

Identification of Etanercept Responsive Biomarkers in Rheumatoid Arthritis





Title: Identification of Etanercept Responsive Biomarkers in Rheumatoid Arthritis

Field of study: Medicine with Industrial Specialization, Biomedicine

Semester: 9th and 10th semester

Project period: Summer exam 2020

Supervisors: Allan Stensballe, Asta Linauskas

Laboratory Technician: Ditte Bech Laursen

ECTS: 60

A handwritten signature in black ink, which appears to read "Jacob".

20154240 - Jacob Skallerup Andersen

Number printed: 43

Pages: 36

Keystrokes: 84688 (from introduction up to and including conclusion and future perspectives)

Appendix: 4 pages. Appendix 1: Complete list of significantly regulated proteins, Appendix 2: Pearson correlation

Preface

Thanks to my supervisor Allan Stensballe and co-supervisor Asta Linauskas for providing guidance during the making of this project. Thanks to the laboratory technician Ditte Bech Laursen for help in setting up all experiments. Thanks to Aalborg University for providing the resources, facilities, and the opportunities to develop my skills in the laboratory through the making of this project.

Abbreviations

ACPA - Anti-citrullinated protein antibodies

ACR - American College of Rheumatology

APC - Antigen-presenting

APOM – Apolipoprotein M

BCA - Bicinchoninic acid

BSA - Bovine serum albumin

CAA - Chloroacetamide

CD - Cluster of differentiation

cf – Cell free

CRP - C-reactive protein

csDMARD - Conventional synthetic disease-modifying antirheumatic drug

bDMARD - Biologic disease-modifying antirheumatic drug

EDTA - Ethylenediaminetetraacetic acid

ESR - Erythrocyte sedimentation rate

EULAR - European League Against Rheumatism

Fab - Antigen-binding fragment

FASP – Filter aided sample preparation

Fc - Fragment crystallizable

GM-CSF - Granulocyte-macrophage colony-stimulating factor

HLA - Human leukocyte antigen

HST - Health Science and Technology

IFN - Interferon

Ig - Immunoglobulin

IL - Interleukin

iRT - Indexed retention time

JAK - Janus kinase

LC - Liquid chromatography

LFQ - Label-free quantification

LT- α - Lymphotoxin- α

M-CSF - Macrophage colony-stimulating factor

MAC - Membrane attack complex

MHC - Major histocompatibility complex

MMP - Matrix metalloproteinase

NETs - Neutrophil extracellular traps

m – mass

MMP - Matrix metalloproteinase

NETs - Neutrophil extracellular traps

m – mass

MRI - Magnetic resonance imaging

MS - Mass spectrometry

NF- κ B - Nuclear factor kappa-light-chain-enhancer of activated B cells

PAD - Peptidyl-arginine-deiminase

PASEF - Parallel accumulation-serial fragmentation

PCA - principal component analysis

PEA - Proximity Extension Assay

PTM - Posttranslational modification

PZP - Pregnancy zone protein

qPCR - Quantitative polymerase chain reaction

QToF - Quadrupole time of flight

RA – Rheumatoid arthritis

RANK - Receptor activator of nuclear factor kappa-

RANKL - Receptor activator of nuclear factor

kappa-B ligand

RBP4 - Retinol-binding protein 4

RF - Rheumatoid factor

SDC - Sodium deoxycholate

SLRP - Small leucine-rich repeat proteoglycans

sTNFR – Soluble tumor necrosis factor receptor

TCEP - tris-(2- carboxyethyl)-phosphine

TCR - T-cell receptor

TEAB - triethylammonium bicarbonate

TFA - trifluoroacetic acid

TIMS - trapped ion mobility spectrometry

TNF – Tumor necrosis factor

TNFR - Tumor necrosis factor receptor

ToF - Time of flight

Tris-HCl - Tris-hydrochloride

tsDMARD - targeted synthetic disease-modifying antirheumatic drug

VDBP - Vitamin D-binding protein

z - charge

Abstract

Introduction: Rheumatoid arthritis (RA) causes inflammation of the joints which if left untreated or inadequately treated may result in joint- and bone destruction and facilitate development of systemic co-morbidities such as cardiovascular diseases and inflammatory lung diseases. The introduction of bDMARDs have revolutionized the treatment options for RA. One of these bDMARDs is etanercept, a fusion protein made by genetic recombination. Etanercept works as a decoy receptor for TNF- α resulting in decreased cellular signaling from TNF- α binding to its target receptor. However, about 30% to 40% of RA patients do not achieve remission or low disease activity in response to treatment with anti-TNF- α treatment.

Aim: The aim of this study was to investigate the effects of long-term treatment of RA with etanercept and to identify biomarkers to monitor treatment response in these patients.

Methods: Serum samples collected from 12 etanercept-treated RA patients before treatment, and after three, six, and 12 months of treatment, respectively were analyzed proteomically by LC-MS/MS and analyzed by quantification of cfDNA levels.

Results: 79 proteins were found to have a significantly altered abundance as a result of etanercept treatment either for three, six or 12 months. Out of these, five potential etanercept responsive biomarkers were identified. Furthermore, a continuous, significant fall in cfDNA concentration was observed in the serum samples.

Conclusion:

Five candidate biomarkers of etanercept treatment response in RA patients were identified. Especially vitamin D-binding protein, pregnancy zone protein, and plasma kallikrein were found to be interesting potential biomarkers. Furthermore, lowered cfDNA concentrations may have the ability to be used as a biomarker of RA remission. Future studies are needed to expand the knowledge of the potential biomarkers and their relation to inflammation in RA.

Resumé

Introduktion: Reumatoid artrit (RA) eller leddegigt forårsager inflammation i kroppens led, og hvis sygdommen forbliver ubehandlet eller utilstrækkeligt behandlet kan der opstå irreversibel ledskaide samt komorbiditeter såsom kardiovaskulære sygdomme eller inflammatoriske lungesygdom. Introduktionen af bDMARDs har revolutioneret behandlingsmulighederne for patienter med RA. En af disse bDMARDs er etanercept; et fusionsprotein, der er skabt ved genetisk rekombination. Etanercept udøver sine effekter ved at inhibere binding af TNF til sin tilsigtede receptor, hvilket resulterer i nedsat cellulær respons. Dog opnår 30% til 40% af RA-patienter ikke tilstrækkelig remission eller lav sygdomsaktivitet som resultat af anti-TNF-behandling.

Formål: Formålet med dette studie var at undersøge effekterne af langtidsbehandling af RA med etanercept, og at identificere biomarkører, der kan benyttes til at indikere behandlingsrespons i disse patienter.

Metoder: Serumprøver indsamlet fra 12 etanercept-behandlede RA-patienter før behandling, samt efter henholdsvis tre, seks og 12 måneders behandling blev analyseret for ændringer i deres proteom med LC-MS/MS. Herudover blev koncentrationen af cfDNA målt i prøverne.

Resultater: 79 proteiner blev fundet ændret som resultat af behandlingen, enten efter tre, seks eller 12 måneders behandling. Ud af disse blev fem potentielle biomarkører for etanercept-respons identificeret. Ydermere identificerede vi et kontinuerligt, signifikant fald i cfDNA koncentrationen i prøverne

Konklusion: Fem potentielle biomarkører for etanercept-respons i RA patienter blev identificeret. Især vitamin D-binding protein, pregnancy zone protein, og plasma kallikrein var interessante som mulige biomarkører. Herudover er det muligt at et fald i cfDNA-koncentration kan benyttes som biomarkør for opnåelse af remission i RA-patienter. Yderligere studier er nødvendige for at få mere viden om de potentielle biomarkører og deres relation til inflammation i RA.

Table of contents

Introduction	1
<i>Etiology of rheumatoid arthritis</i>	1
<i>Pathophysiology of rheumatoid arthritis</i>	2
Triggering stage.....	3
Targeting stage.....	4
Chronic inflammatory stage and joint damage.....	4
<i>Diagnosis of rheumatoid arthritis</i>	6
<i>Treatment of rheumatoid arthritis</i>	7
Etanercept.....	9
Aim and strategy	11
<i>Aim</i>	11
<i>Strategy</i>	11
Methods	14
<i>Ethics, patients included, sample collection and biobanking</i>	14
<i>Protein quantification by a bicinchonic acid assay</i>	14
<i>Protein digestion by filter-aided sample preparation</i>	15
<i>Fractioning of samples</i>	15
<i>Loading of samples prior to mass spectrometry</i>	16
<i>Data acquisition by LC-MS/MS</i>	16
<i>cfDNA quantification</i>	16
Results	18
<i>Proteomic analysis of serum</i>	18
Principal component analysis.....	19
Analysis of changes in protein expression.....	21
Functional enrichment analysis.....	23
<i>Serum concentration of cfDNA</i>	25
Discussion	27
<i>Proteomic investigation of the alterations caused by etanercept</i>	27
<i>Possible protein biomarkers of etanercept response</i>	28
Apolipoprotein M.....	28
Retinol-binding protein 4.....	28
Vitamin D-binding protein.....	29
Pregnancy zone protein.....	29
Plasma kallikrein.....	30
<i>Lowered cfDNA concentrations indicates successful treatment response</i>	31
<i>The need for inflammatory profiling</i>	32
Conclusion and future perspectives	33
References	34

Introduction

Rheumatoid arthritis (RA) is a systemic autoimmune disease which affects around 0.5% to 1% of the population (1,2). This makes RA one of the most commonly occurring autoimmune diseases (1). RA occurs due to an interplay between genetic and environmental factors (3). In the majority of RA patients this autoimmune response is mediated by the autoantibodies rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPAs) (4). RF has affinity for the fragment crystallizable (Fc) region of immunoglobulin (Ig) G antibodies, and ACPAs, as the name indicates target citrullinated proteins (4). RA manifests by inflammation in the joints causing arthritis (5). A process that often starts as inflammation of the small joints of fingers and feet and as the disease progresses often spreads to larger joints such as shoulder-, knee- and ankle joints resulting in polyarthritis (5). Eventually, RA progression may result in irreversible joint damage and systemic consequences to several vital systems of the body such as heart, lungs, brain, bones, and skeletal muscles by causing diseases and conditions such as cardiovascular disease, inflammatory lung disease, fatigue and cognitive deficiencies, osteoporosis, and sarcopenia (2). Although no curative treatment for RA exists, these long-term consequences may be hindered by early diagnosis and treatment (6). Current treatment is focused on relieving pain, halting the progression of the disease, and maximizing joint function (6). Within the last few decades, the approval of biologic disease-modifying antirheumatic drugs (bDMARDs) such as etanercept has revolutionized the treatment of RA and has improved the prognosis for patients that did not achieve a satisfying treatment response with conventional treatment options (7,8). However, 30% to 40% of patients treated with bDMARDs experience an inadequate response to treatment (7). Therefore, there is a need to expand the knowledge of how these drugs affect inflammatory processes and identify biomarkers to monitor or predict the treatment response of a patient.

Etiology of rheumatoid arthritis

The mechanisms of RA development are not completely understood but is known to be the result of an interplay between genetic and environmental factors (3,9).

Twin studies have shown that RA has a heritability of ~60%, meaning that genetics makes up for 60% of the variations making an individual more liable to develop RA in comparison to another individual (10). More than 100 different genetic loci have been associated with development of RA in genome-wide association studies (3). In particular, variations to the major histocompatibility complex (MHC) region on chromosome 6 has been observed to be pose a distinctive risk of RA development, and accounts for approximately 30% to 50% of genetic susceptibility to RA development (3). More specifically, having some specific alleles of the human leukocyte antigen (HLA)-DRB1 is the major genetic risk factor associated to RA development (3). The HLA-DRB1 gene encodes the protein HLA-DR β 1 which is the most prevalent β -subunit of HLA-DR, an MHC class II cell surface receptor (3). One theory of the mechanism behind specific HLA-DRB1 alleles resulting in increased risk of RA development is the so-called *shared epitope hypothesis*. This hypothesis contributes the increased risk of RA to a sequence of shared amino acids found in the protein product encoded the by some HLA-DRB1 alleles (3). This sequence consists of the amino acids DEERA and is found from amino acid 70 to 74 (3). It has been found that around 80% of RA positive patients carry at least one of the shared epitope alleles of HLA-DRB1 (11,12). Furthermore, a study has shown that the

transcription levels of the HLA-DRB1 gene in RA patients is highly dependent of which allele a person has. Individuals with two or more of the risk HLA-DRB1 alleles showed upregulated HLA-DRB1 transcript levels compared to RA patients with only one of no risk alleles (11).

Another genetic risk factor associated with RA susceptibility is a polymorphism of the PTPN22 gene (9,12). The PTPN22 loci is related to T- and B-cell activation and the effectiveness of their receptors (9). This polymorphism has been found to lower the threshold for activation of these cells, increasing the likelihood of developing autoimmune diseases (9,13). Furthermore, this polymorphism is thought play a role in interactions causing hypercitrullination of proteins, an important mechanism in APCA-related RA, that will be described in more detail in *Pathophysiology of rheumatoid arthritis* (9).

Cross-trait studies have identified that some genes associated with RA is likely to have pleiotropic effects, meaning that single genes are associated with several, unrelated phenotypic traits (3). These genes may result in RA patients experiencing co-development of allergies, cancer, or other autoimmune diseases such as coeliac disease (3).

Environmental factors also play a major role in the risk of developing RA (3,9). As an example, exposure to noxious agents such as dust and particularly cigarette smoke has been found to increase the risk of RA development (2,14). Smoking is estimated to account for 20% to 30% of all environmental risk for developing RA. Studies have shown that smoking is especially related to RA development of patients that are ACPA positive and carry high risk HLA-DRB1 alleles (9). Furthermore, smoking has been related to presence of RF even if the individual has not developed RA (9). Studies hypothesize that smoking result in increased citrullination of proteins which may lead to an immune response resulting in production of ACPAs (9). Other risk factors include obesity, and being female; two thirds of all RA patients are female which may be caused by female-specific conditions such as early menopause, pre-eclampsia, and polycystic ovary syndrome (9).

Pathophysiology of rheumatoid arthritis

RA pathology can be divided in two based on the presence or non-presence of ACPAs (2). ACPAs are antibodies against citrullinated proteins such as fibrin, fibronectin and vimentin (2,15). Citrullination is a process by which the amino acid arginine in a protein is changed to a citrulline by a post-translational modification (PTM) process catalyzed by the enzymes peptidyl-arginine-deiminases (PADs) (2,15). ACPAs can be detected in approximately two third of RA patients and their presence have been related to a more complex and aggressive disease picture (2). These autoantibodies are regularly the marker used for clinical diagnosis of RA because of their high specificity in that 97% of all patients that are tested to be ACPA-positive do in fact suffer from RA (2).

ACPA-negative RA differentiates from ACPA-positive RA in that different genetic associations with disease development have been observed, and also treatment with conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) such as methotrexate and bDMARDs such as rituximab has been proven less effective (2). Because the majority of RA cases involve ACPA-positive patients, and because this study is investigating treatment with a bDMARD, this review of RA pathophysiology will focus on ACPA positive RA (4). Pathophysiology of ACPA positive RA can be divided into three stages; the triggering stage, the targeting stage, and the chronic inflammatory stage (2,4).

Triggering stage

ACPA-positive RA is triggered by a combination of hypercitrullination of proteins combined with an immune response targeted towards these proteins (3,13). Citrullination is an, as far as it is known, irreversible process by which a positively charged arginine is replaced by a neutral citrulline (2,15). This makes the protein less charged which can cause the protein to become more hydrophobic, alter its structure, and affect its molecular interactions. Often these alterations will result in loss of function for the protein but in some cases function will remain unchanged or even increased (15). However, citrullination is a normal physiological process that contributes to the physiological function of some proteins, and the mechanism behind citrullination causing an autoimmune response in some individuals is not fully understood (15). The reason for this autoimmunity is thought to be a combination hypercitrullination as a result of overactivity of the PAD enzymes and an individual liability to produce an autoimmune response against these caused by the genetic factors leading to RA that were described in *Etiology of rheumatoid arthritis* (15).

Proteomic investigations have identified more than 100 different citrullinated proteins in the synovial fluid of RA patients, however, so far, ACPAs have been only been confirmed to target a few of these proteins e.g. vimentin and fibrinogen (15). In RA patients, several factors are thought to result in hypercitrullination (9). Smoking is one of the factors able to facilitate hypercitrullination (9,13). Smoking has been shown to result in upregulation of PAD2 in lung tissue which result in hypercitrullination of proteins in the lungs (13). Furthermore, some microorganisms have the ability to induce citrullination of proteins (9,13). One of these is *Porphyromonas gingivalis* which is a bacterium that can cause periodontal disease (9). This bacterium produces a bacterial PAD called PPAD that has the ability to citrullinate human or bacterial proteins (9,13). However, the citrullination done by PPAD only occurs at the C-terminal arginine while human PADs are able to citrullinate arginine within the proteins (13). This result in citrullination that differs significantly from self-citrullinated antigens and therefore some doubt exist of the degree to which citrullination done by PPADs is able to trigger an autoimmune response (13). Another microorganism that has been shown to be capable of triggering citrullination is *Aggregatibacter actinomycetemcomitans* that is also a bacterium related to periodontal disease (9,13). This bacterium releases a pore forming toxin called leukotoxin A that binds cluster of differentiation (CD)18 on neutrophils which facilitates the influx of calcium to the cells resulting in activation of calcium dependent PADs(13). Therefore, this bacterium induces citrullination using the body's own PAD enzymes, which causes the citrullinated sites to mimic the citrullination done by PADs in physiological conditions better than the citrullinations caused by bacterial PADs (13).

Under normal circumstances, the body will not produce autoantibodies against citrullinated proteins (13). However, as a result of the genetic factors predisposing to RA development, tolerance for these proteins may be broken, resulting in production of ACPAs and autoimmunity (13). As described in *Etiology of rheumatoid arthritis*, the alleles of HLA-DRB1 encoding proteins carrying the shared epitope amino acid sequence DEERA is the major genetic predisposition for RA (9,13). This amino-acid sequence is present in the antigen-binding probe of the proteins and it has been shown that this sequence has a greater affinity for citrullinated proteins than versions of HLA-DRβ1 transcribed from non-shared epitope alleles (13). This is hypothesized to result in an increased amount of complexes of HLAs and citrullinated antigens on the surface of antigen-presenting cells (APCs) which lead to a

T-cell response against these proteins (13). Once APCs have engulfed and digested the proteins, they present the neoantigens on their MHC class II molecule and are able to activate naïve CD4+ T-cells with a T-cell receptor (TCR) that has affinity for that specific antigen (2,14). In turn, these activated CD4+ T-cells may migrate to nearby lymph tissue and activate B-cells resulting in secretion of antibodies against the citrullinated proteins: ACPAs (2,14).

Targeting stage

This point of RA pathogenesis is early and individuals may have ACPAs circulating in their bloodstream for years without developing RA symptoms, and some people may never develop RA (2,14). The autoantibodies do not affect the joints at this point as the disease originated from proteins found in the triggering sites such as lungs and gums (14). The mechanisms causing the disease to target the joints are not fully understood, but one of the major theories is that ACPAs are able to induce changes to osteoclast activity (2,14).

Studies have found that ACPAs are able to promote osteoclast differentiation and activation resulting in bone resorption (2,14). This differentiation and activation of osteoclasts as a result of ACPAs occurs through two mechanisms 1) because osteoclasts show affinity of the antigen-binding fragment (Fab) of ACPAs, and 2) because osteoclasts carry Fc receptors on their surface (14). Furthermore, PADs have been found to play a major role in osteoclast differentiation (14). The mechanism behind this is not understood but inhibition of PAD function will result in osteoclast precursor cells not differentiating into mature osteoclasts (14). Studies suggest that osteoclasts and their precursor cells are the cell type with the highest density of citrullinated antigens present on their surface as a result of the importance of PADs for differentiation of the cells (14). Thus, ACPAs can induce activation and differentiation of osteoclasts by specific binding of citrullinated proteins with their Fab part (14). When osteoclasts are exposed to ACPAs, they will secrete interleukin (IL-)8 which exerts autocrine functions on the osteoclasts and induce their activation (14). IL-8 also plays a role stimulation of nociceptive nerve-endings and therefore ACPA-activation of osteoclasts may cause some of the pain related to RA (14).

Chronic inflammatory stage and joint damage

The presence of ACPAs can cause joint inflammation in variety of ways and through a lot of mechanisms (2). These different activation pathways and mechanisms of the immune system can affect each other and can take place at the same time or at different times (2). Therefore, this walkthrough of immune system activities in RA is not necessarily chronologic.

At some point in the disease, activated T-cells and APCAs will enter the joint space (2). Due to their affinity and reactivity to citrullinated proteins, the T-cells will secrete cytokines such as interferon (IFN)- γ and IL-17 which will result in recruitment and activation of macrophages and neutrophils to the joint space (2,16). The macrophages will secrete proinflammatory cytokines such as IL-1, IL-6 and TNF- α which results in activation of fibroblast-like synoviocytes which are cells found in the synovium of the joint (2,14). As a result of the microenvironment caused by RA, the activated fibroblast-like synoviocytes will aggressively proliferate causing the synovium to thicken and form what is termed the pannus (17). Fibroblast-like synoviocytes secretes IL-6 and chemoattractants which further enhances inflammation, as well as granulocyte-macrophage colony-stimulating factor (GM-

CSF) and macrophage colony-stimulating factor (M-CSF) (14,17). GM-CSF and M-CSF has the ability to cause macrophage progenitor cells recruited to the joint to differentiate into osteoclasts increasing the potential bone resorption in the joint (17). Furthermore, fibroblast-like synoviocytes secrete large amounts of receptor activator of nuclear factor kappa-B ligand (RANKL) which interact with receptor activator of nuclear factor kappa-B (RANK) on osteoclasts to initiate breakdown of bone (14,17). As the disease progresses, the pannus can grow into the cartilage, tendons, and bones surrounding it which enhance joint destruction further (17). Additionally, fibroblast-like synoviocytes will secrete proteases such as matrix metalloproteinases (MMPs) with MMP-1, MMP-3, and MMP-13 being the most important for joint destruction in RA, directly causing cartilage break-down (17). Finally, fibroblast-like synoviocytes has the ability to migrate to other joints initiating inflammation at a foreign site (2). Studies have also shown that the epigenetic regulation of the fibroblast-like synoviocytes differs from joint to joint which may be able to explain why RA only affects certain joint and why RA spreads as symmetrically between joints as it does (14).

IL-8 from osteoclasts are both involved in neutrophil recruitment and in inducing neutrophils to enter a state where they release neutrophil extracellular traps (NETs)(14). NETs are a network of fibers mainly consisting of highly condensed neutrophil DNA material that is secreted to trap and kill pathogens (13). NETs are interesting in RA pathogenesis in that the release of intracellular contents will expose autoantigens such as citrullinated proteins from within the neutrophils to the autoantibodies produced as a result of RA development (13). This allows ACPAs to form immune complexes with i.e. citrullinated histones from neutrophils, and by binding of their Fc part to osteoclasts these complexes can trigger osteoclast differentiation and activation, or by binding to macrophages induce secretion of TNF- α and IL-6 thus increasing inflammation (14).

Another role of autoantibodies in RA is complement system activation (13,18). Autoantibodies have been shown to bind several sites in the joints e.g. citrullinated proteins in the articular cartilage, and directly on the synovial membrane following injury or infections which expose the citrullinated proteins (14). In RA patients, it has been found that complement protein levels are reduced while the cleavage products of these proteins are increased which indicate use of the complement system (13). The complement system is a cascade of proteins cleaving and activating each other following initiation (18). In RA, initiation of the complement cascade may happen by the classical pathway, where initiation originates from the immune complexes formed by ACPAs binding to citrullinated proteins, and from complement factors interacting with small leucine-rich repeat proteoglycans (SLRPs) such as fibromodulin, osteomodulin, and chondroadherin that are present in the joint as a product of joint damage (18,19). However, activation of the complement system in RA also takes place through the alternative pathway (18). The activation through this pathway is thought to be a result of collagen type II on its own, as well as IgA antibodies such as RF being able to initiate alternative complement activation (18). Nonetheless, no matter which way the complement system is activated, the active biproducts are the same with C3a, C5a, and C5b being the most important (18). C5a and C3a will act as chemoattractants for various immune cells such as neutrophils, monocytes and macrophages (18). C5b will be the initiator for assembly of a membrane attack complex (MAC) that can further damage cells of the joint or even lyse cells of the joint (18).

In combination, all these factors contributing to RA pathology will cause damage to the joint which may lead to pain, deformity and disabilities in the joints(18). Additionally, over the years, 40% of RA patients experience extraarticular symptoms such as heart attacks, atherosclerosis, glomerulonephritis, neuropathy, and fibrotic or inflammatory lung disease (2,20). These complications are most likely the result of long-standing inflammation and therefore vasculitis is especially a challenge in cases where patients are poorly treated(20). This makes early diagnosis and proper treatment important of RA very important to ensure the best prognosis possible for each patient.

Diagnosis of rheumatoid arthritis

Until recently, the classification of RA was based on criteria from American College of Rheumatology (ACR) developed in 1987 (21). In 2010, ACR and European League Against Rheumatism (EULAR) proposed new criteria for RA classification which focused on the ability to diagnose RA in its early stages to avoid disease progression and bone destruction as a result of earlier initiation of treatment (21). According to the 2010 ACR/EULAR criteria, RA is the probable diagnosis if an individual obtains a certain score in a point system including four categories; joint involvement, serology (RF/ACPA), C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR), and duration of symptoms, see Table 1 (21).

Table 1 ACR/EULAR classification criteria for rheumatoid arthritis. If an individual obtains a score of six or more, the criteria classifies the person as having “definitive” rheumatoid arthritis. Table inspired from Kourilovitch et. al - Diagnosis and classification of rheumatoid arthritis (22).

Joint distribution	Points
1 large joint	0
2-10 large joints	1
1-3 small joints (large joints not counted)	2
4-10 small joint (large joints not counted)	3
More than 10 joints (at least one small joint)	5
Serology	
Negative RF and negative ACPA	0
Low positive RF or low positive ACPA	2
High positive RF or high positive ACPA	3
Symptom duration	
Shorter than 6 weeks	0
Over 6 weeks	1
Acute phase reactants	
Normal CRP and normal ESR	0
Abnormal CRP or abnormal ESR	1
“Definitive” Rheumatoid arthritis	≥6

These classification criteria for RA can be used as a guideline in the clinic to diagnose RA but may not be enough, often further evidence of RA is necessary to support an RA diagnosis (21). Therefore, the use of the classification criteria may be accompanied by ultrasound or magnetic resonance imaging (MRI) which are superior to clinical examination in determining inflammation of specific joints (21).

Treatment of rheumatoid arthritis

As of today, no curative treatment for RA exists. Treatment of RA is focused on remission or lowering disease activity in an attempt to prevent further disease progression leading to bone or joint destruction and disabilities (22,23). Medical treatment of RA is based on recommendations from EULAR and the primary focus of each treatment phase is to achieve remission or low disease activity within six months of initiation. At the first stage of treatment following diagnosis, the patient is treated with a conventional synthetic disease-modifying antirheumatic drug (csDMARD). The mechanisms of action of csDMARDs are not fully understood as they came into the clinic as a result of empiric observations of their effects, and was not developed specifically for any disease (22). Methotrexate is the recommended first choice drug in phase one of treatment as it is a csDMARD that has been extensively used and has a history of good effects both administered as monotherapy and in combination with other DMARDs or drugs (22). Opposed to the use of methotrexate for oncology purposes, it has been found that methotrexate treatment of RA is effective even at lower doses (23). The mechanisms of action of methotrexate in high doses used to treat e.g. cancers is to antagonize purine synthesis and therefore the synthesis of DNA resulting in cell cycle arrest in the S-phase followed by cellular apoptosis (23). When treating RA with lower doses of methotrexate it has been found that administering folate to avoid adverse effects of methotrexate's purine antagonistic effects, it does not result in lowered therapeutic effects on RA (23). This suggests that the antirheumatic effects of methotrexate relies on other mechanisms of actions than high-dose cancer treatment (23). The mechanism of action of low-dose methotrexate in RA is not completely understood but is mainly thought to be a result of its ability to facilitate adenosine signaling through receptors that have increased expression levels as a result of RA (22). In turn, this increased adenosine signaling leads to a lowered production of tumor necrosis factor(TNF)- α and nuclear factor kappa-light-chain-enhancer of activated B cells (NF- κ B) (22). Additionally, adenosine is thought to play a role in downregulating activation and proliferation of T-cells (22). These effects of methotrexate should result in dampening of the immune response. However, some patients do not experience remission or low disease activity from low-dose methotrexate after six months, and therefore the second phase of treatment is initiated (22). Choice of drug for phase two of treatment varies depending on presence of factors indicating disease progression such as autoantibodies, early joint damage, and high disease activity. If these factors are not present, another csDMARD should be attempted for another six months. Other csDMARDs used for RA treatment include leflunomide and sulfasalazine (22). However, if these factors indicating disease progression are present, a biologic DMARD (bDMARD) or a targeted synthetic DMARD (tsDMARD) should be administered instead (22).

bDMARDs and tsDMARDs are treatments targeting specific cellular processes to combat the autoimmune response in RA (7). It is recommended that these agents are administered in combination with csDMARDs but if the patient did not show any signs of response to treatment in phase one,

bDMARDs or tsDMARDs may be used in monotherapy (7). EULAR does not specify a hierarchical order on which of these drugs should be preferred for phase two treatment (22). This is mainly due to very scarce research comparing the effects of the different bDMARDs and tsDMARDs, and the research that has been conducted find very similar effects of the different drugs (7). These drugs have different targets, and in total, bDMARDs with five different mechanisms of action have been approved for treatment of RA; inhibition of TNF, inhibition of the IL-6 receptor, depletion of B cells, inhibition of IL-1, and blocking of T-cell co-stimulation, see table 1. Furthermore, only tsDMARDs with one mechanism of action has been approved for RA treatment; inhibition of Janus kinase (JAK) (7). If treatment in phase two is unsuccessful, treatment with a different bDMARDs or tsDMARDs should be initiated after every six months without reaching the treatment goal of remission or low disease activity (22).

Table 2 Some of the bDMARDs and a tsDMARD approved for treatment of rheumatoid arthritis. Table uses information from Rein, Philipp et. Al. – Treatment with Biologicals in Rheumatoid Arthritis: An Overview(7).

Agent	Type of DMARD	Mechanism of action
Abatacept	bDMARD	Blocking of T-cell co-stimulation
Adalimumab	bDMARD	TNF inhibitor
Anakinra	bDMARD	IL-1 receptor antagonist
Etanercept	bDMARD	TNF inhibitor
Golimumab	bDMARD	TNF inhibitor
Infliximab	bDMARD	TNF inhibitor
Rituximab	bDMARD	Depletion of B-cells
Tocilizumab	bDMARD	IL-6 receptor antagonist
Tofacitinib	tsDMARD	Janus kinase inhibition

Etanercept

Etanercept is a bDMARD that exerts its antirheumatic effects by TNF inhibition (7). It is a fusion protein made by genetic recombination combining two identical domains of the extracellular part of TNF receptor type 2 (TNFR2) and a the Fc part of a human IgG1 antibody, see figure 1 (8). The fact, that the TNFR 2 molecules are linked to the Fc part of an IgG antibody elongates the serum half-life of etanercept (8).

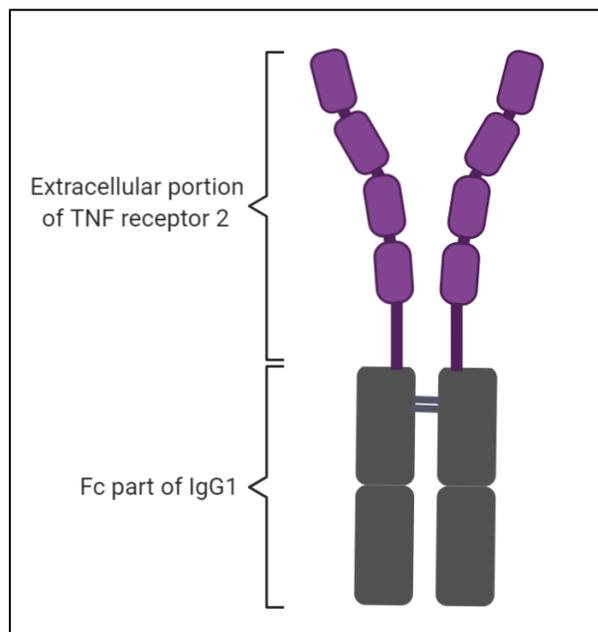


Figure 1 The composition of etanercept; two identical domains of the extracellular part of TNF receptor type 2 (TNFR2) and a the Fc part of a human IgG1 antibody.

Under normal conditions, TNF binds to two types of receptors found on all nucleated cells, TNF receptor type 1 (TNFR1) and TNFR 2 (8). Furthermore, naturally cleaved monomers of the TNFRs extracellular portion named soluble TNF receptors (sTNFR) works as natural antagonists to TNF by binding it without exerting any cellular signaling and thereby depleting some TNF molecules from reaching their cell-bound receptor (8). Just like sTNFR, Etanercept works as a decoy receptor, as it will bind to TNF efficiently but is not capable of initiating the signaling process that would normally occur when TNF binds to its receptor, see figure 2 (8).

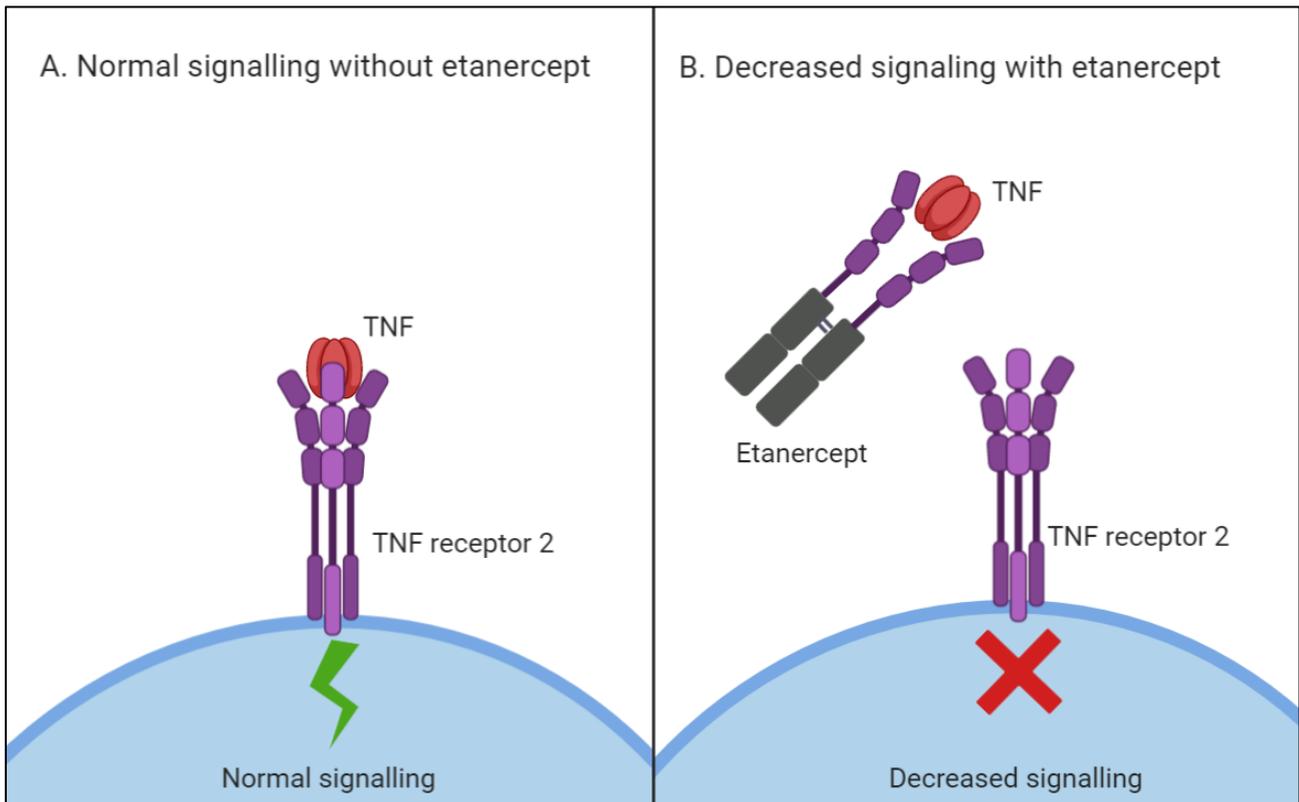


Figure 2 Mechanism of action of etanercept. A) Normal signaling without etanercept: TNF is free to bind to its receptor which results in normal signaling within the cell. B) Etanercept works as a decoy receptor to inhibit the functions of TNF on its target cells resulting in decreased intracellular signaling.

Etanercept can bind both TNF- α and TNF- β (also named lymphotoxin- α (LT- α)). Thus, etanercept decreases the effects that TNF stimulation has on the immune system under normal conditions (8). Usually, TNF is secreted by activated macrophages and other immune cells such as CD4⁺ T-cells and it has the ability to activate T-cells, B-cells, macrophages, neutrophils and endothelial cells (24). Furthermore, TNF can work as a chemoattractant for other immune cells enhancing the immune response (8). TNF is cytotoxic and can induce apoptosis in cells, while TNF- β in particular has the ability to induce proliferation of fibroblasts which may play a role in the formation of the pannus (8,24).

Despite good outcomes of etanercept in the patients that responds well to treatment, 30% to 40% of patients receiving anti-TNF bDMARD treatment do not achieve adequate response to treatment (25). As mentioned, the current treatment guidelines recommend administering another bDMARD if remission or low disease activity do not occur after six months, however, to avoid progression of disease and severe co-morbidities, the identification of biomarkers predicting or monitoring treatment response to these drugs is crucial (22).

Aim and strategy

Aim

The aim of this study is to investigate the molecular effects of etanercept in patients diagnosed with RA in an attempt to identify biomarkers for monitoring of treatment response.

Strategy

To assess the long-term effects of etanercept, the serum of 12 patients diagnosed with RA and treated with etanercept will be analyzed. Each patient had their serum isolated from blood samples collected at four different timepoints during treatment; before initiation of treatment, after three months, after six months and after 12 months. The average age of the participants at the time of enrollment was 60.93 (± 11.85), 10 of the participants are female, and four of the participants are male. The patients had to meet a list of inclusion- and exclusion criteria and was only included if they were over 18 years of age and had the willingness and ability to give written informed consent of participation. On the other hand, patients were excluded from the study if they met any of the exclusion criteria presented in table 3:

Table 3 The exclusion criteria by which possible candidates for the study was excluded. Participation in the study was denied if an individual met just one of these criteria.

Exclusion criteria

1. Other inflammatory, autoimmune, or rheumatic disease than rheumatoid arthritis.
2. Infection that has required hospitalization or administration of intravenous antibiotics within the last 30 days
3. Infection that has required administration of oral antibiotics within the last 14 days.
4. Treatment with glucocorticoids within the last 12 weeks.
5. Severe heart disease.
6. Demyelinating disease such as amyotrophic lateral sclerosis and multiple sclerosis.
7. Abuse of drugs or alcohol
8. Pregnancy and lactation

The serum proteome will be analyzed in two ways; by discovery-based mass spectrometry and by inflammatory profiling. Additionally, the cfDNA concentrations in the samples will be quantified to assess whether treatment has an effect on cfDNA levels. The overall workflow can be seen on figure 3.

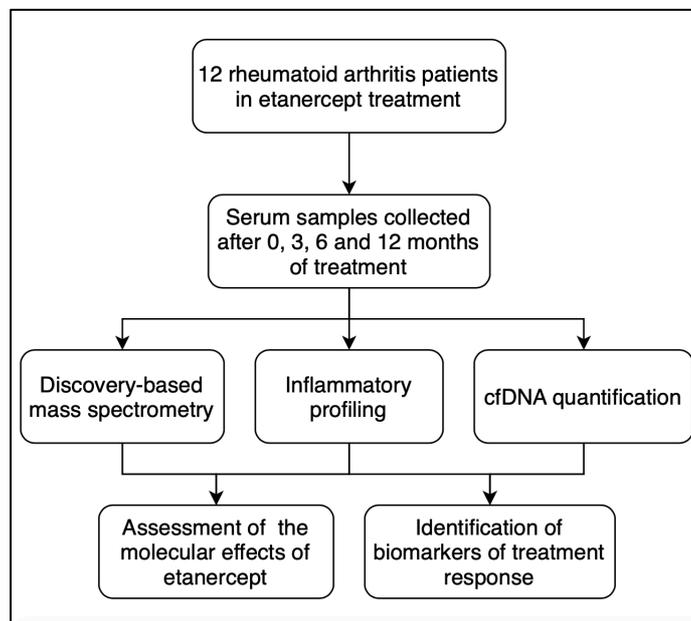


Figure 3 Overall workflow of the study. Serum samples from 12 patients with rheumatoid arthritis will be collected at four different time points; before treatment and after three, six and 12 months of treatment. These samples will be analyzed by discovery-based mass spectrometry, inflammatory profiling and by quantification of cfDNA levels. The results of these analyses will be used to assess the molecular effects of etanercept and to identify biomarkers of treatment response.

Discovery-based mass spectrometry is a non-targeted approach that attempts to map the proteome of a sample (26,27). This approach of proteome analysis allows for identification and quantification of a large number of proteins present in a biofluid sample and is therefore often used in an attempt to identify new biomarkers of disease activity or treatment response (26). Thus, this approach may allow for observation of changes in the abundance of specific proteins in response to an intervention such as treatment with etanercept (26). Because the aim of this study is to identify novel biomarkers of treatment response in patients treated with etanercept, such a discovery-based mass spectrometry approach will be used in this study. For this study, the samples will be analyzed on a TIMS-ToF Pro (Bruker Daltonics, Bremen, Germany) mass spectrometer, where the 10 most intense peptide signals for each scan cycle will be further analyzed by tandem mass spectrometry. However, one of the challenges of an approach like this is that only the proteins with the highest intensities for each scan, and therefore the most abundant proteins, are discovered. Since serum contains some proteins like immunoglobins and serum albumin in extremely high quantities, it can have a negative effect on the discovery of less abundant proteins that are actually of interest (26). As an example, serum contains approximately 35-50 mg serum albumin per mL, while some proteins that may be of interest can be present in quantities that are 1-million fold lower and in the pg per mL range (26). This range of protein quantities, especially with a mass spectrometry setup like in this study where high intensity proteins are prioritized, exceeds the capabilities of mass spectrometry and may hinder our ability to identify some proteins of interest (26). To increase the amount of protein identifications obtained

from the analysis, a reference sample containing serum, and therefore protein, from 12 different samples will be fractionated according to the hydrophobicity of the proteins using a high pH protein fractionation kit. This will result in 10 different fractions with a decreased complexity, each containing a different ‘subtype’ of proteins. All these fractions will be analyzed on the mass spectrometer and when analyzing these we should observe a higher coverage of the serum proteome than if the samples were analyzed without fractioning. A study investigating the effects of hydrophobicity fractioning on *E. coli* samples found that without fractioning they identified 163 proteins but after fractioning they identified 365 proteins (28). The mass spectrometry data of the fractions will be used as a spectral library to increase the number of protein identifications found in the samples. The identification of proteins will be conducted with the software program MaxQuant (version 1.6.12.0), where the mass (*m*) to charge (*z*) ratios will be compared to a UniProt human proteome database. In MaxQuant, the function *match between runs* will be enabled to allow for spectral comparison between the fractions and each of the samples. Furthermore, MaxQuant’s label-free quantification (LFQ) algorithm will be used to assign an arbitrary value to the intensities of each protein identified in each sample. These LFQ-intensities will be used for further data analysis in Perseus (version 1.6.12.0). LFQ algorithms are based on the fact that the concentration of a peptide in a sample has been found to be proportional to the peak intensity in a high resolution and high-mass-accuracy mass spectrometer setup such as the one used in this study (29).

Another approach by which the proteome will be investigated in this study is by inflammatory profiling. For the purpose of this study, a multiplex assay measuring the levels of 10 different cytokines that are important for inflammatory response and regulation of the immune system will be applied. Given their role in inflammation and RA pathology, alterations in the expression of these cytokines are of great interest. However, cytokines are low-abundance proteins and with a discovery-based mass spectrometry approach, the detection and quantification of such proteins is unlikely (30). Therefore, the use of an antibody-based assay to detect these low-abundant is of great benefit and may help to identify changes to the proteome that we are not able to detect using just mass spectrometry.

(Disclaimer: Sadly, this analysis was not finished at the time of handing in this project. This is described in more detail in the section *The need for inflammatory protiling.*)

Finally, an experiment measuring the concentration of cell free (cf) double-stranded DNA will be conducted. cfDNA has been found to be a marker of inflammation, and quantification of cfDNA concentrations in the serum of RA patients treated with etanercept may provide insight of the effectiveness of the treatment (31).

Methods

Ethics, patients included, sample collection and biobanking

This project was approved by the regional ethics committee of the Capital of Denmark (H-2-2014-086). In total, 26 patients diagnosed with RA using the EULAR/ACR criteria, all in treatment with etanercept were included in the study. However, 14 participants were excluded either due to inconsistent sample collection, pausation of treatment due to ongoing infection, or because they were non-responsive to treatment. This left 12 participants included in the study. Blood samples from the patients were collected at Danish Rheumatic Biobank located at Department of Clinical Biochemistry, North Denmark Regional Hospital, Hjørring. From each patient four samples were drawn; the first before initiation of treatment, and the following samples after three, six and 12 months of treatment, respectively.

To isolate serum, the blood samples were put in covered test tubes with no anticoagulative agent, after which they were left to clot for 15-30 minutes. After clotting, the samples were centrifuged for 15 minutes at 2000g to separate the clot from serum, and the serum layer was separated and transferred to cryotubes by pipetting. After collection, the samples were relocated to the Department of Health Science and Technology (HST) at Aalborg University for analysis. The serum sample of each patient was divided into several aliquots to avoid unnecessary freeze-thaw cycles for each experiment conducted, and the samples and stock cryotubes were kept at -80°C until use.

Protein quantification by a bicinchonic acid assay

To determine protein concentrations in the samples, a bicinchonic acid (BCA) assay was carried out using the ThermoFisher Scientific Pierce™ BCA Protein Assay Kit and guidelines. Firstly, a series of dilutes consisting of Milli-Q water containing a known concentration of bovine serum albumin (BSA) were loaded in a sample plate. Additionally, Milli-Q water alone was loaded to the plate to determine the background noise picked up by the analysis aperture. Each sample was loaded to the plate as well, and a working reagent provided in the kit was mixed in a ratio of 50 parts of reagent A to 1 part of reagent B. The working reagent was added to each of the wells, and the plate was incubated for 30 minutes to activate the color change reaction. After incubation, the samples were left at room temperature for 10 minutes to cool and halt the reaction.

Subsequently, the absorbance of each well was measured at 570 nm using a microchannel plate reader (TECAN Sunrise). The absorbance from the blank samples containing just Milli-Q water was deducted from each sample to correct for background noise. The absorbances measured of the BSA samples were used to make a linear model describing protein concentration as a function of absorption, and this model was applied to the absorbances measured in the actual biofluid samples to determine their protein concentration. To ensure a result as correct as possible, all samples and standards used to conduct this protein quantification were loaded in triplicates and the average absorbance values were used to calculate the protein concentrations.

Protein digestion by filter-aided sample preparation

To isolate, digest, and remove detergents from the proteins in the serum samples, the samples went through a filter aided sample preparation (FASP) protocol. Firstly, the volume of serum containing 100 µg of protein was determined using the results of the BCA assay, and was transferred to a spin filter tube (VWR North America Filter). The spin filter tubes were placed in a heat block pre-heated to 95°C for 5 minutes to denature the proteins. Following this, a digestion buffer containing 0.5% sodium deoxycholate (SDC) and 50 mM triethylammonium bicarbonate (TEAB) was mixed and added to each of the samples and the samples were centrifuged at 14,000g for 10 minutes. The flow-through was discarded, and the step was repeated again. To reduce and alkylate the samples, a mixture of digestion buffer, tris-(2- carboxyethyl)-phosphine (TCEP) (final TCEP concentration: 10 mM), and chloroacetamide (CAA) (final CAA concentration: 50 mM) was added to each sample, and the samples were vortexed and incubated at 37°C for 30 minutes. Following incubation, the samples were centrifuged at 14,000g for 10 minutes, digestion buffer was added to the spin filters, and the samples were centrifuged once again at 14,000g for 10 minutes. Subsequently, the spin filters were transferred to new Protein LoBind tubes, and digestion buffer as well as 1 µg trypsin per 100 µg protein was added to the filters. The samples were incubated in a plastic bag with a wet napkin at 37°C overnight. The following day, the samples were centrifuged at 14,000g for 10 minutes, followed by addition of digestion buffer and another centrifugation at 14,000g for 10 minutes. This left the digested proteins in the flow-through and the spin-filters were discarded.

To remove excessive detergents from the samples of digested protein, phase separation was performed using ethyl acetate and trifluoroacetic acid (TFA). Ethyl acetate and TFA were added to the samples, and the samples were vortexed thoroughly and centrifuged for five minutes at 14,000g. After centrifugation, the samples were left with an upper phase containing the detergents. This phase was removed by pipetting, and phase separation was performed once more but just adding ethyl acetate this time. Following separation of digested proteins, the samples were dried in a vacuum centrifuge and prepared for MS.

Fractioning of samples

Fractioning of the peptides based on hydrophobicity was done using the Thermo Fisher Scientific Pierce High pH Reversed-Phase Peptide Fractionation Kit. Firstly, the spin column was centrifuged at 5,000g for two minutes to pack the peptide-binding resin in the filter. After this, the spin column was washed in four cycles, first twice by adding acetonitrile to the filter, and then twice with a 0.1% TFA solution, centrifuging at 5,000g for 2 minutes between each cycle.

100 µg of peptides was collected from 12 random samples; three from each of the treatment groups (baseline, three months, six months and 12 months), and the peptides were dissolved in a 0.1% TFA solution. Following this, the peptides were separated in 10 fractions by adding different solutions to the filter and centrifuging each time at 3,000g for 2 minutes. Between each addition of another liquid, the spin column was transferred to different Protein LoBind tubes, and after centrifugation, the flow-through was kept. First, a flow-through fraction was made without adding anymore solution to the column. Subsequently, a wash-fraction was made by adding MilliQ water to the spin column. For the remaining eight fractions a mixture of acetonitrile and 0.1% triethylamine of varying ratios were

added, starting with a mixture of 5% acetonitrile for the first fraction (for each step the remaining part of the mixture is 0.1% triethylamine), and ending up with 50% acetonitrile for the last fraction. Each of the fractions were dried in a vacuum centrifuge and prepared for MS.

Loading of samples prior to mass spectrometry

After the samples had been dried in the vacuum centrifuge, they were resuspended in a mass spectrometry loading buffer consisting of ultrapure water, formic acid, TFA and acetonitrile. The samples were then ultrasonicated for five minutes followed by centrifugation at 14,000g for 10 minutes. The samples were loaded in a randomized order (obtained using www.random.org) for mass spectrometry in well-plates. Before the first sample and after every 10 samples and an indexed Retention Time (iRT) standard sample was loaded and analysed on the mass spectrometer. The spectra of the iRT standards were analyzed to monitor system stability during the analysis of the samples.

Data acquisition by LC-MS/MS

Each sample was sequenced on a hybrid trapped ion mobility spectrometry (TIMS) quadrupole time of flight (QToF) mass spectrometer, in this case TIMS-ToF Pro (Bruker Daltonics, Bremen, Germany). The mass spectrometer was coupled to a modified nano-electrospray ion source (CaptiveSpray, Bruker Daltonics, Bremen, Germany) which was applied a voltage of 1500 V. Furthermore, liquid chromatography (LC) was done by the use of a RSLC Proflow ultra-high-performance liquid chromatography (UHPLC) (Dionex, Thermo Scientific, Waltham, USA) LC-system. The digested peptide material diluted in mass spectrometry loading buffer was injected into a C18 reversed-phase pre-column (μ PAC, PharmaFluidics, Gent, Belgium) and onto a 50 cm analytical C18 Micro-pillar array column (μ PAC, PharmaFluidics, Gent, Belgium) at 30°C with a constant flow rate of 0.75 μ L per minute for further separation. Mobile phase A was 0.1% formic acid in HPLC-grade water (v/v) and mobile phase B was 0.1% formic acid in HPLC-grade acetonitrile (v/v). The loading was performed at 2% B while separation was done by a linear increase from 2% B up to 30% B over a gradient of 35 minutes. Between each injection, the column was flushed with 98% B for two minutes, and then reequilibrated by injection of 2% A for 6 minutes. The mass spectrometer was operated in the default application TIMS-On and parallel accumulation-serial fragmentation (PASEF) positive mode. The radio frequency (RF) field for funnel 1 was set to 300 Vpp and for funnel 2 to 200 Vpp. Trapped ion mobility separation was done by scanning a mobility range of 0.6 to 1.6 1/k0. Further experiments were acquired using a 100 ms ramp and 10 PASEF tandem mass spectrometry scans per top 10 with a cycle time of 1.10 seconds. The mass range for tandem mass spectrometry spectra was set to 100-1700 m/z.

cfDNA quantification

To measure the concentration of cfDNA in each of the samples, the kit Quant-iT™ PicoGreen™ dsDNA Assay Kit was used following the instructions of the manufacturer. A TE buffer containing 10 mM Tris-hydrochloride (Tris-HCl), 1 mM ethylenediaminetetraacetic acid (EDTA) with a pH of 7.5 was mixed. Furthermore, a Picogreen-buffer consisting of 1:200 Picogreen:TE was mixed. To make a reference standard curve to later determine DNA concentrations in the samples, a series of dilutes consisting of TE buffer containing a known concentration of Lambda DNA standard was

loaded in a sample plate as well as a sample containing only TE buffer to determine the background noise picked up by the analysis aperture. 10 μL of each sample was diluted in TE and loaded in the sample plate as well. Picogreen-buffer was added to all samples in an amount equal to the amount already present in the wells. The plate was incubated at room temperature in darkness for five minutes and the absorbances of all samples were measured at 260 nm using an ELISA-reader.

The absorbance from the blank samples containing just TE was subtracted from the absorbance of each sample to eliminate background noise. The reference samples containing a known concentration of DNA and their absorbances were used to make a linear model which was applied to all serum samples to determine their cfDNA concentration. All samples were loaded in duplicates and if the absorbance value of the duplicates of a single sample varied more than 10% from each other, the sample was excluded from the experiment.

Results

The proteomic investigation of serum samples from patients treated with etanercept revealed a total of 79 proteins with a significantly altered expression across all groups compared as a result of treatment. This includes proteins identified when comparing patients before treatment to patients after three, six and 12 months of treatment. Of these 79 proteins, 12 were found to be regulated significantly in the same direction across all groups compared. A functional enrichment analysis of the proteins significantly downregulated in expression when comparing samples before treatment and after 12 months of treatment showed that the proteins had significantly more interactions between each other than if it was a similarly sized group of randomly selected proteins. Also, the functional enrichment analysis revealed that the downregulated proteins are involved in biological processes such as acute inflammation, complement activation, and inflammatory response. Furthermore, a quantification of cfDNA in the serum of the patients revealed a continuous fall in cfDNA concentration as a result of treatment with etanercept.

Proteomic analysis of serum

Protein groups data from from data processing in MaxQuant was imported to Perseus where the LFQ intensities were utilized for statistical analysis of the proteome. A total of 369 proteins were identified across all samples before any data filtration. The data analysis workflow in Perseus is shown on figure 4.

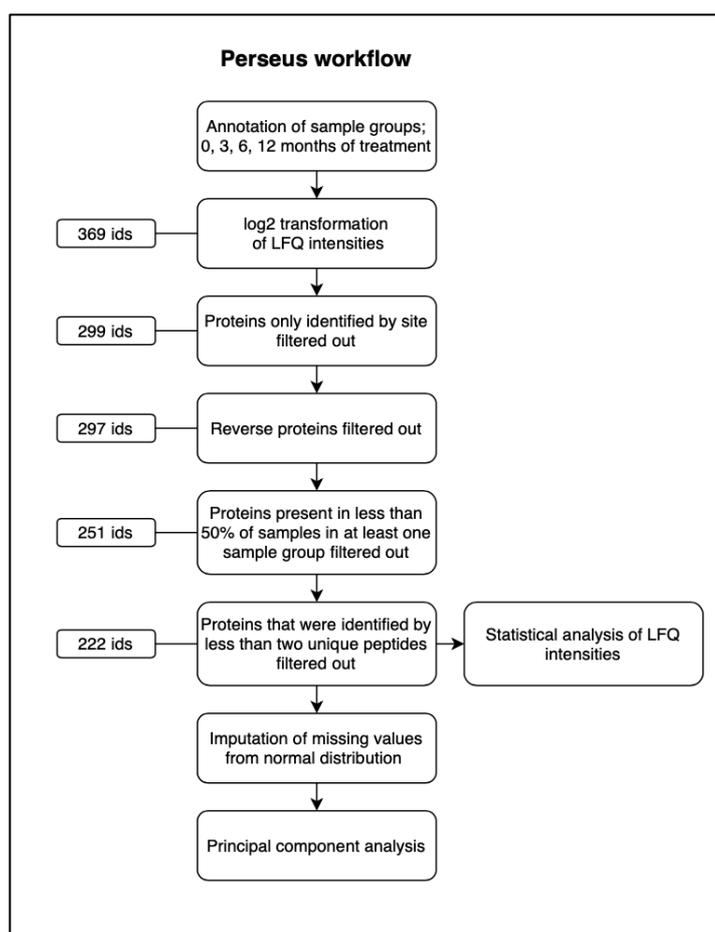


Figure 4 The chronologic workflow including filtrations and tests performed in Perseus to analyze the proteome in the samples. For each step filtering out protein identifications, the amount of proteins left after the step is shown to the left.

In Perseus, the samples were annotated to groups according to their duration of etanercept treatment. Then, the LFQ intensities of proteins were log₂ transformed for the purpose of later statistical analysis. Proteins that were only identified by site and reverse protein identifications were filtered out. Following this, proteins that were identified in less than 50% of the samples in at least one of the annotated groups were filtered out as well as proteins that were identified by less than two unique peptides. This filtration process resulted in 222 remaining protein identifications that were analyzed further. The log₂ transformed LFQ intensities were utilized to compare the changes in expression of proteins between the treatment groups. All statistical analyses were done with paired tests where the two groups compared contains multiple pairs of data from the same patients. This should allow for a better statistical foundation as the biological intervariability is taken into consideration. Finally, the missing values from normal distribution were imputed because a complete data set of valid values is required to conduct a principal component analysis.

Principal component analysis

A principal component analysis (PCA) was made to analyze whether any major proteome changes occurred over time as a result of treatment with etanercept, see figure 5.

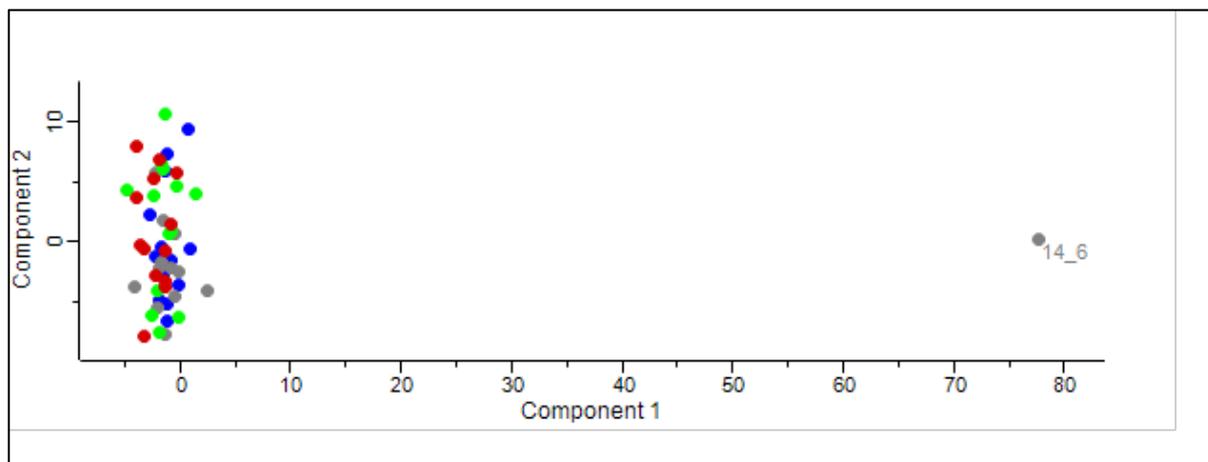


Figure 5 Plot showing a two-dimensional principle component analysis of serum samples of RA patients treated with etanercept. The red dots represent patients before treatment, the green dots after three months of treatment, the grey dots after six months of treatment, and the blue dots after 12 months of treatment. On the figure the one sample from patient 14 collected after six months of treatment (14_6) stands out as having a proteome that varies greatly from the other samples.

From figure 5 it can be seen that the sample collected from patient 14 after six months (14_6) stands out from all the other samples. This indicates that this sample has a vastly different proteome composition than all the other samples seen in the denser population on the left of the figure. To further assess the variance of sample 14_6, a Pearson correlation coefficient was calculated between this sample and all the other samples. This calculation describes the correlation coefficient, r , of the LFQ intensities measured for each protein in sample 14_6 and all the other samples. The Pearson correlation coefficient has a value between -1 and 1, where a value of 1 means that the two samples have total positive linear correlation, a value of 0 means no linear correlation, and a value of -1 means total

negative linear correlation (32). Usually, r-values above 0.70 can be regarded as having high positive linear correlation, while values above 0.90 are regarded as very high positive linear correlation (32). The Pearson correlation coefficient of sample 14_6 and the 51 other samples varied from 0.57 to 0.65, while the correlation coefficient between other samples were mostly above 0.90 but varied from 0.85 to 0.98. Thus, all samples, apart from sample 14_6, had high or very high positive linear correlations between each other. The Pearson correlation coefficients can be found in appendix 2.

The cause of the variation of this sample is unknown but could be a result of instrument issues, improper sample handling or an ongoing infection in the patient. Nonetheless, as a result of this variation, all samples from this patient will be excluded from the data for the rest of the analysis. All samples from the patient are withdrawn as that will allow for all further statistical tests to still be carried out in pairs while keeping an incomplete sample set from one patient would hinder this.

A PCA plot after the exclusion of patient 14 can be seen on figure 6:

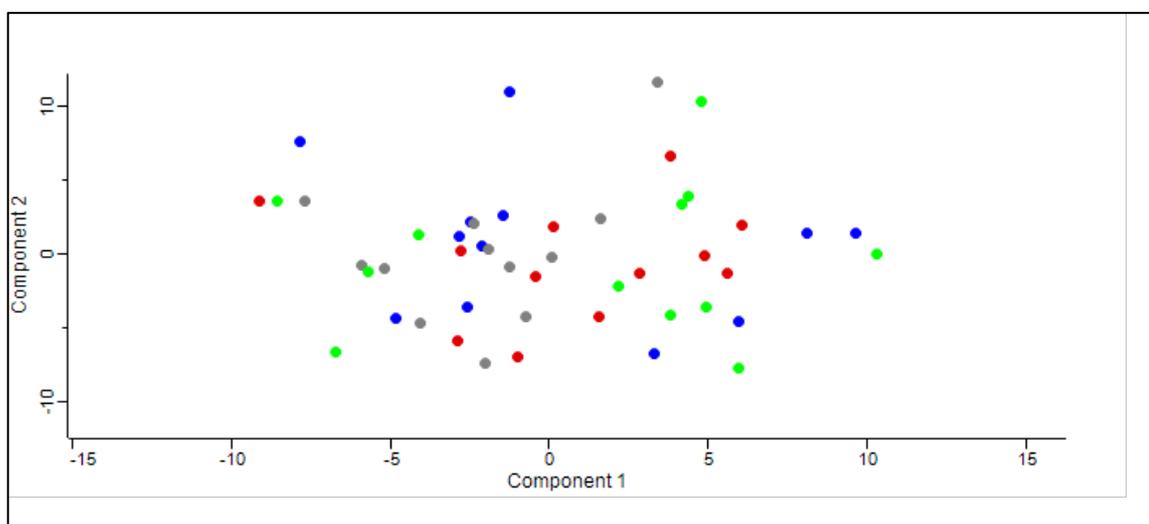


Figure 6 Plot showing a two-dimensional principle component analysis of serum samples of RA patients treated with etanercept. The red dots represent patients before treatment, the green dots after three months of treatment, the grey dots after six months of treatment, and the blue dots after 12 months of treatment. No distinct sample populations seem to occur on the plot as a result of treatment.

On figure 6 each of the patient samples is represented by a dot which is colored according to time of sample collection; red at baseline, green after three months, grey after six months, and blue after 12 months. All samples seem to be present in a mixed population on the PCA plot and thus no proteome changes strong enough to shift the positions of the populations in a PCA plot seem to have happened as a result of etanercept treatment.

Analysis of changes in protein expression

After all filtering processes conducted as part of data analysis, a total of 222 proteins were identified across all four groups. To investigate the changes in expression of these proteins over the treatment period, separate paired *t*-tests were performed comparing the baseline protein intensity values to the samples after three, six, and 12 months, respectively. A summary of the proteins found to have a significantly changed expression as a result of etanercept treatment can be seen in table 4:

Table 4 An overview of the proteins with significantly changed expression levels found in the serum of rheumatoid arthritis patients treated with etanercept. The table shows the total amount of significant protein regulations, the amount of these that were downregulated, and the amount of these that were upregulated when comparing the data from before treatment to after three, six and 12 months of treatment, respectively.

	Total	Direction of change	
		Downregulated	Upregulated
Baseline vs 3 months	47	42	5
Baseline vs 6 months	29	20	9
Baseline vs 12 months	40	36	4

As it can be seen in Table 4, 47 proteins were found to have a significantly regulated expression when comparing samples before treatment to samples after three months of treatment, and out of these, 42 were downregulations while 5 were upregulations. When comparing samples before treatment and samples after six months of treatment, 29 significant regulations were seen, 20 downregulations and 9 upregulations. Finally, a total of 40 significant protein regulations were seen after 12 months of treatment compared to the baseline, 36 of these regulations were a lowering in protein expression, while 4 were an increased protein expression.

To illustrate how these significantly changed protein expressions overlap between the groups compared, a Venn diagram was made, see figure 7.



Figure 7 A Venn diagram showing the overlap of the proteins with a significantly altered expression as a result of etanercept treatment. It can be seen that a total of 12 proteins are found to be significantly regulated both when comparing before treatment to after three, six and 12 months of treatment, respectively.

As seen in the Venn-diagram on figure 7, 12 proteins were found to be significantly regulated across all *t*-tests. Furthermore, the direction of change of all these proteins is also the same. These 12 proteins, the fold-changes and p-values can be seen in table 5. A complete list of proteins found to have a significantly changed expression and the p-values and fold-changes can be found in appendix 1.

Table 5 The 12 proteins found to be significantly regulated as a result of etanercept treatment both when comparing serum samples before treatment to after three, six and 12 months of treatment.

Protein name	Gene name	Fold-change			p-value		
		0 vs 3	0 vs 6	0 vs 12	0 vs 3	0 vs 6	0 vs 12
Trypsin	N/A	1.42	1.72	1.77	0.0246	0.0042	0.0220
Apolipoprotein M	APOM	0.90	0.88	0.83	0.0272	0.0179	0.0005
Ig lambda chain V-III region SH	IGLV3-19	1.33	1.34	1.45	0.0147	0.0379	0.0132
Beta-2-glycoprotein 1	APOH	0.82	0.89	0.77	0.0056	0.0154	0.0001
Retinol-binding protein 4	RBP4	0.79	0.82	0.71	0.0228	0.0125	0.0025
Vitamin D-binding protein	GC	0.91	0.89	0.78	0.0199	0.0000	0.0000
Plasma kallikrein	KLKB1	0.91	0.92	0.89	0.0276	0.0360	0.0380
Coagulation factor XIII B chain	F13B	0.92	0.77	0.80	0.0274	0.0426	0.0032
Complement C1s	C1S	0.86	0.90	0.89	0.0037	0.0157	0.0259
Complement component C6	C6	0.76	0.77	0.70	0.0002	0.0001	0.0000
Pregnancy zone protein	PZP	0.69	0.74	0.45	0.0060	0.0094	0.0160
Hyaluronan-binding protein 2	HABP2	0.82	0.83	0.73	0.0095	0.0151	0.0000

To visualize the significantly regulated proteins, three volcano plots were made; one for changes between the protein expressions before treatment and three months after treatment, one for changes from before treatment and six months after initiation, and lastly, one comparing protein expressions before and after 12 months of treatment. These volcano plots can be seen in figure 8 A), B) and C).

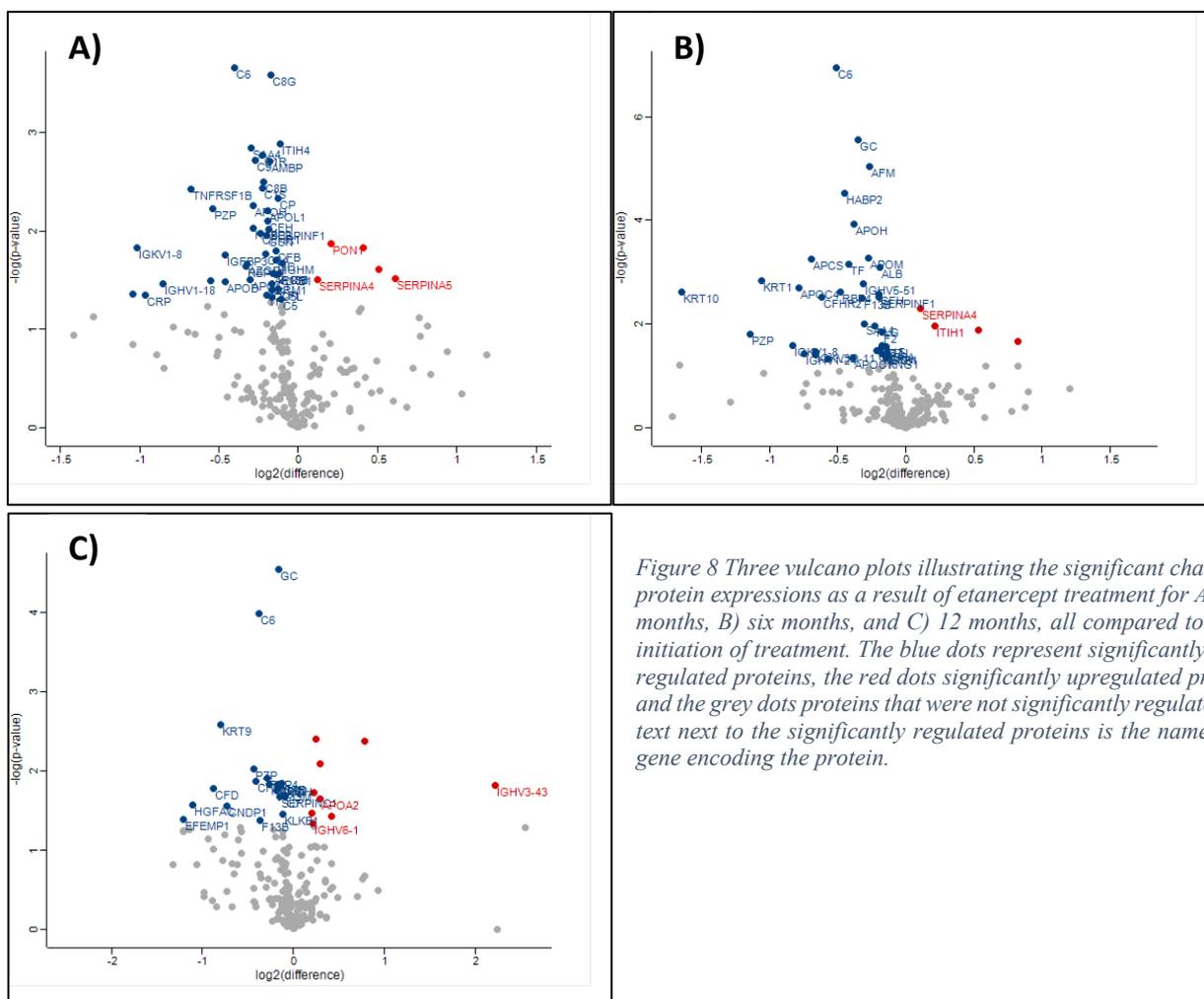


Figure 8 Three volcano plots illustrating the significant changes in protein expressions as a result of etanercept treatment for A) three months, B) six months, and C) 12 months, all compared to before initiation of treatment. The blue dots represent significantly down-regulated proteins, the red dots significantly upregulated proteins, and the grey dots proteins that were not significantly regulated. The text next to the significantly regulated proteins is the name of the gene encoding the protein.

Functional enrichment analysis

To further assess what type of proteins are being regulated as a result of etanercept treatment, a functional enrichment analysis was made. Using www.string-db.org, the roles and interactions of the proteins found to be downregulated as a result of treatment with etanercept was identified, see figure 9. Another analysis was made for the upregulated proteins, but those proteins were not found to have significantly more interactions than if it was a set randomly selected proteins. For these analyses, the protein regulations found after 12 months of treatment was chosen.

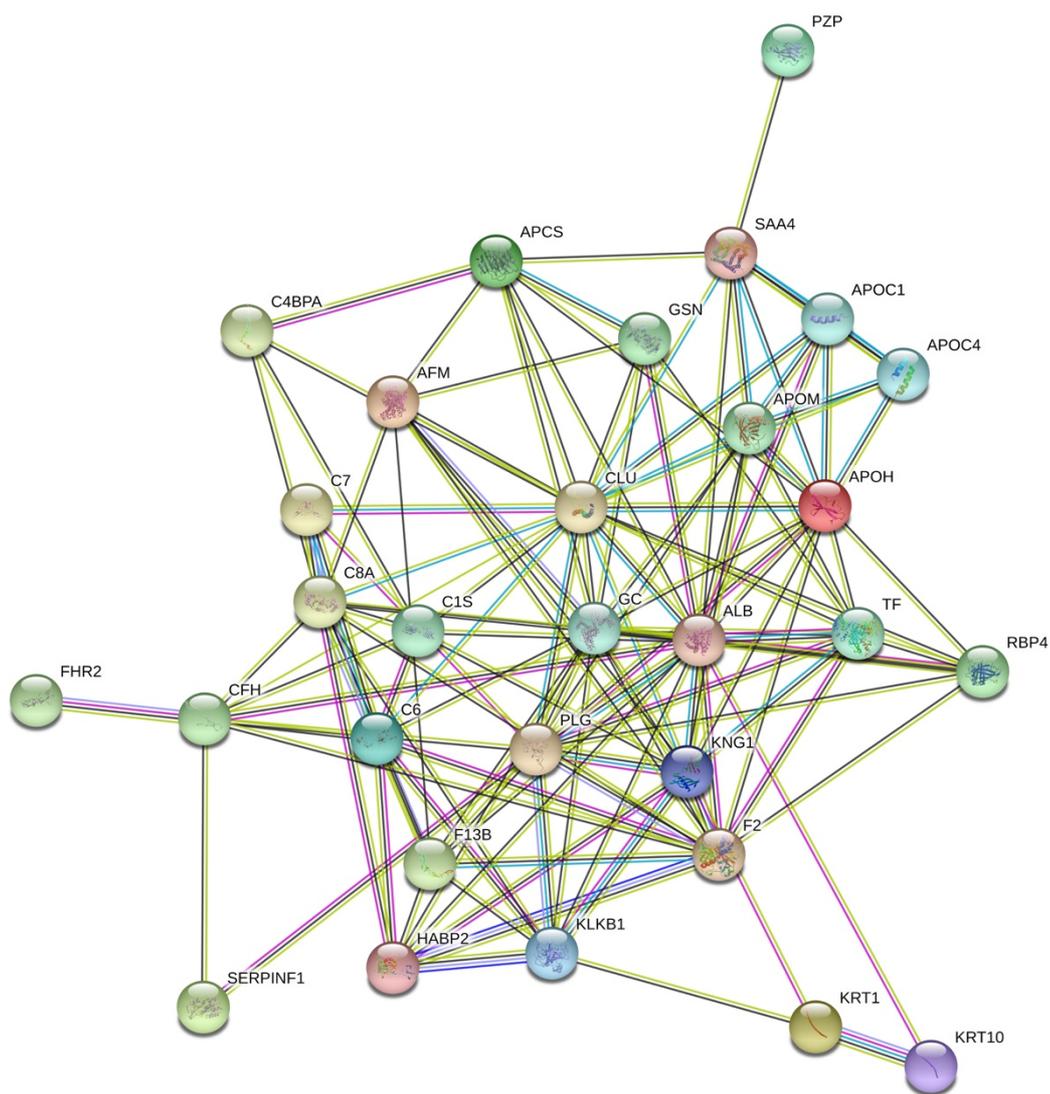


Figure 9 Visualization of a functional enrichment analysis of proteins downregulated in RA patients after 12 months of treatment with etanercept. Each dot represents a protein, and each colored line between the dots describe a different source that describes an interaction between these proteins or a way that these proteins interact.

The functional enrichment analysis of proteins with downregulated expression after 12 months of etanercept treatment showed that the proteins had significantly more interactions than if it had been a random set of proteins ($p=0.0000$). Furthermore, when looking into which biological processes these proteins are a part of, it can be seen that some examples include protein activation, regulation of acute immune response, complement activation, regulation of inflammatory response, regulation of proteolysis, and humoral immune response. A more comprehensive list can be found in table 6.

Table 6 Some of the biological processes that the proteins found to be significantly downregulated after 12 months of etanercept treatment were found to be involved in by a functional enrichment analysis.

Biological Process (GO)			
GO-term	description	count in gene set	false discovery rate
GO:0072376	protein activation cascade	13 of 74	1.41e-20
GO:0002673	regulation of acute inflammatory response	11 of 92	1.51e-15
GO:0070613	regulation of protein processing	11 of 116	1.10e-14
GO:0030449	regulation of complement activation	9 of 52	4.95e-14
GO:2000257	regulation of protein activation cascade	9 of 54	5.62e-14
GO:0032101	regulation of response to external stimulus	16 of 732	2.69e-13
GO:0006956	complement activation	8 of 49	2.34e-12
GO:0050727	regulation of inflammatory response	12 of 338	6.96e-12
GO:0030162	regulation of proteolysis	14 of 742	1.32e-10
GO:0006959	humoral immune response	10 of 252	3.74e-10
GO:0080134	regulation of response to stress	16 of 1299	8.66e-10
GO:0061045	negative regulation of wound healing	7 of 66	1.24e-09
GO:0006958	complement activation, classical pathway	6 of 34	2.29e-09
GO:0002697	regulation of immune effector process	10 of 362	7.88e-09
GO:0030195	negative regulation of blood coagulation	6 of 48	1.14e-08
GO:0006952	defense response	14 of 1234	4.86e-08
GO:0072378	blood coagulation, fibrin clot formation	5 of 27	5.70e-08
GO:0019538	protein metabolic process	22 of 4194	8.54e-08
GO:0019835	cytolysis	5 of 32	1.06e-07
GO:0032102	negative regulation of response to external stimulus	8 of 281	3.24e-07
GO:0051917	regulation of fibrinolysis	4 of 14	4.72e-07

Serum concentration of cfDNA

cfDNA concentration in the serum of the patients was measured before treatment, after three months, after six months, and after 12 months of treatment. The measurements were done in technical duplicates, and samples that had duplicate values varying more than 10% from each other were excluded. The average concentration of cfDNA in serum was falling for every timepoint of measurement, starting with an average concentration of 1542.9 ng/mL cfDNA in serum before initiation of treatment, after three months of treatment the average was down to 1450.0 ng/mL, after six months 1395.7 ng/mL, and finally, after twelve months the average had fallen to 1312.3 ng/mL, see table 7.

Table 7 Average concentrations of cfDNA in the serum of RA patients treated with etanercept. Samples from before treatment (baseline), after three months of treatment, after six months of treatment, and after 12 months of treatment were analyzed.

Time of measurement	Average cfDNA concentration (ng/mL)
Baseline	1542.9
3 months	1450.0
6 months	1395.7
12 months	1312.3

To assess whether the consistent fall in the average concentration of cfDNA as a result of etanercept treatment was likely to be a random incidence, three paired *t*-tests were done; one comparing the means of cfDNA concentration before and after three months of treatment, one comparing before and

after six months of treatment, and one comparing before and after 12 months of treatment, see table 8.

Table 8 Two-tailed paired t-test comparison of cfDNA levels in serum at different timepoints before and after treatment with etanercept.

	No. of paired sample sets	Direction of change	Two-tailed p-value
Baseline vs 3 months	11	Fall	0.0479
Baseline vs 6 months	10	Fall	0.0347
Baseline vs 12 months	9	Fall	0.0008

As it can be seen in table 8, cfDNA levels were significantly lowered after all treatment durations. A visualization of the fall in cfDNA levels can be seen on the boxplots in figure 11. Here it can be seen that the median cfDNA concentration falls continuously as a function of etanercept treatment duration.

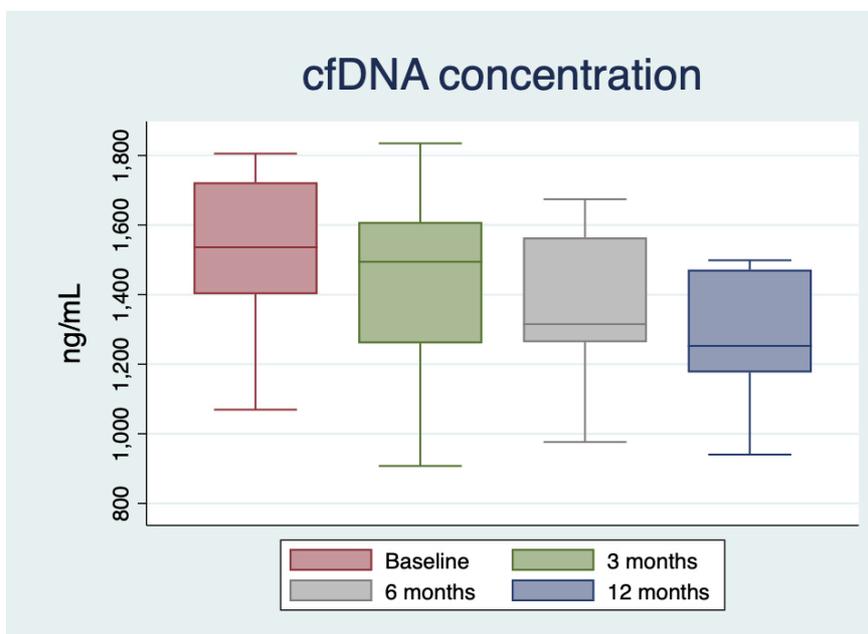


Figure 10 Boxplots illustrating the fall in cfDNA concentration as a result of etanercept treatment. The red boxplot shows the concentrations before initiation of treatment, the green after three months of treatment, the grey after six months of treatment, and the blue after 12 months of treatment.

Discussion

Proteomic investigation of the alterations caused by etanercept

After etanercept treatment, a total of 79 serum proteins were found to have a significantly altered expression across the groups comparing baseline to treatment for three, six and 12 months. Out of these 79 proteins, 12 were found to be significantly regulated after treatment at all of these treatment lengths (see table 5). Since the purpose of this study is to identify biomarkers that may indicate treatment response of etanercept, these 12 proteins will be investigated further.

These proteins all present the ability to be good biomarkers of treatment response as their alteration in expression was observable relatively early in the treatment period after three months but also work as indicators of treatment response after as long as 12 months of treatment. Potentially, as patients that did not show response to treatment with etanercept were excluded from the study, all these 12 proteins could be seen as biomarkers for successful treatment response. However, the proteins will be examined further by conducting a search in existing literature to expand the knowledge on how etanercept treatment may influence the serum of patients with RA on a proteomic level. Further analysis of the functions or relations of the proteins may also reveal specific usefulness of some proteins as biomarkers.

By researching published literature, five out of the 12 proteins were found to have an established relation to RA or to have been found regulated from healthy individuals compared to individuals with RA in other studies, see table 9.

Table 9 The five proteins that were identified in literature to be regulated from healthy individuals to rheumatoid arthritis (RA) patients AND found to be etanercept responsive in this study across all lengths of treatment. The table also shows the direction of regulation from healthy individuals to RA patients, which biofluid this regulation was observed in and the direction of regulation in abundance identified in our study.

Protein	Regulation from healthy to RA (biofluid tested)	Regulation from RA to etanercept treated RA
Apolipoprotein M	Up (plasma)(33)	Down
Retinol-binding protein 4	Up (serum)(34)	Down
Vitamin D binding protein	Up (serum)(27)	Down
Plasma kallikrein	Up (synovial fluid)(35)	Down
Pregnancy zone protein	Up (serum) (36).	Down

Possible protein biomarkers of etanercept response

In this section, the five candidate etanercept responsive biomarkers from table 9 will be described in regard to their association to RA pathology, and their likely ability to predict etanercept response based on current knowledge will be discussed.

Apolipoprotein M

Apolipoprotein M (APOM) is a protein involved in metabolism of high-density lipoproteins which may not seem associated to RA (33). However, APOM has been found to be related to inflammation and infection, and in obese patients a fall in APOM concentration is associated with increased levels of proinflammatory cytokines such as TNF- α and IL-6 (33,37). Based on that, the fact that we observe a lowered abundance of APOM in the serum of RA patients treated with etanercept ($p_{3\text{months}}=0.0272$, $p_{6\text{months}}=0.0179$, $p_{12\text{months}}=0.0005$) may indicate that etanercept does not inhibit the inflammatory process to cause lowered disease activity. However, the study associating falls in APOM concentrations to increased inflammation in obese patients is not related to RA in any way. More knowledge about the role of the protein and its relation to RA is therefore needed to draw any conclusion about its regulation in response to RA severity. Nonetheless, the patients included in this study responded well to etanercept, and therefore APOM regulations in RA patients may have different outcomes than the regulation seen in obese patients. Also, as shown in table 9, a study found that APOM was upregulated in the plasma of RA patients compared to healthy individuals which again is contradictory to the findings of other study which indicate that APOM levels should decrease as a result of inflammation (33). These inconclusive findings and lack of in-depth knowledge of the functions of APOM and how it contributes to the inflammatory process calls for further research of the relationship between APOM, inflammation, and RA.

Retinol-binding protein 4

In our study we found retinol-binding protein to be significantly downregulated in response to etanercept treatment ($p_{3\text{months}}=0.0228$, $p_{6\text{months}}=0.0125$, $p_{12\text{months}}=0.0025$). Retinol-binding protein 4 (RBP4) is responsible for transportation of retinol (vitamin A) to the tissues of the body (38). In relation to RA, a study has been able to relate higher serum concentrations of retinol-binding protein 4 to higher severity of RA (34). The same study also concludes that high levels of RBP4 is related to development of insulin resistance in obese patients with RA but not in patients that are not obese (34). In conclusion, based on the finding that high levels of RBP4 is associated with RA severity, the finding that this protein was reduced in abundance in our study could indicate a lowering of the severity of RA in response to etanercept. Limited knowledge exists about how exactly the protein is related to RA and what causes it to be regulated in response to the severity of disease but it does present the possibility to indicate successful etanercept treatment, and may even be able to predict systemic complications of RA in form of development of insulin resistance in obese patients allowing for early intervention against this. Further knowledge about how RBP4 is affected by etanercept treatment, how it is related to RA pathology, and its relation to development of systemic consequences of RA is needed to draw any conclusion about the utility of RBP4 as a biomarker.

Vitamin D-binding protein

We found vitamin D binding protein (VDBP) to be a potential etanercept responsive biomarker as it was significantly downregulated in the serum of RA patients after three, six and 12 months of etanercept treatment ($p_{3\text{months}}=0.0199$, $p_{6\text{months}}=0.0000$, $p_{12\text{months}}=0.0000$). Another study attempting to identify RA biomarkers by comparing the serum of RA patients to the serum of healthy individuals found VDBP to be upregulated in patients with RA, and concluded that VDBP had a great potential to be a biomarker for RA (27). The role of VDBP in RA is not completely understood but the protein is thought to play a role in maintaining joint homeostasis in response to RA pathogenesis (27). VDBP have been shown to combat F-actin polymers secreted by cells with increased permeability due to RA-associated inflammation(27). If not combated, these actin polymers cause blood vessels to block leading to microthrombosis (27). VDBP is able to bind to the F-actin polymers and convert them into monomers after which they are transferred to the liver and broken down (27). VDBP moves to tissues where it is needed, and therefore the fact that we identify a lowered abundance of VDBP in serum following treatment could be interpreted as a case of VDBP moving to joint tissue as a result of more severe RA (27). However, the study that also identified VDBP as a biomarker suggests that VDBP gets overexpressed even in the serum whenever it is needed in a tissue (27). Therefore, a downregulation of VDBP may indicate an overall lowered inflammatory response in the body or specifically in the joint as a result of RA remission (27).

Pregnancy zone protein

Pregnancy zone protein (PZP) was found to be downregulated in RA patients after etanercept treatment ($p_{3\text{months}}=0.0060$, $p_{6\text{months}}=0.0094$, $p_{12\text{months}}=0.0160$). The protein has been identified in a number of tissues and biofluids including the brain, uterus, liver, serum, synovial fluid, and cerebrospinal fluid (39). The name of the protein is associated with the high concentrations it is found at during pregnancy where the concentration of the protein can be detected up to 100-fold its normal concentration, hence part of the reason why pregnancy is among the exclusion criteria for participation in this study (39). The function of PZP is thought to be anti-inflammatory and it is upregulated i.e. during pregnancy where it has been found to exert synergic functions with another protein, placental protein-14, to combat CD4+ T-cell activation and thereby inhibiting an immune response against the fetus (39).

Very scarce literature exists about the function of pregnancy zone protein in RA and how its regulation changes in response to developing RA or being treated for RA. However, one study identified that concentrations of pregnancy proteins in individuals with RA were higher than in controls (36). Curiously and contradictory to the finding in our analysis, the same study mentions that the high levels of PZP during pregnancy causes remission in cases where the pregnant individual has RA indicating that higher levels of PZP is associated with RA remission (36). However, a study investigating sputum content of patients with severe bronchiectasis found that these patients had an increased concentration of PZP in their sputum compared to healthy individuals (40). They also found that the PZP concentration was lowered in response to disease remission after treatment with antibiotics, and lastly, they identified that the one of the sources of the protein was NETs (40). In conclusion, it occurs as if PZP is an anti-inflammatory protein which rise in concentration as a part of the body's response to inflammation, and that lowered values of PZP indicate a lowered inflammatory

state. A theory behind the fall of PZP that we observe in this study is that due to a relieved state of inflammation, NETosis occur more rarely and therefore the serum content of this protein is lowered. PZP could be a good biomarker candidate for the treatment response of etanercept in RA patients.

Plasma kallikrein

Another potential biomarker of successful etanercept treatment of RA patients identified in this study is plasma kallikrein which was shown to be significantly less abundant after three, six and 12 months of treatment ($p_{3\text{months}}=0.0276$, $p_{6\text{months}}=0.0360$, $p_{12\text{months}}=0.0380$). The protein has been shown to be present in synovial fluid of RA patients in increased concentrations compared to healthy individuals, and high concentrations of the protein has been related to increased pain and inflammation in RA (35).

Plasma kallikrein is part of a hormonal system known as the kinin–kallikrein system. This system is initiated by the activation of plasma kallikrein by coagulation factor XII (35). Following this activation, plasma kallikrein will cleave kininogen which results in the release of bradykinin (35). Bradykinin is a proinflammatory peptide which by binding to its receptors causes an increase in pain, vasodilation facilitating the recruitment of more immune cells to the site, and is theorized to cause endothelial cells to release CXCL5, a chemoattractant for neutrophils (35).

Furthermore, we found that kallistatin, another protein associated with the kinin-kallikrein system, was significantly regulated in response to etanercept. Kallistatin was upregulated 1.09-fold after three months ($p=0.0316$) and a 1.08-fold after 12 months ($p=0.0051$), see appendix 1. Kallistatin is a protein that exerts effects that may be important to the remission of RA patients, and the protein has been found to be present in lowered concentrations in the serum of individuals with RA but also in various other pathological conditions such as liver disease, cancer, pneumonia, sepsis, and hypertension (27,41). Kallistatin has two important structural elements; an active site and a heparin-binding domain (41). Both of these sites have effects that may be important for relief of pathologic events increasing the severity of RA. In example, the active site of kallistatin has been shown to suppress cytokine signaling in macrophages and to inhibit plasma kallikrein, thereby lowering the release of bradykinins (41). The heparin-binding domain functions by binding to heparin sulfate proteoglycans on cell surfaces and have been found to antagonize effects such as angiogenesis induced by vascular endothelial growth factor, vascular permeability, TNF- α induced inflammation, oxidative stress, and apoptosis (41).

RA pathology and effects of etanercept seems to be related to the kinin–kallikrein system, and especially one of its components, plasma kallikrein, shows signs of being a very promising biomarker for etanercept efficiency as it is provenly associated to RA pathology and consistently lowered following etanercept treatment for either three, six or 12 months.

In summary, all of these proteins could in theory work as markers of etanercept treatment response. However, the known relation of some of these markers to RA is very scarce, and more knowledge about these proteins and their relations to RA pathology or etanercept mechanisms are needed to fully confirm their usefulness as biomarkers. Nonetheless, especially VDBP, PZP, and plasma kallikrein occurs to be good biomarker candidates as their relation to RA pathology is somewhat well-documented.

Lowered cfDNA concentrations indicates successful treatment response

The concentration of cfDNA found in the serum of the RA patients treated with etanercept was significantly lowered at all timepoints of sample collection compared to the baseline concentration ($p_{3\text{months}}=0.0479$, $p_{6\text{months}}=0.0347$, $p_{12\text{months}}=0.0008$). Usually, elevated cfDNA levels are associated to inflammation, and cfDNA levels in serum have been observed to rise in response to acute inflammation, rise further with chronic inflammation, and even further with severe infection (31). A rise in cfDNA levels has also been observed to occur with RA when comparing healthy controls to RA patients (42). One of the mechanisms that is theorized to cause this rise is the release of NETs by neutrophils as described in *Pathology of rheumatoid arthritis* (42). Yet, some studies find that cfDNA levels are lower in patients with established RA than it is in patients with early RA (42). This is however thought to be a result of successful treatment of RA, and therefore a phenomenon that would not occur if the patients were left untreated (42). Different studies present different patterns in how cfDNA levels are altered as a result of bDMARD treatment. One study observed that cfDNA levels were elevated eight weeks post bDMARD treatment but lowered after 12 weeks (43). The study correlated an early spike in cfDNA levels to lowered disease activity and found that patients who did not show this early elevation had an overall worse treatment response (43). Another study correlated improvement in clinical and laboratory findings of RA patients to lowered levels of cfDNA in plasma after three months of treatment with bDMARDs but the fall was not significant until six months after treatment (44). In our study we found significantly lowered cfDNA levels after three months after treatment, and with the p-value falling and increasing the confidence in the lowered concentration further after six and 12 months of treatment.

Studies have contradicting understandings regarding the alterations of cfDNA levels early in the treatment process of RA with bDMARDs and how these changes are related to improvement of disease activity (42–44). However, studies agree that after a prolonged period of treatment, cfDNA concentration in RA patients falls and this fall is associated with improved clinical findings and lowered disease activity (43,44).

The findings in our study suggest that a significant fall in cfDNA levels can be detected as soon as three months post bDMARD treatment with etanercept. However, due to uncertainties of the correlation of early treatment response and cfDNA levels, using a fall in cfDNA levels as a biomarker for successful etanercept treatment response may not be viable until three to six months into the treatment process. Furthermore, based on the findings in the study observing an early elevation of cfDNA in patients with a good treatment response to bDMARDs, future studies regarding the changes cfDNA levels early in the etanercept treatment period may prove that a rise in cfDNA levels within the first few weeks of treatment can predict whether the patient will end up having the desired treatment response (43).

The need for inflammatory profiling

As mentioned in *Aim and strategy*, inflammatory profiling of specific proteins of interest would complement the proteomic analysis by discovery-based mass spectrometry well. Such an experiment is planned to be conducted for this study, however, due to the development of the COVID-19 pandemic, student laboratories have been closed since March 13th, 2020, and therefore we have not been able to finish this experiment. Hopefully, the results of the analysis will be available for the oral examination of this thesis.

During the work of this thesis, we also worked on a project regarding the proteomic effects of treatment of osteoarthritis with intraarticular gold micro-implants. This project has a similar workflow as seen in figure 3, and synovial fluid and serum samples were analyzed proteomically using discovery-based mass spectrometry and inflammatory profiling, and cfDNA quantification was done as well. To give an example of the usefulness of inflammatory profiling for investigation of the effects of a treatment and biomarker discovery, the results of the inflammatory profiling of the synovial fluid of these OA patients will be described briefly and the impact such an analysis could have on this project in the investigation of treatment with etanercept will be discussed.

In the study investigating the effects of gold on osteoarthritis patients, inflammatory profiling was done using a kit based on Proximity Extension Assay (PEA) technology, where antibodies are bound to the protein of interest and if the protein is bound by the two right antibodies, they hybridize and are extended by DNA polymerase and amplified. The amplification of the DNA associated to each protein marker is analyzed through quantitative polymerase chain reaction (qPCR) technology, and a relative abundance of each protein can be assessed from this readout (45).

We found that 24 cytokines were significantly downregulated in the synovial fluid of OA patients treated with intraarticular gold micro-implants for eight weeks. Further analysis of these cytokines revealed that 19 of these proteins were involved in positive regulation of the immune system indicating an anti-inflammatory effect of gold treatment.

Expanding the knowledge on how cytokine levels are altered after treatment with etanercept could provide valuable information on how exactly the drug affects the inflammatory process in RA and may be helpful in identifying biomarkers for treatment response in patients.

Conclusion and future perspectives

The analysis of serum from rheumatoid arthritis patients in treatment with etanercept resulted in identification of different 79 proteins which had a significantly regulated abundance either after three, six or 12 months of treatment. Out of these proteins, 12 were found to be regulated significantly in the same direction after treatment for three, six and twelve months. Out of those 12 proteins, five had been identified in previous studies as significantly regulated when comparing RA patients to healthy individuals, all of them in the opposite direction of how we found them to be regulated after RA patients were treated with etanercept. Out of these proteins, vitamin D-binding protein, pregnancy zone protein, and plasma kallikrein were especially interesting as candidate biomarkers as these could be related fairly well to RA pathology and confirm some of the effects of etanercept. The consistent regulation across all treatment lengths makes these proteins particularly interesting for the use as biomarkers as it seems they may be regulated early in the treatment process but still maintain the regulation after a longer period of treatment. Future studies regarding the exact mechanisms by which the candidate biomarkers identified in this study influence RA pathology and are influenced by etanercept treatment may help reveal useful information elaborating their potential as biomarkers. Furthermore, the concentration of cfDNA was found to be lowered significantly when comparing the baseline concentration to concentrations after three, six and twelve months of treatment with etanercept. Additionally, the concentration was found to be continuously falling over time. Based on that, lowered concentrations of serum cfDNA was identified as a potential biomarker for etanercept response that may even allow for monitoring of the grade of remission or relapse a patient is experiencing. However, when comparing to the findings of other studies, some doubts exist of how cfDNA levels are regulated in the early response to treatment with bDMARDs but it is plausible that an early rise in cfDNA may indicate that the patient will have a better response than patients not experiencing this rise. Future studies regarding cfDNA concentrations in the early response to etanercept treatment could provide valuable information on the degree to which an early spike in cfDNA concentration can actually predict the long-term treatment response of etanercept.

Finally, the planned inflammatory profiling of the serum was not performed yet. However, results from another study we conducted with OA patients treated with intraarticular gold micro-implants suggest that an analysis like this could provide valuable findings of how etanercept affects the low-abundance proteins like cytokines and may facilitate the identification of new biomarkers of treatment response to etanercept.

References

1. Sardu C, Cocco E, Mereu A, Massa R, Cuccu A, Marrosu MG, et al. Population based study of 12 autoimmune diseases in Sardinia, Italy: prevalence and comorbidity. 2012;7(3):e32487.
2. Guo Q, Wang Y, Xu D, Nossent J, Pavlos NJ, Xu J. Rheumatoid arthritis: Pathological mechanisms and modern pharmacologic therapies . Vol. 6, Bone Research. 2018.
3. Okada Y, Eyre S, Suzuki A, Kochi Y, Yamamoto K. Genetics of rheumatoid arthritis: 2018 status Vol. 78, Annals of the Rheumatic Diseases. 2019. p. 446–53.
4. Song YW, Kang EH. Autoantibodies in rheumatoid arthritis: Rheumatoid factors and anticitrullinated protein antibodies. Vol. 103, QJM. 2009. p. 139–46.
5. Suresh E. Diagnosis of early rheumatoid arthritis: what the non-specialist needs to know. J R Soc Med. 2004. ;97(9):421–4.
6. Bullock J, Rizvi SAA, Saleh AM, Ahmed SS, Do DP, Ansari RA, et al. Rheumatoid arthritis: A brief overview of the treatment. Vol. 27, Medical Principles and Practice. S. Karger AG; 2019. p. 501–7.
7. Rein P, Mueller RB. Treatment with Biologicals in Rheumatoid Arthritis: An Overview. Rheumatol Ther. 2017 Dec;4(2):247–61.
8. Haraoui B, Bykerk V. Etanercept in the treatment of rheumatoid arthritis. Vol. 3, Therapeutics and Clinical Risk Management. Dove Press; 2007. p. 99–105.
9. Deane KD, Demoruelle MK, Kelmenson LB, Kuhn KA, Norris JM, Holers VM. Genetic and environmental risk factors for rheumatoid arthritis. Vol. 31, Best Practice and Research: Clinical Rheumatology. Bailliere Tindall Ltd; 2017. p. 3–18.
10. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum. 2000;43(1):30–7.
11. Kerlan-Candon S, Combe B, Vincent R, Clot J, Pinet V, Eliaou JF. HLA-DRB1 gene transcripts in rheumatoid arthritis. Clin Exp Immunol. 2001;124(1):142–9.
12. Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. Vol. 22, Current Opinion in Rheumatology. NIH Public Access; 2010. p. 293–8.
13. Derksen VFAM, Huizinga TWJ, van der Woude D. The role of autoantibodies in the pathophysiology of rheumatoid arthritis. Vol. 39, Seminars in Immunopathology. Springer Verlag; 2017. p. 437–46.
14. Malmström V, Catrina AI, Klareskog L. The immunopathogenesis of seropositive rheumatoid arthritis: From triggering to targeting. Vol. 17, Nature Reviews Immunology. 2017. p. 60–75.
15. Darrah E, Andrade F. Rheumatoid arthritis and citrullination. Vol. 30, Current Opinion in Rheumatology. 2018. p. 72–8.
16. Zhang JM, An J. Cytokines, inflammation, and pain. Vol. 45, International Anesthesiology Clinics. 2007. p. 27–37.
17. Yoshitomi H. Regulation of immune responses and chronic inflammation by fibroblast-like synoviocytes. Vol. 10, Frontiers in Immunology. Frontiers Media S.A.; 2019.
18. Okroj M, Heinegård D, Holmdahl R, Blom AM. Rheumatoid arthritis and the complement system. Ann Med ;39(7):517–30.
19. Happonen KE, Sjöberg AP, Mörgelin M, Heinegård D, Blom AM. Complement inhibitor C4b-binding protein interacts directly with small glycoproteins of the extracellular matrix. J Immunol. 2009
20. Cojocar M, Cojocar IM, Silosi I, Vrabie CD, Tanasescu R. Extra-articular Manifestations in

- Rheumatoid Arthritis. *Maedica (Buchar)*. 2010 Dec;5(4):286–91.
21. Kourilovitch M, Galarza-Maldonado C, Ortiz-Prado E. Diagnosis and classification of rheumatoid arthritis. *J Autoimmun*. 2014;48–49:26–30.
 22. Köhler, Günther, Kaudewitz, Lorenz. Current Therapeutic Options in the Treatment of Rheumatoid Arthritis. *J Clin Med*. 2019 Jun 28;8(7):938.
 23. Brown PM, Pratt AG, Isaacs JD. Mechanism of action of methotrexate in rheumatoid arthritis, and the search for biomarkers. Vol. 12, *Nature Reviews Rheumatology*. 2016. p. 731–42.
 24. Agger R, Nielsen CH, Leslie G, Aasted B. *Immunologi*. 3. Munksgaard Danmark; 2015. 220–221 p.
 25. Thomson TM, Lescarbeau RM, Drubin DA, Laifenfeld D, de Graaf D, Fryburg DA, et al. Blood-based identification of non-responders to anti-TNF therapy in rheumatoid arthritis. *BMC Med Genomics*. 2015 Dec 3;8(1):26.
 26. Street JM, Dear JW. The application of mass-spectrometry-based protein biomarker discovery to theragnostics. Vol. 69, *British Journal of Clinical Pharmacology*. 2010. p. 367–78.
 27. Mun S, Lee J, Park A, Kim HJ, Lee YJ, Son H, et al. Proteomics approach for the discovery of rheumatoid arthritis biomarkers using mass spectrometry. *Int J Mol Sci*. 2019 Sep 1;20(18).
 28. Wang Z, Ma H, Smith K, Wu S. Two-dimensional separation using high-pH and low-pH reversed phase liquid chromatography for top-down proteomics. *Int J Mass Spectrom*. 2018;427:43–51.
 29. Beer LA, Liu P, Ky B, Barnhart KT, Speicher DW. Efficient quantitative comparisons of plasma proteomes using label-free analysis with MaxQuant. In: *Methods in Molecular Biology*. 2017. p. 339–52.
 30. Lynn AB, Liu P, Ky B, Barnhart KT, Speicher DW. Bioanalytical Chemistry of Cytokines-A Review. *Physiol Behav*. 2016;176(1):139–48.
 31. Frank MO. Circulating Cell-Free DNA Differentiates Severity of Inflammation. Vol. 18, *Biological Research for Nursing*. 2016. p. 477–88.
 32. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012 Sep ;24(3):69–71.
 33. Huang Y, Liu Y, Jiang L, Sun R, Zhang H, Liu R, et al. Apolipoprotein m (APOM) levels and APOM rs805297 G/T polymorphism are associated with increased risk of rheumatoid arthritis. *Jt Bone Spine*. 2014 Jan;81(1):32–6.
 34. Wei Y, Xia N, Zhang W, Huang J, Ren Z, Zhu L, et al. Serum retinol-binding protein 4 is associated with insulin resistance in patients with early and untreated rheumatoid arthritis. *Jt Bone Spine*. 2019 May 1;86(3):335–41.
 35. Dai J, Agelan A, Yang A, Zuluaga V, Sexton D, Colman RW, et al. Role of plasma kallikrein-kinin system activation in synovial recruitment of endothelial progenitor cells in experimental arthritis. *Arthritis Rheum*. 2012 Nov;64(11):3574–82.
 36. Cater JH, Kumita JR, Abdallah RZ, Zhao G, Bernardo-Gancedo A, Henry A, et al. Human pregnancy zone protein stabilizes misfolded proteins including preeclampsia- and Alzheimer's-associated amyloid beta peptide. *Proc Natl Acad Sci U S A*. 2019;116(13):6101–10.
 37. Li T, Yang L, Zhao S, Zhang S. Correlation between apolipoprotein M and inflammatory factors in obese patients. *Med Sci Monit*. 2018 Aug 15;24:5698–703.

38. Baeten JM, Richardson BA, Bankson DD, Kreiss JK, Lavreys L, Mandaliya K, et al. Use of serum retinol-binding protein for prediction of vitamin A deficiency: effects of HIV-1 infection, protein malnutrition, and the acute phase response 1-3. Vol. 79, *Am J Clin Nutr*. 2004
39. Wyatt AR, Cater JH, Ranson M. PZP and PAI-2: Structurally-diverse, functionally similar pregnancy proteins? Vol. 79, *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd; 2016. p. 113–7.
40. Finch S, Shoemark A, Dicker A, Keir H, Smith A, Fardon T, et al. T5 Pregnancy zone protein is released into neutrophil extracellular traps in severe bronchiectasis. In: *Thorax*. BMJ; 2019. p. A3.1-A3.
41. Chao J, Li P, Chao L. Kallistatin: Double-edged role in angiogenesis, apoptosis and oxidative stress. *Biol Chem*. 2017;398(12):1309–17.
42. Duvvuri B, Lood C. Cell-free DNA as a biomarker in autoimmune rheumatic diseases. Vol. 10, *Frontiers in Immunology*. Frontiers Media S.A.; 2019.
43. Hashimoto T, Yoshida K, Hashimoto N, Nakai A, Kaneshiro K, Suzuki K, et al. Circulating cell free DNA: a marker to predict the therapeutic response for biological DMARDs in rheumatoid arthritis. *Int J Rheum Dis*. 2017 Jun 1;20(6):722–30.
44. Lauková L, Konečná B, Vlčková B, Mlynáriková V, Celec P, Šteňová E. Anti-cytokine therapy and plasma DNA in patients with rheumatoid arthritis. *Rheumatol Int*. 2018 Aug 1;38(8):1449–54.
45. Olink Proteomics (PEA). Proximity Extension Assay (PEA) Technology | Olink. 2017 [cited 2020 May 20]; Available from: <https://www.olink.com/data-you-can-trust/technology/#>