis. Three subclasses of receptors (A,B and C) are known for FcγRI and FcγRII, and two subclasses A and B are described for FcγRIII[7]. All Fcγ receptors have high degrees of sequence identity but differ significantly in their cytoplasmic domains[8] which are decisive in directing downstream cellular signaling.

1.2.2 Complement receptors

The complement system encompasses a biochemical cascade which promotes the ability of IgG antibodies to clear pathogens. Phagocytic cells have a receptor for complement component such as C3b. Binding of C3b-coated bacteria to this receptor results in enhanced phagocytosis and stimulation of the respiratory burst. Activation of the complement cascade promotes clearance of dead cells or antibody complexes. These complement components are deposited on the cell surface of the pathogen and aid in its destruction[9-12].

1.2.3 Toll-like receptors

Phagocytes have a variety of Toll-like receptors (TLRs) which are a class of Pattern Recognition Receptors (PRRs) able to recognize broad three-dimensional molecular patterns termed PAMPs (pathogen-associated molecular patterns) on infectious agents[13-15]. PAMP structures recognized by TLRs include lipids, lipoproteins, proteins and nucleic acids derived from a wide range of microbes for instance bacteria, viruses and parasites. The TLRs are divided into two subsets depending on their cellular localization and the specificity of their PAMP recognition. The first group TLRs are expressed on cell surfaces and comprise of TLR1, TLR2, TLR4, TLR5, TLR6 and TLR11, those primarily recognizing microbial membrane components[14,16,17]. The other group TLRs are expressed on intracellular vesicles (endoplasmic reticulum, endosomes, and lysosomes) and includes TLR3, TLR7, TLR8 and TLR9, recognizing microbial nucleic acids[18-20]. Binding of these exogenous agents via Toll-like receptors results in activation of phagocytosis, the respiratory burst and the release of inflammatory cytokines such as IL-1, TNF-alpha and IL-6 by the phagocytes[16,21].

1.2.4 Macrophages

The mononuclear phagocytic system consists of blood monocytes and tissue macrophages. Monocytes circulate in the blood for up to 48 hours, whereas macrophages have a longer life span, about 90 days, and display a wide heterogeneity of activities. AMs represent the initial line of protection against inhaled microorganisms, and other toxins agents. They express a broad range of plasma membrane receptors which mediate interaction with microorganisms based on the recognition of conserved PAMPs[16,22]. PAMP recognition leads to activation of intracellular signaling cascades and ultimately results in a pro-inflammatory response and further activation of both the innate and adaptive immune systems[23]. Indeed, AMs play a role in the induction and regulation of adaptive immune responses[24], e.g. they work as competent antigen presenting cells (APCs) in the lungs[25] and have been linked to the pathogenesis of multiple lung diseases such as COPD and sarcoidosis[26-29].

1.2.5 Neutrophils

Neutrophils, also known as granulocytes, are essential in host defense against microbial pathogens[30]. They constantly circulate in the blood and are recruited to the infection site by gradient of chemokines (complement C5a and interleukin-8) and recognize pathogens using TLRs. Upon activation neutrophils produce superoxide anion (O·2), hydrogen peroxide (H2O2) and other reactive oxygen species (ROS) together with microbicidial peptides and proteases.

1.2.6 Dendritic cells

Dendritic cells (DC) are bone-marrow-derived leucocytes present in most tissues and mainly in the epithelium (known as Langerhans cells). They are one link between the innate and adaptive systems. DCs are specialized APCs capable of stimulating primary T-cell response to specific antigens. The main function of DCs is to phagocytose pathogens, migrate to peripheral lymph nodes and present pathogen antigens to T-cells using the MHC II molecules.

1.3 ADAPTIVE IMMUNITY

The adaptive immune response is antigen-specific and complementary to the innate response. It is active approximately after 4 to 5 days of an infection. Although this may seem as a slow response, it is specific to the invading pathogen and very powerful. The effector cells of the adaptive immune response include antigen presenting cells (macrophages and dendritic cells (DCs)) which interact with lymphocytes (T and B lymphocytes). Accordingly, the adaptive response can be subdivided into a cell-mediated immune response (including T lymphocytes and APC) and humoral immune mediated response (including B cells and soluble secreted antibodies). T and B cells work as complementary mechanisms and differ in the way they recognize antigens. While the B cell-mediated humoral response is directed against soluble exogenous proteins (antigens), T lymphocytes mainly detect protein fragments (peptides) as they are presented by specialized molecules called major histocompatibility complex (MHC).

1.3.1 MHC and antigen presentation

Major histocompatibility complex (MHC) is a large genomic region on chromosome 6 of about 3.6 Mb (3 600 000 base pairs) including 140 genes, where half of them encode for human leukocyte antigen (HLA) genes. The classical MHC molecules are also referred to as HLA molecules in humans. There are two classes of MHC molecules, MHC class- I and class II.

MHC class I molecules are expressed on every nucleated cell and has the function to bind peptides and display them on the plasma membrane. MHC I loads peptides of intracellular origin generated in the proteasome system. MHC I molecules consist of a heterodimer polypeptide, the α-chain and β2-microglobulin. There are three class I genes (HLA-A, HLA-B, and HLA-C), so it means that any cell can express at most 6 different MHC-I molecules. Due to central and peripheral tolerance mechanisms cells presenting endogenous proteins will be ignored whereas infected cells containing foreign proteins will be destroyed. MHC I bound to a microbial peptide (about 8-10

amino acids), interact with cytotoxic CD8+ T Lymphocyte (CTLs) and this lead to cell destruction. Hence, CTLs are important in clearing intracellular infections.

MHC class II molecules are found on professional antigen-presenting cells (APC) such as macrophages, dendritic cells (DCs), B-cells and also activated T cells. They are specialized to exclusively present exogenous peptides to CD4+ T cells. MHC II is a heterodimer composed of α- and a β- chain. There are three MHC II genes used for antigen presentation; HLA-DP, HLA-DQ, HLA-DR, and two other genes used for loading antigens on the MHC molecule (HLA-DM, HLA-DO). The size of the peptide loaded on the MHC II molecule is about 13-20 amino acid. HLA-genes which are highly polymorphic within a population, have been linked to different disorders, for instance it has been shown that the MHC genes are important to the disease outcome in sarcoidosis.

1.3.2 T cells

T lymphocytes play a central role in cell mediated immunity. They interact with other cells by using the cell surface T cell receptor (TCR). There are several types of T cells and can be divided based on their specific function; T helper cells (Th cells), T regulatory cells (Treg cells) and T cytotoxic cells (CTLs).

Th-cells also known as CD4+ T-cells help other cells through the release of cytokines. They are able to assist the maturation of B cells as well as activation of CTLs and macrophages. Depending on the cytokines released they are categorized as Th1, Th2 and Th17 cells, as well as Treg cells. Th1 release pro-inflammatory cytokines including IL-2 and IFN-gamma that activate macrophages, thereby increasing the antimicrobial action, and stimulate the release of IgG important for the opsonisation and clearance of pathogens. Th2 cytokines include IL-4, IL-5, IL-10, and IL-13 and mediate class switching in B cells to IgE production as well as eosinophil activation. Th17 cells are more recently discovered and considered important for various autoimmune conditions. Treg cells negatively regulate other T cells through the release of IL-10 and TGF-β, and/or through direct cell-cell interaction.

Cytotoxic T cells or killer T cells induce death of infected somatic cells. CTLs uses T-cell receptors (TCR) and the CD8 receptor to recognized specific antigens bound to MHC class I molecules. When exposed to infected cells, CTLs release perforin, granulysin and granzymes to disrupt the plasma membrane of the target cell creating pores and allowing serine and cysteine proteases to enter the cell, activate the caspase cascade and lead to programmed cell death. In addition, CTLs also express FASL that binds to Fas on target cells to induce apoptosis.

1.4 SARCOIDOSIS

1.4.1 Epidemiology

Sarcoidosis is a systemic inflammatory disease of unknown etiology, most commonly involving the lungs, skin, lymph nodes and eyes. Sarcoidosis affects people of all ethnic groups, often middle-aged adults. The incidence of sarcoidosis varies globally, probably because of differences in environmental exposures and predisposing genetic factors such as HLA alleles among others. The highest incidence of sarcoidosis has been observed in northern European countries (5-40 cases per 100,000 people)[31,32]. In Japan the annual incidence ranges from 1-2 cases per 100,000 people[33]. In USA the incidence among black Americans is about three times higher than that among white Americans (35.5 versus 10.9 cases per 100,000 people)[34]. In Sweden the incidence of sarcoidosis is about 15-20 per 100,000 individuals per year and it usually affects individuals between the ages of 20-40 years-old[35]. Epidemiological studies support the idea of environmental exposure to be important for developing sarcoidosis. Work-related risks for sarcoidosis for nurses[36,37] and season-related clustering[38,39] support a transmissibility aspect in sarcoidosis.

1.4.2 Etiology

Sarcoidosis has primarily been suspected to be an infectious disease. Although there are studies demonstrating the presence of microbial proteins, or DNA in granuloma of patients, none has been convincingly proven to be causative. Studies in Japanese population have reported the presence of Propionibacterium acnes in the bronchoalveolar lavage of sarcoidosis [40 41] as well as in granulomas biopsy samples of lymph nodes [42,43]. Mycobacterium tuberculosis has also been suggested as a possible candidate[44,45]. Systemic cellular responses to M. tuberculosis have been reported in sarcoidosis[46], and it has recently been reported that alveolar T-cells respond to different mycobacterial antigens[47,48]. In particular, the mycobacterial mkatG protein which was identified in granulomatous tissue of sarcoidosis patients[48] has been shown to be associated with both circulating anti-mKatG antibodies as well as anti-mKatG specific T cell responses in about 50% of sarcoidosis patients[45]. A meta-analysis of biopsies further support the presence of mycobacterial DNA in over 26% of sarcoidosis cases[49]. Additionally, autoimmune reactivity has also been reported; one study from our own group demonstrated immune reactivity to the self-protein vimentin in sarcoidosis patients[50].

1.4.3 Genetics

The genetic background is important for the development of sarcoidosis and also for the disease outcome. A genome wide scan analysis including 63 German sarcoidosis families was performed and identified the most prominent peak at the chromosome 6p, i.e. in the MHC region[51] Also, several studies showed associations between sarcoidosis and various HLA gene polymorphisms. In an American study, HLA-DRB1*11 and HLA-DPB1*01 were reported as risk factors for both black and white Americans[52]. In our own group, we have found HLA-DRB1*03 positive patients to associate with a self-resolving disease within 2 years time[53], whereas patients having HLA-DRB1*14 and 15 are instead to associated with a prolonged unresolved disease[54].

In addition to HLA genes, there are other genes that appear as a strong factors for sarcoidosis, such as annexin A11[55,56] as well as other more controversial BTNL2[57,58].

1.4.4 Pathology

A T cell alveolitis is characteristic for sarcoidosis; especially the increase of CD4+ T-cells with a subsequently increased CD4/CD8 ratio >3.5 is specific for sarcoidosis[59]. These immune cells are recruited from the peripheral blood using a chemokines gradient of among others IP-10 and RANTES being, shown to be highly elevated in active sarcoidosis[60]. In addition the IL-2 autocrine production is also highly elevated in sarcoidosis, leading to further local T-cell proliferation[61]. Alveolar macrophages also accumulate in the lung and have the capacity to function as specialized antigen presenting cells[62,63].

The interaction between AM acting as professional antigen presenting cells and CD4+ T-helper cells leads to an adaptive response with a characteristic Th1 profile; and in the lung tissue the formation of noncaseating granulomas in sarcoidosis[64,65] (figure 2). The formation of granulomas occurs as a result of an immune response against unresolved antigen(s) with the aim of limiting inflammation and protecting the surrounding tissue. A sarcoid granuloma is formed by the recruitment of macrophages and primarily CD4+ T-cells but also some CD8+ T-cells are present on the affected organs. The macrophages aggregate and differentiate into epithelioid and multinucleated giant cells in the core of the granuloma that is surrounded by T-cellsmainly CD4+ T-cells[66].

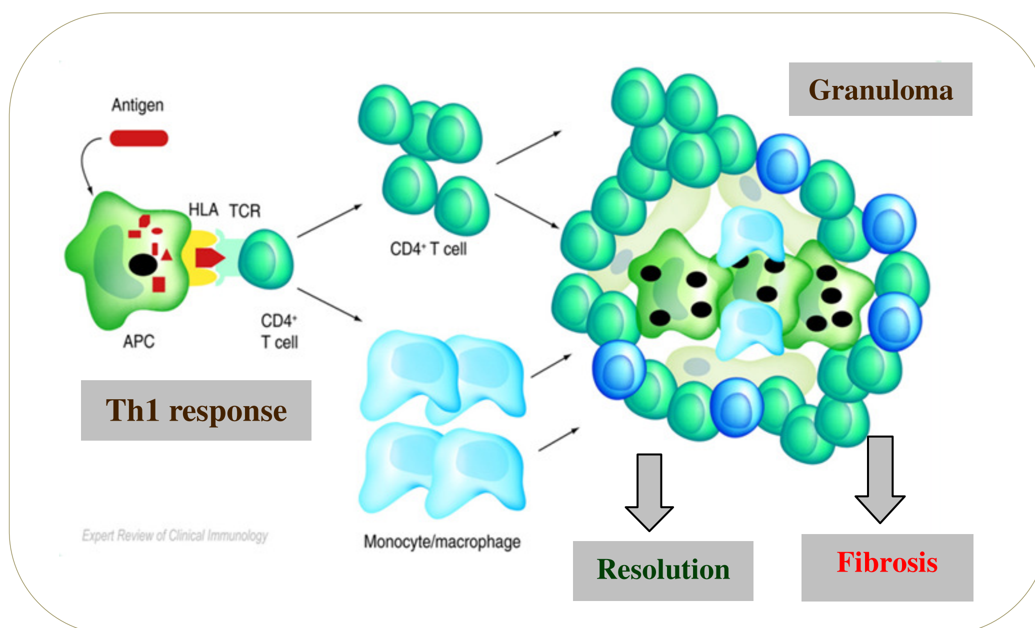

Figure 2 Schematic view of noncaseating granulomas and possible outcomes (picture was taken from Grunewald et.al, *Expert Review of Clinical immunology*)

Several cytokines play a role in the pathogenesis of sarcoidosis. The sarcoid granulomatous inflammation is characterized by dominant expression of Th1 cytokines such as IFN-γ and interleukin IL-2[67,68]. One key cytokine for the induction of Th1 response is IL-12 which is secreted by AM and found elevated in the lungs and lymph nodes of sarcoidosis patients[69]. After activation, T-cells release IL-2 that contributes to the proliferation of activated T-cells, as well as the production IFN-γ that induces macrophage activation. IL-12 and IL-18 are also increased in the lungs of sarcoidosis and they further stimulate IFN-γ production[68,70]. Tumor necrosis factor (TNF)-α plays a major role in the formation and maintenance of the granuloma[71]. In sarcoidosis the spontaneous release of the cytokines TNF-α, IL-1 by activated AM is apparent[72]. The granulomas get either self-resolved or remain in the tissue with a persistent inflammatory milieu. The mechanism(s) that result in progression to chronic disease and fibrosis are unclear but may be linked to host susceptibility and genetic factors.

1.4.5 Diagnosis

There is no single diagnostic marker for sarcoidosis as the hallmark of the disease, the epithelioid granuloma formation, is unspecific. The diagnosis is based on a combined evaluation of clinical symptoms, chest radiographic findings, noncaseating granulomas proven in tissue biopsies, and pulmonary function test. The evaluation should be in concordance with the WASOG statement on sarcoidosis[65]. If there is a CD4/CD8 T-cell ratio >3.5 in bronchoalveolar lavage fluid this will support the diagnosis. BALF findings will also be helpful in discriminating sarcoidosis from other inflammatory lung disease[59]. HLA genotyping may be valuable in predicting the prognosis of the disorder as HLA-DRB1*03 strongly associates with good prognosis[73] whereas HLA-DRB1*14 and-*15 often associates with prolonged disease course [54,73].

Chest radiographic findings have been used for decades as a means of classifying sarcoidosis patients (Table I)

Table I: Radiological stage for diagnosis of sarcoidosis

Stage	Finding
0	Normal lungs
I	Bilateral hilar lymphadenopathy (BHL)
II	BHL + parenchymal infiltration
III	Parenchymal infiltrates without BHL
IV	Signs of fibrosis with volume reduction

1.4.6 Löfgren's and non-Löfgren's sarcoidosis patients

Lung sarcoidosis patients can be sub-divided into those with Löfgren's syndrome and those without Löfgren's. Patients with Löfgren's syndrome have an acute onset with erythema nodosum and/or bilateral ankle arthritis, fever and bilateral hilar lymphadenopathy with or without parenchymal infiltrates (stage I or II), and usually resolve spontaneously within a two-year period[74-76]. The non-Löfgren's patients have a more insidious disease onset with unproductive cough, dyspnea, fatigue and later weight-loss. They commonly have radiological lung changes classified as stage (II-IV) [28,35,77].

1.5 PROTEOMICS

Quantitative proteomics is the field of global protein identification and quantification from complex samples such as cells, tissues and other biological material, where several complementary strategies are available today. Two-dimensional gel electrophoresis (2DE) is a method that has been used since 1975 to resolve hundreds of proteins simultaneously. The other complementary technique is a gel-free approach termed 'shotgun proteomics' that uses multidimensional liquid chromatography coupled to tandem mass spectrometry. This is also known as bottom-up proteomics approach since the protein identification and quantification is performed at the peptide level.

1.5.1 Two-dimensional gel electrophoresis (2DE)

Complex protein samples can be separated using 2DE technique, whereby hundreds of proteins may be resolved in a single experiment. Although 2DE has been used since the 1970s it is still a powerful method, especially for the analysis of Post-Translational Modifications (PTMs)[78]. 2DE in quantitative proteomic provides a comprehensive view of a proteome state where alteration in protein levels, splicing and PTMs can be investigated [79]. 2DE allows protein separation of complex samples according to their net charge (isoelectric point) and molecular weight. Using immobilized pH gradient strips all proteins are separated (focused) in a electric field according to the isoelectric point[80].

The used of 2DE in quantitative proteomics has been used to the study differences in protein patterns of whole proteomes. However quantification has been hampered by intrinsic factors related to the 2DE technique such as the low reproducibility due to gel-to-gel variation (up to 30% variance). In addition, the silver staining normally used to visualize proteins is not suitable for quantitative analysis since it has relative low dynamic range of two orders of magnitude and inappropriate for biological sample. Taken together, these factors have made 2DE unsuitable for accurate quantification analysis until the end of twenty century when a new approach, the Differential Gel Electrophoresis (DIGE), with significant improvements became available. DIGE is the gel-based approach used for reliable quantification of intact proteins and has been available since 2002[81]. The detection limit of the DIGE approach is in the range of 150-500pg with linear signal over four orders of magnitude of protein concentrations[82]. Besides the high dynamic range of DIGE the introduction of internal standards used for normalization was the most important improvement in order to perform reliable quantitative expression analyses using gel-based proteomics[83,84].

1.5.2 Mass Spectrometry

Mass spectrometry (MS) is a technique that measure masses of charged peptides with the aim to determine the peptide mass following identification. The principle in mass spectrometry is based on the ionization of molecules (peptides) to be able to measure their mass-to-charge ratios. Matrix-assisted Laser Desorption Ionization (MALDI) [85] and electrospray ionization (ESI) [86] are central for protein ionization in MS and their development led to award of the Nobel prize in chemistry in 2002. Using the MALDI technique, as in our study, the trypsin digested peptides are mixed with a matrix (MALDI) that is a low molecular weight that co-crystalizes and become irradiated with

a laser beam of appropriate wavelength. Here matrix sublimes and leaves the peptides in gas phase. The analytes undergo molecule collision and become often single charged protonated H+ and are then accelerated into the high vacuum mass analyzer, Time of Flight (ToF)[87] to finally reach a detector which measure the ions masses (see Figure 3). The MS data are obtained as a peptide mass fingerprint data (PMF) from individual proteins. The PMF is then compared with the in silico (theoretical) protein cleavage products for the corresponding protein sequence obtained from a protein database e.g. the National Center for Biotechnology Information (NCBInr) for protein identification.

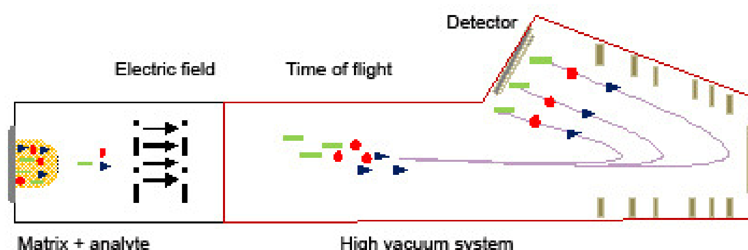

Figure 3. MALDI-ToF (Matrix-assisted Laser Desorption Ionization time of flight)

The mechanism of MALDI-ToF technique consists of three basic steps:

(i) Formation of a “matrix-analyte cake”: The matrix is always in excess to keep the analytes isolated from each-other.

(ii) Analyte in gas phase: The laser causes excitation so that the analyte and matrix is ejected from the metal surface into the gas-phase.

(iii) Analyte Ionisation: The photo-excited matrix molecules transfer protons to the analyte and create ions (mostly positive charged). The ions fly into the vacuum system time-of-flight device (TOF), and their masses (m/z) are measured.

1.5.3 Tandem Mass spectrometry LC-MS/MS

Tandem mass spectrometry (MS/MS) appeared in the early 2000s and was a breakthrough for protein identification in proteomics. These techniques also known as bottom-up proteomics or shotgun proteomics are new alternatives becoming more widely used to perform differential proteomics with the advantage to obtain huge amount data[87]. In these approaches the total protein content is solubilized and usually trypsin-digested into peptides named as precursor ions. This complex peptide soup created is then subjected to nano-scale Liquid Chromatography to allow optimal peptide resolution for identification thousands of proteins[88].

The peptides are first separated on the basis of relative hydrophobicity using reverse phase C18 column chromatography. Next, in the MS-mode the precursor ions get first resolved and then selected for further fragmentation using collision-induced dissociation energy (CID) and inert gas. The energy applied here is enough to break peptide bonds (weakest bond) and at the same time inefficient to create peptide fragments of different lengths. One peptide fragmented sequence yields enough

information to match it to a peptide sequence in a protein database on the basis of observed and expected fragmentation ions[87]. The strength of tandem MS is that many identified proteins will have multiple matches of MS/MS-derived sequences. The Orbitrap MS, being used in our study (paper IV), is the latest technology for tandem MS. This employs a new physical principle, namely separating ions in an oscillating electric field, and gives greater proteomic coverage due to their much higher resolution, thus being ideal for quantification [89,90].

1.5.4 Isotope Tags for relative and absolute quantification

Global protein quantification can be performed using a peptide-labeling approach. One of the most robust approach available today for peptide labeling techniques is Isotope Tags for Relative and Absolute Quantification (iTRAQ)[91] which requires tandem MS. An advantage is that iTRAQ labels all peptides at the N-terminal and at lysine residues, and allows use of 8-plex to simultaneously analyze 8 individual samples in one LC-MS analysis. As these tags are isobaric, in the MS-mode the precursor-ions from eight differentially labeled samples appear as a single peak thus allowing joint identification. Subsequently, in the MS/MS-mode the precursor-ions are fragmented into number peptides with different mass to be used for protein sequence analysis. Furthermore, single charged reporter ions 113-119 and 121 Da are released and located at regions of the mass spectra that do not contain information of the peptide sequencing (at the 110-130 m/z) which can be used for relative quantitation.

1.5.5 Multivariate Analysis

Multivariate statistical analysis is a group of methods used to reduce the dimensionality of the data, and is very useful in omics methods where the large number of variables tested at ones poses the risk of high false positive rates when using traditional univariate methods such as t-tests. Principal component analysis PCA was invented in 1901 by Karl Pearson Karl[92]. This is an unsupervised method that involves a mathematical transformation to convert a set of observations and correlated variables into a set of principal components[93].

Partial least squares (PLS) also known as projection to latent structures was developed in 1960 by Sture Wold[94]. PLS regression is a supervised multivariate method for assessing relationships between two set of variables; the effector variables (matrix X) and the response variable Y. In our studies the X matrix was conformed of all proteins identified in proteomics and the Y variable was the diagnosis (healthy or sarcoidosis). PLS-regression is particularly suited when the matrix of predictors has more variables than observations, and when there are multi-collinearities among X variables with emphasis on predicting the response.

The orthogonal PLS (OPLS) method is a recent modification of the PLS method developed by Johan Trygg and Svante Wold[95]. OPLS separates the systematic variation of X matrix into two parts, one that is related to Y-variable and one that is unrelated (orthogonal) to Y variable. This partitioning of the X-data provides improved model transparency and interpretability, but does not change the predictive power.

2 CURRENT INVESTIGATION

2.1 AIMS OF THE THESIS

The overall aim of this thesis was to investigate alterations in protein profiles and to discover specific diagnostic markers for pulmonary sarcoidosis by applying different proteomics techniques. Thereby biological mechanisms related to sarcoidosis pathogenesis would be elucidated.

The specific aims of individual sections were:

I. to characterize protein expression profiles in bronchoalveolar lavage fluids from patients with two granulomatous disorders, sarcoidosis patients (HLA-DRB1*15 positive) and chronic beryllium disease.

II. to quantify different sources of variance related to 2DE techniques and to validate the levels of objectivity from the 4th generation image analysis software SameSpotsTM.

III. to further elucidate new proteins and mechanisms in lung sarcoidosis by studying alveolar macrophages using complementary gel-based DIGE and gel-free LCMS/MS proteomics approaches.

2.2 COMMENTS ON METHODOLOGIES

2.2.1 Patient data and sampling collection

In the present thesis the experimental work was based on human clinical samples including proteins from bronchoalveolar lavage fluids, isolated alveolar macrophages and sera from sarcoidosis patients, chronic beryllium disease patients and healthy controls, respectively.

Sarcoidosis patients

All sarcoidosis patients were investigated following a strict protocol at the department of respiratory Medicine at Karolinska University Hospital, Stockholm, Sweden. The diagnosis and disease phenotype was assessed in concordance with the World Association of Sarcoidosis and Other Granulomatous disorders (WASOG) criteria [65]. Thus the patients had symptom (see 1.4.6) chest radiographic and, pulmonary function tests compatible with sarcoidosis. In addition, BALF findings can support the diagnosis. Furthermore, the patients were genotyped assessment of HLA class II (paper I). In papers III and IV the sarcoidosis patients were also classified according to the presence or absence of Löfgren's syndrome. Löfgren's syndrome is common (>30%) in sarcoidosis patients with Caucasians origin [96]. Four patients had Löfgren's syndrome and four were classified as non-Löfgren. Additionally, 18 Löfgren's and 18 non-Löfgren's patients were recruited for ELISA validation. At the time of bronchoscopy all patients had active disease and did not receive any steroid treatment.

Chronic Beryllium Disease patients

BAL samples from five CBD patients were recruited (paper I) from the National Jewish Medical and Research Center, Denver, USA. The diagnosis was established by using defined criteria including a history of beryllium exposure, presence of a positive blood BeLPT-test[97] (Beryllium Lymphocyte Proliferation test), the presence of granulomatous formation and/or mononuclear cell infiltration in lung biopsies[98].

Healthy Controls

All healthy controls recruited were never-smokers and had normal pulmonary function test and chest radiography. Paper I included 5 healthy subjects, papers II and III 7 subjects each and paper IV six subjects. In addition, 14 healthy subjects were added for ELISA to validate the results. All subjects had to be free from airway infections at least four weeks prior to investigation. Informed consents were obtained, and all studies were approved by the local Ethics committee.

2.2.2 Bronchoscopy

BAL samples were obtained using a standardized bronchoscopy procedure performed at Karolinska University Hospital, Stockholm, Sweden[99]. BAL samples from CBD patients were obtained at the National Jewish Medical and Research Center, Denver, CO, USA. Briefly, a bronchoscope was inserted into the right middle lobe under local nasobronchial anaesthesia. Aliquots (5 x 50 ml) of sterile saline solution at 37°C were instilled. The fluid was then gently recovered by aspiration into a siliconized bottle and kept at 4°C. BAL samples were filtered to remove debris using a Dracon net (Miipore,

Ireand) and then centrifuged to separate BAL cells (papers II, III, & IV) from the supernatant soluble proteins (paper I) (figure 4). Cell pellets were suspended in PBS buffer to determine cell viability (using trypan blue exclusion) and to do differential cell counting.

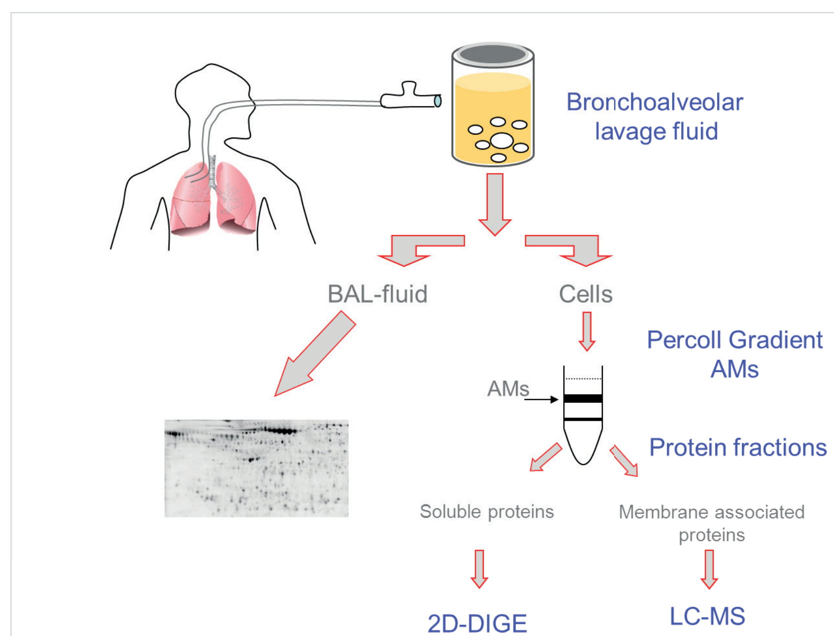

Figure 4. Workflow of clinical sample used in the studies

2.2.3 BAL fluid concentration (paper I)

In order to use BAL soluble proteins in 2DE analyses (paper I), desalting and protein concentration was required. For this purpose Amicon Ultra® (Millipore, Bedford, USA) centrifugal filters with an exclusion filter cutoff of >5000 Da were used. First, aliquots of 15ml were centrifuged using Amicon Ultra® 10 (45 min at 4000g). Further BAL concentration was performed with Amicon Ultra® 4 ml devices (70 min, 4000g). Finally, BAL samples were desalted and concentrated from 50 ml to a final volume of 100µl. Protein concentration determination was performed prior to 2DE.

2.2.4 Alveolar macrophage (papers III & IV)

Immune cells from the BAL fluid were collected through centrifugation. AMs were then isolated using a PercollTM density gradient (GE Healthcare Uppsala, Sweden). A two-phase gradient solution with density $\rho 1= 1.062$ g/ml, and $\rho 2= 1.08$ g/ml was created in 5ml polyprolylene tubes. Samples were gently layered on the top of the system and cells separated through centrifugation (400xg for 20 min) with a recovery of >80% and purity of 89.4 ±4.2%. Cells were suspended in RPMI/10% DMSO and store at -70°C.

2.2.5 Microsomal preparation (papers II, III, IV)

Cell pellets containing >4x106 AMs were suspended in 1 mL 0.1 M potassium phosphate buffer, pH 7.4, 0.25 M sucrose and protease inhibitors (Complete Mini, Roche Diagnostics). The homogenate was sonicated (probe tip sonicator, power: 30%,

Bandelin Sonopuls, Buch & Holm) for 4 x 10 seconds on ice. A centrifugation was performed at 10000 rpm at 4°C for 10 minutes and the pellet was discarded. The protein suspensions were then centrifuged at 100 000 x g at 4°C for 1 hour. The resulting supernatants containing soluble proteins were stored at -20°C and the pellets were suspended in 500µL 2.5 M NaBr for 45 min on ice with shaking. Another centrifugation was performed at 4°C for 1 hour at 100 000 x g. The supernatant containing membrane-associated proteins (MAPs) and the pellets containing microsomes were stored separately at -20°C until further analysis.

2.2.6 Differential Gel Electrophoresis (papers I, II & III)

DIGE utilizes three fluorescent Dyes (Cy2, Cy3, and Cy5) that covalently bind to Lysine residues on intact proteins using NHS ester chemistry (figure 5). Proteins from three samples can be labeled and run simultaneously in one 2DE gel. Classically, Cy3 and Cy5 are used for sample labeling whereas Cy2 is used as an internal standard[100]. These dyes have identical mass (500Da) and charge +1, but they have distinct emission spectra that make it possible to discriminate between each sample used in multiplexing[83]. Proteins coordinates in the 2DE gels are the same for all three samples and emission data can be collected using laser scanning. The protein expression levels were measured using software analysis Decyder (paper I) and Samespots (papers II and III).

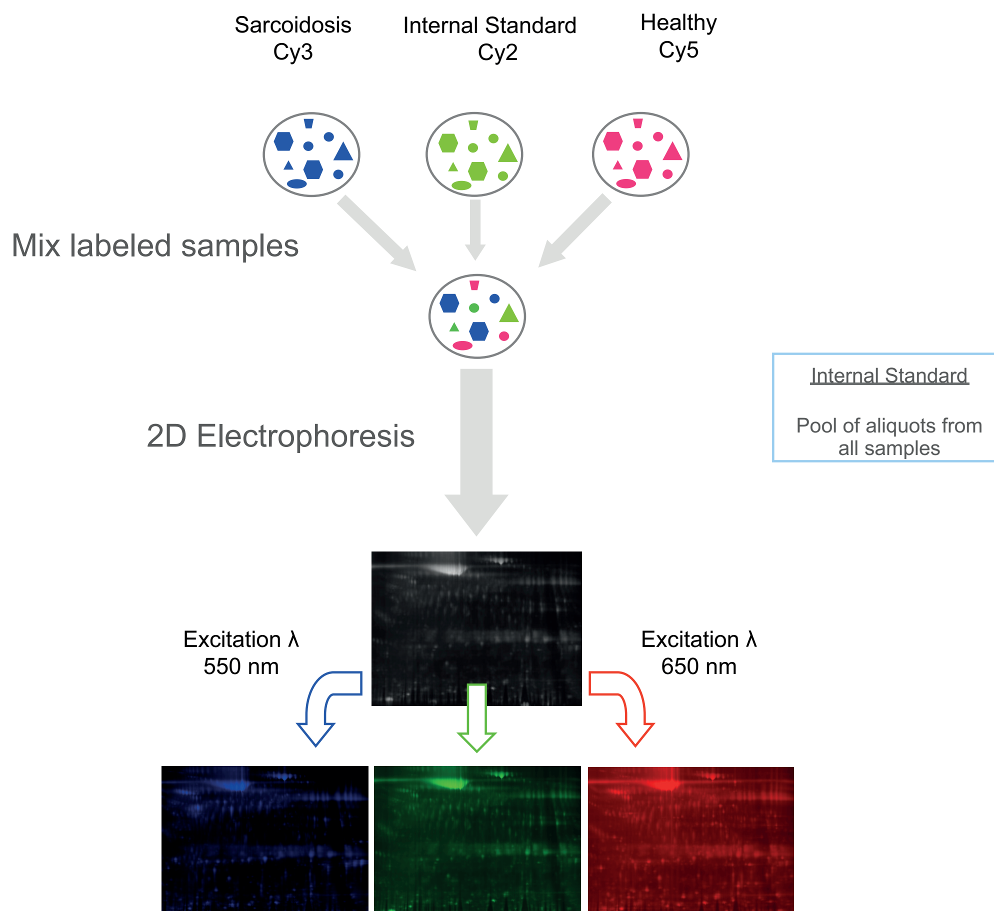

Figure 5. Principle and workflow of DIGE

Preparative Gels

Deep Purple™ Total Protein Stain (paper I): gels were first fixed overnight (7.5% acetic acid/10% methanol) and then incubated in working solution (Deep purple 1:200 dilution). Gel images were acquired using the Typhoon 9400 laser scanner, with emission Argon laser (□ex = 457nm) and band pass emission filter at 610nm (±15nm). Coomassie Brilliant Blue staining (paper I): Gels were fixed (50 % ethanol, 10% acetic acid for 1 hr), and then stained (Coomassie Brilliant Blue R-250 0.2%, acetic acid 10%, ethanol 50%) overnight followed by destaining (10% acetic acid, 10% ethanol for 6h).

2.2.7 Mass Spectrometry (papers I and III)

Peptides of interest were trypsin-digested prior to mass spectrometry analysis using MALDI-ToF. The instrument was calibrated (autodigestive tryptic peptides from trypsin) and individual ions were assigned with accurate masses to obtain peptide mass fingerprint data (PMF) from individual proteins. The mass spectra were compared with the protein sequence database National Center for Biotechnology information (NCBInr) having tripsin-digested proteins in silico for protein identification. The statistical probability threshold and Z-score are used to accept protein identification as a correct match [101].

2.2.8 Peptide labeling iTRAQ (paper IV)

Membrane-associated proteins were solubilized in 0.5% SDS to determine protein concentrations using the DC protein assay (Bio-Rad Laboratories, Hercules, CA, USA). Samples were then reduced with DTT (final concentration 5 mM, 30 min at 56° C) and alkylated with iodoacetamide (final concentration 0.015 M, 30 min, in the dark) before trypsin digestion. Modified sequencing grade Trypsin, (Promega, Madison, USA) was used and the samples were incubated at 37°C overnight. For each sample 20 µg of were labeled with iTRAQ (8-plex; Applied Biosystems, Foster City, CA, USA) according to the experimental design. An internal standard composed of aliquots from each sample used was in the study, assuring that all proteins present were represented in the internal standard. Excess reagent was removed from the pooled sample using an SCX-cartridge (StrataSCX, Phenomenex, Torrence, CA, USA) and the eluate was dried in a vacuum centrifuge to further fractionate with narrow IPG-strips.

Peptide fractionation using narrow IPG-strips

A single dimension separation is not enough to resolve the tens of thousands peptides presented in these biological samples. Consequently, narrow IPG 3.5-4.5 offline can lead to successful separation and identification of many proteins in complex samples[90], [102]. The integration of narrow IPG-strip and the automatic fractionation into 72 fractions is high reproducible.

2.2.9 LC-MS/MS analyses (paper IV)

3µl of each fraction sample were injected on the hybrid LTQ-Orbitrap Velos mass spectrometer (Thermo Fischer Scientific, San Jose, USA). Online reversed-phase nano-LC at 0.4 µl/min for peptide resolution was performed using Solvent A (3 % ACN, 0.1 % FA); and solvent B (95 % ACN, 0.1 % FA). The gradient created was from 2 % B up

to 40 % B in 45 min, followed by a steep increase to 100 % B in 5 min. Samples were installed onto the nano electrospray ionisation (NSI). Acquisition proceeded in ~3.5 s scan cycles, starting by a single full scan MS at 30000 resolution (profile mode), followed by two stages of data-dependent tandem MS (centroid mode): the top 5 ions from the full scan MS were selected firstly for collision induced dissociation (CID, at 35 % energy) with MS/MS detection in the ion trap, and finally for higher energy collision dissociation (HCD, at 45 % energy) with MS/MS detection.

2.3 RESULTS AND DISCUSSION

The aim of this thesis was to study the protein expression in the lungs of sarcoidosis patients with the purpose of identifying relevant proteins markers that could help to better understand this disease. Accordingly, previous proteomic studies have reported the discovery of altered proteins in sarcoidosis. However, the lack of disease-specificity of these proteins has not allowed the identification of disease markers for sarcoidosis.

2.3.1 Paper I

A quantitative proteomic analysis of soluble bronchoalveolar fluid proteins from patients with sarcoidosis and chronic beryllium disease (paper I)

In paper I the aim was to analyze whole protein expression profiles from the alveolar lumen in HLA-DRB1*1501 positive sarcoidosis patients, and to compare these to healthy controls but also to chronic beryllium disease patients which is another granulomatous disease. Both sarcoidosis and CBD patients have several similarities regarding the immune response including; the granulomas formation, infiltration of T-cells in lung parenchyma, and the characteristic Th1 shifted immune response in the lungs[103]. Sarcoidosis patients with HLA-DRB1*1501 tend to have a prolonged, chronic inflammation and are at risk of developing fibrosis.

Using two different controls, i.e. CBD and a healthy group, we searched for specific protein markers in sarcoidosis that could differentiate from other granulomatous diseases. Aiming at identifying proteins that differ between groups, the Student's T-test with $p < 0.05$ was considered as significant in our analysis and a total of 14 proteins were identified.

Oxidative Stress and Inflammation response

Oxidative stress is implicated in the pathogenesis of sarcoidosis[104], CBD[105,106] and other respiratory disorders e.g. lung fibrosis, asthma, and chronic obstructive lung disease (COPD)[107,108]. From our results we have identified a set of proteins that correlate to the anti-oxidative system (peroxiredoxin 5, hemopexin, α1-antitrypsin, and superoxide dismutase) that are altered in granulomatous diseases. These proteins have been reported previously in sarcoidosis and CBD[106,109,110]. Scandinavian sarcoidosis patients (Finland) were shown to have peroxiredoxin V highly expressed in their lungs[104]. The other altered mechanism was the inflammatory response where levels of β2-microglobulin, annexin II, Complement C3, apolipoprotein A1, IgG kappa chain and heat shock protein 70 were also found significantly altered.

In summary, the proteins identified reflect an ongoing inflammatory response[111,112,113] as well as an up-regulation in oxidative homeostasis[104 114-116] in the lungs of these patients. Another conclusion was that the type of proteins identified in BAL samples was mainly high abundant proteins. This is a frequent problem when analyzing complex samples with a wide range in protein concentration[117-119]. Therefore, in the following studies we tried to avoid this by sample treatment enabling us to target deeper into the proteome.

2.3.2 Paper II

In the eye of the beholder: does the master see the SameSpots as the novice?

Two-dimensional gel electrophoresis (2DE) has been widely used for separation of complex protein samples and subsequent comparison of protein profiles. 2DE has been though disadvantaged in quantitative proteomics because of the intrinsic variance introduced through the workflow[80,120,121]. The introduction of internal standard in DIGE technique, have reduced most of the gel-to-gel variance of 2-DE, thus allowing 2DE technology to be a robust platform for quantitative proteomics[100,122,123]. With these improvements new problems such as software-related variance as well as subjectivity of the analysis has become more prominent[124-126] and need to be addressed. Although the background subtraction is meant to allow more accurate protein quantification, it had been shown that these algorithms introduced additional variance to the analysis[127].

The aim of this study was to measure software related variance from available background algorithms (average on boundary; AoB, lowest on boundary; LoB and mode of none spot; MoNS) as well as the experimental variance from traditional single stained technique, respectively. As showed in figure 6 different variance from background subtraction algorithms were measured performed by two independent users. The default algorithm LoB introduced the least amount of variance (Lab-user: CVmean=2.5%; Expert: CVmean=3.1%), the algorithm AoB showed (Lab-user: CVmean=2.8%; Expert: CVmean=3.8%) and finally, the algorithm MoNS introduced the highest variance (Lab-user: CVmean=5.5%; Expert: CVmean=7.2%). Comparing these results with early software these variance shows significant improvements[127,128].

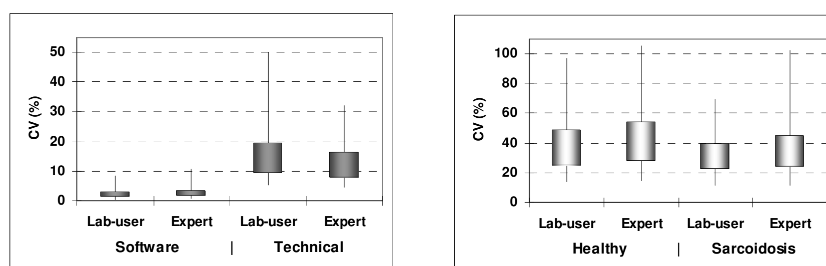

Figure 6: Software, technical and biological variance measured in 2DE

In addition, single-stained gel variance was measured as a technical related variance. For single gels, spot-based ratio normalization was used whereas for DIGE the ratiometric normalization (including an internal standard) was applied. As showed in the figure 6 the normalization method of choice had a much larger effect in variance introduced to the analysis. Finally, we correlated the total spot-volumes identified from expert and lab-user, using partial least squares of latent structures (PLS) analysis to see levels of similarities and objectivity of the analysis. Accordingly, the PLS demonstrates a strong correlation (correlation coefficients near to 1.0) for the contribution of proteins between the two analyses, showing that the outcome of the analysis is largely user-independent and highly objective.

2.3.3 Paper III

A proteomic approach to study alveolar macrophages from sarcoidosis patients

Herein we studied alterations in the expression levels of AM soluble proteins in pulmonary sarcoidosis. Using 2DE couple with MS and MSMS we identified a set of new proteins that were differentially expressed ($p<0.05$) in patients with sarcoidosis as compared to healthy controls (Table II).

Improvements in sample preparation including the isolation of AMs from other BAL cells[129] and subsequent protein fractionation[129] has allowed identification of new set proteins not previously reported in sarcoidosis[130-133].

Table II

Protein name	Analysis	Accession(*)	Score	Theoretical		Observed		coverage	Match
			(†)	*Mw Kda*(‡)	*pI*(‡)	*Mw Kda*(‡)	*pI*(‡)	(%)	Peptides
sly 1p	MALDI-TOF	AAP97146	*p <0.0001*	70.27	5.6	82	5.3	12	6
Eph-like Receptor tyrosine kinase	MALDI-TOF	AAD02030	*p <0.0001*	111.51	5.9	103	6.0	12	7
Myeloid differentiation factor 88	MALDI-TOF	AAB49961	*p=0.0064*	33.73	5.6	38	5.3	34	7
Diacylglycerol Kinase isoform1	MALDI-TOF	NP_690874	*p <0.0001*	130.07	5.9	87	5.2	9	7
Dipeptidylpeptidase	MALDI-TOF	AAQ83119.1	*p <0.0001*	99.05	6.0	87	6.2	7	5
Zing finger protein isoform 1	MALDI-TOF	NP_991331	*p <0.0001*	60.49	5.4	88	6.2	19	9
Phospholipase D	MALDI-TOF	AAA36444	*p <0.0001*	93.28	5.9	95	5.4	9	6
Glutathione synthetase GSH-S	MALDI-TOF	NP_000169	*p=0.0068*	52.54	5.7	78	5.8	25	10
Vimentin	MALDI-TOF	AAC05002	*p <0.0001*	53.69	5.1	55	5.6	31	16
Rab-11 family interacting protein 1B	MALDI-TOF	AAQ18786	*p <0.0001*	137.55	5.3	115	5.2	10	7
OSBP-related protein 1	MALDI-TOF	AAG53407	p=*0.009*	107.13	6.1	98	5.9	14	9
Adaptor-related protein complex 2	MALDI-TOF	AAH06201	*p <0.0001*	106.59	5.2	110	5.1	11	8
Gelsolin isoform b	MALDI-TOF	NP_937895	*p <0.0001*	80.91	5.6	95	5.9	22	13
Dynein light chain-A	MALDI-TOF	AAD44481	*p <0.0001*	56.89	6.1	58	6.2	12	7
Valosin-containing protein	MALDI-TOF	NP_009057	*p <0.0001*	89.99	5.2	94	5.3	19	12
Topoisomerase IIb	MALDI-TOF	AAB01982	*p=0.009*	38.55	5.1	34	5.3	31	6
Aldehyde Dehydrogenase 1	MALDI-TOF	AAC51652	*p <0.0001*	55.44	6.3	85	6.7	20	7
Na+/K+ ATPase alpha 1 subunit	MALDI-TOF	NP_000692	*p <0.0001*	114.19	5.3	93	5.1	11	6
Adenylate Kinase 5	MALDI-TOF	AAH36666	*p*=0.0012	63.83	5.0	69	4.9	14	4
RhoA/RAC/CDC42 exchange factor	MALDI-TOF	AAAH12860	*p <0.0001*	53.58	5.3	114	5.2	22	7
golgi integral membrane protein 4	MALDI-TOF	NP_055313	*p <0.0001*	81.92	4.7	87	5.1	18	9
Actin, beta	MALDI-TOF/TOF	CAA45026	54	41.79	5.2	43	6.0	5	2
Actin, beta	MALDI-TOF/TOF	AAH08633	173	41.32	5.6	43	6.1	11	5
Gamma-interferon thiol reductase	MALDI-TOF/TOF	AF097362	92	28.5	4.7	25	5.5	4	1
Keratin 10, type I, cytoskeletal	MALDI-TOF/TOF	KRHU0	173	41.32	5.6	43	6.1	5	3
Vimentin	MALDI-TOF/TOF	AAA61279	69	53.68	5.0	81	5.1	5	8
Heat Shock 70kDa protein 9 (Mortalin)	MALDI-TOF/TOF	AAH00478	494	73.68	5.9	90	5.7	13	6
VIM (Vimentin)	MALDI-TOF/TOF	CAG28618	301	53.55	5.1	81	5.1	14	5
Vimentin	MALDI-TOF/TOF	AAA61279	557	53.68	5.0	75	4.6	25	8

(*) Abbrevation and accession number to the NCBI database
(†) For MALDI-TOF/MS analysis, ProFound probability and for MALDI-TOF/TOF analysis, MASCOT score is shown.
(‡) Theoretical and experimental values: molecular weight (*Mw*) and isoelectric point(*pI*)

Pathway analysis

We found that the Fcγ receptor-mediated phagocytosis in macrophages (p=1.3*10-5), and the clathrin-mediated endocytosis signalling (p=0.0019) were both upregulated in sarcoidosis. The proteins rab11, phospholipase D, diacylglycerol kinase, dynein and gamma interferon lysosome transferase were also up-regulated, and they are associated with Fcγ receptor-mediated phagocytosis in macrophages. Several studies show the functionality of these proteins in phagocytosis[134-137]. Dynein has been reported to be involved in internalization of exogenous pathogens during macropinocytosis[138].Gamma interferon-inducible lysosomal thiol reductase (gilt) was found to be, upregulated in sarcoidosis; it cleaves disulphide-bounds of antigens and is implicated in MHC-II antigen processing[139,140].

In the clathrin-mediated endocytosis signalling pathway AP-2, actin and rab-11 were up-regulated: AP-2 is highly active in endosomal trafficking involved in MHC class II antigen presentation[141] and is related to the antigen presenting function on AMs in sarcoidosis[29]. The Rab family, approximately composed of 70 related proteins identified in humans, regulates steps during membrane traffic, vesicle formation and membrane fusion[142]. Rab11 has been reported to be associated with the recycling compartment, post-Golgi vesicles, and the trans-Golgi network in early endosomes[143]. So , the Fcγ receptors (FcγR) play a crucial role in the innate response as cellular phagocyte effector machinery[2,144]. In sarcoidosis, AMs have been shown to be activated antigen-presenting cells able to phagocyte foreign antigens and present epitopes on HLA-molecules to the T-cells and induce a potent adaptive response[29].

Multivariate analysis

Multivariate modeling was applied to identify protein candidate markers for sarcoidosis. When response variable diagnosis was correlated with protein-spot abundances, we constructed a model able to explain 86% of the variance with a predictive power of 41% (R2Y=0.86, Q2=0.41) (figure7). When using variables important to the projection (VIP-plot), to rank the important variable according to their contribution to the model; valosin containing protein (vcp) came out as the most relevant variable for discriminating between sarcoidosis and the healthy group.

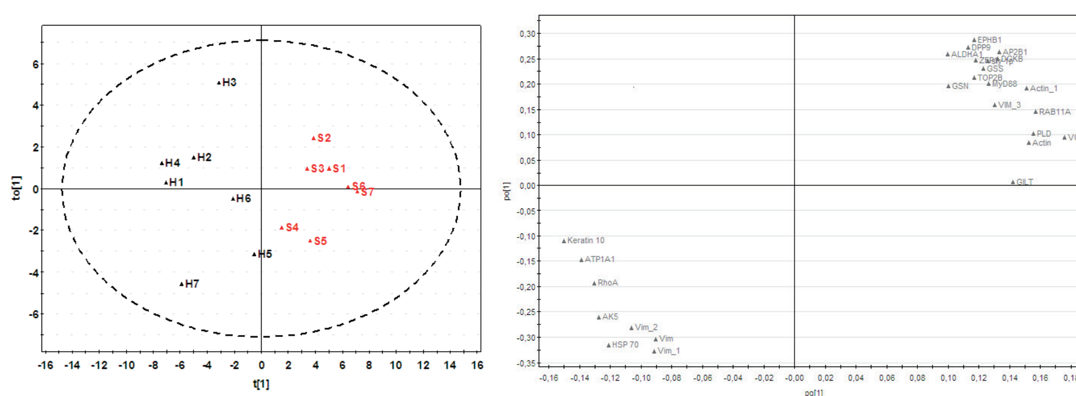

Figure 7. Score-plot (left graph) shows the group clustering for sarcoidosis (S1-7) and healthy (H1-7) respectively. The load-plot (righ panel) shows the proteins and their correlation to the groups. Proteins located at the right area of the plot correlated directly with sarcoidosis group, whereas those proteins at the left area correlated inversely.

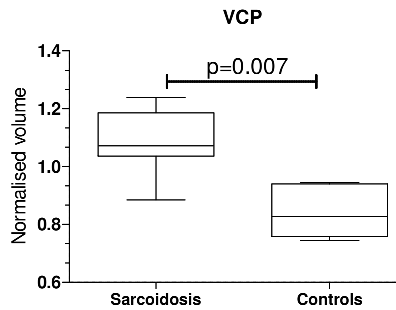

Figure 8: Valosin containing protein upregulated in DIGE expression analysis

2DE image analysis revealed valosin contained protein (vcp) to be up-regulated in sarcoidosis (p=0.007) (figure 8). vcp is involved in several relevant functions such as membrane fusion[145], regulation of transcriptional activation NFκB[146], apoptosis[147,148] and the proteosomal degradation system (ERAD)[149 150,151].

In summary, we reported here the identification of 25 unique proteins from the AMs altered in sarcoidosis. In the network analysis 11 proteins were found linked to the 26s proteasome and the NFkB complex nodes; 8 up regulated and 3 down regulated in sarcoidosis. Two pathways associated to phagocytosis were upregulated in sarcoidosis; the Fcγ receptor mediated phagocytosis in macrophages, and the clathrin-mediated endocytosis signalling. Additionally, the up regulation of myd 88 and gilt in AMs from sarcoidosis patients may suggest the possible presence microorganism agents. Myd is in tight connection with TLR signalling and gilt strictly specific for MHC II antigen processing. These findings represent a step forward in understanding the complexity of sarcoidosis pathogenesis, and help to better understand mechanisms that could be of relevance in the development of sarcoidosis. Further validation of these data in larger sample material is required in order to be able to suggest them as potential markers.

2.3.4 Paper IV

Proteomic analysis of membrane-associated proteins in alveolar macrophages from patients with pulmonary sarcoidosis

Results from Paper III revealed a differential analysis of soluble proteins from AMs using the gel-based technique DIGE and highlighted the discovery of a new set of 25 proteins and pathway alterations. Several proteomics studies have reported the benefits in applying sample fractionation prior to expression analysis in order to reduce protein complexity and to enable deeper insight into the proteome for discovery of protein biomarkers or to identify relevant associated mechanisms.

In this paper we performed a Shotgun proteomics approach using LTQ-Orbitrap Velos mass spectrometer on membrane associated proteins (MAP). This new technique is capable to identify and in a relative way quantify large numbers of proteins simultaneously. Accordingly, 1650 proteins were identified from which 423 proteins (expressed in all samples) have been possible to perform differential analyses using iTRAQ labeling.

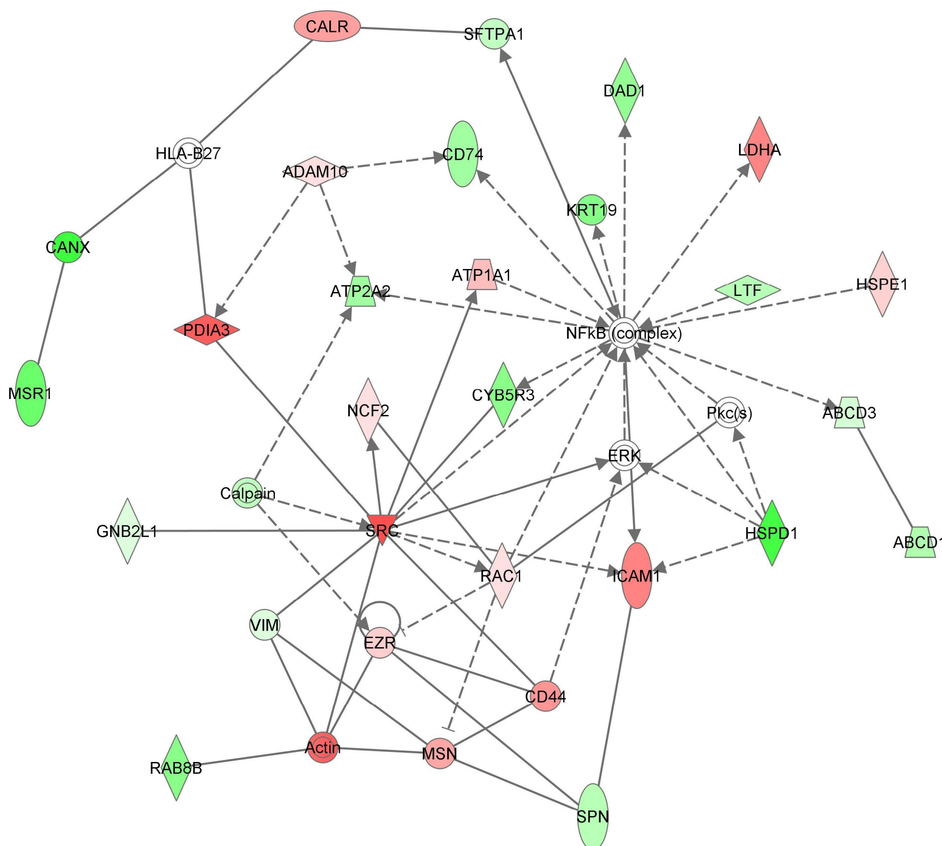

Figure 9. Network analysis showed proteins altered in sarcoidosis linked to the NFkB complex. Green represents decreased level and red an increased level in sarcoidosis. A total of 31 proteins, 14 up-regulated and 17 down-regulated in sarcoidosis were linked directly or indirectly to NFkB.

Pathway analysis:

Mitochondrial dysfunction in Non-Löfgren's sarcoidosis patients

Ingenuity system analysis permits the characterization of pathway alterations in sarcoidosis samples. Down- regulation of the oxidative phosphorylation pathway in the mitochondrial respiration system was observed in sarcoidosis (p-value 1.02x10-10) (figure 10). Interestingly, when the patients were sub grouped according to their clinical phenotype into Löfgren's and non-Löfgren's, we found this pathway up-regulated in Löfgren's patients but down-regulated in non-Löfgren's patients.

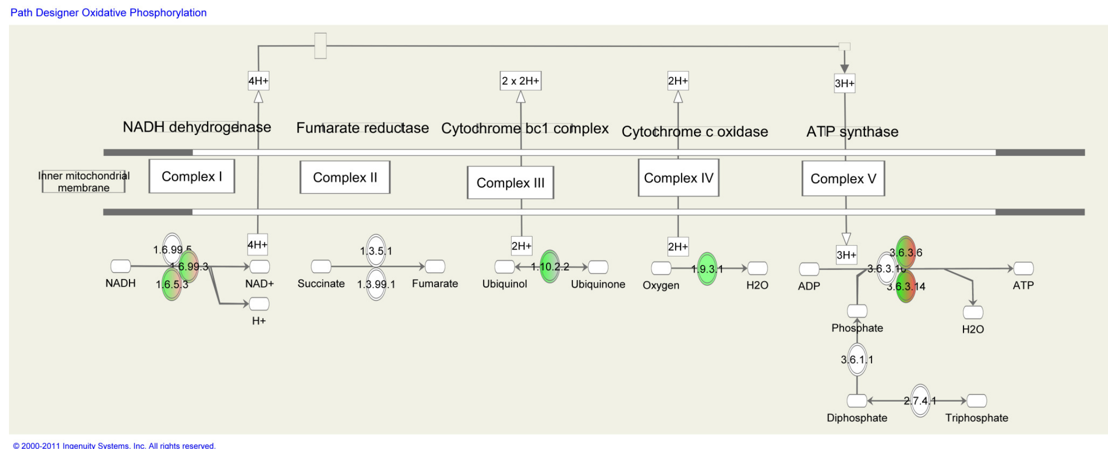

Figure 10. Mitochondrial dysfunction in the electron transport in Non-Löfgren's sarcoidosis patients

Non-Löfgren's sarcoidosis patients have a tendency to associate with unresolved chronic inflammation, with elevated levels of reactive oxygen species (ROS). They often present a more advanced radiological stage (III-IV) and have an increased risk of developing fibrosis.

It is known that excessive levels of ROS cause tissue damage, and redox imbalance has also been associated with other lung disease such as, chronic inflammation in COPD[152], asthma [153], lung fibrosis and cancer [107]. In sarcoidosis, ROS species such as superoxide anion in lungs has shown to be highly elevated. Even a systemic oxidative stress has been reported in pulmonary sarcoidosis[154].

Furthermore, this redox imbalance has been shown to cause mitochondrial proteins damage leading to increased protein degradation and dysfunction in mitochondria [155,156]. These results suit well with the clinical phenotype of these non-Löfgren's patients. This down-regulation of oxidative phosphorilation and mitochondrial dysfunction is correlated to high level of ROS in lungs; therefore we suggest that this could be of therapeutic importance. Since it is known that systemic steroids do not help in advanced stage of sarcoidosis, it may be that antioxidant drugs may help to reduce the chronic inflammation in lungs.

Up-regulation of the phagocytosis pathways in sarcoidosis

Phagocytosis is complex and comprises among other events particle binding, actin rearrangement, pseudopod extension, membrane recycling and phagosome closure. In concordance with the previous study (paper III); here we also found the up-regulation of Clathrin-mediated endocytosis signaling in AMs from sarcoidosis patients (p-value 9.64x10-9). Clathrin-dependent endocytosis is a process by which cells internalize molecules through the inward budding of plasma membrane vesicles containing proteins with receptor sites specific to the molecules being internalized. This mechanism is relevant to the MHC II-peptide mechanism[157,158]. Ingenuity system showed clathrin-mediated endocytosis signaling increased levels in sarcoidosis of the adaptor-related protein (complex 1 beta 1 subunit AP1B1; complex 2 alpha 1 subunit, AP2A1; and complex 2 sigma 1 subunit AP2S1), member RAS oncogene family (RAB11A, RAB4A, RAB4B, RAB5C, RAB7A), clathrin (light-chain CLTA, and heavy-chain CLTC), actin like protein-3; ARP-3, actin beta, ACTB; and actin alpha, ACTA1. In addition, heat shock 70kDa protein 8, HSPA8; and transferrin; TF, were decreased in sarcoidosis.

Additionally, we have also found similar up regulation on the Fcγ receptor-mediated phagocytosis in macrophage pathway in sarcoidosis (p-value 1.47x10-5) but in this study with even more proteins identified in this pathway. Here nine proteins were upregulated (Actin alpha-1, ezrin, actin-beta, actin related protein-3 homolog, myosine VA, rab-11, ras-related C3 botulinum toxin substrate 1, v-src sarcoma viral oncogene homolog, talin-1) and two were down-regulated (vesicle-associated membrane protein-3,heme oxigenase-1) in sarcoidosis.

Accordingly, these two complementary pathways were up-regulated in both the Löfgren's and non-Löfgren's sarcoidosis patients, as reported in paper III. These pathways are part of the mechanism that mediates internalization of pathogens, for instance bacteria, into membrane-derived vacuoles and the phagosome system, finally leading to degradation of pathogens and the presentation of specific antigens to memory T-cells in the lungs[6]. This results suggest the hypothesis of an infectious agent in sarcoidosis, where AMs as APC constantly phagocyte, process and present the antigens T cells in the lungs.

Predictive OPLS model to discriminate sarcoidosis from the healthy group

Using 423 proteins the orthogonal projection to latent structures (OPLS) was performed to identify, if possible, a model able to discriminate sarcoidosis patients from healthy controls. We constructed a model with a subset of 13 putative biomarker proteins that resulted in a highly predictive model (Q2 = 0.72) with an overall model p-value of 0.0009 (figure 11).

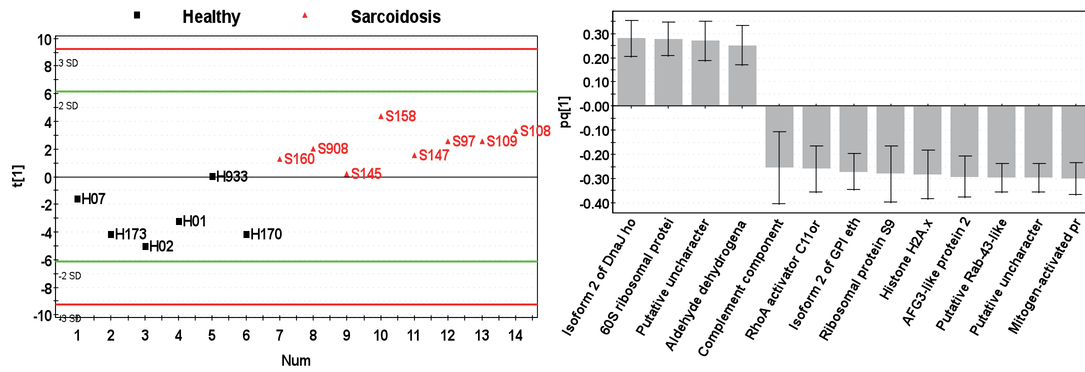

Figure 11a.OPLS Scoring-plot showing group clustering in first component t[1]. X-axis included all subjects (sarcoidosis n=8, and healthy n=6) without outliers (±2SD). Figure 11b. Loading-plot shows the proteins group that discriminate sarcoidosis and healthy. The plot shows the direct correlated variables (bars above zero-level; n=4) and indirect correlated variables (n=9).

2.4 GENERAL CONCLUSIONS

This thesis has focused on searching for new proteins altered in the lung of sarcoidosis patients using proteomics technologies in order to increase the understanding of the pathology behind the disease. Conclusions can be summarized as follow:

- In the technical study aimed to measure source of variance in a 2DE approach we reported improvements in reduced variance introduced to 2DE experiments from background algorithms; the removal of experimental variance when using DIGE approach and overall improvements in the objectivity of the analysis using this 4th generation software analysis software, SameSpots.

- Looking at soluble proteins from the lungs of sarcoidosis and CBD patients we could observe alterations of a set of proteins which in general reflected the activation of the inflammatory process. The up-regulation of complement C3, β2-microglobulin and IgG may indicate a pronounced opsonisation mechanism against microbial pathogens in the innate immune response. In addition, the anti-oxidants peroxiredoxin 5, α1-antitrypsin were also up regulated, reflecting the high levels of reactive oxygen species (ROS) in the lungs of these granulomatous disorders.

By applying complementary proteomic techniques on different protein fractions from isolated AMs, we could identify and quantify a new set of proteins that have not been described previously in sarcoidosis pathology.

- In a DIGE study investigating the soluble proteome of AMs, we identified 25 proteins to be altered in sarcoidosis. When applying pathway analyses we found the Fcgamma receptor-mediated phagocytosis pathway and the clathrin-mediated endocytosis signaling pathway to be up regulated in sarcoidosis. In the complementary shotgun proteomics analyses, investigating the AM membrane associated proteome, we found 80 proteins altered and accordingly the same pathways were identified. These pathways were also up regulated in sarcoidosis compared to healthy controls.

Sarcoidosis can be divided into Löfgren´s syndrome and non-Löfgren´s syndrome patients, having different clinical phenotype and outcome. Patients with Löfgren's often recover spontaneously within 2 years, whereas non-Löfgren's usually having a prolonged unresolved inflammation in the lungs. When comparing these subgroups we found:

- Down-regulation of the oxidative phosphorylation pathway from the mitochondrial respiration system in non-Löfgren's, but up-regulation in Löfgren's patients. This mitochondrial dysfunction could lead to an imbalance in the optimal AM function as APC in non-Löfgren's sarcoidosis patients. In addition, this mitochondrial dysfunction only observed in non-Löfgren's patients could partly be related to the unresolved inflammation and prolonged disease course observed. Furthermore, this imbalance in oxidative homeostasis

could potentially be associated with the increased risk of development of fibrosis observed in this sarcoidosis phenotype.

- We observed in both studies on AMs that the expression pattern in Löfgren's patients was more homogeneous as compared to the non-Löfgren. This could be translated clinically as Löfgren's syndrome is a distinct phenotype presenting with a typical acute onset, whereas non-Löfgren's patients is a more heterogeneous group involving different clinical manifestations and may accordingly involve multiple mechanisms.

Searching for molecular markers we performed multivariate analysis and found a set of proteins in a model able to discriminate sarcoidosis from healthy. A subset of 13 putative biomarker proteins resulted in a highly predictive model (Q2 = 0.72) with an overall model p-value of 0.0009.

- Three of these proteins are involved in endocytic transport; mitogen-activated protein kinase scaffold protein 1, regulator complex protein PDRO and ras related protein Rab-7a. One of the proteins, the complement component 1 Q subcomponent-binding protein, is involved in host-virus interactions [159]. The rest of the 13 proteins are involved in DNA repair, alcohol metabolism, protein folding and ribosomal activity.

Taken together, these studies have identified new proteins potentially involved in the molecular mechanisms of sarcoidosis disease. A limitation in numerous proteomics studies has been the inability to verify the biological relevance of the findings, which may hamper the utility of the results. The similarity of the findings from the complementary proteomics approaches presented in this dissertation may serve as a confirmation of the biological relevance of our results.

3 ACKNOWLEDGEMENTS

I am thankful for all support I got during my PhD and I would like to take this opportunity to acknowledge people around me:

My supervisor Professor **Johan Grunewald** for the guidance and encouragement I got and also for the efforts to connect me with proteomic people outside academia and at Karolinska.

Anders Eklund my big gratitude for having the possibility to work at the lung lab, but specially thanks for always being disposable to help me with any subject during these years.

Serhiy Souchelnytskyi for providing your expertise in proteomics, good scientific advice and all your help in the lab.

My co-supervisors Fariba for introducing me to the proteomics field and for all your help during my time at GE-Healthcare. Åsa for helping me with the writings and support in proteomics.

Co-authors Hanna Eriksson, Rui Mamede Branca, Janne Lehtiö at KBC for your expertise in mass spectrometry and a nice collaboration.

Eva Marie Karlsson for taking care of administration.

Helene Blomqvist, Marghita Dahl and Gunnel De Forest for always having a smile and being kind and helpful.

All the members at the lung research lab. Micke for being a good friend and for nice conversations, Lukas for your special personality, Janne, Carolin, Benita E, Annsofi, Farah, Marija, Bettina, Karin, Cecilia, Ulrika, Abraham, Maria, Mahyar, Charlotta, Luca, Helena, Kie, Benita D, Montasir, Pernilla, for creating a good atmosphere.

My brother Fidel for being a good person. My parents and family in Peru for all their support and love. My polish family for always being generous with me. My finnish family for the nice time in Finland fishing and at the sauna, thanks all. My good friends at the churches for always having an attitude to help others. To Orlando, Betto, Benigna and Yallito for being good examples.

My wife Suvi for all support, understanding and encouragement during these years and to my children Jacob and Paula for the energy and unconditional love I always get.

4 REFERENCES

1. Diakonova, M., Bokoch, G. & Swanson, J.A. Dynamics of cytoskeletal proteins during Fcgamma receptor-mediated phagocytosis in macrophages. *Mol Biol Cell* **13**, 402-11 (2002).
2. Joshi, T., Butchar, J.P. & Tridandapani, S. Fcgamma receptor signaling in phagocytes. *Int J Hematol* **84**, 210-6 (2006).
3. Tjelle, T.E., Lovdal, T. & Berg, T. Phagosome dynamics and function. *Bioessays* **22**, 255-63 (2000).
4. Swanson, J.A. & Hoppe, A.D. The coordination of signaling during Fc receptor-mediated phagocytosis. *J Leukoc Biol* **76**, 1093-103 (2004).
5. Henry, R.M., Hoppe, A.D., Joshi, N. & Swanson, J.A. The uniformity of phagosome maturation in macrophages. *J Cell Biol* **164**, 185-94 (2004).
6. Daeron, M. Fc receptor biology. *Annu Rev Immunol* **15**, 203-34 (1997).
7. Anderson, C.L., Shen, L., Eicher, D.M., Wewers, M.D. & Gill, J.K. Phagocytosis mediated by three distinct Fc gamma receptor classes on human leukocytes. *J Exp Med* **171**, 1333-45 (1990).
8. Ravetch, J.V. & Kinet, J.P. Fc receptors. *Annu Rev Immunol* **9**, 457-92 (1991).
9. Medzhitov, R. & Janeway, C., Jr. Innate immune recognition: mechanisms and pathways. *Immunol Rev* **173**, 89-97 (2000).
10. Gasque, P. Complement: a unique innate immune sensor for danger signals. *Mol Immunol* **41**, 1089-98 (2004).
11. Frank, M.M. & Fries, L.F. The role of complement in inflammation and phagocytosis. *Immunol Today* **12**, 322-6 (1991).
12. Ehlers, M.R. CR3: a general purpose adhesion-recognition receptor essential for innate immunity. *Microbes Infect* **2**, 289-94 (2000).
13. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol* **2**, 675-80 (2001).
14. Kang, J.Y. et al. Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* **31**, 873-84 (2009).
15. Haas, T. et al. The DNA sugar backbone 2' deoxyribose determines toll-like receptor 9 activation. *Immunity* **28**, 315-23 (2008).
16. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124**, 783-801 (2006).
17. Jin, M.S. et al. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**, 1071-82 (2007).
18. Brinkmann, M.M. et al. The interaction between the ER membrane protein UNC93B and TLR3, 7, and 9 is crucial for TLR signaling. *J Cell Biol* **177**, 265-75 (2007).
19. Hoebe, K. et al. CD36 is a sensor of diacylglycerides. *Nature* **433**, 523-7 (2005).
20. Bell, J.K., Askins, J., Hall, P.R., Davies, D.R. & Segal, D.M. The dsRNA binding site of human Toll-like receptor 3. *Proc Natl Acad Sci U S A* **103**, 8792-7 (2006).
21. Jin, M.S. & Lee, J.O. Structures of the toll-like receptor family and its ligand complexes. *Immunity* **29**, 182-91 (2008).
22. Beutler, B. & Rietschel, E.T. Innate immune sensing and its roots: the story of endotoxin. *Nat Rev Immunol* **3**, 169-76 (2003).
23. Pasare, C. & Medzhitov, R. Toll-like receptors: linking innate and adaptive immunity. *Microbes Infect* **6**, 1382-7 (2004).
24. Medzhitov, R. & Janeway, C.A., Jr. Innate immunity: impact on the adaptive immune response. *Curr Opin Immunol* **9**, 4-9 (1997).
25. Ramachandra, L., Simmons, D. & Harding, C.V. MHC molecules and microbial antigen processing in phagosomes. *Curr Opin Immunol* **21**, 98-104 (2009).
26. Agostini, C. et al. Regulation of alveolar macrophage-T cell interactions during Th1-type sarcoid inflammatory process. *Am J Physiol* **277**, L240-50 (1999).

27. Agostini, C. et al. Pulmonary alveolar macrophages in patients with sarcoidosis and hypersensitivity pneumonitis: characterization by monoclonal antibodies. *J Clin Immunol* **7**, 64-70 (1987).
28. Wiken, M. et al. No evidence of altered alveolar macrophage polarization, but reduced expression of TLR2, in bronchoalveolar lavage cells in sarcoidosis. *Respir Res* **11**, 121.
29. Wahlstrom, J. et al. Phenotypic analysis of lymphocytes and monocytes/macrophages in peripheral blood and bronchoalveolar lavage fluid from patients with pulmonary sarcoidosis. *Thorax* **54**, 339-46 (1999).
30. Hampton, M.B., Kettle, A.J. & Winterbourn, C.C. Inside the neutrophil phagosome: oxidants, myeloperoxidase, and bacterial killing. *Blood* **92**, 3007-17 (1998).
31. Forsen, K.O., Milman, N., Pietinalho, A. & Selroos, O. Sarcoidosis in the Nordic countries 1950-1987. *Sarcoidosis* **9**, 140-1 (1992).
32. Milman, N. & Selroos, O. Pulmonary sarcoidosis in the Nordic countries 1950-1982. Epidemiology and clinical picture. *Sarcoidosis* **7**, 50-7 (1990).
33. Pietinalho, A. et al. The frequency of sarcoidosis in Finland and Hokkaido, Japan. A comparative epidemiological study. *Sarcoidosis* **12**, 61-7 (1995).
34. Baughman, R.P. et al. Clinical characteristics of patients in a case control study of sarcoidosis. *Am J Respir Crit Care Med* **164**, 1885-9 (2001).
35. Hillerdal, G., Nou, E., Osterman, K. & Schmekel, B. Sarcoidosis: epidemiology and prognosis. A 15-year European study. *Am Rev Respir Dis* **130**, 29-32 (1984).
36. Bresnitz, E.A., Stolley, P.D., Israel, H.L. & Soper, K. Possible risk factors for sarcoidosis. A case-control study. *Ann N Y Acad Sci* **465**, 632-42 (1986).
37. Edmondstone, W.M. Sarcoidosis in nurses: is there an association? *Thorax* **43**, 342-3 (1988).
38. Wilsher, M.L. Seasonal clustering of sarcoidosis presenting with erythema nodosum. *Eur Respir J* **12**, 1197-9 (1998).
39. Hosoda, Y., Yamaguchi, M. & Hiraga, Y. Global epidemiology of sarcoidosis. What story do prevalence and incidence tell us? *Clin Chest Med* **18**, 681-94 (1997).
40. Ichikawa, H. et al. Quantitative analysis of propionibacterial DNA in bronchoalveolar lavage cells from patients with sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* **25**, 15-20 (2008).
41. Hiramatsu, J. et al. Propionibacterium acnes DNA detected in bronchoalveolar lavage cells from patients with sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* **20**, 197-203 (2003).
42. Yamada, T. et al. In situ localization of Propionibacterium acnes DNA in lymph nodes from sarcoidosis patients by signal amplification with catalysed reporter deposition. *J Pathol* **198**, 541-7 (2002).
43. Ishige, I., Usui, Y., Takemura, T. & Eishi, Y. Quantitative PCR of mycobacterial and propionibacterial DNA in lymph nodes of Japanese patients with sarcoidosis. *Lancet* **354**, 120-3 (1999).
44. Chen, E.S. & Moller, D.R. Etiology of sarcoidosis. *Clin Chest Med* **29**, 365-77, vii (2008).
45. Song, Z. et al. Mycobacterial catalase-peroxidase is a tissue antigen and target of the adaptive immune response in systemic sarcoidosis. *J Exp Med* **201**, 755-67 (2005).
46. Milman, N. & Andersen, A.B. Detection of antibodies in serum against M. tuberculosis using western blot technique. Comparison between sarcoidosis patients and healthy subjects. *Sarcoidosis* **10**, 29-31 (1993).
47. Oswald-Richter, K.A. et al. Multiple mycobacterial antigens are targets of the adaptive immune response in pulmonary sarcoidosis. *Respir Res* **11**, 161.
48. Chen, E.S. et al. T cell responses to mycobacterial catalase-peroxidase profile a pathogenic antigen in systemic sarcoidosis. *J Immunol* **181**, 8784-96 (2008).
49. Gupta, D., Agarwal, R., Aggarwal, A.N. & Jindal, S.K. Molecular evidence for the role of mycobacteria in sarcoidosis: a meta-analysis. *Eur Respir J* **30**, 508-16 (2007).

50. Wahlstrom, J. et al. Autoimmune T cell responses to antigenic peptides presented by bronchoalveolar lavage cell HLA-DR molecules in sarcoidosis. *Clin Immunol* **133**, 353-63 (2009).
51. Schurmann, M. et al. Familial sarcoidosis is linked to the major histocompatibility complex region. *Am J Respir Crit Care Med* **162**, 861-4 (2000).
52. Rossman, M.D. et al. HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. *Am J Hum Genet* **73**, 720-35 (2003).
53. Grunewald, J. & Eklund, A. Lofgren's syndrome: human leukocyte antigen strongly influences the disease course. *Am J Respir Crit Care Med* **179**, 307-12 (2009).
54. Grunewald, J. et al. Different HLA-DRB1 allele distributions in distinct clinical subgroups of sarcoidosis patients. *Respir Res* **11**, 25.
55. Hofmann, S. et al. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. *Nat Genet* (2008).
56. Li, Y., Pabst, S., Kubisch, C., Grohe, C. & Wollnik, B. First independent replication study confirms the strong genetic association of ANXA11 with sarcoidosis. *Thorax* **65**, 939-40.
57. Li, Y. et al. BTNL2 gene variant and sarcoidosis. *Thorax* **61**, 273-4 (2006).
58. Li, Y., Pabst, S., Lokhande, S., Grohe, C. & Wollnik, B. Extended genetic analysis of BTNL2 in sarcoidosis. *Tissue Antigens* **73**, 59-61 (2009).
59. Costabel, U. CD4/CD8 ratios in bronchoalveolar lavage fluid: of value for diagnosing sarcoidosis? *Eur Respir J* **10**, 2699-700 (1997).
60. Agostini, C. et al. Involvement of the IP-10 chemokine in sarcoid granulomatous reactions. *J Immunol* **161**, 6413-20 (1998).
61. Pinkston, P., Bitterman, P.B. & Crystal, R.G. Spontaneous release of interleukin-2 by lung T lymphocytes in active pulmonary sarcoidosis. *N Engl J Med* **308**, 793-800 (1983).
62. Grunewald, J., Eklund, A., Wigzell, H., Van Meijgaarden, K.E. & Ottenhoff, T.H. Bronchoalveolar lavage cells from sarcoidosis patients and healthy controls can efficiently present antigens. *J Intern Med* **245**, 353-7 (1999).
63. Zissel, G., Ernst, M., Schlaak, M. & Muller-Quernheim, J. Accessory function of alveolar macrophages from patients with sarcoidosis and other granulomatous and nongranulomatous lung diseases. *J Investig Med* **45**, 75-86 (1997).
64. Newman, L.S., Rose, C.S. & Maier, L.A. Sarcoidosis. *N Engl J Med* **336**, 1224-34 (1997).
65. Costabel, U. & Hunninghake, G.W. ATS/ERS/WASOG statement on sarcoidosis. Sarcoidosis Statement Committee. American Thoracic Society. European Respiratory Society. World Association for Sarcoidosis and Other Granulomatous Disorders. *Eur Respir J* **14**, 735-7 (1999).
66. Agostini, C., Basso, U. & Semenzato, G. Cells and molecules involved in the development of sarcoid granuloma. *J Clin Immunol* **18**, 184-92 (1998).
67. Agostini, C., Adami, F. & Semenzato, G. New pathogenetic insights into the sarcoid granuloma. *Curr Opin Rheumatol* **12**, 71-6 (2000).
68. Grunewald, J. & Eklund, A. Role of CD4+ T cells in sarcoidosis. *Proc Am Thorac Soc* **4**, 461-4 (2007).
69. Moller, D.R. et al. Enhanced expression of IL-12 associated with Th1 cytokine profiles in active pulmonary sarcoidosis. *J Immunol* **156**, 4952-60 (1996).
70. Shigehara, K. et al. IL-12 and IL-18 are increased and stimulate IFN-gamma production in sarcoid lungs. *J Immunol* **166**, 642-9 (2001).
71. Kunkel, S.L., Chensue, S.W., Strieter, R.M., Lynch, J.P. & Remick, D.G. Cellular and molecular aspects of granulomatous inflammation. *Am J Respir Cell Mol Biol* **1**, 439-47 (1989).
72. Fehrenbach, H. et al. Alveolar macrophages are the main source for tumour necrosis factor-alpha in patients with sarcoidosis. *Eur Respir J* **21**, 421-8 (2003).
73. Berlin, M., Fogdell-Hahn, A., Olerup, O., Eklund, A. & Grunewald, J. HLA-DR predicts the prognosis in Scandinavian patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* **156**, 1601-5 (1997).

74. Hedfors, E. & Lindstrom, F. HLA-B8/DR3 in sarcoidosis. Correlation to acute onset disease with arthritis. *Tissue Antigens* **22**, 200-3 (1983).
75. Wiken, M., Grunewald, J., Eklund, A. & Wahlstrom, J. Higher monocyte expression of TLR2 and TLR4, and enhanced pro-inflammatory synergy of TLR2 with NOD2 stimulation in sarcoidosis. *J Clin Immunol* **29**, 78-89 (2009).
76. Mana, J. et al. Lofgren's syndrome revisited: a study of 186 patients. *Am J Med* **107**, 240-5 (1999).
77. Voorter, C.E., Drent, M. & van den Berg-Loonen, E.M. Severe pulmonary sarcoidosis is strongly associated with the haplotype HLA-DQB1*0602-DRB1*150101. *Hum Immunol* **66**, 826-35 (2005).
78. Rabilloud, T.e.a., Chevallet, M., Luche, S. & Lelong, C. Two-dimensional gel electrophoresis in proteomics: Past, present and future. *J Proteomics* **73**, 2064-77 (2010).
79. Gorg, A. et al. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**, 1037-53 (2000).
80. Gorg, A., Drews, O., Luck, C., Weiland, F. & Weiss, W. 2-DE with IPGs. *Electrophoresis* **30 Suppl 1**, S122-32 (2009).
81. Karp, N.A., Feret, R., Rubtsov, D.V. & Lilley, K.S. Comparison of DIGE and post-stained gel electrophoresis with both traditional and SameSpots analysis for quantitative proteomics. *Proteomics* **8**, 948-60 (2008).
82. Karp, N.A. & Lilley, K.S. Maximising sensitivity for detecting changes in protein expression: Experimental design using minimal CyDyes. *Proteomics.* **5**, 3105-3115 (2005).
83. Lilley, K.S. & Friedman, D.B. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev Proteomics* **1**, 401-9 (2004).
84. Marouga, R., David, S. & Hawkins, E. The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem* **382**, 669-78 (2005).
85. Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **60**, 2299-301 (1988).
86. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. & Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71 (1989).
87. Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5**, 699-711 (2004).
88. Duncan, M.W., Aebersold, R. & Caprioli, R.M. The pros and cons of peptide-centric proteomics. *Nat Biotechnol* **28**, 659-64 (2010).
89. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212-7 (2006).
90. Yates, J.R., Ruse, C.I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* **11**, 49-79 (2009).
91. Ross, P.L. et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **3**, 1154-69 (2004).
92. Pearson, K. On lines and planes of closest fit to system of points in space. *Philosophical Magazine* **6**, 559-572 (1901).
93. Wold, S., Esbensen, K. & Geladi, P. Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems* **2**, 37-52 (1987).
94. Wold, H. Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, 391-420 (1966).
95. Trygg, J. & Wold, S. Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics* **16**, 119-128 (2002).
96. Grunewald, J. & Eklund, A. Sex-specific manifestations of Lofgren's syndrome. *Am J Respir Crit Care Med* **175**, 40-4 (2007).
97. Newman, L.S. Significance of the blood beryllium lymphocyte proliferation test. *Environ Health Perspect* **104 Suppl 5**, 953-6 (1996).
98. Williams, W.J. Diagnostic criteria for chronic beryllium disease (CBD) based on the UK registry 1945-1991. *Sarcoidosis* **10**, 41-3 (1993).

99. Eklund, A. & Blaschke, E. Relationship between changed alveolar-capillary permeability and angiotensin converting enzyme activity in serum in sarcoidosis. *Thorax* **41**, 629-34 (1986).
100. Alban, A. et al. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* **3**, 36-44 (2003).
101. Perkins, D.N., Pappin, D.J., Creasy, D.M. & Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-67 (1999).
102. Pernemalm, M. et al. Use of narrow-range peptide IEF to improve detection of lung adenocarcinoma markers in plasma and pleural effusion. *Proteomics* **9**, 3414-24 (2009).
103. Johan Grunewald, A.E. Det är svårt att skilja sarkoidos från beryllios. *Läkartidningen* **97**, 5915-18 (2000).
104. Kinnula, V.L. et al. Cell specific expression of peroxiredoxins in human lung and pulmonary sarcoidosis. *Thorax* **57**, 157-64 (2002).
105. Sawyer, R.T. et al. Beryllium-stimulated reactive oxygen species and macrophage apoptosis. *Free Radic Biol Med* **38**, 928-37 (2005).
106. Comhair, S.A., Lewis, M.J., Bhathena, P.R., Hammel, J.P. & Erzurum, S.C. Increased glutathione and glutathione peroxidase in lungs of individuals with chronic beryllium disease. *Am J Respir Crit Care Med* **159**, 1824-9 (1999).
107. MacNee, W. Oxidative stress and lung inflammation in airways disease. *Eur J Pharmacol* **429**, 195-207 (2001).
108. Jeffery, P.K. Comparison of the structural and inflammatory features of COPD and asthma. Giles F. Filley Lecture. *Chest* **117**, 251S-60S (2000).
109. Young, R.C., Jr. et al. Alpha 1 antitrypsin levels in sarcoidosis: relationship to disease activity. *Chest* **64**, 39-45 (1973).
110. Sabounchi-Schutt, F., Mikko, M., Eklund, A., Grunewald, J. & J, A.S. Serum protein pattern in sarcoidosis analysed by a proteomics approach. *Sarcoidosis Vasc Diffuse Lung Dis* **21**, 182-90 (2004).
111. Yoo, C.G. et al. Anti-inflammatory effect of heat shock protein induction is related to stabilization of I kappa B alpha through preventing I kappa B kinase activation in respiratory epithelial cells. *J Immunol* **164**, 5416-23 (2000).
112. Pittet, J.F., Mackersie, R.C., Martin, T.R. & Matthay, M.A. Biological markers of acute lung injury: prognostic and pathogenetic significance. *Am J Respir Crit Care Med* **155**, 1187-205 (1997).
113. Hagiwara, S. et al. A case of renal sarcoidosis with complement activation via the lectin pathway. *Am J Kidney Dis* **45**, 580-7 (2005).
114. Rosell, F.I., Mauk, M.R. & Mauk, A.G. pH- and metal ion-linked stability of the hemopexin-heme complex. *Biochemistry* **44**, 1872-9 (2005).
115. Bakker, W.W. et al. Protease activity of plasma hemopexin. *Kidney Int* **68**, 603-10 (2005).
116. Bowler, R.P. et al. Furin proteolytically processes the heparin-binding region of extracellular superoxide dismutase. *J Biol Chem* **277**, 16505-11 (2002).
117. Lee, Y.H., Tan, H.T. & Chung, M.C. Subcellular fractionation methods and strategies for proteomics. *Proteomics* **10**, 3935-56.
118. Bjorhall, K., Miliotis, T. & Davidsson, P. Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. *Proteomics* **5**, 307-17 (2005).
119. Echan, L.A., Tang, H.Y., Ali-Khan, N., Lee, K. & Speicher, D.W. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* **5**, 3292-303 (2005).
120. Van den Bergh, G. & Arckens, L. Fluorescent two-dimensional difference gel electrophoresis unveils the potential of gel-based proteomics. *Curr Opin Biotechnol* **15**, 38-43 (2004).
121. Gorg, A., Weiss, W. & Dunn, M.J. Current two-dimensional electrophoresis technology for proteomics. *Proteomics.* **4**, 3665-3685 (2004).
122. Unlu, M., Morgan, M.E. & Minden, J.S. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071-2077 (1997).

123. Chakravarti, B., Gallagher, S.R. & Chakravarti, D.N. Difference gel electrophoresis (DIGE) using CyDye DIGE fluor minimal dyes. *Curr Protoc Mol Biol* **Chapter 10**, Unit 10 23 (2005).
124. Wheelock, A.M. & Goto, S. Effects of post-electrophoretic analysis on variance in gel-based proteomics. *Expert Rev.Proteomics.* **3**, 129-142 (2006).
125. Dowsey, A.W., Dunn, M.J. & Yang, G.Z. The role of bioinformatics in two-dimensional gel electrophoresis. *Proteomics.* **3**, 1567-1596 (2003).
126. Marengo, E. et al. Numerical approaches for quantitative analysis of two-dimensional maps: a review of commercial software and home-made systems. *Proteomics.* **5**, 654-666 (2005).
127. Wheelock, A.M. & Buckpitt, A.R. Software-induced variance in two-dimensional gel electrophoresis image analysis. *Electrophoresis* **26**, 4508-4520 (2005).
128. Wheelock, A.M. & Wheelock, C.E. Bioinformatics in gel-based proteomics. in *Plant Proteomics: Technologies, Strategies and Applications* (eds. Rakwal, R. & Agrawal, G.) 107-125 (John Wiley & Sons Inc., Hoboken, NJ, USA, 2008).
129. Silva, E. et al. In the eye of the beholder: does the master see the SameSpots as the novice? *J Proteome Res* **9**, 1522-32.
130. Sabounchi-Schutt, F., Astrom, J., Hellman, U., Eklund, A. & Grunewald, J. Changes in bronchoalveolar lavage fluid proteins in sarcoidosis: a proteomics approach. *Eur Respir J* **21**, 414-20 (2003).
131. Rottoli, P. et al. Cytokine profile and proteome analysis in bronchoalveolar lavage of patients with sarcoidosis, pulmonary fibrosis associated with systemic sclerosis and idiopathic pulmonary fibrosis. *Proteomics* **5**, 1423-30 (2005).
132. Kriegova, E. et al. Protein profiles of bronchoalveolar lavage fluid from patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* **173**, 1145-54 (2006).
133. Silva, E. et al. A quantitative proteomic analysis of soluble bronchoalveolar fluid proteins from patients with sarcoidosis and chronic beryllium disease. *Sarcoidosis Vasc Diffuse Lung Dis* **24**, 24-32 (2007).
134. May, R.C. & Machesky, L.M. Phagocytosis and the actin cytoskeleton. *J Cell Sci* **114**, 1061-77 (2001).
135. Castellano, F., Chavrier, P. & Caron, E. Actin dynamics during phagocytosis. *Semin Immunol* **13**, 347-55 (2001).
136. Chimini, G. & Chavrier, P. Function of Rho family proteins in actin dynamics during phagocytosis and engulfment. *Nat Cell Biol* **2**, E191-6 (2000).
137. Caron, E. & Hall, A. Identification of two distinct mechanisms of phagocytosis controlled by different Rho GTPases. *Science* **282**, 1717-21 (1998).
138. Yang, Z., Vadlamudi, R.K. & Kumar, R. Dynein light chain 1 phosphorylation controls macropinocytosis. *J Biol Chem* **280**, 654-9 (2005).
139. Maric, M. et al. Defective antigen processing in GILT-free mice. *Science* **294**, 1361-5 (2001).
140. Hastings, K.T., Lackman, R.L. & Cresswell, P. Functional requirements for the lysosomal thiol reductase GILT in MHC class II-restricted antigen processing. *J Immunol* **177**, 8569-77 (2006).
141. Dugast, M., Toussaint, H., Dousset, C. & Benaroch, P. AP2 clathrin adaptor complex, but not AP1, controls the access of the major histocompatibility complex (MHC) class II to endosomes. *J Biol Chem* **280**, 19656-64 (2005).
142. Stenmark, H. & Olkkonen, V.M. The Rab GTPase family. *Genome Biol* **2**, REVIEWS3007 (2001).
143. Wilcke, M. et al. Rab11 regulates the compartmentalization of early endosomes required for efficient transport from early endosomes to the trans-golgi network. *J Cell Biol* **151**, 1207-20 (2000).
144. Strzelecka, A., Kwiatkowska, K. & Sobota, A. Tyrosine phosphorylation and Fcgamma receptor-mediated phagocytosis. *FEBS Lett* **400**, 11-4 (1997).
145. Kondo, H. et al. p47 is a cofactor for p97-mediated membrane fusion. *Nature* **388**, 75-8 (1997).
146. Dai, R.M., Chen, E., Longo, D.L., Gorbea, C.M. & Li, C.C. Involvement of valosin-containing protein, an ATPase Co-purified with IkappaBalpha and 26 S proteasome, in ubiquitin-proteasome-mediated degradation of IkappaBalpha. *J Biol Chem* **273**, 3562-73 (1998).

147. Asai, T. et al. VCP (p97) regulates NFkappaB signaling pathway, which is important for metastasis of osteosarcoma cell line. *Jpn J Cancer Res* **93**, 296-304 (2002).
148. Dai, R.M. & Li, C.C. Valosin-containing protein is a multi-ubiquitin chain-targeting factor required in ubiquitin-proteasome degradation. *Nat Cell Biol* **3**, 740-4 (2001).
149. Yeung, H.O. et al. Insights into adaptor binding to the AAA protein p97. *Biochem Soc Trans* **36**, 62-7 (2008).
150. Bays, N.W., Wilhovsky, S.K., Goradia, A., Hodgkiss-Harlow, K. & Hampton, R.Y. HRD4/NPL4 is required for the proteasomal processing of ubiquitinated ER proteins. *Mol Biol Cell* **12**, 4114-28 (2001).
151. Davies, J.M., Brunger, A.T. & Weis, W.I. Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure* **16**, 715-26 (2008).
152. Repine, J.E., Bast, A. & Lankhorst, I. Oxidative stress in chronic obstructive pulmonary disease. Oxidative Stress Study Group. *Am J Respir Crit Care Med* **156**, 341-57 (1997).
153. Henricks, P.A. & Nijkamp, F.P. Reactive oxygen species as mediators in asthma. *Pulm Pharmacol Ther* **14**, 409-20 (2001).
154. Koutsokera, A. et al. Systemic oxidative stress in patients with pulmonary sarcoidosis. *Pulm Pharmacol Ther* **22**, 603-7 (2009).
155. Riedl, M.A. & Nel, A.E. Importance of oxidative stress in the pathogenesis and treatment of asthma. *Curr Opin Allergy Clin Immunol* **8**, 49-56 (2008).
156. Bulteau, A.L., Szweda, L.I. & Friguet, B. Mitochondrial protein oxidation and degradation in response to oxidative stress and aging. *Exp Gerontol* **41**, 653-7 (2006).
157. McCormick, P.J., Martina, J.A. & Bonifacino, J.S. Involvement of clathrin and AP-2 in the trafficking of MHC class II molecules to antigen-processing compartments. *Proc Natl Acad Sci U S A* **102**, 7910-5 (2005).
158. Walseng, E., Bakke, O. & Roche, P.A. Major histocompatibility complex class II-peptide complexes internalize using a clathrin- and dynamin-independent endocytosis pathway. *J Biol Chem* **283**, 14717-27 (2008).
159. Beatch, M.D. & Hobman, T.C. Rubella virus capsid associates with host cell protein p32 and localizes to mitochondria. *J Virol* **74**, 5569-76 (2000).