Structure Determination of Autoimmune Disease – Related Proteins by High Performance Liquid Chromatography – Mass Spectrometry

Dissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
an der Universität Konstanz

vorgelegt von

Irina Perdivara

Konstanz 2009

Tag der mündlichen Prüfung: Donnerstag, den 06. August 2009
1.Gutachter: Prof.Dr.Dr.h.c. Michael Przybylski
2.Gutachter: Prof.Dr. Martin Scheffner
3.Gutachter: Prof.Dr. Valentin Wittmann
Vorsitzender der Prüfungskomission: Prof.Dr. Gerhard Müller

In the United States there's a Puritan ethic and a mythology of success. He who is successful is good. In Latin countries, in Catholic countries, a successful person is a sinner.

Umberto Eco

I dedicate this work to my great parents Sanda and Minia Perdivara, and to Adi, the love of my life.

The present work has been carried out from June 2006 to January 2009 in the Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry at the University of Konstanz, under supervision of Prof. Dr. Dr. h.c. Michael Przybylski. Part of this work has been performed in the Laboratory of Structural Biology, at the National Institutes of Environmental Health Sciences, Research Triangle Park, North Carolina, USA, under supervision of Dr. Kenneth B. Tomer and Dr. Leesa J. Deterding.

I would like to express my special gratitude and appreciation to:

Prof. Michael Przybylski, for guiding my steps throughout these years and for the continuous encouragement of my research. More so, I am thankful to him for giving me the opportunity to pursue the scientific goals of my work in the research group of Dr. Kenneth B. Tomer at NIEHS, an experience which greatly added to my scientific development.

Prof. Dr. Martin Scheffner (University of Konstanz), for writing the second evaluation of my thesis.

Dr. Kenneth B. Tomer and Dr. Leesa J. Deterding (NIEHS, North Carolina, USA), for mentoring and supporting my research projects, and for the great scientific discussions.

Prof. Dr. Alexandru Cecal (Al. I. Cuza University, Iasi, Romania), for the opportunity he gave me in 2001, to start the study of Chemistry at the University of Konstanz.

Dr. Bernadette Allinquant (Center for Neuroscience, Paris, France) for providing our laboratory with the cDNA encoding for sAPP695 that was used in the present work.

Dr. Robert Petrovich (NIEHS), for the help provided with the expression and purification of the amyloid precursor protein.

Dr. Frederick Miller (NIEHS), for providing the plasma samples for the twin-sibling clinical study.

Dr. Ronald Mason and Dr. Marylin Ehrenschaft (NIEHS), for the interesting discussions regarding the tryptophan oxidation.

Prof. Dr. Alina Zamfir (Aurel Vlaicu University, Arad, Romania), for a wonderful collaboration throughout these years, in particular for sharing her valuable expertise in the field of glycoconjugate analysis.

Dr. Suraj Dhungana, Dr. Jason Williams, Dr. Jeffrey Kuhn, Dr. James Smedley III (NIEHS), for the helpful discussions and for the pleasant working atmosphere.

Ms. Katina Johnson (NIEHS), for our wonderful friendship outside the laboratory.

My dear friends from Konstanz, Claudia Cozma and Marius Iurascu.

Finally, I would like to thank to my parents, my aunt, my uncle and my cousin who supported me throughout these years with their love and good thoughts, and to Adi, the person who believed in me more than anybody else.

This dissertation has been published in part, and presented at the following Conferences:

Peer-reviewed publications

1. Stefanescu R, Iacob RE, Damoc EN, Marquardt A, Amstalden E, Manea M, **Perdivara I**, Maftei M, Paraschiv G, Przybylski M (**2007**), "Mass spectrometric approaches for elucidation of antigen – antibody recognition structures in molecular immunology", *Eur J Mass Spectrom (Chichester, Eng)*, 13(1): 69-75

2. Iacob RE, **Perdivara I**, Przybylski M, Tomer KB (**2008**), "Mass spectrometric characterization of glycosylation of hepatitis C virus E2 envelope glycoprotein reveals extended microheterogeneity of N-glycans", *J Am Soc Mass Spectrom,* 19(3): 428-44

3. **Perdivara I**, Sisu E, Sisu I, Dinca N, Tomer KB, Przybylski M, Zamfir AD (**2008**), "Enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry of long-chain polysaccharides", *Rapid Commun Mass Spectrom,* 22(6): 773-82

4. **Perdivara I,** Deterding L, Moise A, Tomer KB, Przybylski M (**2008**), "Determination of primary structure and microheterogeneity of a β-amyloid plaque specific antibody using high performance LC-tandem mass spectrometry", *Anal Bioanal Chem*, 391 (1): 325-36

5. **Perdivara I**, Petrovich R, Alliquant B, Deterding L, Tomer KB, Przybylski M (**2009**), "Elucidation of O-glycosylation structures of the β-amyloid precursor protein by liquid chromatography – mass spectrometry using electron transfer dissociation and collision induced dissociation", *J Proteome Res*, 8(2): 631-42

6. Ehrenshaft M, Silva SdO, **Perdivara I**, Bilski P, Sik RH, Chignell CF, Tomer KB, Mason RP (**2009**), "Immunological detection of N-formylkynurenine in oxidized proteins", *Free Radic Biol Med*, 46(9): 1260-6

7. **Perdivara I**, Deterding L, Cozma C, Tomer KB, Przybylski M (**2009**), "Glycosylation profiles of epitope-specific anti-β-amyloid antibodies revealed by liquid chromatography – mass spectrometry", *Glycobiology*, in press.

Book chapters

Perdivara I, Iacob RE, Przybylski M, Tomer KB. (**2008**), Site specific identification of N-linked glycosylation in proteins by liquid chromatography – electrospray ionization tandem mass spectrometry, *Applications of Mass Spectrometry in Life Safety*, ISBN 978-1-4020-8811-7, Springer Verlag

Conference presentations

Oral presentations

1. "Identification of *N*-linked glycosylation in proteins by liquid chromatography – mass spectrometry", 1*st Advanced Workshop on Applications of Mass Spectrometry in Life Safety under NATO-Auspices*, Herculane, Romania **2007**

2. "ESI-FTICR MS analysis of modified polysaccharides", *European Fourier Transform Mass Spectrometry Meeting (EFTMS)*, Moscow, Russia **2007**

3. "Mass spectrometric identification of oxidative modifications of tryptophan residues in monoclonal antibodies: artifact or true post-translational event?", *Oxidative Post-Translational Modifications in the Cardiovascular System*, Boston, USA **2008**

4. "Fragmentation pathways of antibody peptides: charge and side chain modification dependence", *56th ASMS Conference on Mass Spectrometry, Denver*, USA **2008**

5. "Mass spectrometric approaches for *de novo* identification of glycosylated structures in proteins", *42nd Annual Meeting of the German Society for Mass Spectrometry*, Konstanz, Germany **2009**

6. “Biopolymer sequencing and antibody proteomics between NIEHS and University of Konstanz”, *20 Years Analytical Chemistry and Mass Spectrometry at the University of Konstanz*, Anniversary Symposium, Germany **2009**

7. “What triggers formation of the $b_{n-1}+H_2O$ ion?”, *42nd Annual Meeting of the German Society for Mass Spectrometry*, Konstanz, Germany **2009**

8. “*De novo* identification of O-glycosylation structures of ß-amyloid precursor protein using electron transfer dissociation and collision induced dissociation”, *42nd Annual Meeting of the German Society for Mass Spectrometry*, Konstanz, Germany **2009**

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Biochemical and immunological features of antibody molecular recognition

1.1.1 Structural and functional properties of the immune system

The immune system represents the body's line of defence against foreign invaders, such as viruses, bacteria, fungi and parasites. The most remarkable property of the immune system is the ability to distinguish between "self" and "non-self", to remember previous challenges and to provide suitable responses. The immune system consists of a vast complexity of finely tuned and organized cells that pass information back and forth to generate a complex network of checks and balances that ultimately result in a specific and effective immune response [1-4].

The central organs of the immune system, the lymphoid organs, are localized throughout the entire body, and include the bone marrow and thymus, as well as the lymph nodes, spleen, tonsils and adenoids, and the appendix. The key players of the immune system are the white blood cells – the lymphocytes – produced in the bone marrow. There are two major types of white blood cells: B cells, which are produced and complete their maturation in the bone marrow, and T cells, which migrate to the thymus where they multiply and become mature, and are "programmed" to distinguish "self" and "non-self" [1].

Two distinct operational mechanisms of the immune system have been described, as schematically depicted in Figure 1.1: the B cell- and T cell- mediated immune response. The first process, also known as humoral immunity, is accomplished by soluble antibodies secreted by the B cells into the body fluids that can interact with circulating antigens, such as bacteria or toxic foreign molecules, but are unable to penetrate living cells. By contrast, in the T cell driven cellular immunity, T lymphocytes directly target cells that have been pathophysiologically transformed , e.g. by invading viruses or by malignancy [1].

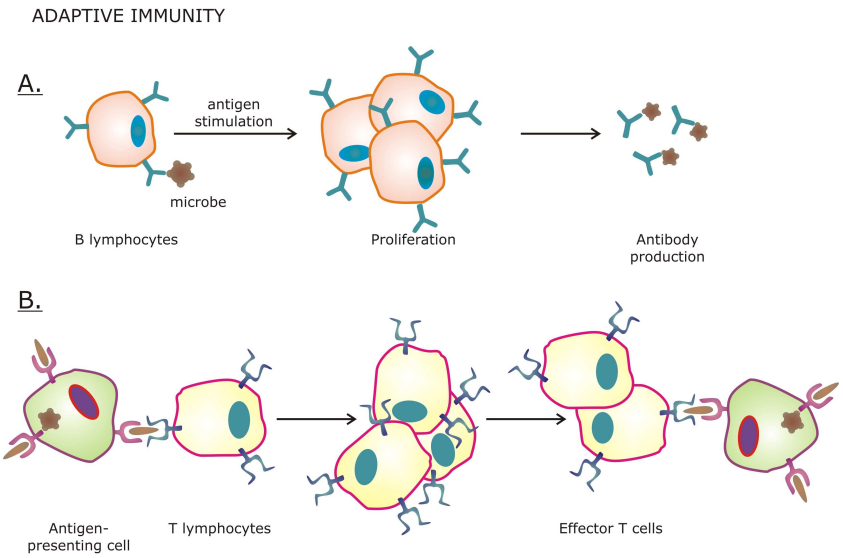


Figure 1.1: Mechanisms of the immune response: (A) humoral immunity, mediated by secreted antibodies vs. (B) cellular immunity in which T cells interact with cells attacked by pathogen. Antigen-presenting cells display peptides of the processed antigen on their surface, via proteins from the major histocompatibility complex.

1.1.2 Immunoglobulin structure and generation of antibody diversity

Humoral immune responses occur when soluble antigens attack the body. Each B cell is programmed to make one specific antibody with a unique protein sequence. When a triggering antigen is encountered, and with participation of helper T cells, a B cell will give rise to many large plasma cells, producing millions of identical copies of a given antibody, that are released into the blood stream (see Figure 1.2).

Antibodies are antigen binding glycoproteins belonging to the immunoglobulin family, divided into five distinct isotypes (IgA, IgM, IgG, IgE and IgD, respectively). The amino acid sequence of the heavy chain constant region determines the antibody isotype, each having distinct structural and functional features. Immunoglobulin G (IgG) represents the most abundant isotype in human plasma, making up to 80 % of the secreted immunoglobulins.

Four distinct polypeptide chains, two identical heavy chains of 440 amino acids each and two identical light chains of 220 amino acids each are assembled in a Y-shaped IgG molecule displaying two identical antigen binding domains on each arm.

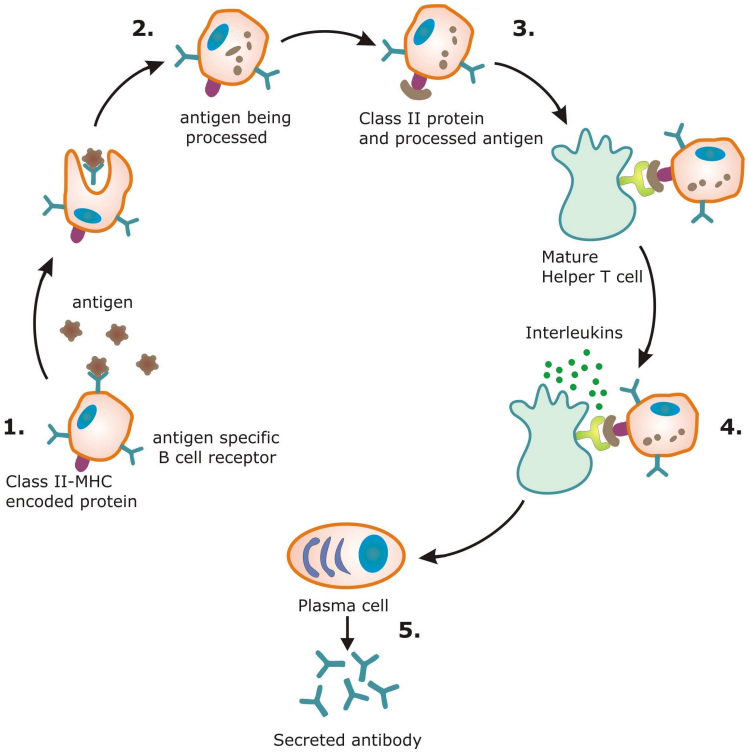


Figure 1.2: Activation of B cells for antibody production: B-cell receptors bind to the circulating antigen (step 1), which is engulfed, processed (step 2) and presented on the cell surface by a MHC II protein (step 3). This complex is bound by a mature helper T cell that releases interleukins (step 4), transforming the B cell into an antibody secreting plasma cell (step 5) [1].

Each light chain contains a variable region (V_L), located between amino acids 1-110, and a constant region between amino acids 111-220 (C_L). Similarly, heavy chains contain a variable region located at the N-terminus (V_H, amino acids 1-110) and three distinct constant regions (C_H1, C_H2, and C_H3, amino acids 111-440). Heavy and light chains are connected by intermolecular disulfide bridges, while within each polypeptide chain intramolecular disulfide bridges give rise to the immunoglobulin fold. Within each variable region, heavy and light chains incorporate three antigen binding sites – the paratope – spanned by antiparallel ß-strands, also termed complementary determining regions (CDRs), as illustrated in Figure 1.3 [4-6].

The immune system has the capability to recognize and respond to approximately 10^7 different antigens. For each antigen the assembly of CDRs represents unique combinations of amino acids, achieved with only a limited repertoire of genes. The heavy chain germ line DNA contains three types of antibody encoding genes: up to 200 variable (V), 12 diversity (D) and 4 joining (J) region genes, scattered along a chromosome. The antibody diversity is the result of three events occurring during B-cell development and maturation, schematically represented in Figure 1.4: *(i)* rearrangement of the V-D-J genes; *(ii)* imprecision at the recombination sites, and *(iii)* somatic hypermutation (SMH) of the

rearranged V_H and V_L genes, which leads to substitution of nucleobases and consequently to amino acid substitutions in the protein [7].

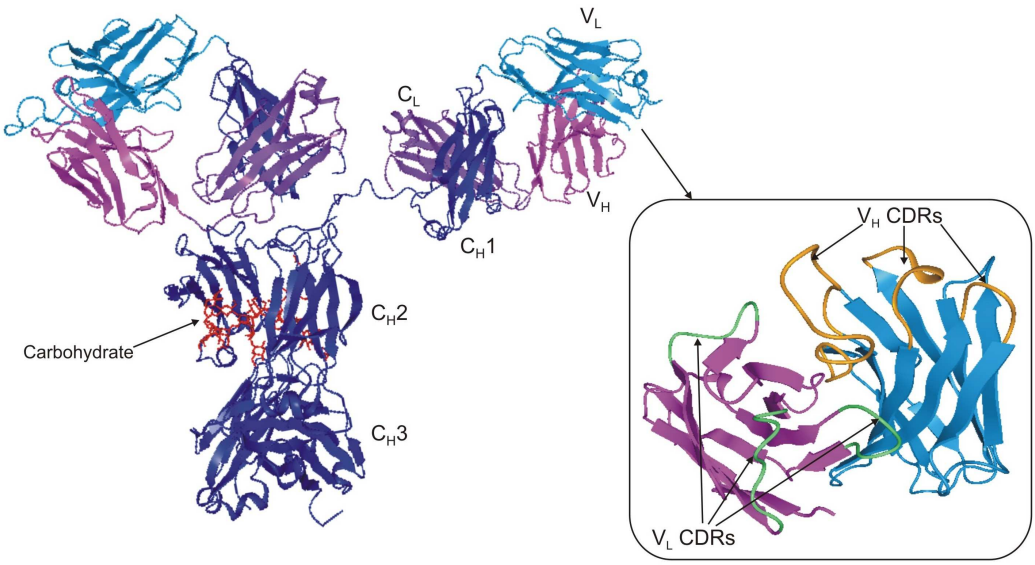


Figure 1.3: Cartoon representation of the IgG architecture, showing the topology of the secondary structure elements of heavy and light chains, and their assembly into individual variable and constant domains. Colour code: dark blue – heavy chain constant regions, C_H1, C_H2 and C_H3; light blue – heavy chain variable region, V_H; dark purple – light chain constant region, C_L; light purple – light chain variable region, V_L. The carbohydrate attached in the C_H2 domain is shown in red. The enlarged view shows the loop structure of the CDRs that alternate with the conserved ß-strands of the variable region. The representation of the antibody with the PDB [8] accession number 1HZH was created with the software BAIIView 1.1.

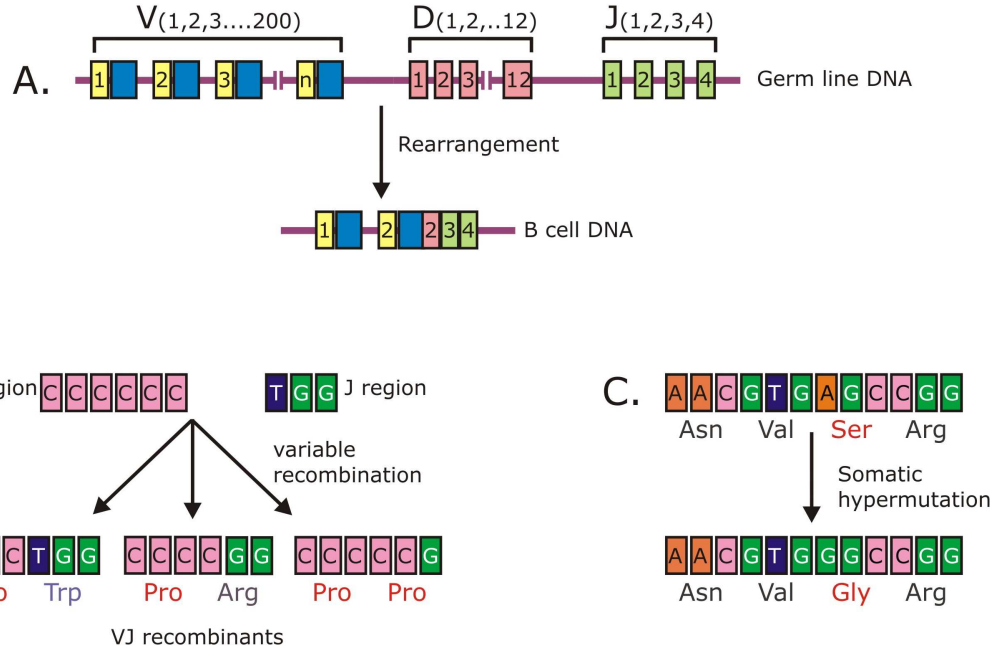


Figure 1.4: Overview on mechanisms generating antibody diversity; (A) V-D-J recombination that takes place during B-cell development; (B) imprecision at the recombination sites of the genes; (C) somatic hypermutation (SHM) occurring in the process of B-cell maturation in the already recombined VDJ genes.

One of the characteristics of immunoglobulins is glycosylation of a conserved Asn residue in the C_H2 domain of the heavy chain constant region - one of the sources of molecular heterogeneity in antibodies (see Figure 1.3). Each heavy chain contains one glycan moiety. The carbohydrate attached at this conserved Asn residue from the F_c region is an essential component required for high affinity receptor binding, representing one of the pathways developed during the immune response [9, 10]. In addition, glycans help to stabilize the immunoglobulin fold by making contact with residues on the protein backbone and with each other within the same molecule [6, 9, 11-14].

1.2 Pathophysiological aspects of immune diseases

When its functionality is altered, the immune system fails to defend the body even against less harmful agents. Hence, in persons born with a defective immune system, an infection with a parasite may be fatal. For example, in severe combined immunodeficiency, the lack of an enzyme leads to formation of toxic products inside the cells of the immune system that ultimately destroy them [15-17], whereas the improper development of the thymus resulting in a reduced number of T-cells represents the basis of the DiGeorge syndrome [18, 19]. Another example is represented by the immunodeficiency with hyper-IgM, characterized by increased production of IgM and by the inability to produce IgG and IgA [20, 21]. Ultimately, the most threatening immune disease of our age – the acquired immunodeficiency syndrome (AIDS) – is the final result of infection with the human immunodeficiency virus (HIV); HIV attacks the CD4+ lymphocytes – a type of helper T cells, and progressively destroys the body's immune system [22, 23]. Most immune diseases result from either an excessive immune response (such as asthma, familial Mediterranean fever, Chron's disease), or from an autoimmune attack [24-26].

1.3 Current state of knowledge in the field of autoimmunity

A key property of a healthy immune system is to prevent autoimmunity, defined as the reaction of the body against the body itself. This capability is enabled by specific marker proteins, carried by each cell of an organism, and encoded by genes of the major histocompatibility complex (MHC). Each individual carries a distinct set of marker MHC proteins, which dictates to his/her immune system not to react with own organs or tissues. When the distinction between "self" and "non-self" is not successfully performed, the

immune system becomes self-reactive, launching attacks against body's own components [27].

The effector mechanisms of autoimmunity are analogous to those used to combat exogenous agents and include soluble products such as autoantibodies recognizing self proteins – autoantigens –, as well as cell mediated autoresponses (see Figure 1.5). In most cases, autoantibodies elicit a damaging cascade of events, leading to inflammations of organs or tissues. During the 1990s it has been recognized that similar immune mechanisms were involved in more than one autoimmune disease. For example, activation of the CD4+ type 1 helper T cells (see Figure 1.5) was shown to be important in the pathogenesis of rheumatoid arthritis (RA), type I diabetes and multiple sclerosis (MSc), although the antigenic specificities of these cells are quite different. Furthermore, similar genetic risk factors have been observed for RA, MSc, lupus erythematosus and inflammatory bowel disease [28, 29].

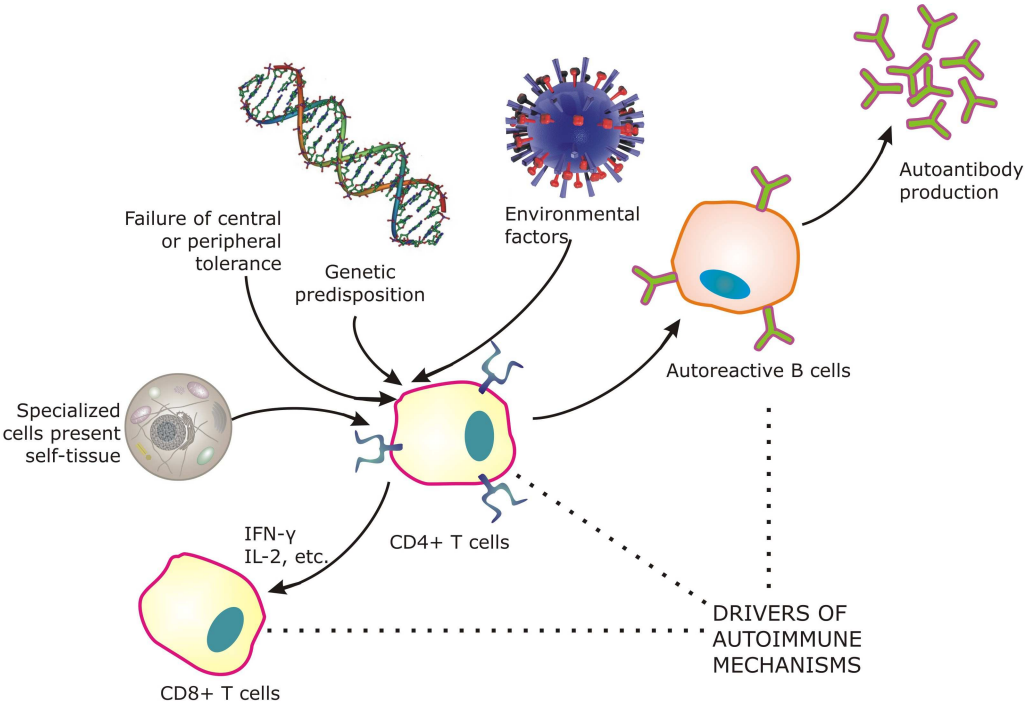


Figure 1.5: Overview on events involved in autoimmunity: both humoral and cell-mediated autoimmune responses result from a combination of genetic and environmental triggers, as well as from failure in body's self tolerance. The products of humoral autoimmunity are autoantibodies reactive to autoantigens leading to a cascade of inflammatory processes.

Autoimmune diseases comprise a broad variety of phenotypes and set of symptoms – roughly 80 separate disorders –, each with an unique genetic-environmental trigger and each working through a specific pathological mechanism. Up to 8 % of the population, of which 78 % are women suffer from autoimmune syndromes [30], the overall cost for their

treatment approximating that of cancer and heart diseases. Most importantly, this group of diseases lacks the immediate recognition impact, as non-specific symptoms often resemble manifestations of unrelated disorders; hence they are frequently misdiagnosed. Their social impact is enormous, life quality of individuals and their families being profoundly affected, while treatment still remains a challenge. Figure 1.6 shows the incidence by sex for some most common autoimmune diseases.

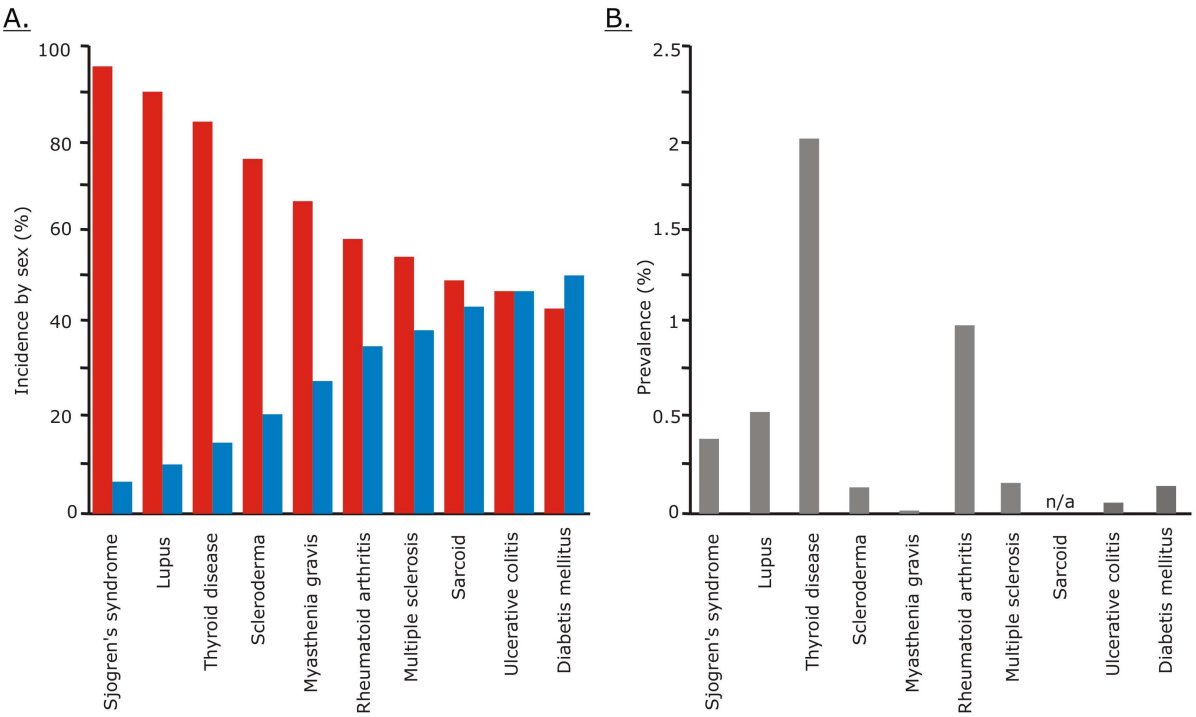


Figure 1.6: (A) The sex distribution of the major autoimmune diseases, indicating that females are more affected than males [28]; (B) Prevalence of major autoimmune diseases in North America, according to National Institute of Allergy and Infectious Diseases (NIAID/NIH). Once diagnosed, these autoimmune syndromes represent life-long conditions, whereas in the case of sarcoidosis, the symptoms disappear spontaneously within two years after onset, i.e. prevalence can not be determined.

For some autoimmune diseases, such as RA and MSc, remission of disease activity has been observed in pregnancy [31-35]. The observation that the immune response is different between males and females [36, 37], the increased prevalence of autoimmune diseases in women and the modulatory effects of sex steroids on the immune response has focused attention on the role of these hormones as mediators of the sex differences.

To date, no single mechanism has been described to account for the diversity of autoimmune responses, or production of autoantibodies. The general hypothesis is that the onset of an autoimmune condition occurs in genetically predisposed individuals exposed to environmental risk factors [38].

In the following parts, the current state of knowledge in the field of two distinct autoimmune-type diseases will be described: *(i)*, an overview of Alzheimer's disease and its controversial autoimmune features (Chapters 1.4 and 1.5), and *(ii)*, pathophysiological characteristics of idiopathic inflammatory myopathies (Chapter 1.6).

1.4 Age-related neurodegeneration of autoimmune nature: Alzheimer's disease

> "Every 70 seconds, someone in America develops Alzheimer's disease. By mid-century, someone will develop Alzheimer's every 33 seconds." [39]

As the life expectancy of individuals has continued to increase, there has been a concomitant increase in diseases primarily associated with appearance late in life. Age-related dementia is a major category of such diseases and Alzheimer's disease (AD) is one of the most widely known and most widely feared neurodegenerative diseases. The dramatically increased life span has promoted AD to the 6th leading cause of death across all ages in the United States in 2006, whereas in Europe, AD is rated as the 3rd cause of death after heart diseases and cancers in 2009. While the total numbers attributed to other major causes of death such as heart diseases, breast and prostate cancer, and stroke, have declined over the past several years, those due to AD have continued to increase, with recent statistics indicating the rate of death has increased by 47 % from 2000 to 2006 [39].

Alzheimer's disease is characterized by the gradual loss of memory and other cognitive abilities. Memory difficulties, apathy and depression are often early clinical symptoms, while later symptoms include impaired judgement, disorientation, confusion, and behaviour changes. In advanced Alzheimer's people need help with bathing, dressing, eating and other daily activities [40].

Although the causes of AD are still poorly understood, most experts agree that Alzheimer's develops as a result of multiple factors rather than a single cause. The greatest risk factor for AD is advancing age. Most individuals with AD are aged 65 or older, and the risk of developing AD has been reported to increase 50 % by the age of 85 [39] (see following paragraphs for a discussion of genetic mutations of the amyloid precursor protein). A

genetic factor related to increased risk of AD late in life is apolipoprotein-E4 (APOE-e4), one of three common forms of APOE gene, which provides the blueprint for a protein that carries cholesterol in blood. While everyone inherits one of the APOE genes from each parent, individuals that inherit one or both APOE-e4 genes are at higher risk [41, 42].

There is currently no treatment available to stop or slow the cognitive decline in AD. The U.S. Food and Drug Administration has approved five drugs that temporarily alleviate the symptoms of neurodegeneration for 6 to 12 months for about half of the patients. These are acetylcholinesterase inhibitors (Aricept®, Razadyne® and Exelon®) and regulators of glutamate activity (Namenda®), two small molecules involved in learning and memory [43, 44]. Hence, the present lack of effective therapeutic agents requires the exploration of alternative approaches for treatment and prevention.

Post-mortem examination of the brains of AD patients shows dramatic losses of brain mass, most severe in hippocampus, temporal and parietal lobes. Neuronal loss, intra- and extracellular protein accumulation, as well as microvascular angiopathy are histopathological characteristics of AD. Intraneuronal neurofibrillary tangles consisting of filaments of hyperphosphorylated forms of *tau* proteins twisted around microtubules, cause disintegration and hinder propagation of electric signals [45, 46].

A major pathological feature of AD is the accumulation of extracellular plaques that contain aggregates of the neurotoxic ß-amyloid (Aß) polypeptide as major components, surrounded by astrocytes and activated microglia [47]. Beta-amyloid comprises a group of polypeptides of 38-43 residues with partially N- and C-terminally truncated sequences, which are formed by proteolytic cleavage of the amyloid precursor protein (APP), a transmembrane protein involved in synaptogenesis and neuronal plasticity. As schematically shown in Figure 1.7, overproduction and/or accumulation of Aß results in the formation of extracellular deposits.

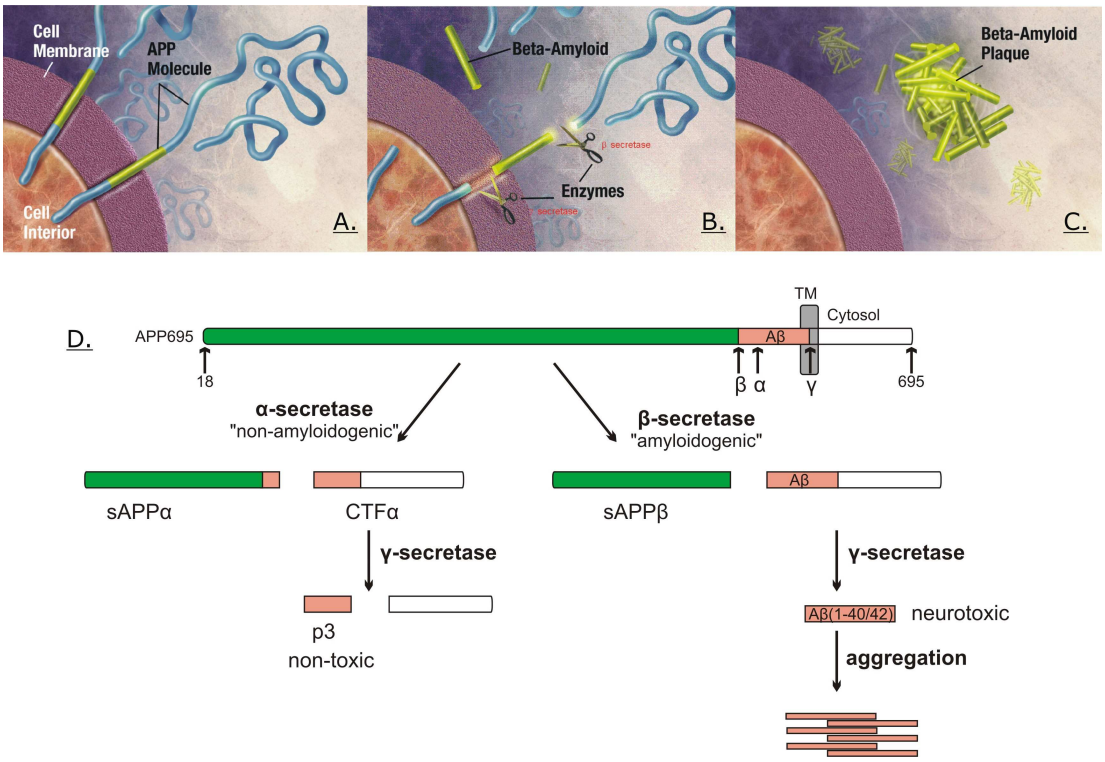


Figure 1.7: Proteolytic processing of the amyloid precursor protein (APP695) and formation of the ß-aggregates in AD: (A) transmembrane amyloid precursor protein, (B) proteolytic formation of ß-amyloid peptide, (C) extracellular accumulation and aggregation of Aß-peptide, and (D) schematic representation of proteolytic processing pathways of APP695: "non-amyloidogenic" pathway producing sAPPα and CTFα fragments (left), and "amyloidogenic" pathway, releasing the neurotoxic, plaque forming Aß-peptides (right).

APP was not initially discovered for its physiological role in many organs and cells, but rather because of the characteristic deposition of Aß-containing plaques in AD [48]. The function of APP is only partially understood at present. There are three major isoforms of APP, containing 695, 751 and 770 amino acid residues (designated as APP695, APP751 and APP770, respectively) [49], which are derived from alternative splicing of the mRNA of a single gene located on chromosome 21 [50]. In nerve cells, the predominantly expressed isoform is APP695, while isoforms APP751 and APP770 predominate in other cell types [51]. Three proteolytic enzymes, (denoted α-, ß- and γ-secretases), are involved in the proteolytic degradation of APP and the secretion of soluble APP (sAPP) forms [52, 53] (see Figure 1.7 D). In the "non-amyloidogenic" pathway, cleavage by α-secretase occurs at the Lys-16 residue downstream of the N-terminal of Aß, releasing the sAPPα fragment and the C-terminal transmembrane fragment of APP, CTFα [54, 55]. The soluble extracellular domain sAPPα derived from the non-amyloidogenic processing pathway, acts as a growth factor in many cell types and promotes neuritogenesis in post-mitotic neurons [49]. In vivo, infusion of sAPPα into the brain increases synaptic density, protects hippocampal neurons against ischemic injury and enhances memory performance [56-58]. Alternatively, cleavage of APP by ß-secretase(s) at the C-terminal end of sAPPß (Lys-Met

in human physiological APP) leads to the formation of the neurotoxic Aß peptide(s), and has been denoted as the “amyloidogenic” pathway. Several pathogenic mutations in the APP gene leading to increased Aß production and early onset of non-age related AD (around the age of 50) have been identified. The "Swedish" mutations comprise the simultaneous substitutions Lys670Asn and Met671Leu in full-length APP770, which were found to enhance the ß-secretase cleavage rate yielding elevated levels of sAPPß. These mutations represented the basis for the development of corresponding transgenic mouse models of AD [59-61].

During transit through the intracellular protein secretory pathway, APP has been shown to undergo multiple post-translational modification, such as by *N*- and *O*-glycosylation, phosphorylation and tyrosine sulfation [62-64]. It has been suggested that *N*- and *O*-glycosylation of the extracellular domain of APP are prerequisites for phosphorylation of Thr668 of the cytoplasmic domain during neuronal differentiation [65], and for the proteolytic cleavage by secretases [66]. Moreover, mutants defective in *O*-glycosylation have been reported to exert an altered cellular metabolism compared to wild type, physiological APP695 [66]. The structure-function relation of APP and Aß as its key AD peptide, have not been elucidated in molecular detail. The pathophysiological importance of this relation to Aß and amyloid deposits in neurodegeneration is underlined by the above-mentioned early-onset cases of AD, with mutations in the APP gene chromosome 21, and in presenilin genes that lead to increased production of Aß [67-69]. Furthermore, patients with Trisomy 21 (Down syndrome), having an extra copy of the APP gene, develop large number of plaques at early age [70-72]. These results and the present lack of knowledge of the molecular structure of APP have been a major motivation for the structural studies of APP in this thesis.

1.5 Immunotherapeutic and diagnostic approaches using Aß specific antibodies in Alzheimer's disease

Recent data from transgenic mouse models of AD suggest that Aß clearance via immune mediated pathways may have a major role on the development of plaques [73]. Over the past decade, Aß and the senile plaques have been the target of several immunotherapeutic approaches aimed at prevention of subsequent amyloid deposition

[74-78]. Initially, therapeutically active antibodies produced by active immunization with protofibrillar Aß(1-42) were found to reduce the amyloid burden and to restore cognitive functions in transgenic mouse models of AD [79]. These antibodies against Aß in the immunized TgCRND8 mice recognize with high specificity a short epitope sequence located at the N-terminus of ß–amyloid (FRHDSGY), as first identified by proteolytic epitope excision and high resolution mass spectrometry [79] (Figure 1.8). However, a therapeutic trial by immunization of AD patients with Aß(1-42) was discontinued because patients developed severe meningo-encephalitic inflammations [80, 81]. Supportive data has shown that passive immunization with antibodies against Aß also reduces amyloid deposition and may even clear existing plaques [82].

The analytical approach of proteolytic epitope excision, discussed in Chapter 1.7.1, is based on the protection of molecular recognition structures in antigen-antibody complexes against chemical reagents. This feature enables differential proteolytic degradation of the non-binding regions of an affinity-bound antigen, whereas the antigenic determinant – the epitope – remains shielded, and can be subsequently dissociated and identified by mass spectrometry.

Recently, physiologic antibodies against Aß were found in the cerebrospinal fluid (CSF) and plasma of AD patients and healthy controls [83-86]. These Aß-autoantibodies are produced without specific immunization, independent of exposure to foreign antigens. There is evidence that in the context of neurodegenerative diseases naturally occurring antibodies may be involved in physiologic clearance of misfolded proteins. Preliminary data suggests that fibrillation and neurotoxicity of Aß peptides is reduced in presence of Aß-antibodies as they might contribute to peripheral and central degradation of Aß and to inhibition of plaque formation [87]. A comparative representation of immune-mediated amyloid clearance pathways of “plaque – specific” and “plaque – protective” Aß autoantibodies is shown in Figure 1.8 A.

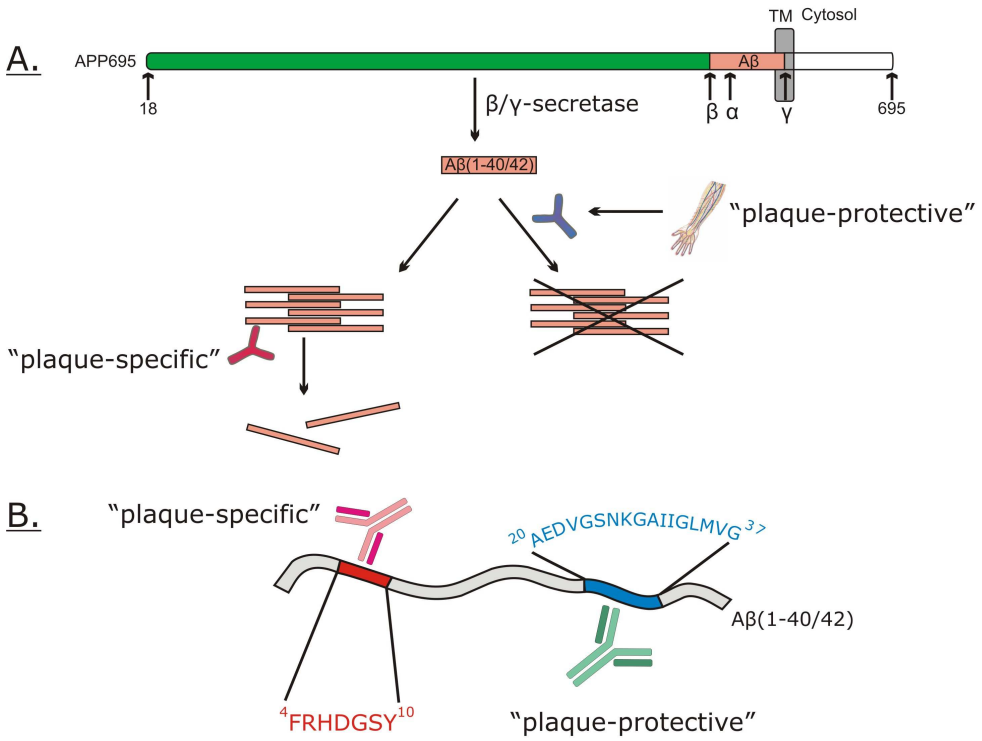


Figure 1.8: (A) Differential epitope recognition of "plaque-specific" (red) and "plaque-protective" antibodies (blue). Whereas therapeutic (plaque specific) antibodies recognize Aβ in plaques and trigger disaggregattion, natural autoantibodies (plaque-protective) bind soluble Aβ monomers and inhibit fibrillation (B) Molecular details and localization of Aβ-epitopes recognized by therapeutic antibodies (red) and natural autoantibodies (blue).

A key property of Aß-autoantibodies is derived from their epitope specificity, recently elucidated by epitope excision – mass spectrometry in our laboratory [88-91]. The antigenic determinant recognized by natural autoantibodies was determined as the middle region of Aß(1-40), residues 21-37, involved in fibrillation and plaque formation. Binding of this epitope in soluble Aß monomers by autoantibodies is thought to exert a protective role in healthy individuals, by preventing its oligomerization. The CSF titer of anti-Aß antibody was found lower for AD patients than for non-demented individuals, suggesting that decreased levels of these antibodies may be a factor contributing to pathogenesis of AD [83, 92]. As a result of these observations, commercial intravenous immunoglobulin preparations (IVIg) containing naturally occurring Aß-autoantibodies have received increased attention as a potential therapeutic concept for AD [93]. In December 2008 a phase III clinical trial using 10 % IVIg was initiated in patients with mild to moderate AD. The protective role of natural Aß-antibodies, possibly derived from their defined epitope specificity, as well as their potential use as therapeutic and early-diagnostic tool reinforces their detailed molecular characterization, which represented a major goal of this thesis.

1.6 Idiopathic inflammatory myopathies: polymyositis and dermatomyositis

The idiopathic inflammatory myopathies (IIM) are a heterogeneous group of systemic autoimmune syndromes characterized by chronic muscle weakness and skeletal muscle inflammation. Although initially described more than a century ago, these diseases are rare and heterogeneous, that to date there is only a limited understanding of their molecular cause and treatment. Of the IIM, the best recognized subsets of diseases are polymyositis (PM); dermatomyositis (DM) and inclusion body myositis (IBM); together these have an incidence of 0.001 % (DM>IBM>PM), among which females are more affected [94, 95].

Polymyositis affects predominantly adults who present subacute or chronic proximal muscle weakness, elevated creatine kinase and mononuclear cell infiltration [96] (see Figure 1.9 A-B). Dermatomyositis, affecting both children and adults, causes a purple discoloration of the eye lids, edema around the eyes and the mouth, skin rashes on the face and upper body extremities, muscle pain and weakness.

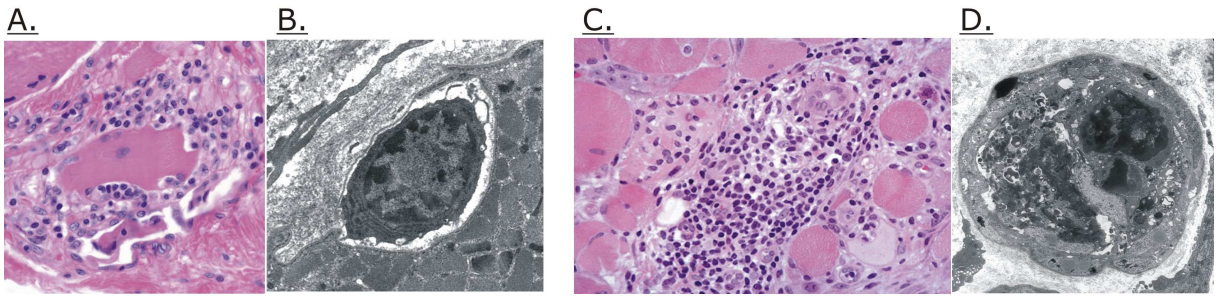


Figure 1.9: (A) Muscle biopsy in polymyositis (PM) showing the muscle fibers (pink) being attacked by inflammatory cells (purple) ; (B) Lymphocyte infiltration in myofiber in PM; (C) Muscle biopsy in dermatomyositis (DM); (D) Necrotic capillary in DM.

The key pathological feature in DM is a vasculitis which attacks the capillaries and arterioles in the connective tissue around and within the muscle fibres. Capillary loss is derived from endothelial swelling and necrosis (see Figure 1.9). This inflammatory condition is caused by circulating anti-endothelial antibodies which interact with vascular antigens, activating the complement and leading to formation of the membranolytic attack complex [96] that ultimately destroys endothelial cells.

The borders separating myositis from related syndromes within the category of rheumatologic diseases are not sharply defined. For example, systemic lupus erythematosus (SLE) shares both clinical and serological manifestations with myositis;

furthermore, several conditions treated as myositis represent rare, or as-yet unrecognized genetic metabolic disorders, characterized by deficient production of enzymes vital to the energy production of a muscle cell [96]. The diagnosis relies on histological features of muscle biopsy indicating muscle degeneration [97], such as mononuclear inflammatory infiltrates in PM, whereas in case of DM, additional heliotrope sign or papules represent characteristic manifestations. Furthermore, a strong family history of autoimmune diseases [98] and a clinical response to immunosuppressive therapy [99] may help confirm the diagnosis.

1.6.1 Humoral and cellular autoimmunity in myositis

The autoimmune nature of myositis was derived from the presence of autoantibodies to nuclear and cytoplasmic antigens occurring in 80-90 % of the patients with DM and PM. From these, almost 50 % have defined autoantibodies of recognized specificities, referred to as myositis specific autoantibodies (MSA), which occur almost exclusively in myositis. By contrast, myositis associated autoantibodies (MAA) lack disease specificity, as they are present in overlapping syndromes and also in healthy individuals; hence their clinical utility is limited [100, 101]. Myositis autoantibodies are usually present from the earliest stage of the disease and persist over time even when the disease is controlled or in remission [102, 103].

The most common established MSA is anti-Jo-1, present in approximately 20 % of patients with myositis [104]. Anti-Jo-1 reacts with the cytoplasmic enzyme histidyl-transfer-ribonucleic acid (-tRNA) synthetase, which catalyzes binding of histidine to its tRNA [105]. A recent study using a mouse model of myositis suggests that production of anti-Jo-1 is a secondary event, appearing as a consequence of myositis, rather than requiring a primary stimulus, such as a virus, directing the immune response to Jo-1 antigen [106]. Other MSAs include autoantibodies to the signal recognition particle, anti-SRP, detected in 4 % of the myositis patients [107], and anti-Mi-2, in 5-10 % of the patients [107-109]. Anti-Mi-2 autoantibodies have high myositis specificity, and 90-95 % of the patients have the characteristic rashes of dermatomyositis. Mi-2 is a nuclear protein, involved in the chromosomally mediated regulation of transcription by an ATP-dependent mechanism [110]. Clinical data suggests that patients with different MSAs show distinct features of muscle histology [108], although it is unclear whether myositis specific autoantibodies play

a role in disease pathogenesis. One hypothesis for the occurrence of these autoantibodies against cellular components is that autoantigens undergoing proteolytic cleavage during apoptosis may generate fragments containing new epitopes which may be displayed on the cell surface in apoptotic blebs inducing an autoimmune response [111, 112].

Cell-mediated immune responses are believed to play a pathogenic role in myositis [113]. Histopathologic examination of muscle biopsy in patients with dermatomyositis showed increased number of T helper cells (CD4+) and B lymphocytes in perivascular regions of the muscle. In contrast, muscle biopsy of individuals with polymyositis showed infiltration of cytotoxic T cells (CD8+) and macrophages in the endomysial region [114]. Recent studies documented a characteristic pattern of T-cell receptor gene expression in clinical groups with myositis, suggesting that an antigen-driven immune response might account for the restricted repertoire of the expressed genes [115].

The pathogenesis of myositis is largely unknown at the present. Recent findings suggest that environmental agents acting on genetically susceptible persons lead to physiologic responses involving immunological activation and subsequent tissue damage in myositis syndromes [96]. Primary therapy, aimed at reducing the muscle inflammation, involves the use of corticosteroids and immunomodulatory agents; however, many factors have to be considered in order to achieve successful therapy responses [116]. Hence, further research is necessary in this field for a better understanding of both genetics and environmental risk factors of autoimmunity. A clinical study recently initiated at the National Institutes of Environmental Health Sciences in Bethesda is aimed at delineating contributing factors from these two categories, by analyzing myositis patients, their unaffected twins/siblings, as well as unrelated controls. The design of the twin-sibling study is presented in Chapter 2.3. Driven by the current state of knowledge in the field of myositis, and in an effort to compare its immunological features with other autoimmune disorders, part of the clinical study, and a major goal of this thesis was to determine the subclass specific glycosylation profiles of total plasma IgG in myositis patients.

1.7 Analysis of biopolymer structure and post-translational modifications by mass spectrometry

Unveiling structure – function relationships of biologically relevant molecules has been a main theme of analytical biopolymer chemistry for decades, and an ultimate goal of all recent analytical developments is to contribute to the understanding of life. The classic techniques of structure determination, X-ray crystallography and NMR, have established important milestones in life sciences. Some outstanding examples of results in X-ray crystallography are noteworthy, such as the structural elucidation of the ribosome, nucleasome, photosystem I, GroEL – GroES, a bacterial potassium ion channel [117-121]. NMR has revealed information about dynamic processes that occur in solution, for example when a ligand binds to a protein; these results owe to the Nobel Prize-awarded observation that the Nuclear Overhauser Effect can be exploited to map networks of near-by atom pairs that are not connected through covalent bonds [122-124].

Over the past two decades, mass spectrometry (MS) has become an essential tool in structural biology, due to its ability to provide molecular structure information and to complement other analytical methods. Milestones in MS have been recently achieved with the introduction of soft-ionization techniques, electrospray (ESI) [125, 126] and matrix assisted laser desorption ionization (MALDI) [127], that enabled analysis of large biopolymers in gas-phase with high sensitivity and accuracy. Factors such as high sensitivity, low sample consumption, low analysis time and applicability to mixtures make MS a method of choice for many analytical problems in life sciences, where conventional methods reach their limits. In the electrospray process (reviewed in reference [126]) multiply charged analyte molecules are produced by spraying a solution containing the molecules of interest through a thin needle that has a potential difference applied to it, with respect to the counter electrode. The analyte can be an intact protein, small peptide, DNA, lipid or carbohydrate. As very low residual energy is retained on the analyte upon ionization, ESI represents a soft ionization technique capable of preserving the tertiary structure of proteins in the gas phase. This enables the analysis of intact protein complexes [128-130] and, most recently, whole virus assemblies [131-133]. Moreover, in the absence or complementary of NMR or X-ray data, information about surface topology, folding pathways of proteins and protein – protein interactions can be derived from mass spectrometric data. Several approaches have been developed, including differential chemical modification of specific amino acid residues in proteins [134-138] and oxidative

protein footprinting [139-143]. These are based on the assumption that the rate of modification of a specific amino acid is dependent on the inherent reactivity of a side chain and on its solvent accessibility, i.e. surface exposed residues react at a higher rate than residues berried in a hydrophobic pocket. In addition, the use of bifunctional cross-linking reagents can provide information regarding distances between residues in protein [144-146], reviewed in [147].

1.7.1 Mass spectrometric approaches for structural characterization of antibodies

Affinity – mass spectrometry methods, initially developed by our laboratory [136, 148, 149] have now been successfully established by many research groups for the identification of molecular recognition structures in proteins and other biopolymers, such as antigenic determinants – epitopes – recognized by antibodies paratope regions [150-152] (reviewed in [153]). The method, referred to as proteolytic epitope excision, involves the covalent immobilization of a biopolymer, e.g. an antibody or a ligand on a stationary phase. The antigen or a mixture containing the antigen is presented to the column, resulting in an affinity-bound antibody – antigen complex. The basis for this approach is that the molecular recognition structures involved in the antigen – antibody interactions – the epitope and the paratope – are shielded against chemical reagents, such as proteolytic enzymes. Differential proteolytic digestion of the immuno-complex will result in antibody bound peptide epitope(s) that can be subsequently dissociated and identified by MS (see Figure 1.10). In the Aß-specific antibodies, this methodology was applied to identify the ß-amyloid epitopes recognized by Aß-antibodies with distinct serological and therapeutic properties [79, 91]. The results explained the major differences between plaque-resolving and plaque-protective antibodies, whereas follow-up studies pursued the development of immunotherapeutic agents containing the epitope lead structures defined by MS [154, 155].

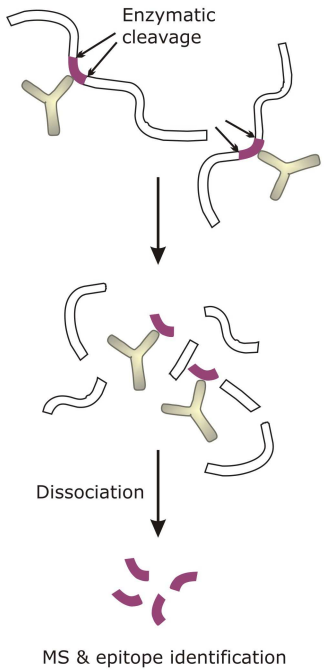


Figure 1.10: Principle of mass spectrometric epitope identification by proteolytic excision of epitopes in immune complexes: the antigen-antibody complex is subjected to enzymatic degradation, leaving the antigenic determinant affinity-bound to the antibody complementary determining regions. Following the dissociation step of the complex, MS analysis provides the molecular mass information of the epitope.

Most recently, affinity-MS has been employed for the first time in the reversed fashion to identify the paratope of the camel anti-lysozyme antibody cAbLys3 [156]; camel antibodies lack the light chains and the heavy chain CDR3 is considerably larger than the corresponding region in human or mouse immunoglobulins [157]. In this experimental setup an intact antibody (paratope excision) or its enzymatic mixture (paratope extraction) is presented to an affinity column containing the covalently immobilized epitope peptide. Several challenges are associated with this approach, such as *(i)* the proteolytic stability of immunoglobulins, which may require their partial denaturation for an efficient enzymatic degradation, *(ii)* the simultaneous involvement of all six antibody CDRs in binding, that may result in loss of affinity of the peptides containing individual regions derived from enzymatic procedures during paratope extraction; and *(iii)* the lack of knowledge of most antibody sequences, derived in part from the complexity of the molecular mechanisms generating antibody diversity. Furthermore, mutations, truncations and post-translational modifications may complicate the analysis.

An alternative of the paratope excision/extraction approach is the primary structure determination of antibodies by mass spectrometry. Several MS-based approaches,

including so called “top-down” and “bottom-up” [158-160], are now routinely applied in the pharmaceutical industry for quality control purposes [161] (see Figure 1.11).

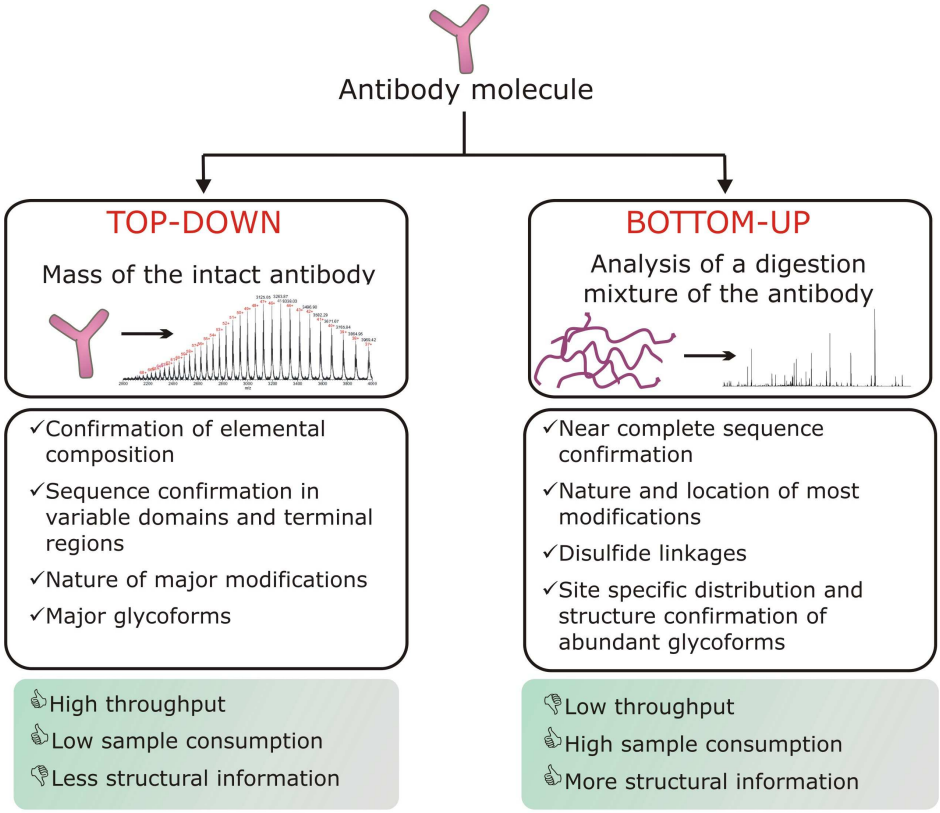


Figure 1.11: Principles of “top-down” (left) and “bottom-up” analysis (right) for structural characterization of antibodies.

Top-down MS refers to an approach where the intact molecule is introduced into the mass spectrometer without enzymatic or chemical fragmentations in-solution, and structural information is obtained from the fragmentation pattern of the intact molecule inside the mass spectrometer. Structural information can be obtained for variable and terminal regions of the antibody; however, its size represents a major challenge for this method. In the bottom-up approach, the protein is digested into small peptides by a protease, followed by liquid chromatography – tandem MS analysis (LC-MS/MS). The antibody sequence is assembled based on the molecular mass and sequence information of individual peptides and predicted amino acid sequence. A representative example of recombinant immunoglobulin is rituximab, a chimeric anti-CD20 mouse/human IgG_1 monoclonal antibody produced in CHO cells – the first therapeutic antibody approved for the treatment of non-Hodgkin’s lymphoma. When the antibody's cDNA information is unknown, as is the case of immunoglobulins produced by hybridoma technology, primary structure determination may become a challenging task. Recently, we showed that *de novo*

interpretation of tandem MS data in combination with N-terminal Edman sequencing of proteolytic antibody peptides can be successfully employed for primary structure determination of antibodies of unknown sequence [162].

Following protein biosynthesis in the ribosomes, a large diversity of post translational modifications (PTMs) may occur expending the proteome diversity. These events finely tune proteins to effect specific biological functions. They usually occur at sub-stoechiometric levels and may be highly heterogeneous. Certain PTMs are permanent, such as *N*-glycosylation of asparagine residues, whereas others are transient, such as acetylation and methylation, phosphorylation or *O*-linked *N*-acetyl glucosamine, serving in the cellular signalling pathways. Knowledge of sites and nature of such modifications in proteins represent critical information for understanding of protein structure and function. Whereas unmodified proteins can be studied by X-ray crystallography and NMR, these methods often fail to provide satisfactory information about post-translational modified proteins. For example, attachment of a carbohydrate moiety on a protein may lead to increased flexibility of the backbone around the attachment site, such that glycans often need to be trimmed to obtain decent diffracting crystals [163].

In the last two decades, tandem mass spectrometry using collision induced dissociation (CID) and electron capture/transfer dissociation (ECD/ETD) [164-167], described in detail in the following chapters, has emerged as a powerful tool for identification and molecular characterization of PTMs, providing information about their site-occupancy and site-specific microheterogeneity. Hyphenated methods combining separation techniques and (tandem) mass spectrometry are particularly suitable for analysis of complex mixtures containing modified and non-modified peptides, as they offer an additional “degree of freedom” to the analytical measurement. In addition, the dynamic range, i.e. the ratio of highest to lowest sample concentration that can be detected under identical analytical conditions, has been significantly extended.

1.8 Scientific goals of the thesis

The current state of research in the field of neurodegenerative and autoimmune syndromes suggests a potential role of both humoral and cell-mediated immune system in disease pathogenesis. Autoantibodies have been observed in both Alzheimer’s disease

and myositis; however, their etiology appears to be quite different, as Aß-autoantibodies exert neuroprotective properties, whereas myositis specific autoantibodies are associated with a broad array of clinical manifestations. With respect to the autoantigens in both disorders, the mechanisms leading to the autoantigenic characteristics of ß-amyloid and of the cytoplasmic and nuclear proteins in IIM remain unclear. Consequently, the development of efficient therapies will rely on a more complete understanding of the structural basis of the immune response. In the present thesis, methods of high performance liquid chromatography in combination with tandem mass spectrometry have been employed for structure determination of Aß- and myositis specific proteins.

The scientific goals of the present dissertation are summarized as follows:

1. Analytical development and bioanalytical application of high performance liquid chromatography – mass spectrometry to the complete primary structure determination of Aß-specific antibodies. In this part, tandem-MS and *de novo* sequence analysis have been employed for *(i)* determination of the amino acid sequence and sequence microheterogeneities, and *(ii)* determination of glycosylation structures of anti-ß-amyloid autoantibodies and of several related Aß-monoclonal antibodies, with biomedical relevance for immune therapy in Alzheimer's disease.

2. Primary structure characterization of the secreted amyloid precursor protein, with focus on *de novo* identification and determination of *O*-glycosylated structures of sAPP695, using high performance liquid chromatography in combination with electron transfer dissociation and collision induced dissociation mass spectrometry.

3. Isolation and structural characterization of the plasma immunoglobulin fraction from myositis patients. In this part, bioanalytical applications of high performance liquid chromatography – mass spectrometry are described *(i)* for clinical determination of the subclass specific glycosylation of total plasma IgG from patients and controls, and *(ii)* for antibody biomarker discovery using statistical analysis of mass spectrometric data in a twin-sibling clinical study.

2 RESULTS AND DISCUSSION

2.1 Primary structure determination of Aß-specific antibodies

2.1.1 Methods of high performance liquid chromatography – tandem mass spectrometry for proteome analysis

High performance liquid chromatography (HPLC) and mass spectrometry (MS) represent two powerful techniques in modern analytical chemistry amenable for hyphenation. Liquid chromatography using reversed stationary phases (RP-HPLC) enables separation of components in a mixture based on differences in their relative hydrophobic character. Depending on the pump delivery system and on the dimensions of the separation media determining the flow rate through the column, liquid chromatography can be performed at preparative (>20 mL/min), analytical (1-10 mL/min), micro (50-1000 µL/min), capillary (0.4-200 µL/min) and nano (24-4000 nL/min) scales. The eluted fractions containing individual sample components can be collected and subjected to “off-line” analytical characterization by mass spectrometry and/or N-terminal Edman sequencing.

The “on-line” coupling of liquid chromatography and mass spectrometry benefits of the combined advantages provided by the LC-separation and by the sensitivity, resolution and accuracy of MS, and renders the LC-MS combination to a highly efficient technique in proteome analysis [168]. In “on-line” LC-MS the analytes eluting from the reversed phase column *(i),* flow through the electrospray needle, *(ii)* are transformed into gaseous molecular ions via electrospray ionization, and *(iii)* are analyzed by mass spectrometry, which detects their mass-to-charge ratio (*m/z*). Furthermore, nearly all modern MS instruments have the capability to perform tandem MS experiments (MS/MS), hence introducing an additional dimension for structure determination. A major advantage of LC-MS/MS is the extended dynamic range, i.e. complex mixtures containing a wide range of concentrations of the individual components can be simultaneously analyzed, down to femto- and even lower levels. With the development of bioinformatic tools for data acquisition and analysis and the expansion of genomic and proteomic data bases, LC-MS can now be routinely carried out and interpreted in a widely automated fashion. The increasing demands for sensitivity, low sample consumption and improved dynamic range in proteome analysis have led to continuous miniaturization of the stationary phase

systems. The concept of "microfluidics" refers to the behavior, precise control and manipulation of fluids that are geometrically constrained to a small, typically sub-millimeter scale. The properties of fluids at microscale can differ from the macrofluidic behavior, in that factors such as surface tension, energy dissipation and fluidic resistance may become dominant. Advances in microfluidics technology have revolutionized the fields of molecular biology and proteomics. One of the key application areas in proteomics involves the use of continuous flow microfluidics, such as nano-scale chromatographic columns and chips containing microfabricated channels, in which the solvent flow is delivered by external pressure sources. These devices proved efficient for analysis of complex mixtures, typically in combination with mass spectrometry.

Structure determinations of Aß-antibodies and amyloid precursor protein presented in the following sections were performed on two different types of mass spectrometers, *(i)* high resolution quadrupole time-of-flight (Q-Tof), and *(ii)* ion trap MS. Both instruments were interfaced with high performance liquid chromatography (LC) systems which enabled nano-scale separation of sample components (typical column flow 300 – 400 nL/min). The microfluidic separation devices employed as stationary phases for nano-liquid chromatography were *(i)* a nano – C_{18} reversed phase column of 100 mm length and 75 µm internal diameter, and *(ii)* an Agilent chip system containing a 40 nL enrichment column followed by a 43 mm × 75 µm analytical column packed with C_{18} particles. In the following sections, the principles of operation of both instruments are described.

2.1.1.1 Principles of quadrupole time-of-flight mass spectrometry

The most important component of any mass spectrometer, the mass analyzer, represents the region of high vacuum in which ions generated from an ion source are brought to motion by static or oscillating electromagnetic fields, and are subsequently separated according to their mass-to-charge ratio (*m/z*). The most common mass analyzers include the quadrupole mass filter (Q), the flight tube (time-of-flight, Tof), and several trapping devices, including the ion trap (IT), the ion cyclotron resonance cell (ICR) and the recently invented orbitrap [169, 170]. The selection of a mass analyzer depends

upon resolution[1], available mass range[2], scan rate[3] and detection limit required for an application. Each analyzer has very different operating characteristics and the selection of an instrument involves important tradeoffs. These analyzers can be used either as stand-alone devices, or combined in series in order to expand and exploit their individual capabilities.

The most common mass analyzer, the quadrupole mass filter, consists of four straight and parallel rods arranged such as the beam of ions coming from source is directed axially between them. A voltage comprising a direct-current (DC) and a radio frequency (RF) component is applied between adjacent rods, opposite rods being electrically connected. The ions entering the quadrupole will oscillate in the (x) and (y) directions as a result of the high frequency field. For ions of specific *m/z*, these oscillations are stable for defined values of DC and RF voltages, while for other values these ions strike the rods and are no longer detected (see Figure 2.1). A mass spectrum is obtained by "scanning the quadrupole", i.e. by monitoring the ions passing through the quadrupole filter as the DC and RF voltages on the rods are varied [171].

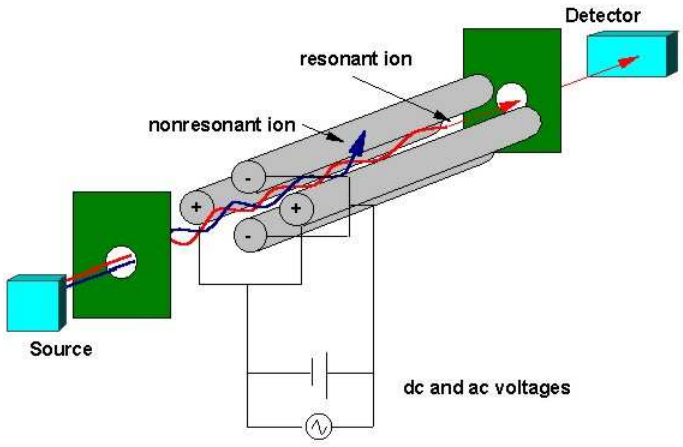


Figure 2.1: Schematic representation of the quadrupole mass filter: four parallel rods, opposite rods being connected. A voltage comprising a dc and a rf component is applied between adjacent rods. Ions entering the quadrupole have stable oscillating trajectories for certain values of dc and rf (red trajectory), whereas for the same values other ions become unstable (blue trajectory), strike the rods and are dissipated.

[1] Resolution (R) in mass spectrometry refers to the separation of two ions as R=m/Δm, where Δm is the width of the peak at a height that is a specified fraction of the maximum peak height. Typical value is 50% of the height, which indicates the value full width at half maximum, FWHM.

[2] Mass range refers to the highest mass-to-charge ratio transmitted by the mass spectrometer.

[3] The scan rate of a mass spectrometer refers to how fast it scans a mass spectrum. This is important in the LC-MS applications where the entire mass spectrum must be scanned faster than the elution time of the chromatographic peak.

Quadrupoles are compact devices, capable of fast scan rate, high transmission efficiency[4] and modest vacuum requirements, and are ideal for small inexpensive instruments.

The time-of-flight (Tof) mass analyzer [172] uses an electric field of typically 2-25 kV between two plates to accelerate ions to the same kinetic energy *(E_c)* that are subsequently separated in time as they travel down a field free region (see Figure 2.2 A). The kinetic energy is determined by the acceleration voltage of the instrument *(V)* and by the charge of the ion *(e×z)*, implying that the ion velocity is directly proportional to *V* and inversely proportional to its *m/z* ratio. The time-of-flight *(t)*, related to the ion velocity *(v)* by the length *(L)* of the flight tube, $L = v \times t$, is given by:

$$t = \frac{L}{\sqrt{2V \times e}} \sqrt{m/z}$$

Equation 1

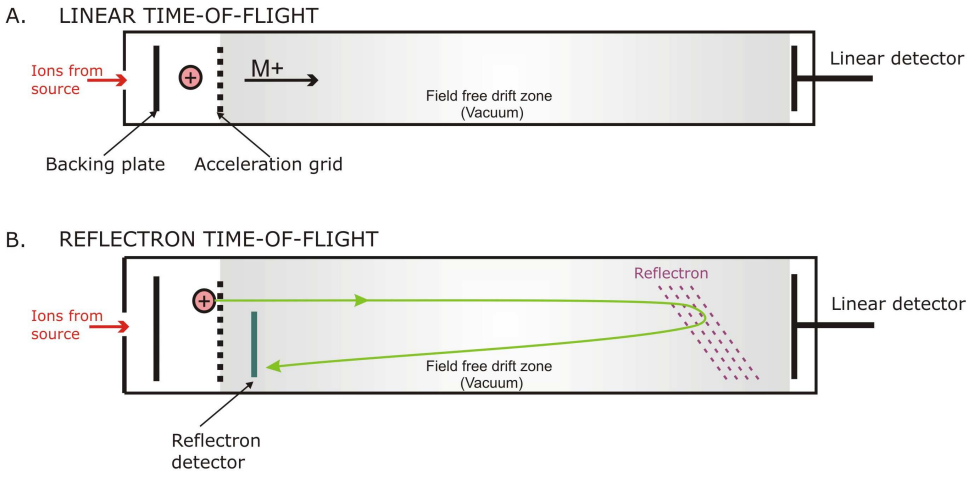


Figure 2.2: Schematic representation of a time-of-flight mass analyzer: (A) linear mode, and (B) reflectron mode. The voltage applied between the backing plate and the acceleration grid imparts the same kinetic energy to all the ions which are then drifting in the field free region. In the reflectron mode the ions are refocused by voltages applied on a series of lenses (reflectron), are allowed to drift back in the reversed direction, and subsequently detected.

Equation 1 shows that ions of low *m/z* will reach the detector first. The mass spectrum is obtained by measuring the detector signal as a function of time for each pulse of ions produced in the source region. Because all the ions are detected, this analyzer has high transmission efficiency, no upper *m/z* limit, fast scan rates and very low detection limit.

[4] Transmission efficiency refers to how many ions produced in the source region reach the detector. This is an important measure of sensitivity for mass spectrometers.

However the low mass resolution, usually less than 500, represents a major drawback. This is derived from the fact that in practice not all ions experience the acceleration pulse *(V)* with the same intensity, such as a distribution of kinetic energies exists for each discrete *m/z* value [173]. This limitation can be corrected with the use of a reflectron at the end of the drift zone, consisting of a series of electric fields which repulse the ions back along the flight tube, usually at a slightly displaced angle, resulting in the refocusing of the ions with the same *m/z* value on the reflectron detector (see Figure 2.2 B) [174].

The *m/z* value of a molecular ion alone does not provide structural information about the analyte. This can be obtained in a tandem MS (MS/MS) experiment in which the molecular ion is dissociated inside the mass spectrometer into fragments that are put together to decipher its structure. Technically, fragmentation of a molecular ion can be obtained if at least two mass analyzers are connected in series such as ions selected for MS/MS in the first analyzer are subjected to fragmentation in the second one and these fragments are subsequently detected. In the simplest way, this was achieved in a triple quadrupole instrument ($Q_1q_2Q_3$), containing three quadrupoles connected back-to-back, in which each component is set up as follows: the first quadrupole (Q_1) is a mass filter, which allows selection of precursor ions by scanning across preset *m/z* values; the second quadrupole (q_2) – the collision cell – is non-mass filtering (RF-only, DC voltage set at 0), and contains an inert gas to a pressure of approximately 10 mTorr that is injected into the flight path of the accelerated ions transmitted through the first quadrupole, hence inducing their fragmentation. The third quadrupole (Q_3) is also a mass filter that is set to scan the entire *m/z* range, giving information about the *m/z* values of the fragments formed in Q_2 [175].

Quadrupole time-of-flight (Q-Tof) mass spectrometers have become commercially available in 1995 and have been rapidly recognized by the analytical community as powerful and robust instruments with unique capabilities. In the simplest way, a Q-Tof mass spectrometer can be described as a triple quadrupole mass spectrometer which has the last quadrupole (Q_3) replaced by the time-of-flight analyzer. In the Q-Tof configuration an additional RF quadrupole (Q_0) is placed in front of the first mass filter (Q_1) to provide collisional cooling of the ions injected from the high pressure ionization source; hence, the QTof can be schematically represented as $Q_0Q_1q_2$Tof. The ions transmitted through the last multipole are orthogonally accelerated into the time-of-flight analyzer. In modern instruments, the quadrupoles Q_0 and q_2 are replaced by hexapoles [176], offering an

additional gain in resolution (see Figure 2.3). For normal mass spectra the quadrupole is used in the RF-only mode as a wide band-pass filter to transmit a wide mass range. The collision cell is not pressurized and the ions are transmitted to the Tof for mass analysis. In the MS/MS mode the quadrupole operates in the normal resolving mode, and is able to select precursor ions up to *m/z* 4000 for collisional activation in the hexapole gas cell. The benefits of the hybrid Q-Tof instrument over the triple quadrupole are improved signal-to-noise (S/N), higher mass resolution up to 10000 (FWHM) owing to the reflectron Tof, and higher mass accuracy in both MS and MS/MS modes of operation [176].

An innovative feature in the design of the Q-Tof Premier mass spectrometer, employed for the experiments described in the following chapters, is the so called traveling wave collision cell (T-Wave), which replaces the traditional hexapole collision cell (q_2), as well as the first ion guide element (Q_0) (see Figure 2.4). In a normal hexapole cell (shown in Figure 2.3), multiple collisions of an ion with the inert gas may result in a reduction of the axial ion velocity that may affect performance in modes of operation when fast scanning and switching are required, such as the analysis of a complex biological sample. A T-wave consists of a stack of rings with opposite RF phases applied on adjacent rings. In addition to the RF voltage each ring electrode has a constant DC offset (collision energy voltage) and may also have a transient DC potential applied. The transient DC voltage produces a local change in the electric field causing the ions to move away from the electrode in both forward and reverse directions. To push the ions in one direction along the axis, the DC voltage is switched to an adjacent ring after a given time and so along the entire ion guide. This results in a moving electric field or “traveling wave” which reduces the residence time of the ions in the cell [177] (see Figure 2.4).

Accurate mass measurement is strongly dependent on the calibration of the Tof analyzer, which can be performed with a quality control standard. A specific computer software, MassLynx, uses a polynomial equation to calibrate precisely over a wide mass range. In the present work, external calibration of the Tof was performed using the fragment ion spectrum of the doubly protonated molecule of Glu-Fibrinogen peptide (GluFib) (*m/z* 785.8496), formed by collision induced dissociation. In addition, an external reference spray (GluFib, 300 fmol/µL) (see Figure 2.4), which interchanges with the analyte delivery system during data acquisition, was used to correct for the temperature variations in the laboratory surroundings.

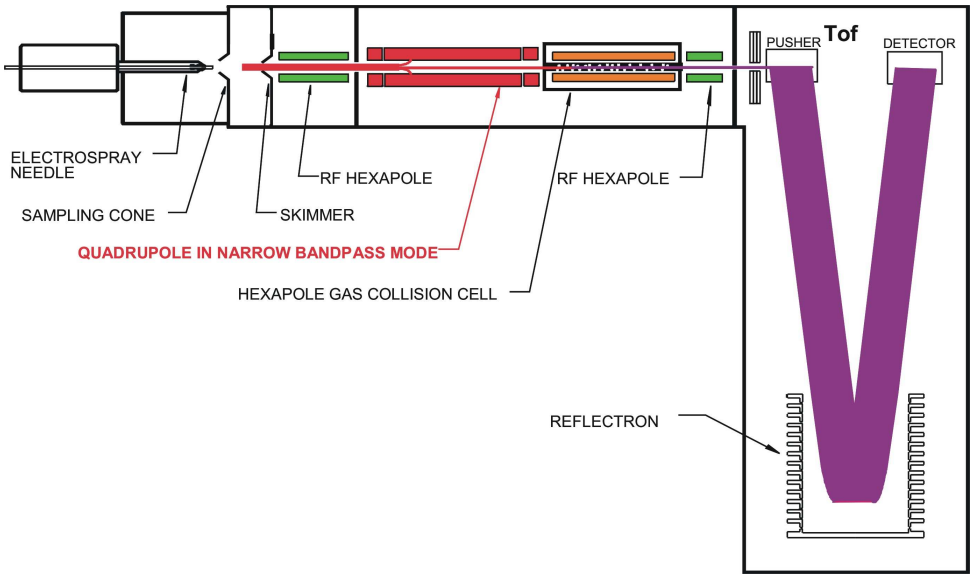


Figure 2.3: Schematic representation of a hybrid Q-Tof mass spectrometer. The ions formed by electrospray ionization are transported through an RF-only hexapole (green rectangle) into the first quadrupole (red rectangle), where precursor ions of interest (red thin line) are isolated, accelerated into the hexapole collision cell (orange rectangle), where collisions with an inert gas (white dots) induce fragmentation into smaller fragment ions (purple line). Fragment ions (purple thick line) are travelling through the time-of-flight mass analyzer operated in the reflector mode and are detected by a photomultiplier.

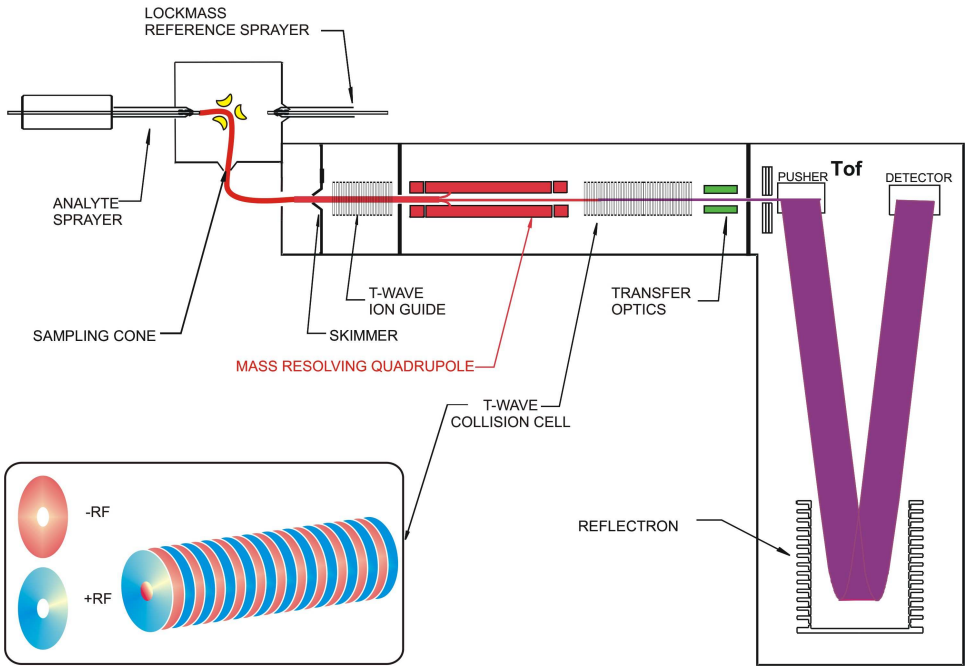


Figure 2.4: Schematic representation of the Q-Tof Premier (Waters). The instrument features the Zspray ion source (i.e. the spray is orthogonally directed into the first ion transport element via a sampling cone positioned perpendicular to the direction of the spray), and the travelling wave T-Wave. T-wave is used both as a transport element of the ions from the electrospray source and as collision cell. The insert shows the T-wave technology, consisting of a stack of rings with alternating RF amplitude.

2.1.1.2 Principles of ion trap mass spectrometry

The quadrupole ion trap consists of a ring electrode and two endcap electrodes. The internal surface shape of these three electrodes follows a three dimensional nearly hyperbolic profile. Holes at the center of the endcaps allow ions to pass in and out the trap. A high voltage RF potential (*V*) is applied to the ring, while the endcaps are held at ground (see Figure 2.5, the schematic at the top). The oscillating potential difference established between the ring and endcap electrodes forms a quadrupolar field, which can be thought as a three dimensional pseudo-potential well. The value of the RF voltage and the mass of an ion determines the depth of the well for that particular ion. An auxiliary voltage (*U*) is applied on the exit end cap, which is involved in precursor ion isolation, fragmentation and mass analysis phases of the scan sequence.

The ions are trapped in the ion trap by the RF voltage (*V*) applied to the ring electrode. The quadrupolar field induces an oscillatory harmonic motion of the ions in both the radial (*x, y*) and the axial direction (*z*), with a secular frequency *f* which is smaller than the frequency of the field. As the ions repel each other in the trap, their trajectories expand as a function of the time. In order to avoid ion losses through the exit end cap by this expansion, the trap is filled with an inert gas, typically helium, to a pressure of approximately 3×10^{-3} mbar, which removes excess energy from the ions by collision, hence reducing their trajectory. The ion movement in the trap is described by the Mathieu equations (reviewed in [178]), the solutions of which allow locating the areas where ions of given *m/z* values have stable trajectories at the operating RF drive level. The stability diagram of the ions resulting from the integration of the Mathieu equation is shown in Figure 2.5 (bottom).

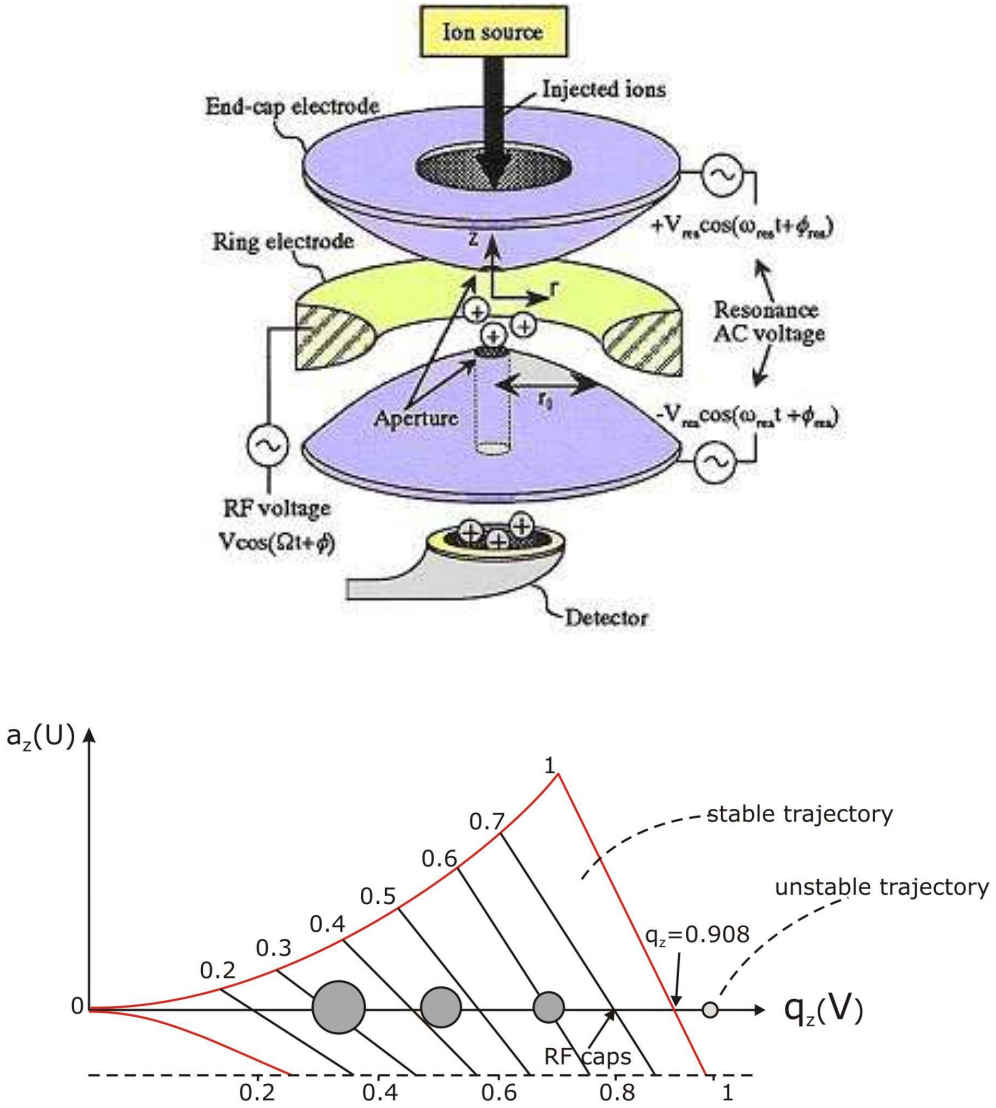


Figure 2.5: Top – Schematic representation of the quadrupole ion trap, consisting of a ring and two endcap electrodes. Ions from an external source enter the trap through the endcap, are trapped by the RF voltage applied to the ring and leave the trap through the opposite endcap; an auxiliary voltage may be applied to the endcaps. Bottom – Stability diagram of the ions inside an ion trap operated in the RF only mode: heavier ions (larger circles) are located left of the lighter ions (smaller circles). Ions located anywhere within the red plot have stable trajectories, whereas those located outside are unstable, and are expelled through the exit endcap.

The parameters of the Mathieu equation a_U and q_U, determining the stability region of ions inside the trap in the axial direction are given by:

$$a_u = a_z = -\frac{16zeU}{m\cdot(r_0^2+2Z_0^2)\cdot\omega^2} \quad \textbf{and} \quad q_u = q_z = \frac{8zeV}{m\cdot(r_0^2+2Z_0^2)\cdot\omega^2} \qquad \textit{Equations 2}$$

where V is the amplitude of the RF voltage applied to the ring, U is the amplitude of the auxiliary voltage applied to the endcaps, e is the elementary charge, r_0 and Z_0 are the

dimensions of the ion trap, ω is the fixed frequency of the field, *m* is the mass, and *z* is the charge of the ion.

The ion trap is operated in the RF only mode, which implies that *U=0* and $a_U=0$. Hence, the ions are located along the horizontal axis q_Z described by the Equation 2. Because heavier masses have lower q_Z values than lighter masses (see Figure 2.5), it follows that there exists a cut-off mass, depending on the level of the RF voltage (V). Ions with q_Z <0.908 are stable in the field, while those with q_Z above this value are unstable and leave the trap. The mass analysis of the ions confined in the trap is performed by resonant ejection, implying that the auxiliary voltage (*U*) is applied to the end caps in the axial direction in which the ions enter and leave the trap. When the frequency of this field is identical with the secular frequency *f* of the ion motion, the ions will resonate, i.e. the amplitude of their oscillation will increase to the point where their trajectories are no longer stable, and are ejected from the trap before reaching the stability limit (q_Z= 0.908).

The resonance phenomenon derived from coupling of additional fields into the ion motion is of major importance for multistage tandem mass spectrometry MS(n). Similar to the mass analysis, a precursor ion can be isolated by applying a broadband radiofrequency spectrum that expels all except for the ion of interest, i.e. the resonance frequency of the parent ion is not contained in the frequency spectrum. To induce fragmentation the energy of the precursor is increased by resonance excitation with the dipole field, which comprises a small frequency band above and below the precise resonance frequency of the ion. Following collisions with the helium gas the parent ion will dissociate into predictable and reproducible daughter ions, which are trapped by the primary RF and subsequently analyzed by resonant ejection. One of the resulting fragments can be selected and fragmented again, in an MS/MS/MS (MS^3) experiment, in a similar manner as described for the first stage MS/MS (MS^2). This cycle can be repeated, in principle, up to the MS^{11} stage, although experimental practice was rarely successful at stages higher than MS^4.

The major difference in the collision induced dissociation (CID) in quadrupoles vs. ion trap instruments is the excitation method. In a quadrupole ions are accelerated and passed through a high pressure collision cell. A fragment ion may undergo further MS(n) fragmentation, because not all the energy is necessarily lost after the first collision. In an ion trap the excitation occurs by irradiation of the precursor at its secular frequency; hence,

only this ion is excited, whereas the fragment ions may be too cold to undergo further fragmentation.

2.1.1.3 Pathways and techniques of peptide fragmentation

Mass spectrometric biopolymer analysis, with emphasis on structural characterization of proteins, relies *(i)* on the accurate mass measurement of the intact protein or its fragments derived from enzymatic degradation (peptide mass fingerprinting), and *(ii)* on the sequence information obtained from the fragmentation pattern of precursor ions of proteolytic peptides subjected to decomposition by collision induced dissociation (CID) and/or electron capture/transfer dissociation (ECD/ETD). Both approaches are strongly dependent on the existence of genomic/proteomic databases that contain the information about the protein sequence. Typically, peptide mass fingerprinting (PMF) is used to rapidly evaluate the correctness and/or identify a protein sequence by comparing the peptides molecular weights to those predicted *in silico* for the sequences in protein and/or translated nucleic acid databases. Of major importance for this approach is the high mass accuracy, which can be used as a constraint in the database search, hence reducing the rate of false-positive results. Of particular interest is the case when limited knowledge of genome data, unknown post-translational modifications and variability in amino acid sequence impose restrictions on the use of the database search based approaches. In these instances, the protein sequence is primarily identified from the fragment ion spectrum of proteolytic peptides in combination with computer algorithms developed for the automated assignment of the fragments. This approach is referred to as *de novo* sequencing [179-181], and was largely employed in the present work.

In the low-energy collision induced dissociation of multiply protonated peptides fragment ions arise from non-specific cleavages of amide bonds across the peptide backbone. Fragmentation is brought about by increasing the ion energy either by resonant excitation or axial acceleration, followed by collision with an inert gas such as helium or argon. During each collision, imparted translational energy is converted to vibrational energy which is rapidly distributed throughout all covalent bonds. Fragment ions are formed when the internal energy of the ion exceeds the activation barrier required for a particular bond cleavage. The charge retaining N-terminal fragments are designated as “b” ions, whereas the complementary C-terminal fragments are termed “y” ions. This nomenclature was

introduced by Roepstroff [182] and subsequently modified by Biemann [183, 184] (see Figure 2.6).

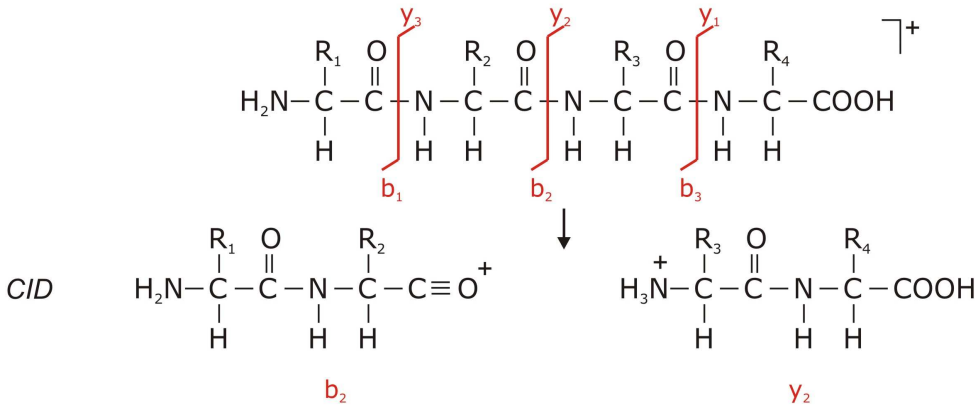


Figure 2.6: Fragmentation of a tetrapeptide by collision induced dissociation (CID). N-terminal fragment ions arising from peptide bond cleavage (indicated with red vertical bars) are termed "b" ions, whereas the fragments containing the C-terminus of the molecule are referred to as "y" ions. Nomenclature of peptide fragmentation after Roepstroff and Biemann [182-184].

MS/MS spectra of protonated peptides can be used to identify amino acid residues from the mass difference of successive fragment ions of the same type (e.g. b_n and b_{n-1}), given that these are contained in the spectrum. In practice, it is rarely the case when a complete series of b or y ions is observed. The abundance of the b and y fragment ions in the CID spectrum of a given peptide is dependent on several factors, such as the charge of the precursor ion, the size of the peptide, the nature of the amino acid residues, the excitation method, the time scale of the instrument and the energy of the fragmentation [185]. The non-specific nature of the amide bond cleavage is largely explained by the “mobile proton” model, explored by Wysocki, Gaskell *et al.* [186-189]. Being multifunctional compounds, peptides can be protonated at multiple sites (terminal amino groups, side chain groups, amide oxygens and nitrogens) leading to various isomers. The basic character of each site dictates the amount of energy required for the attachment of the ionizing proton, such that the side chain of arginine is the most probable moiety for proton attachment, followed by the lysine and histidine side chains, amino terminal group and the amide nitrogen and oxygen, respectively. As the internal energy of ions increases upon excitation, protonation of less favored sites, including those leading to backbone dissociation, becomes accessible.

Alternative soft techniques to induce fragmentation of the peptide backbone are electron capture and electron transfer dissociation (ECD/ETD) [166, 190-192]. In electron capture

dissociation, low energy electrons (< 1 eV) emitted from a heated filament are allowed to react with protonated peptide molecules confined in the Penning trap of a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. Capture of an electron by a protonated peptide is exothermic, causing the peptide backbone to fragment by a nonergodic process, i.e. no redistribution of intramolecular vibrational energy is involved. One pathway in this process involves generation of an odd-electron hypervalent species that dissociates spontaneously to produce N-terminal c ions and C-terminal z• radical cations (Figure 2.7). The technology of ECD has been hitherto unique to FT-ICR mass spectrometers; however, the limited dynamic range of these instruments represents a limitation for its use in conjunction with chromatographic separation.

Alternatively, the attachment of electrons to peptide ions can be achieved in ion trap instruments through ion/ion reactions, which represents the fundamental process in electron transfer dissociation. The ETD technology was recently implemented to ion trap instruments, which are fast, robust and price accessible; hence the field in biopolymer structural analysis has advanced considerably [166]. In ETD, the electron carrier is a polyaromatic hydrocarbon compound, e.g. anthracene or fluoranthene, which is converted into a radical anion by negative ion chemical ionization; these species have sufficiently low electron affinity such that they can function as one-electron donors [193, 194]. The radical anion species reacts with multiply protonated peptide molecules confined in the ion trap, resulting in the transfer of the electron to the N-C_α bond of the peptide backbone, and triggering fragmentation in a similar fashion as in the ECD process (Figure 2.7).

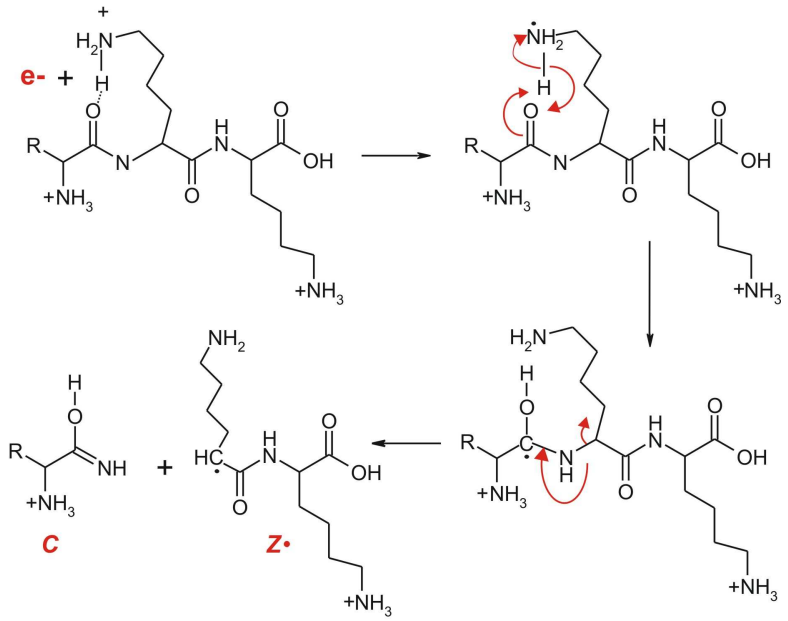


Figure 2.7: Fragmentation scheme for production of c- and z•- type ions after reaction of a low energy electron with a multiply protonated tripeptide. In ETD/ECD fragmentation occurs randomly at the N-C_α bond, producing N-terminal c ions and C-terminal z• radical ions.

Successful peptide identification by either ETD/ECD or CID can be achieved only when product ions from a complete or nearly complete distribution of amide bond cleavages are observed in the corresponding MS/MS spectrum. In this regard, CID fails when multiple basic residues are present, which inhibit random protonation along the peptide backbone. In contrast, ECD is independent of amide bond protonation and occurs on a time scale which is short compared to the internal energy distribution. Hence, multiply protonated peptides fragment randomly across the peptide backbone and are easily sequenced. In the present work an ion trap (Agilent) equipped with a chip cube and an electron transfer dissociation module was employed for the structural characterization of ß-amyloid proteins and their post-translational modifications. The instrument design and the principle of operation are identical with that described in Chapter 2.1.1.2. In addition, the chemical ionization source is positioned such that fluoranthene radical anions are injected in the octapole region of the ion transport system, and subsequently co-injected into the trap through the front endcap, together with the multiply protonated peptide molecules (see Figure 2.8).

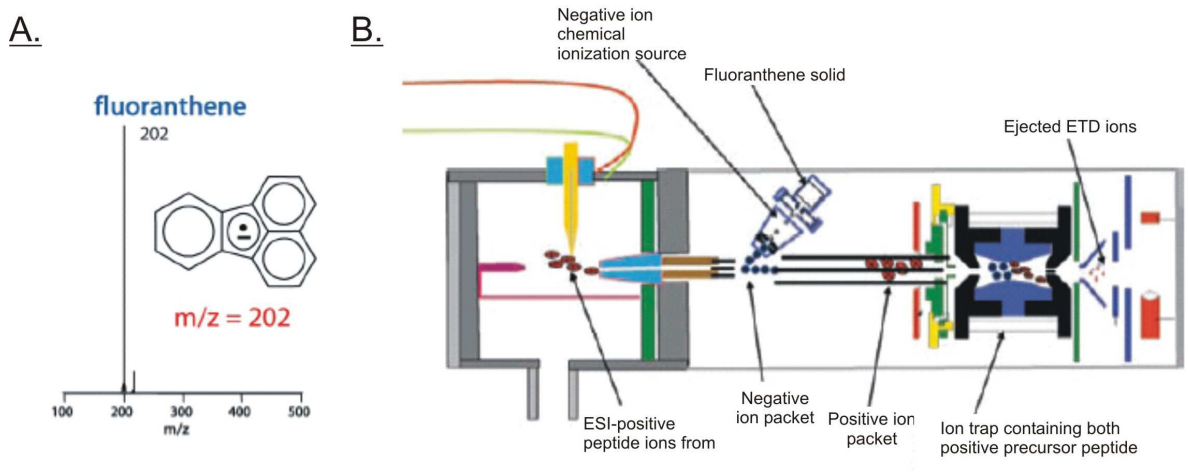


Figure 2.8: (A) Structure of the fluoranthene radical anion, of m/z 202, produced by negative ion chemical ionization, acting as one-electron donor; (B) Schematic representation of the Agilent ion trap instrument, equipped with a chip cube interface (left) and negative ion chemical ionization source (top). Gaseous fluoranthene radical anions (blue bullets) are injected into the octapole region and are confined in the trap together with the multiply protonated peptide molecules (red bullets). Following ion/ion reactions, which results in the transfer of the electron to the peptide, the charge reduced species fragments spontaneously.

2.1.2 Mass spectrometric approaches for structural characterization of immunoglobulins

In the last years antibodies have emerged as one of the most promising classes of biological drugs in the pharmaceutical industry. Albeit their therapeutic properties had been recognized since the end of the 19th century, clinical success was first achieved in 1975, with the development of hybridoma technologies for production of murine monoclonal antibodies [195]. Hybridoma, obtained through the fusion of an antibody secreting B-lymphocyte with a long-lived neoplastic plasma cell, is capable of generating identical copies of that antibody over a long period of time, called monoclonal antibody (mAb). Unless the genome of the immune cell has been previously sequenced, the protein sequences of antibodies produced in hybridoma are unknown. These mAbs with defined antigen specificity have a broad range of immuno-analytical applications, such as Western blot, diagnostic tools and affinity capture of antigens from biological samples. Large-scale production of pharmaceutical-grade antibodies is typically achieved by recombinant DNA technology, i.e. amino acid sequence of the mAb of interest is predicted from its cDNA. Most recombinant mAbs are produced in mammalian cell-expression systems such as Chinese hamster ovary (CHO) or murine lymphoid cell lines to enable proper folding and glycosylation [196, 197]. More than 20 antibody drugs have been currently approved by

the Food & Drug Administration (FDA) for human therapeutic use in oncology, transplantation, infectious diseases and cardiovascular medicine [198, 199].

Nearly all stages in the development of a therapeutic antibody require its detailed structural characterization, comprising analysis of amino acid sequence, disulfide linkages, carbohydrate structure distribution, and other, post-translational modifications. Mass spectrometry has become an essential tool for the analytical characterization of antibodies, due to its superior resolution compared to other analytical methods, sensitivity, low time and sample consumption, and applicability to mixtures. The two "bottom-up" methodologies used for primary structure determination of Aß-antibodies are outlined in Figure 2.9. Several steps common to both approaches involve reduction and alkylation of the antibody in solution, followed by separation of its heavy and light chains by SDS-PAGE and digestion of each protein band with a highly specific enzyme, usually trypsin. Alternative enzymes are α-chymotrypsin, Lys-C, Arg-C, and others proteases.

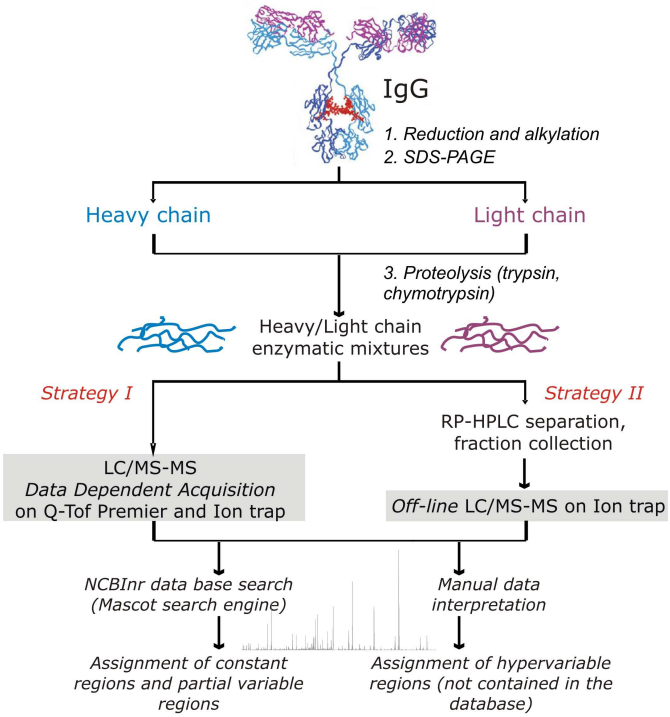


Figure 2.9: Experimental approaches for primary structure determination of Aß-antibodies. Upon in-solution reduction and alkylation, heavy and light chains are separated by SDS-PAGE and the protein bands are digested with a proteolytic enzyme of choice. In Strategy I, the enzymatic mixture is subjected to "on-line" LC-MS/MS analysis, in which mixture components eluted from the reversed phase column are introduced directly into the mass spectrometer. Sequence information is obtained using data dependent acquisition (DDA), in which several most abundant ions at a given time point are selected for MS/MS and their fragment mass spectra are recorded. In Strategy II, the digestion mixture is pre-fractioned using analytical RP-HPLC, the individual fraction are collected, and each is analyzed "off-line" by LC-MS/MS. The data obtained from both approaches are processed and subjected to database search, or manually interpreted.

In *Strategy I*, the enzymatic mixture is analyzed by LC-MS/MS in an "on-line" fashion, i.e. the mixture components eluting from the nano-scale reversed phase column (interfaced with the Q-Tof – MS) or chip system (interfaced with the ion trap – MS) are introduced directly into the mass spectrometer via electrospray ionization. The peptides are analyzed via data dependent acquisition (DDA), which employs a full MS scans recording the analyte ions "eluting" from the column, followed by several MS/MS cycles, in which fragmentation spectra of several most abundant ions at a given time point are recorded. The number of precursor ions is preset and their selection proceeds in a fully automated fashion, based on predetermined ion abundance criteria (threshold). When the last MS/MS cycle was accomplished, a new full MS scan is performed, and the scan sequence repeats. The time elapsed between two successive MS scans has been termed duty cycle. The DDA scan sequence used in the present work is shown in Figure 2.10.

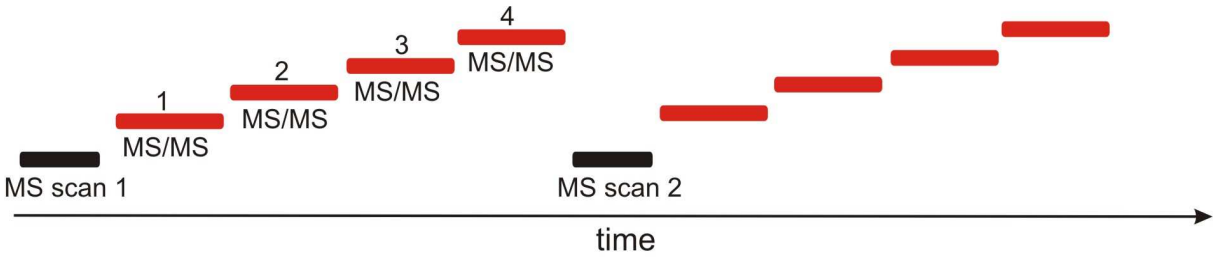


Figure 2.10: Schematic representation of the data dependent acquisition (DDA) approach used in the present work. Each full scan MS (black), typically 1 second, is followed by four separate MS/MS analyses (red) of the four most abundant precursor ions detected in MS (typical duration 1.4 seconds per MS/MS). When the last MS/MS acquisition is completed, the instrument returns to the MS mode, seeking for the next four most abundant ions.

On the Q-Tof Premier, each MS scan has a duration of 1 second, while each MS/MS scan is acquired for 1.4 seconds. Individual scans are separated by an interscan delay of 0.1 seconds. Collectively the duration of a duty cycle is 1 +4×(1.4+0.1) = 7 seconds. These time settings can be varied by the user. The scanning sequence on the ion trap is similar; however, the different principle of mass analysis renders it a faster scanning device compared to the Q-Tof. Typical accumulation time for a microscan on the ion trap is 100-200 milliseconds, while each MS/MS scan can be averaged over several microscans.

Alternatively, in *Strategy II* heavy and light chain digestion mixtures are separated by RP-HPLC; the individual fractions were collected, lyophilized and analyzed "off-line" by LC-MS and LC-MS/MS. The LC-MS analyses were manually inspected for the presence of multiply charged ions. For each precursor ion, a separate LC-MS/MS analysis was

acquired, in which targeted selection of a single parent ion was performed, in order to improve the quality of the MS/MS data with the increasing number of microscans. As the digestion mixture was pre-fractioned by analytical HPLC, a faster LC gradient was employed, so that the analyses could be performed more time effective.

2.1.3 Primary structure determination of Aß-specific antibodies

2.1.3.1 Primary structure determination of a plaque-specific anti-Aß(1-17) monoclonal antibody

In the last years, a mouse monoclonal antibody produced in hybridoma (clone 6E10) has been widely used in AD research. Terai and co-workers have used this N-terminal, plaque specific antibody to characterize the major ß-amyloid species in senile plaques by affinity– mass spectrometry and immunochemistry [200], while Maddalena *et al.* used protein chip technology to capture the Aß-peptides in cerebrospinal fluid (CSF) with 6E10 followed by mass spectrometric characterization of the captured peptides [201]. These results and the most recent clinical studies confirming the epitope specificity [91] emphasize the therapeutic potential of plaque-specific antibodies. Using epitope excision – mass spectrometry and alanine scanning mutagenesis, we have previously shown that the mouse monoclonal 6E10 antibody, directed against ß-amyloid (1-17), recognizes the same short epitope (FRHDSGY) at the N-terminus of Aß, as did the antibodies resulting from active immunization of transgenic mice with Aß(1-42) [202, 203]. Therefore, primary structure details of this plaque-specific antibody (6E10) will provide a better understanding of the antigen recognition process at the molecular level and contribute to the development of more effective vaccines.

In order to determine the amino acid sequence of the 6E10 monoclonal antibody, the heavy and the light chains were separated by SDS-PAGE following in-solution reduction and alkylation. In addition to the heavy and light chain bands, with an apparent molecular weight of 50 kDa and 25 kDa, respectively, a less intense band was observed at 75 kDa which was identified as albumin. Tryptic and chymotryptic digests of the heavy and light chains were analyzed separately by LC/MS/MS on a Q-Tof Premier mass spectrometer using data dependent acquisition (*Strategy I* in Figure 2.9).

The amino acid sequence determined for the heavy chain of the 6E10 antibody (Figure 2.14) resulted in a sequence coverage of 82%. The identified tryptic and chymotryptic peptides were fit to a known homologous IgG heavy chain sequence frame from the database (accession number ABD73933). Typically, γ-heavy chains contain approximately 440 residues, divided in one variable region (V_H) and three constant regions (C_H1, C_H2 and C_H3) with each containing about 110 amino acids. The heavy chain of 6E10 belongs to the IgG_1 isotype, based on the amino acid composition of the constant region. The locations of the three complementary determining regions (CDRs) in the V_H domain, involved in the antigen recognition, were approximated using the Kabat rules [204, 205], and are highlighted with boxes in Figure 2.14.

A higher sequence coverage was obtained by comparison of the MS/MS data in the NCBInr database entries for the heavy chain constant region than for the V_H region. This might be explained by the fact that the constant region is conserved among immunoglobulins while the V_H region is the result of the site-specific recombination of the V-D-J genes and affinity maturation of the antibody [7, 206], leading to a greater extent of variability. The amino acid sequence for the constant region C_H(219-434) was, for the most part, determined from the database with the exception of those amino acids indicated in *italics* in Figure 2.14. The resulting proteolytic peptides containing these residues were possibly too small to be retained on the C18 column or yielded singly charged ions, which would not have been selected for MS/MS analysis based on the predetermined criteria for data dependent acquisition. A partial sequence coverage was obtained for the constant region C_H(149-218). However, as the constant region is conserved among immunoglobulins, it is feasible to assume that the undetermined amino acids are identical to those reported in the database for the IgG_1 isotype. The LC/MS analyses of the trypsin and chymotrypsin digestions provided complementary information; first, by disclosing antibody regions observed with only one of the two enzymes; and secondly, by generating overlapping peptides, which confirmed the correct succession of the amino acids in the sequence frame.

The specificity of an antibody for an antigen is dictated by the amino acids in the complementary determining regions (CDRs). The V_H CDR3 represents a signature of each immunoglobulin, as it is generated somatically, in the process of B-cell maturation. Unless a specific antibody was previously sequenced and the information deposited into the

database, it is not possible to characterize the variable region using the database searching approach. In the case of the 6E10 heavy chain, only the variable region V_H(1-50) spanning the CDR1 (residues 26-35) could be assigned by searching the MS/MS data against the NCBInr database. The fragment ion spectra for peptides containing the regions V_H(1-19), V_H(20-38) and V_H(39-50) (Figure 3.2, Figure 3.3, Figure 3.4 in Chapter 3.9.1) are consistent with the sequence indicated in Figure 2.14. According to the Kabat rules, the V_H CDR1, with a typical length of 10-12 building blocks, is located four residues after the first cysteine of the variable region (CXXX) and is followed always by a tryptophane [204, 205]. In the case of the 6E10 heavy chain, the V_H CDR1 is spanned between the residues V_H(26-35) and contains 10 amino acids (Figure 2.14).

From the LC/MS analysis of the tryptic digest, other ions for which abundant fragmentation was obtained were sequenced *de novo* by using trypsin peptide fragments. As an example, the MS/MS spectrum of the ion of *m/z* 531.278 is shown in Figure 2.11. From these data, the amino acid sequence of the peptide was determined to be FDPVNVNTR based on the nearly complete y ion series (except for y_8) observed in the spectrum.

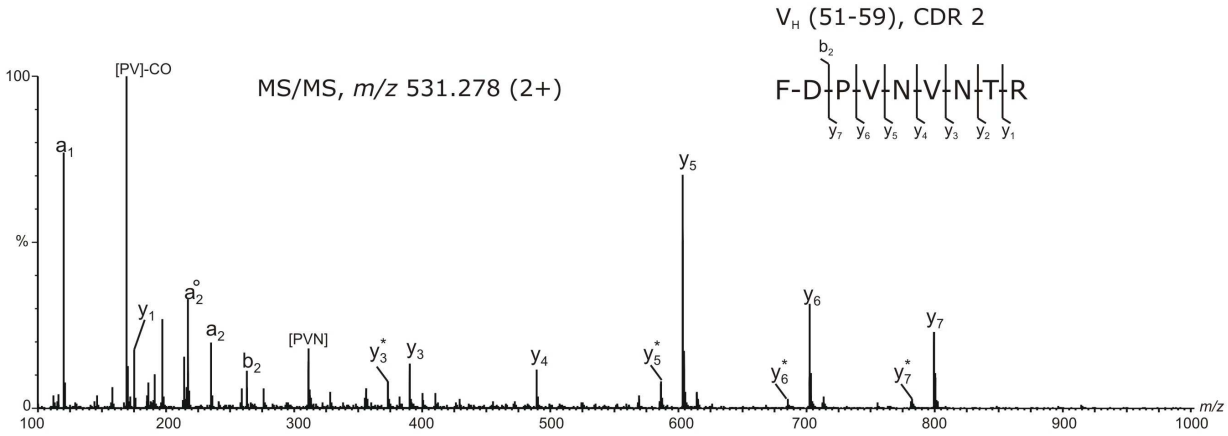


Figure 2.11: Fragment mass spectrum of the precursor ion of m/z 531.278 (2+) from the heavy chain tryptic digest assigned by de novo interpretation to the CDR2 peptide V_H(51-59), indicated at the top (right). C-terminal fragments were assigned as y_1-y_7 and the corresponding cleavages are indicated at the top (right). Internal fragments are indicated between square brackets. The asterisk (*) denotes loss of ammonia. The empty circle (○) denotes loss of water.

The proposed amino acid sequence was confirmed by obtaining the MS/MS spectrum of an overlapping chymotryptic peptide (DPVNVNTRY). This sequence information was inserted into the heavy chain sequence frame based on homology with another heavy chain from a mouse antibody in which the region V_H (51-60) has an amino acid sequence similar to that of the peptide FDPVNVNTRY (NCBI accession number AAA38193). In the

Kabat definition, the V_H CDR2 begins 15 residues after the end of V_H CDR1, being usually preceded by the sequence motif LEWIG (Figure 2.14). The typical length is 16-19 residues [204]. Despite the MS/MS analyses of the two proteases digests the amino acid composition of the region V_H(61-80) could not be determined.

The sequence information of V_H(99-107) CDR3 was determined from the MS/MS spectra of the doubly charged precursor ions of *m/z* 565.811 (Figure 2.12) observed in the chymotryptic digest, and *m/z* 864.952 observed in the trypsin digest (Figure 3.5, Chapter 3.9.1). The MS/MS fragmentation of the $(M+2H)^{2+}$ ion of 565.811 is consistent with the peptide sequence CANVPLPGRF containing an alkylated cysteine, as determined by the observation of a nearly complete series of y ions (except for y_1). This amino acid sequence was not found in the database, but is similar to other heavy chain CDR3 amino acid sequences (accession number AAG25671). This information suggests that this peptide can be assigned as the CDR3 peptide. Usually, the V_H CDR3 begins three residues after a cysteine (a typical sequence motif is CAR) and is followed by the motif WGXG (X can be any amino acid), while it may contain from 3 up to 25 residues [204]. According to this Kabat definition, the residues before the V_H CDR3 of the 6E10 are CAN, those after are WGQG and the determined length is of nine amino acids (Figure 2.14).

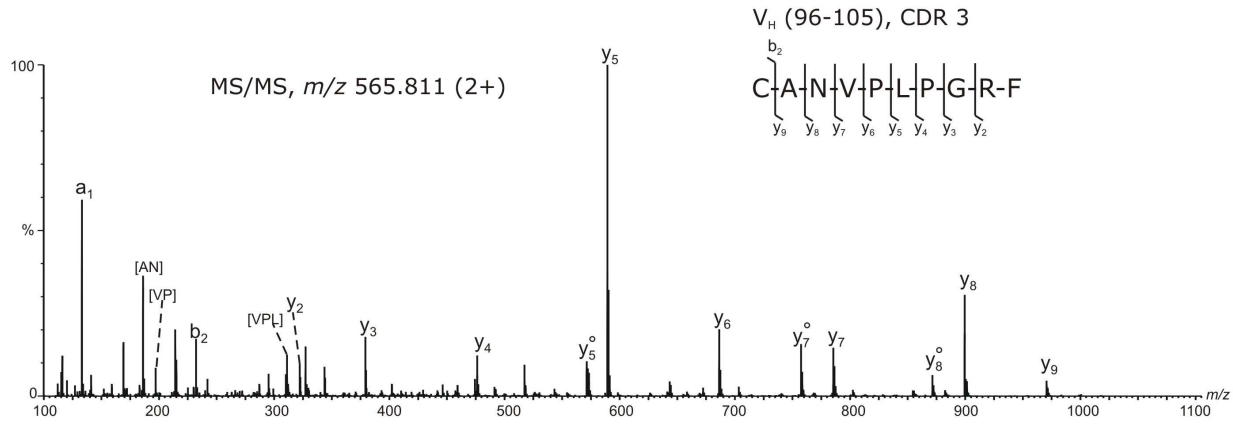


Figure 2.12: Fragment mass spectrum of the precursor ion of m/z 565.811 (2+) from the heavy chain chymotryptic digest assigned by de novo interpretation to the CDR3 peptide V_H(96-105), indicated at the top (right). C-terminal sequence ions are assigned as y_2-y_9 and the corresponding amide bond cleavages are indicated at the top (right). Internal fragments are indicated between square brackets. The empty circle (○) denotes loss of water.

In the tryptic digest of one sequence form of the 6E10 antibody, heterogeneity at the N-terminus of the heavy chain was observed. In addition to the full length N-terminal tryptic peptide (^{1}EVQLQQSGAELVKPGASVK^{19} *m/z* 984.537 (2+)), many other doubly-charged

ions were observed which correspond in mass to the successive loss of two (*m/z* 870.487), three (*m/z* 806.457), four (*m/z* 749.915), five (*m/z* 685.886) and six amino acids (*m/z* 621.857) from the heavy chain N-terminus. The averaged mass spectrum comprising these ions and the peptides assigned to each ion are shown in Figure 2.13 A. The MS/MS spectra for each of these ions are consistent with the assigned peptide structures. One example is illustrated in Figure 2.13 B and shows the fragment ion spectrum of the precursor ion of *m/z* 749.915 (2+). The specific product ions (b and y series) and the mass of the precursor ion enabled the unambiguous identification of the truncated peptide, ^{5}QQSGAELVKPGASVK^{19}, which arises from the loss of the first four amino acids from the N-terminus of the heavy chain. These data suggest that, in addition to the well documented C-terminal lysine clipping [207], antibodies may possibly undergo N-terminal degradation as well. Since this phenomenon was observed only in one of the three antibody forms investigated, the observed N-terminal truncations may arise from non-enzymatic degradation pathways which may be relevant if such a recombinant antibody were investigated as a potential therapeutic agent.

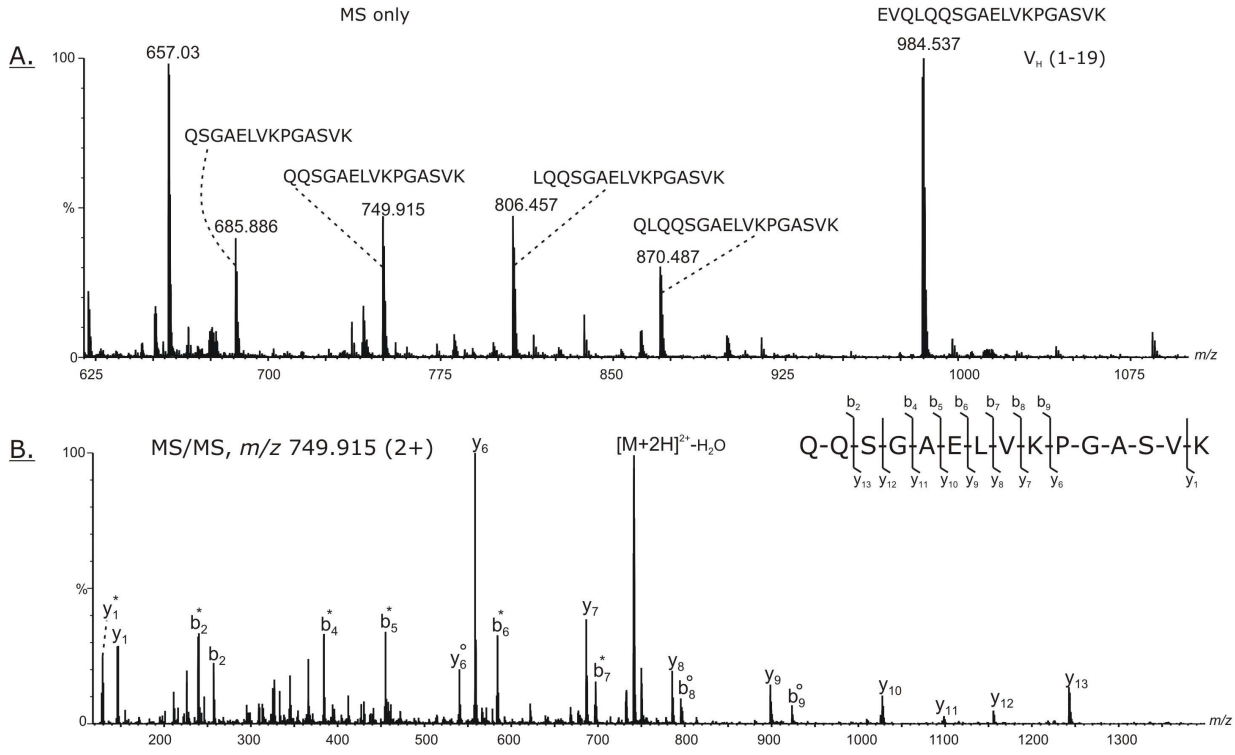


*Figure 2.13: **(A)** Positive ion full MS spectrum showing the doubly charged ions and the corresponding tryptic peptides assigned to each mass which indicate N-terminal truncation of the heavy chain; **(B)** MS/MS of the precursor ion of m/z 749.92 (2+), confirming the indicated peptide sequence. The observed backbone cleavages are indicated at the top (right). The asterisk (*) denotes loss of ammonia. The empty circle (○) denotes loss of water.*

In addition to N-terminal truncation, it was found that approximately 7% of the heavy chain had undergone terminal pyroglutamic acid (pyro-Glu) formation, based on the relative abundances of the ions observed for the native (EVQLQQSGAELVKPGASVK, *m/z* 984.537, 2+) and modified peptide (pyroEVQLQQSGAELVKPGASVK, *m/z* 975.532, 2+). Pyro-Glu and pyro-Gln represent isobaric structures and are formed by intraresidual elimination of water and ammonia, respectively, between the N-terminal amino group and the carboxy and amide moieties, respectively, of Glu and Gln side chains. The observation that the tryptic peptide containing the pyro-Glu residue elutes slightly later (39.55 min) than the native peptide (35.21 min) is in agreement with previously reported studies of pyro-Glu peptides [208-210]. Pyro-Glu formation in antibodies and its identification by MS have been described previously [208-210] and is believed to arise via non-enzymatic pathways after prolonged storage of the samples at high temperature (45°C) and low pH conditions (pH 4) in almost all proteins containing glutamic acid at the N-terminus.

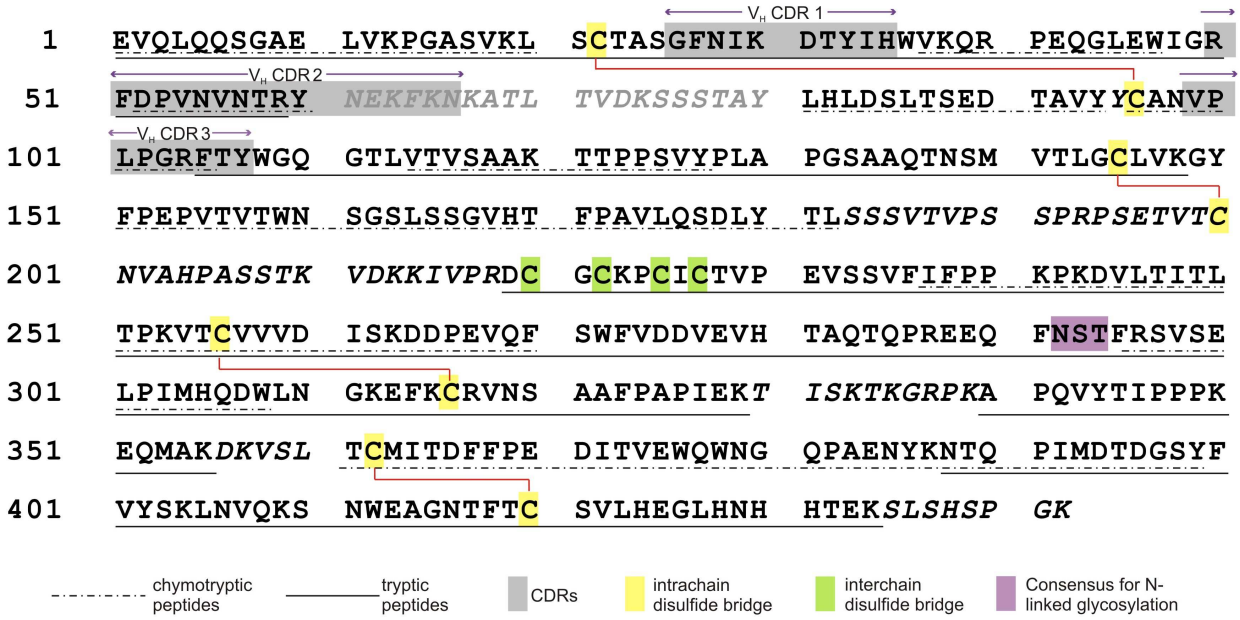


Figure 2.14: Amino acid sequence determined for the 6E10 antibody heavy chain; The CDRs are highlighted with grey boxes; the intra- and interchain disulfide bridges are indicated in yellow and green, respectively; ▬ regions identified using trypsin; ▬·▬·▬ regions identified using chymotrypsin. The amino acids in italic were fitted in the heavy chain sequence based on homology with other antibody sequences and these could not be identified by either method.

The amino acid sequence for the light chain of 6E10 is presented in Figure 2.17, for which the calculated sequence coverage was 95%. Analogous to our approach to sequence identification for the heavy chain, the peptides from the 6E10 light chain were inserted into a template light chain sequence from the database (accession number AAA39162). Light

chains usually contain 220 residues, comprised of the N-terminal variable region (V_L, approximately 110 amino acids) and a C-terminal κ or λ constant region (C_L, 110 amino acids). The complementary determining regions (CDRs) were matched using the Kabat rules [204, 205] and their positions highlighted with grey boxes in Figure 2.17.

The constant region C_H (113-220) was completely ascertained from MS/MS data which revealed that the light chain of 6E10 is a κ chain. A peptide from the variable region, corresponding to the region V_L(103-112) remained unelucidated, despite the use of alternative enzymes. The abundance of the ion *m/z* 628.803 (2+), assigned to the peptide V_L(69-81) was too low for the automatic MS/MS selection, therefore, the assignment was made based only on the mass accuracy of 3 ppm of the experimental mass. This peptide, however, is a feasible candidate for the region V_L(69-81), as the four amino acids at the C-terminus, TLTI (Figure 2.17), are overlapping with those observed at the N-terminus in the peptide V_L(78-89). MS analyses of the trypsin digest provided 100% sequence coverage of the constant region, while both enzymes provided complementary information for the variable region. The region V_L(68-102), determined from the MS/MS data of the chymotryptic digest, lacks potential trypsin cleavage sites, indicating that charge states higher than 3+ or 4+ are expected for the corresponding tryptic peptide in the MS and a complete *de novo* characterization would be challenging and probably unreliable. Upon analyzing the sequence of a large number of antibody light chains in the database, it was observed that the V_L(103-110) region may contain a considerable number of potential chymotrypsin cleavage sites and this could explain why no peptides of comparable length were observed in the LC/MS experiments. In addition, a single potential trypsin cleavage site is located at the C-terminal of V_L (103-110) (residue 109, Figure 2.17) and this could explain the lack of the short region (110-113) in the MS. The amino acid residues spanning the V_L(25-36) CDR1 were not identified from the database search of the MS/MS data, suggesting that the corresponding protein sequence may not be contained in the NCBInr protein database and/or the ion corresponding to this region was not selected for MS/MS analyses. Manual interpretation of the MS/MS data of the ion of *m/z* 610.314 (2+) (Figure 2.15), however, revealed an amino acid sequence of a tryptic peptide resembling the CDR1 containing region V_L(25-36) based on its homology to other mouse light chains (NCBI accession number ABK64007). These data enabled the determination of the peptide sequence ASQS(L/I)(L/I)SSGNQK based on the observation of a complete y ion

series and a partial b ion series. Because Ile and Leu are isobaric, these amino acids could not be discriminated by MS/MS.

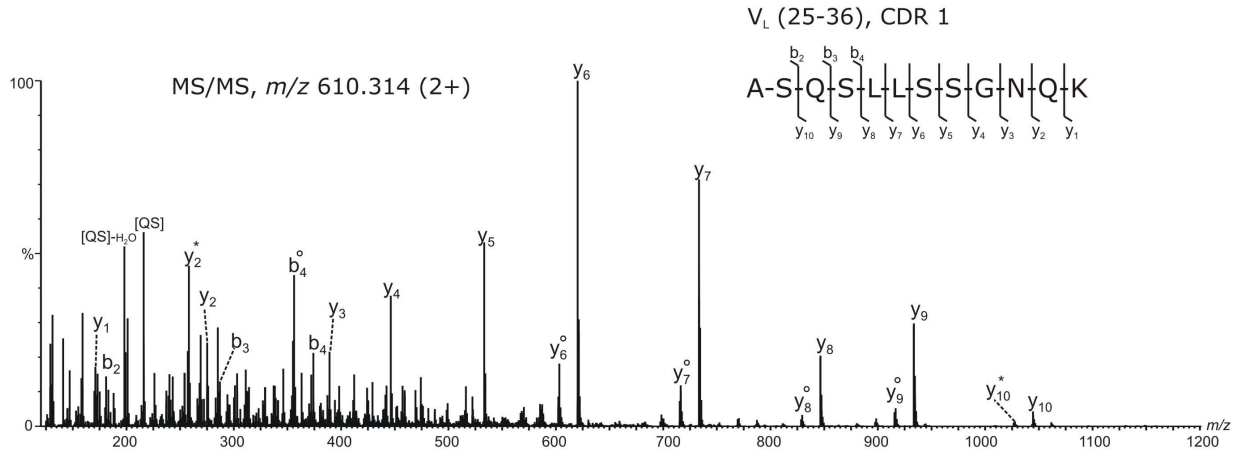


Figure 2.15: Fragment mass spectrum of the precursor ion of m/z 610.33 (2+) from the light chain tryptic digest assigned by de novo interpretation to the CDR1 peptide V_L(25-36), indicated at the top (right). The observed backbone cleavages leading to formation of b and y ions in the spectrum are indicated at the top (right). Internal fragments are indicated between square brackets. The asterisk (*) denotes loss of ammonia, the empty circle (○) denotes loss of water.

The amino acid sequence corresponding to the V_L(97-102) CDR3 was ascertained from the database search using the MS/MS spectrum of the $(M+2H)^{2+}$ ion of *m/z* 611.240, which is consistent with the peptide sequence CQNDYNYPF, where the Asn99 residue was assigned as deamidated and the cysteine in alkylated form. According to the Kabat rules [204], the V_L CDR1 may have a length between 10 and 17 residues and is preceded by the first cysteine of the variable region (Cys23) and followed by the sequence motif WYQ. The V_L CDR3 starts also after a cysteine and is followed by the residues FGXG, where X can be any amino acid. The V_L CDR1 of the 6E10 antibody contains the maximal number of amino acids according to the Kabat definition (see Figure 2.17). The V_L CDR3 fits the definition as well, however, the identity of the amino acid at the C-terminus could not be confirmed from the MS/MS.

The MS analyses of the digest mixtures of the light chain of 6E10 revealed considerable microheterogeneity in the second complementary determining region CDR2 corresponding to V_L(52-60), which has not been found in any previous study. Eight different tryptic peptides were observed by LC/MS/MS which contain the amino acid motif LLIYXXXXR, where X indicates the possible variations (Table 1). Representative MS/MS spectra of three of these ions are shown in Figure 2.16.

Table 1: Amino acid sequence heterogeneity of variable region peptides from antibody light chain V_L (52-60), identified in the LC/MS/MS analysis of the light chain tryptic mixture

Peptide No.	Peptide sequence [a] LLIYXXXXR	Observed mass (Da)	Calculated mass (Da)	Mass accuracy (ppm)
1	LLIYDASIR	1062.6046	1062.6073	3
2	LLIYGLDIR	1074.6450	1074.6437	1
3	LLIYWASTR	1121.6174	1121.6233	5
4	LLIYWASIR	1133,6540	1133.6597	5
5	LLIYWSSTR	1137.6058	1137.6182	11
6	LLIYWVSTR	1149.6492	1149.6546	5
7	LLIYKVSYR	1153.6800	1153.6859	9
8	LLIYWDSTR	1165.6168	1165.6132	3

[a] Peptide sequences identified by database search performed with the MS/MS data; X represents the observed amino acid variations which are underlined within each peptide.

Because these peptides are identical at their N-termini (Leu–Leu–Ile–Tyr-), their fragmentation patterns are very similar with the most abundant product ions belonging to the y series of ions. Two diagnostic product ions observed in the CID spectra of all these ions, *m/z* 199.175 (a_2) and *m/z* 227.176 (b_2), are specific for Leu-Leu N-terminal sequence. The observation of these ions in the MS/MS data was used to identify as many variants as possible of different precursor ions. Ions which produced these fragment ions upon MS/MS analyses are shown in Table 1. Based on manual interrogation of the MS/MS data and/or a subsequent search of the NCBInr database (taxonomy: all species), the peptide sequences were assigned. Of the determined amino acid sequence, the peptide sequence LLIYDASIR was found in human antibody light chains, but not in mouse antibodies. The same set of tryptic peptides was found in all three analyzed antibody batches, suggesting that this observed microheterogeneity was not an isolated random event. This structural diversity may be well explained by somatic hypermutations in Ig V_L genes observed at the protein level.

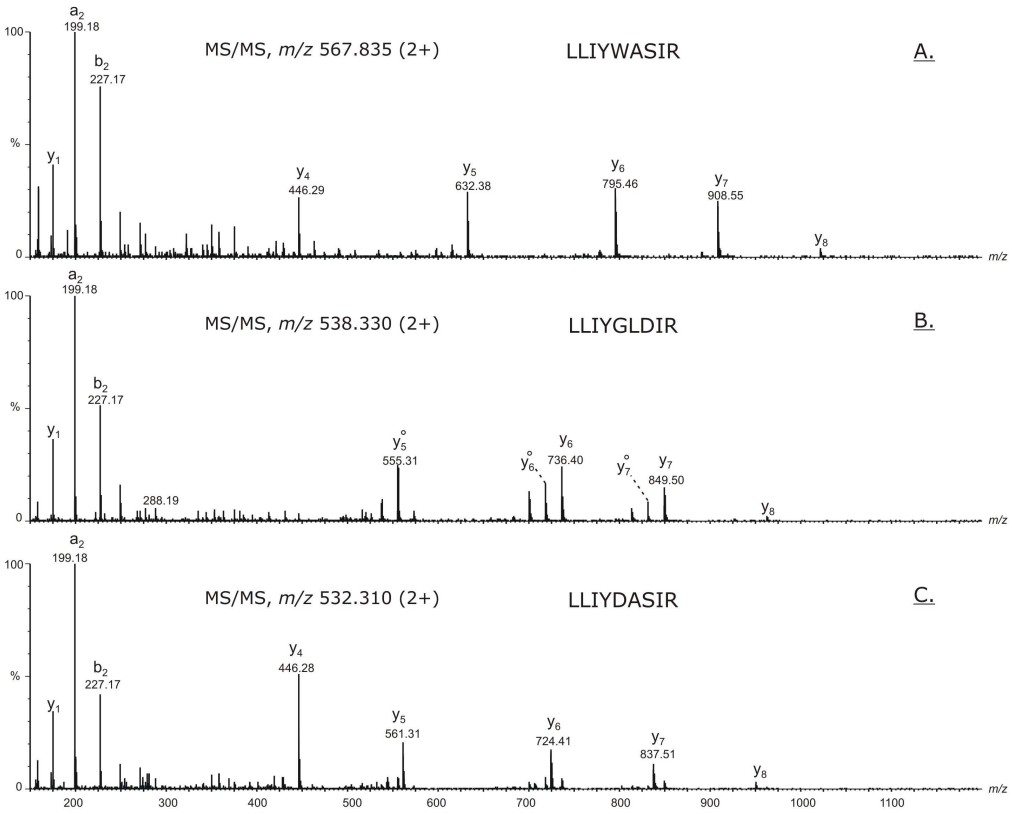


Figure 2.16: Representative product ion spectra of the precursor ions from light chain variable region peptides: (a) m/z 567.835 (2+), (b) m/z 538.330 (2+), and (c) 532.310 (2+), confirming the amino acid sequence of three peptides spanning the light chain region V_L(52-60).

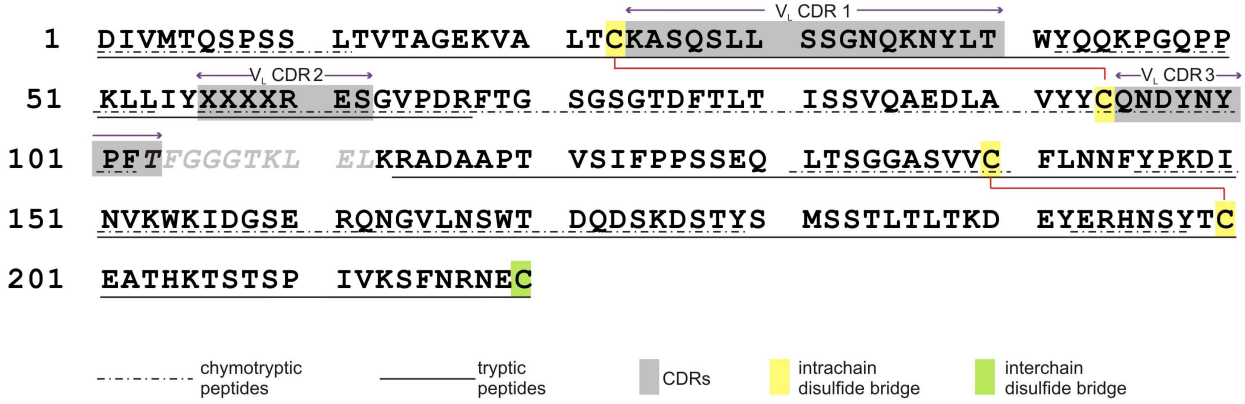


Figure 2.17: Amino acid sequence determined for the 6E10 antibody light chain; The CDRs are highlighted with grey boxes; the intra- and interchain disulfide bridges are indicated in yellow and green, respectively; ▬ regions identified using trypsin; ▬·▬·▬ regions identified using chymotrypsin. The amino acids in italic were fitted in the light chain sequence based on homology with other antibody sequences and these could not be identified by either method.

2.1.3.2 Primary structure determination of a "plaque-protective", anti-Aß(17-24) monoclonal antibody

The murine monoclonal antibody 4G8, raised against residues 17-24 of β-amyloid, was produced by hybridoma technology. Recent studies showed that β-amyloid species from sera of patients with severe AD are less reactive to the 4G8 antibody than sera from patients with mild AD, suggesting that changes of the exposed β-amyloid epitopes may be associated with the progression of disease [211].

The primary structure of the 4G8 antibody was determined from the LC-MS/MS analysis of tryptic and chymotryptic digests of its heavy and light chains separated by SDS-PAGE. The fragmentation mass spectra were obtained using data dependent acquisition of combined electron transfer dissociation on the Agilent ion trap and collision induced dissociation on the Q-Tof Premier. As described for the 6E10 antibody, the MS/MS data were searched against the NCBInr protein database using the Mascot search engine (to assign the antibody constant regions) and were manually interpreted, in an effort to assign antibody variable regions not present in the database.

The antibody was reduced and alkylated in solution and subsequently analyzed by SDS-PAGE using identical experimental conditions as for the 6E10 antibody. A single heavy chain band was observed at an apparent molecular weight of 50 kDa (band 1), whereas two distinct protein bands were observed around 25 kDa, labeled as band 2 and band 3 respectively in Figure 2.18. Their identity was established by LC-MS/MS and data base search which showed that both bands represented antibody light chains. Their differential migration and thickness may be explained by partial glycosylation of the light chain. It has been shown that glycosylation affects the electrophoretic properties of biopolymers, i.e. glycoproteins migrate at higher molecular weight than non-modified proteins. Furthermore, glycosylated proteins often migrate as diffuse bands in SDS-PAGE, as a result of the molecular heterogeneity derived from glycosylation. Antibody heavy chains are entirely glycosylated, such that a single thick band was observed on SDS-PAGE.

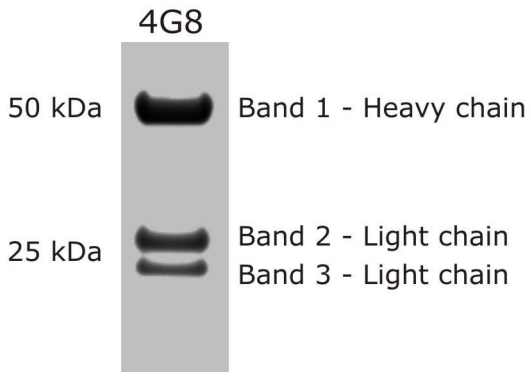


Figure 2.18: SDS-PAGE of reduced and alkylated 4G8 antibody, showing a thick heavy chain band (band 1) and two distinct light chain bands (band 2 and band 3).

The amino acid sequence determined for the heavy chain of mAb 4G8 is shown in Figure 2.22. The peptides determined from the LC-MS/MS analyses of heavy chain tryptic and chymotryptic digests were fit to a known antibody sequence contained in the NCBInr database based on homology. The constant region C_H(111-457), identified by database search, suggested that monoclonal antibody 4G8 belongs to the IgG_{2b} isotype.

The peptides contained in the heavy chain variable region, V_H(1-110) were deciphered from the manual interpretation of ETD and CID data obtained from a tryptic peptide mixture. Collectively, 94% of the variable region V_H(1-110) was identified (see Figure 2.22). The ETD and CID spectra of the doubly protonated peptide of *m/z* 868.455 are shown in Figure 2.19. The observed fragment ions are consistent with the tryptic peptide V_H(24-39) which, according to the Kabat rules [204, 205], spans the V_H CDR1 between residues 31-35. In ETD, 11 of the 15 predicted N-Cα backbone cleavages yielding z• radical ions, were observed, covering the N-terminal and the middle region of the peptide. The observed z• fragments are indicated in Figure 2.19 A. The identity of the C-terminal amino acids 36-39 and of those on positions 29 and 30 in peptide V_H(24-39) was unambiguously determined from the fragment ion spectrum of the same precursor ion obtained by low energy collisions. Multiple amino acid combinations are possible for region 36-39, whereas for residues 29-30 only two isobaric amino acid combinations, Val-Asp and Leu/Ile-Thr, respectively, can be assigned to the mass of 214 Da separating the fragment ions z_9• and z_{11}•. The deconvoluted mass spectrum[5] over the mass range 100-

[5] Deconvolution of mass spectrometric data obtained by electrospray ionization is an algorithm-based process which transforms the recorded charge envelope of a protein/peptide into the corresponding singly charged mass. A fragment mass spectrum containing ions in multiple charge states may be difficult to assign if manual interpretation is employed; hence deconvolution is used to generate the mass spectrum containing exclusively singly charged ions.

1750 is shown in Figure 2.19 B. The entire series of y and b ions (except for the ion b_1, which can not be observed in CID based on mechanistic considerations) confirmed the peptide assignment derived from ETD and identified the residues 29 and 30 as Leu/Ile and Thr, respectively, whereas residues 36-39 were determined as Trp, Val, Gln and Lys.

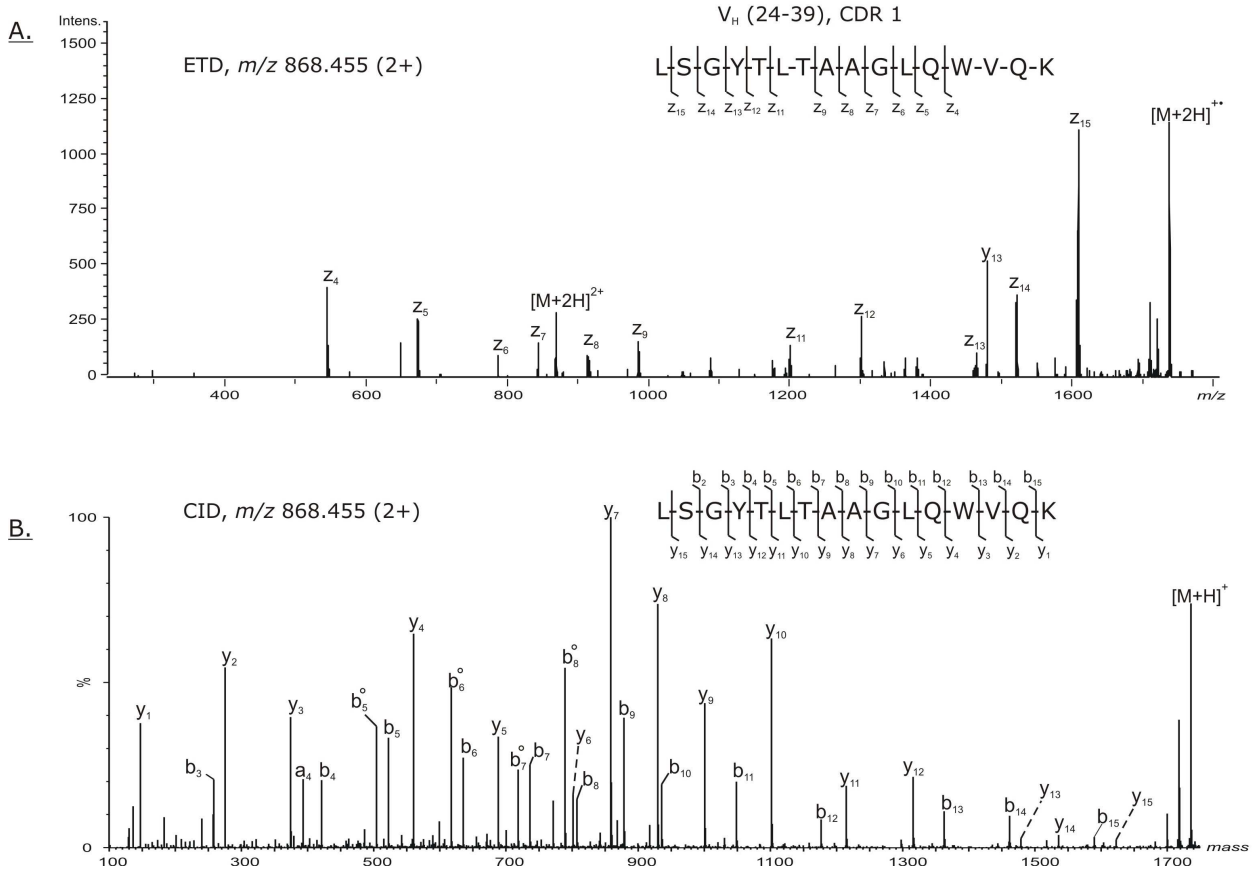


Figure 2.19: (A) Electron transfer dissociation (ETD) and (B) Collision induced dissociation (CID) spectrum of the precursor ion m/z 868.455 (2+), assigned to peptide V_H(24-39) and containing the CDR1 between residues 31-35. For easier interpretation, the CID spectrum was deconvoluted (see Footnote 5). The observed fragment ions derived from backbone cleavage, c and z for ETD, and b and y for CID, respectively, are indicated at the top of each figure. The precursor ion is indicated between square brackets. The empty circle (○) indicates loss of water from the b fragment ions in CID.

The peptide WIGWINTQSGVPR, corresponding to the region V_H(50-59) and containing the partial sequence of the heavy chain CDR2 was determined from the ETD and CID data of the doubly protonated precursor of *m/z* 757.401, shown in Figure 2.20. The series of radical ions z_3• to z_{12}• observed in ETD were consistent with the indicated amino acids (Figure 2.20 A), whereas the information regarding the C-terminal region 56-59 was derived from CID (Figure 2.20 B). Ions arising from amide bond cleavage N-terminal to the Pro residue have characteristic higher abundance in the CID spectra of protonated peptides, as is the y_2 ion observed at *m/z* 272.172. The remaining amino acids of the

CDR2 were determined from the CID spectrum of the doubly protonated molecule of *m/z* 386.681, assigned to the peptide ^{61}YAEDFK^{66} (see Figure 2.22).

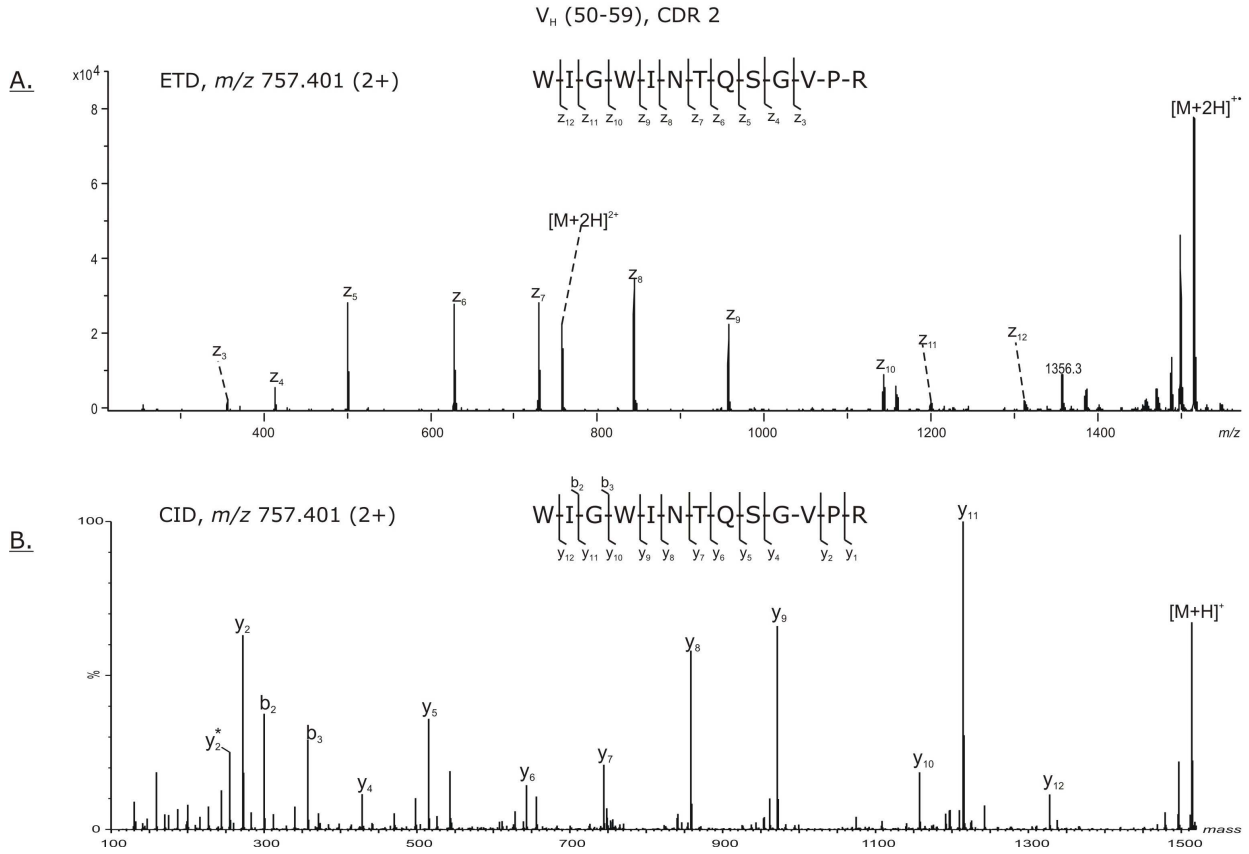


Figure 2.20: (A) Electron transfer dissociation (ETD) and (B) Collision induced dissociation (CID) spectrum of the precursor ion m/z 757.401 (2+), assigned to peptide V_H(50-59). For easier interpretation, the CID spectrum was deconvoluted. The observed fragment ions derived from backbone cleavage, c and z for ETD, and b and y for CID, respectively, are indicated at the top of each figure. The precursor ion is indicated between square brackets. The asterisk (*) indicates loss of ammonia from the y fragment ions in CID.

The V_H CDR3 was found to contain a unique combination of 11 amino acids, based on the ETD/CID spectra of the triply protonated peptide of *m/z* 861.395 showed in Figure 2.21. The peptide spanning the variable region V_H(100-123) was determined as SGHYGNYAMDYWGQGTSVTVSSAK, containing oxidized methionine on position 108. Oxidation of the thioether bond of methionine side chain to methionine sulfoxide, Met(O), is a common artefact which may occur during sample preparation involving gel electrophoresis. The addition of 16 Da to the mass of Met results in a residue mass isobaric with that of phenylalanine, requiring careful interpretation of the mass spectrometric data. The identification of peptides containing oxidized methionine is facilitated by the characteristic loss of methanesulfenic acid (CH_3SOH, 64 Da) from Met(O) under CID conditions. The mechanism of the gas-phase elimination of CH_3SOH has been investigated by O'Hair & Reid using deuterium labelling experiments and MS/MS/MS,

which suggested that both charge-directed and charge-remote decomposition pathways may contribute to the observed neutral loss [212]. In the CID spectrum showed in Figure 2.21 B, the deconvoluted mass of 2518.126 Da corresponds to the neutral loss of 64 Da from the triply protonated molecular ion, hence indicating that at least one oxidized methionine is contained in the peptide. Based on the mass difference between the fragments b_8 and b_9, the position V_H 108 was assigned to the Met residue. Although peptide V_H(100-123) was almost entirely ascertained from the observation of complementary y and b ion series, y_1 - y_{15} and b_6 - b_{21}, respectively, limited information could be determined for the N-terminal part, corresponding to residues 100-105. The identity of these amino acids is of major importance, as these residues resemble the core of the CDR3, which determines the specificity of antibodies for antigens.

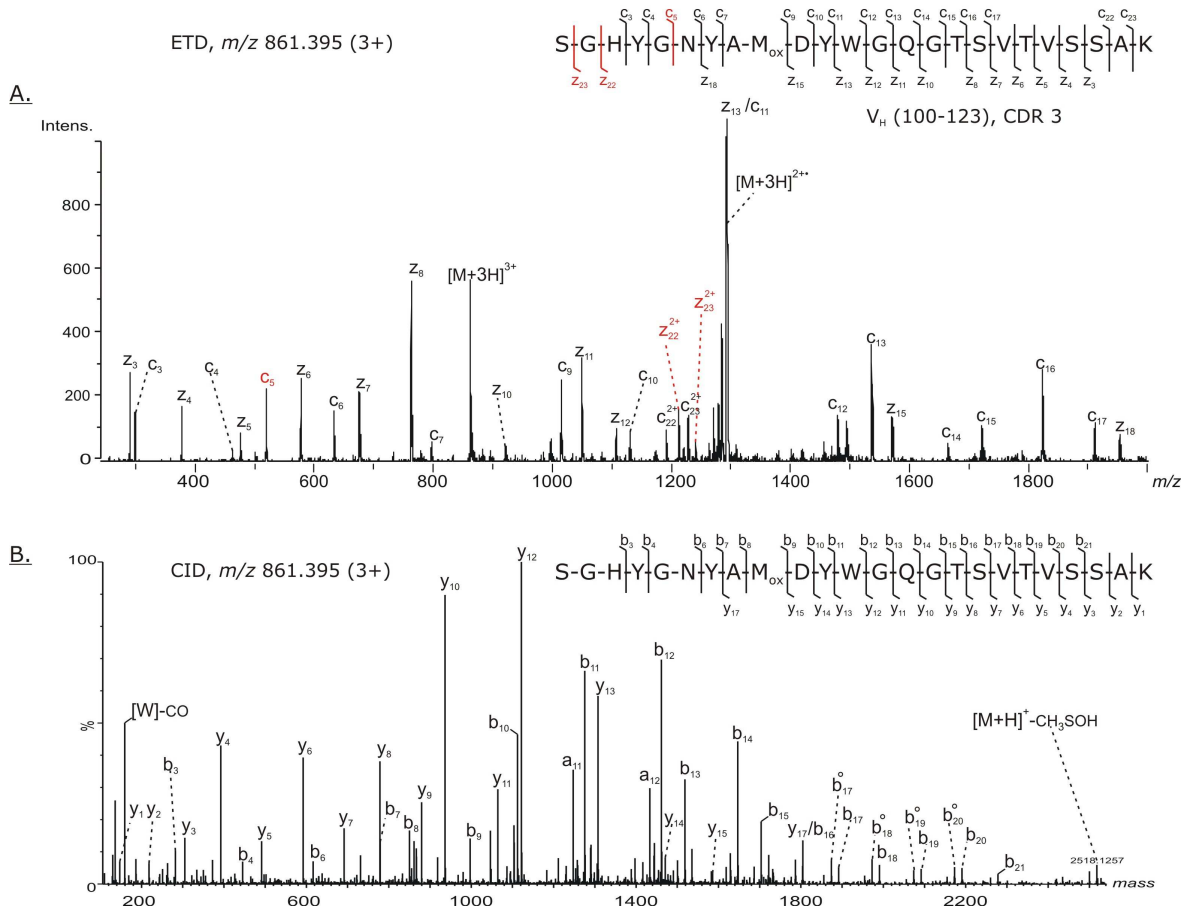


Figure 2.21: (A) Electron transfer dissociation (ETD) and (B) Collision induced dissociation (CID) spectrum of the precursor ion m/z 861.395 (3+), assigned to peptide V_H(100-123). For easier interpretation, the CID spectrum was deconvoluted. The observed fragment ions derived from backbone cleavage, c and z for ETD, and b and y for CID, respectively, are indicated at the top of each figure. Those fragment ions crucial for the complete assignment, not observed by CID, are indicated in red in the ETD spectrum. The precursor ion is indicated between square brackets. The empty circle (○) denotes loss of water from the b ions in CID.

With the use of ETD (Figure 2.21 A) the identity of the remaining residues was unambiguously identified from the successive fragments c_3, c_4, c_5, c_6 and c_7 (*m/z* 299.2, 462.4, 519.4, 633.5 and 796.8, respectively), and from the doubly charged radical ions z_{22}• (*m/z* 1211.5) and z_{23}• (*m/z* 1240.0).

The N-terminal tryptic peptide V_H(1-12) was observed exclusively as the doubly protonated molecule of *m/z* 661.833, corresponding to the tryptic peptide pyroXIQLVQSGPELK containing pyro-Glu or pyro-Gln at the N-terminus. As the mass of these two modified residues is identical, the identity of the N-terminal amino acid could not be unambiguously determined. The full length sequence determined for the 4G8 heavy chain is shown in Figure 2.22.

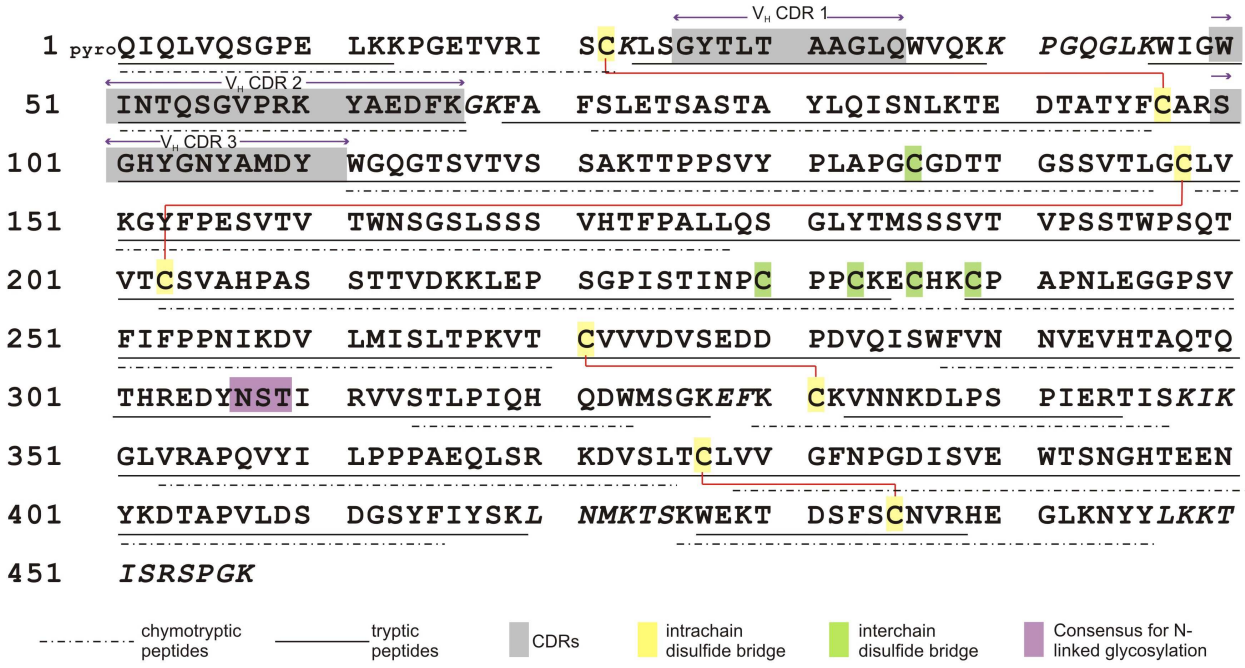


Figure 2.22: Amino acid sequence determined for the 4G8 anti-Aß(17-24) antibody heavy chain (band 1 in SDS-PAGE). The sequence coverage obtained using trypsin and chymotrypsin is underlined. The peptides were fit against a known antibody sequence from the database based on homology. The amino acids in italics could not be identified. The CDRs are highlighted with grey boxes, the consensus sequence for N-glycosylation is highlighted with purple, the intra- and interchain disulfide bridges are indicated in yellow and green, respectively.

These examples describing the identification of variable region peptides containing the antigen binding regions illustrate the importance of high quality MS/MS data for *de novo* sequencing, rendering the combination ETD/CID a powerful tool for primary structure determination of immunoglobulins.

The amino acid sequence of the 4G8 antibody light chain was determined from the LC-MS/MS analyses of enzymatic digests the gel band 3, which was assumed to be non-glycosylated (Figure 2.18). The peptides were fit against a known antibody light chain from the NCBInr database, based on homology (database access number AAA92993). The variable region V_L(1-110) was completely ascertained from the LC-MS/MS analysis of tryptic and chymotryptic digests, whereas only five amino acids from the light chain constant region, indicated in *italics* in Figure 2.25, remained unidentified. The 4G8 light chain belongs to the κ isotypes, based on the amino acid composition of the constant region. Overall, the obtained sequence coverage for the light chain was 98%. The amino acid sequence of the light chain, with the complementary determining regions highlighted with grey boxes, is presented in Figure 2.25. The peptides spanning the regions V_L CDR1 and V_L CDR2 were obtained by database search of the MS/MS data, and were found identical with those contained in light chains of antibodies with distinct antigen specificities (accession numbers: XP_001476729 and XP_001476714, antibodies against melanoma antigen isoform 2). In contrast, V_L CDR3, determined by manual interpretation of the MS/MS data, was not contained in any antibody sequence from the NCBInr database, suggesting that the combination of amino acids in V_L(93-102) may be crucial for the Aß(17-24) recognition by this antibody.

The region V_L(66-96) was determined by manual interpretation of the ETD and CID spectra of the precursor ions *m/z* 858.6 (4+) (Figure 2.23), and *m/z* 874.417 (2+) (Figure 2.24), respectively. The ETD of the quadruply charged ion *m/z* 858.6 contained a series of successive c ions assigned to the N-terminal part of the molecule (c_3 through c_{12}), and several z• radical ions that provided information about the amino acids at the C-terminus (z_2•, z_3•, z_4•, z_6• and z_7•), suggesting that Cys is present in alkylated form. Sequence ions from the middle region of the molecule, residues 78-89 were not observed in the *m/z* scan range 200-2000, suggesting that these may carry a single charge that renders their observable *m/z* ratio outside the scan range, and hence they were not recorded. Another possibility is that the abundance of the corresponding multiply charged species in the spectrum was low and the poor mass resolution of the ion trap did not allow their unambiguous assignment. The partially identified residues in peptide V_L(66-96) indicate that there are several potential chymotryptic cleavage sites in this sequence, such as Phe75, Leu77, Tyr90 and Leu91, respectively. In order to determine the identity of the amino acids from the middle region, chymotrypsin was added to the tryptic mixture

containing the fragment V_L(66-96) and the resulting mixture was analyzed by LC-MS/MS on the Q-Tof Premier. The deconvoluted CID spectrum of the ion *m/z* 874.417 (2+) (Figure 2.24) was assigned to the peptide V_L(76-90), based on the presence of the complete series of b ions (except for b_6) in the spectrum. As for heavy chain of 4G8, the neutral loss of 64 Da from the precursor indicated the presence of oxidized methionine, Met(O), and the assignment of this residue was made based on the mass difference between the successive fragments b_{13} and b_{14}.

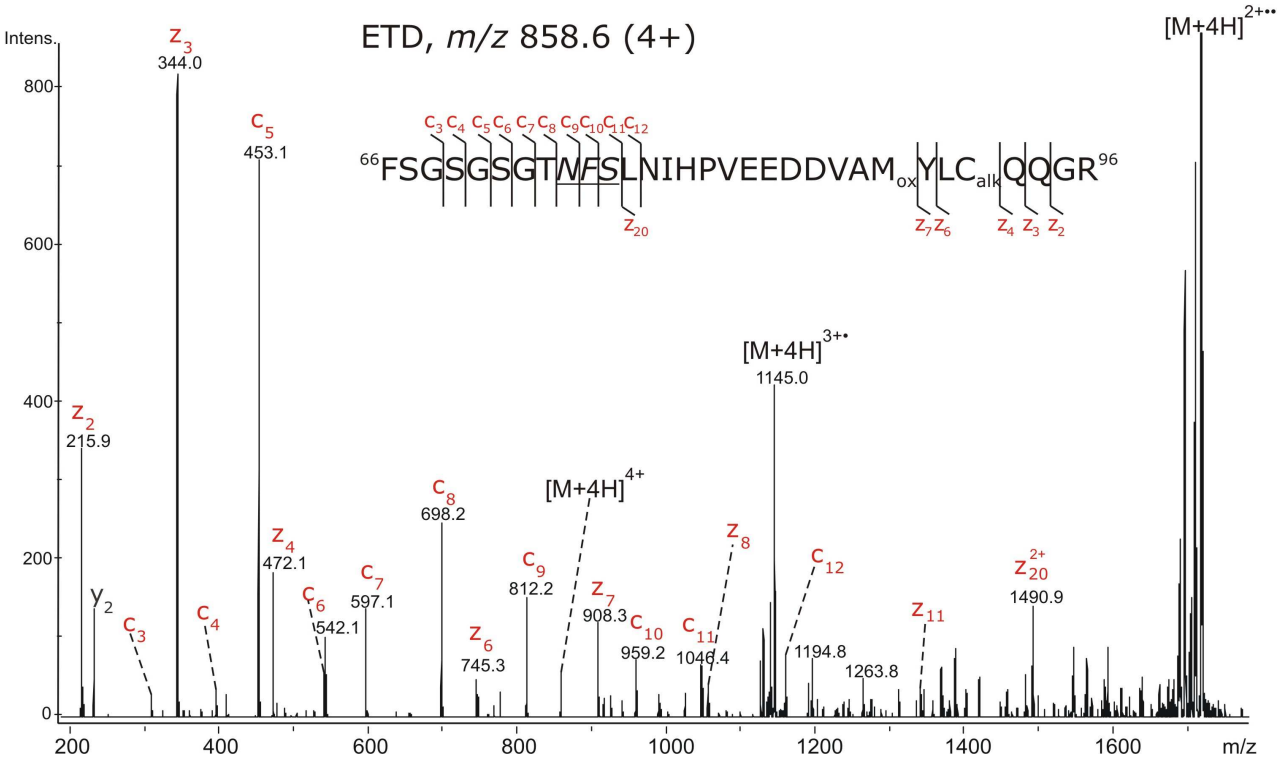


Figure 2.23: Electron transfer dissociation (ETD) of the precursor ion m/z 858.6 (4+), corresponding to tryptic peptide V_L(66-96) from the 4G8 light chain variable region containing alkylated cysteine and oxidized methionine. The observed c and z• fragment ions are indicated in red. The N-terminal region was ascertained from the observation of the successive fragments c_3 through c_{12}, whereas the C-terminal part was determined based on the radical ions z_2•, z_3•, z_4•, z_6• and z_7•, respectively. No information was derived for the middle region of the molecule. The spectrum was obtained on the ion trap using data dependent acquisition.

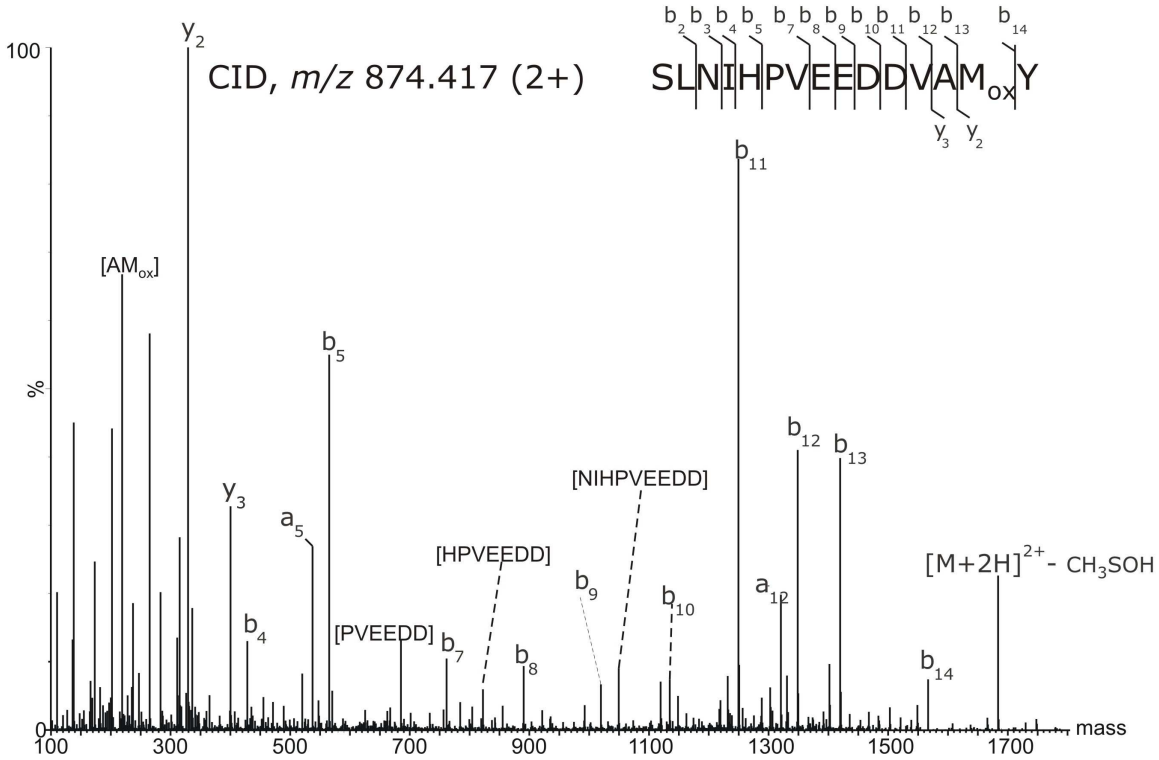


Figure 2.24: Collision induced dissociation (CID) of the precursor ion m/z 874.417 (2+), corresponding to chymotryptic peptide V_L(76-90), containing oxidized methionine. The MS/MS was deconvoluted for easier assignment. The complete series of observable b ions (except for b_6) are contained in the spectrum, complementing the information obtained by ETD for the region V_L(66-96). Internal fragments derived from cleavage of two amide bonds are indicated between square brackets. The spectrum was obtained on the Q-Tof Premier using data dependent acquisition with a collision energy ramp from 20 V to 30 V.

In addition to the b and y sequence ions observed in CID, fragments derived from simultaneous cleavage of two amide bonds were observed, and their composition is indicated between square brackets in Figure 2.24. Although these so called internal fragments are informative by confirming the amino acid composition, their presence rather complicates the interpretation of the spectrum if the sequence of the peptide is unknown. Importantly, the fragment ions observed in ETD (Figure 2.23) revealed a novel consensus sequence for *N*-linked glycosylation spanned between residues V_L(74-76), and this site is probably partially occupied with *N*-glycans, resulting in differential migration of gel bands 2 and 3. The structural features of *N*-glycosylation and the analysis of constant and variable region glycosylation of the 4G8 antibody are discussed in Chapters 2.2.1 and 2.2.4. The full length sequence obtained for the light chain of the 4G8 antibody is presented in Figure 2.25.

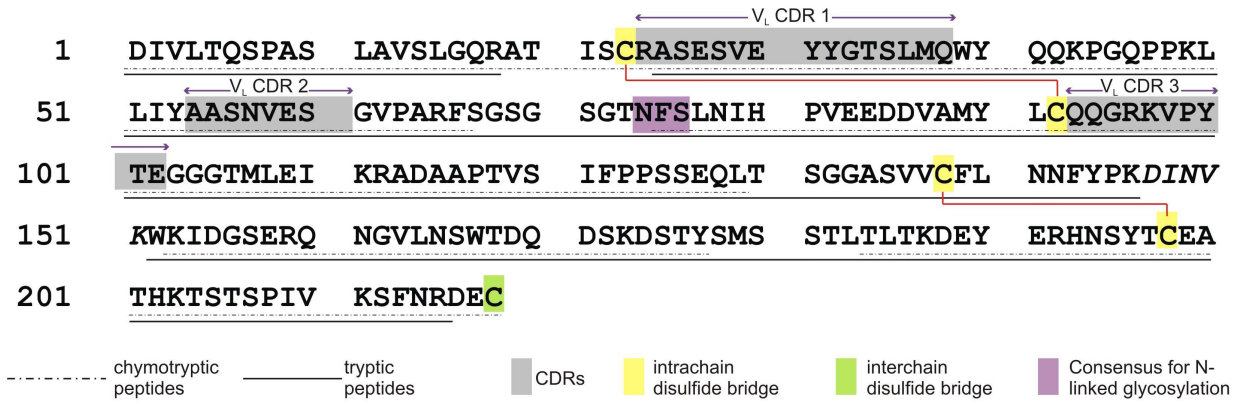


Figure 2.25: Amino acid sequence determined for the 4G8 anti-Aß(17-24) antibody light chain (band 3 in SDS-PAGE). The sequence coverage obtained using trypsin and chymotrypsin is underlined. The peptides were fit against a known antibody sequence from the database based on homology. The amino acids in italics could not be identified. The CDRs are highlighted with grey boxes, the consensus sequence for N-glycosylation is highlighted with purple, the intra- and interchain disulfide bridges are indicated in yellow and green, respectively.

2.1.3.3 Epitope-specific affinity purification of plaque-protective ß-amyloid autoantibodies

Intravenous immunoglobulin (IVIg) is a commercially available product isolated from plasma of healthy individuals and a source of naturally occurring Aß autoantibodies. A phase 3 clinical trial investigating safety and effectiveness of IVIg for the treatment of mild to moderate AD was started in December, 2008. The source of IgG used in the present work was human gamma-globulin from Calbiochem, supplied as a lyophilized powder containing 97% IgG.

The epitope-specific isolation of Aß autoantibodies from IVIg was performed using Cys-Aß(12-40) covalently immobilized on Ultralink® iodoacetyl gel. The attachment of the peptide to the resin occurs via nucleophilic attack of the free sulfhydryl group of the Cys side chain on the primary carbon atom of the iodoacetyl moiety, with formation of a thioether bond between Cys-Aß(12-40) and the resin. The necessity for the elongation of Aß(12-40) with Cys at its N-terminus is derived from the fact that Aß(12-40) contains two Lys residues, Lys16 and Lys28, respectively, and the reactivity of their side chain amino groups is comparable to that of the N-terminus; because the sulfhydryl group is more reactive than the amino group in a nucleophilic substitution reaction, one can assume that the immobilization of Cys-Aß(12-40) on the resin occurs, to a large extent, through the Cys side chain. Upon immobilization, the resin was thoroughly washed to remove the unreacted material, and the remaining iodoacetyl groups were blocked with a solution

containing 50 mM Cys in coupling buffer. IVIg was dissolved in PBS to a concentration of 1.25 mg/mL and was then mixed with the resin containing immobilized Cys-Aß(12-40) with incubation over night at 4℃. After removing of the unbound IgG fraction by washing with PBS, the Aß autoantibodies were eluted under acidic conditions. Their concentration in the elution fraction was determined by the MicroBCA™ Protein Assay Kit. For structural studies, the autoantibody solution was lyophilized.

2.1.3.4 Primary structure determination of plaque-protective Aß-autoantibodies

The primary structure of Aß autoantibodies was characterized using both *Strategy I* and *Strategy II*, depicted in Figure 2.9. As for the monoclonal antibodies described in the previous sections, the lyophilized Aß autoantibody was reduced and alkylated in solution and subsequently separated into heavy and light chains by SDS-PAGE. Each protein band was digested with trypsin and the resulting peptide mixtures were analyzed by LC-MS/MS using data dependent acquisition (DDA), which employs selection for MS/MS of several most abundant ions at a given point in time (*Strategy I*). The number of precursor ions is preset by the user. All LC-MS/MS analyses performed on the Q-Tof Premier involved automatic selection of four most abundant precursor ions. The scan sequence contains a full scan MS, typically 1 second, in which the ions are recorded, followed by four MS/MS cycles × 1.4 seconds, in which fragment ion spectra of the precursors that have met the abundance selection criteria are recorded separately. However, considering the duration of a duty cycle for DDA (7 seconds), a large amount of information is lost during these analyses, i.e. low abundant peptides are not selected for fragmentation. As a result of the polyclonal nature of the Aß autoantibody and of the fact that 95 % of the constant regions are conserved among immunoglobulins, tryptic peptides derived from constant regions have highest abundance, while those derived from the antibody variable regions are extremely heterogeneous and, consequently, each isoform has lower incidence in a digestion mixture.

The Aß autoantibody heavy chain tryptic peptides determined by database search and *de novo* interpretation of the LC-MS/MS data obtained on the Q-Tof Premier using data dependent acquisition are presented in Table 2.

Table 2: Heavy chain tryptic peptides from the Aß-autoantibody determined by LC-MS/MS on the Q-Tof Premier using data dependent acquisition.

Peptide Nº	Residue number	Peptide sequence [f]	MW (exp.)	MW (theor.)	Delta M (ppm)
1[a]	[1-19][e]	pyroXVQLVESGGGLVKPGGSLR	1863.0654	1863.0578	4
2[a]	[1-16][e]	pyroXVQLVESGGGVVQPGR	1591.8596	1591.8682	5
3	[1-19][e]	EVQLLESGGGLVKPGGSLR	1895.0534	1895.0476	3
4	[1-19][e]	EVQLVESGGGLVKPGGSLR	1881.0387	1881.0320	4
5[b]	[1-19][e]	(ZZZZ)LESGGGLVKPGGSLR	1872.0496	1872.0581	5
6	[88-98][e]	AEDTAVYYCAR	1317.5550	1317.5659	8
7	[121-132][d]	GPSVFPLAPSSK	1185.6495	1185.6394	9
8	[133-146][d]	STSGGTAALGCLVK	1320.6834	1320.6708	10
9	[133-146][d]	STSESTAALGCLVK	1422.6932	1422.7025	7
10	[223-252][d]	SCDKTHTCPPCPAPELLGGPSV FLFPPKPK	3333.6506	3333.6348	5
11	[227-252][d]	THTCPPCPAPELLGGPSVFLFP PKPK	2843.4560	2843.4503	2
12	[302-320][d]	VVSVDTVVHQDWLNGKEYK	2215.1386	2215.1273	5
13[c]	[327-334][d]	LPAAPIEK	837.4934	837.4960	3
14	[371-380][d]	NQVSLTCLVK	1160.6284	1160.6224	5
15	[403-419][d]	TTPPVLDSDGSFFLYSK	1872.9064	1872.9145	4

[a] – X can be Glu or Gln;
[b] – The identity of the amino acids marked with Z could not be determined;
[c] – Ala insertion in the motif PAAP (commonly found is ALPAPIEK);
[d] – The residues numbers of the constant regions were determined from reference X;
[e] – Numbering of the peptides were determined based on homology with the heavy chain sequence AAD30396;
[f] – All cysteine residues were observed in alkylated form.

Most of the identified peptides are derived from the antibody constant regions (peptides 7-15 in Table 2), while only a small number was found for the antibody variable regions (peptides 1-6). The residue numbers of the variable region peptides were determined based on homology with the antibody sequence AAD30396 from the NCBInr database. The majority of the peptides from the V_H region span the amino acids V_H(1-19) (peptides 1-5), indicating conservation of the N-terminal residues in Aß autoantibody, while peptide 6 is commonly found in the ß-pleated V_H region preceding V_H CDR3, according to the Kabat rules. Because of the lack of the corresponding fragments in the CID spectrum, the identity of the first four N-terminal amino acids in peptide 5 could not be determined by mass spectrometry. However, the identity of the N-terminal sequences (1-19) and the presence of N-terminal Glu was ascertained by Edman sequence determinations [89]. With the exception of the N-terminal amino acid, pyro-X, which can be assigned to both Gln and Glu, peptide 1 has an identical composition with peptide 4. Peptide 2 has similar amino

acid composition at the N-terminus with peptide 1; it appears that three amino acid deletions from the middle and C-terminal regions shorten its length to 16 amino acids, compared to peptide 1.

The Aß autoantibody light chain tryptic peptides determined by database search and *de novo* interpretation of the LC-MS/MS data obtained by data dependent acquisition are presented in Table 3.

Table 3: Light chain tryptic peptides from the Aß-autoantibody determined by LC-MS/MS on the Q-Tof Premier using data dependent acquisition

Peptide Nº	Residue number	Peptide sequence [c]	MW (exp.)	MW (theor.)	Delta M (ppm)
1	[1-18][a]	EIVLTQSPGTLSLSPGER	1882.9834	1883.0000	9
2	[1-18][a]	EIVLTQSPATLSLSPGER	1897.0240	1897.0157	4
3	[46-54][a]	LLIYGASSR	978.5502	978.5498	1
4	[46-54][a]	LLIYWASTR	1121.6326	1121.6233	8
5	[46-54][a]	LLIYDASTR	1050.5802	1050.5710	8
6	[46-54][a]	LLIYGASTR	992.5670	992.5655	2
7	[46-54][a]	LLIYDASNR	1063.5580	1063.5662	8
8	[46-54][a]	LLIYGASNR	1005.5682	1005.5607	7
9	[46-61][a]	LLIYGASSLQSGVPSR	1646.8905	1646.8992	5
10	[46-61][a]	LLIYAASSLQSGVPSR	1660.9242	1660.9148	6
11	[46-61][a]	LLIYAASTLQSGVPSR	1674.9402	1674.9305	6
12	[62-77][a]	FSGSGSGTDFTLTISR	1631.7858	1631.7791	4
13	[62-74][a]	FSGSGSGTDFTLK	1302.6156	1302.6092	5
14	[109-126][b]	TVAAPSVFIFPPSDEQLK	1945.0280	1945.0197	4
15	[127-142][b]	SGTASVVCLLNNFYPR	1796.8734	1796.8880	8
16	[170-183][b]	DSTYSLSSTLTLSK	1501.7414	1501.7512	7
17	[191-207][b]	LYACEVTHQGLSSPVTK	1888.9236	1888.9353	6
18	[191-207][b]	VYACEVTHQGLSSPVTK	1874.9403	1874.9197	11

[a] Variable region peptides were fit against the light chain sequence CAE54361 from NCBInr database;
[b] Peptides derived from kappa light chain constant region were fit against the sequence ACF34453 from NCBInr data base;
[c] All cysteine residues were observed in alkylated form.

Compared to heavy chain, a larger number of peptides from the variable region V_L were identified (peptides 1 to 13 in Table 3), probably because light chains contain half the number of residues compared to heavy chain, divided into a variable region V_L(1-110) and a constant region C_L(111-120). Peptides 14 to 18 derived from the constant region suggest that these light chains are κ chains. The observed heterogeneity in the N-terminal peptide V_L(1-18) includes a single amino acid mutation, G9A, while for peptides spanning the

amino acids V_L(46-54) and V_L(46-61) several simultaneous mutations were determined. The V_L CDR 2, seven residues long, is partially contained between residues 50-54 in peptides V_L(46-54), and entirely between residues 50-56 in peptides V_L(46-61), explaining the pronounced sequence microheterogeneity.

The complexity of the Aß autoantibody derived from its polyclonal nature, the large dynamic range of variable and constant region peptides, as well as the slow duty cycle of data dependent acquisition represented serious limitations for primary structure determination. Hence, in an attempt to determine a larger number of peptides from the Aß autoantibody variable regions, a different mass spectrometric approach was adopted, which employs separation of the heavy chain in-gel tryptic mixture by analytical RP-HPLC, collection of each fraction based on UV detection (220 nm), and analysis of each separate fraction by LC-MS ("off-line" MS analysis, *Strategy II* in Figure 2.9).

An example of RP-HPLC separation of an Aß autoantibody heavy chain tryptic mixture on C_4 stationary phase is presented in Figure 2.26. The peptides were separated using a linear gradient from 0 to 65 % solvent B (80 % acetonitrile in 0.1 % TFA) over 130 minutes. Nearly eighty-four different fractions were collected, lyophilized and subsequently analyzed by LC-ESI-MS and MS/MS on an ion trap instrument. Each fraction was dissolved in 0.1 % formic acid in a volume sufficient for 4-5 injection cycles. As these fractions were pre-separated on RP-HPLC, a linear gradient (20 to 50 % solvent B over 3 minutes) was employed for their analysis by off-line LC (C_{18})—ESI-MS, such that one analysis was performed in approximately 25 minutes.

The red numbers above the chromatographic peaks in Figure 2.26 indicate the number of the collected fractions in which antibody peptides were identified. Variable and constant region peptides are summarized in Table 4 and Table 5, respectively. The residue numbers of variable region peptides were determined based on homology with the heavy chain sequence AAD30396 from the NCBInr database, while constant region peptides were determined from reference [213]. The observed minimal peptide contained five amino acids (peptides 10, 11 in Table 4 and peptide 26 in Table 5), observed as both doubly and singly protonated molecule. Three peptides containing amino acids from V_H CDR 2 were found in peptides 6, 7 and 8, respectively. Peptide 1 from the heavy chain N-terminus V_H(1-19) was observed in both DDA and off-line LC-MS/MS. The remaining V_H peptides 3

– 5 and 9 – 18 in Table 4 are derived from variable domains flanking the antigen binding regions, which show some degree of conservation among immunoglobulins.

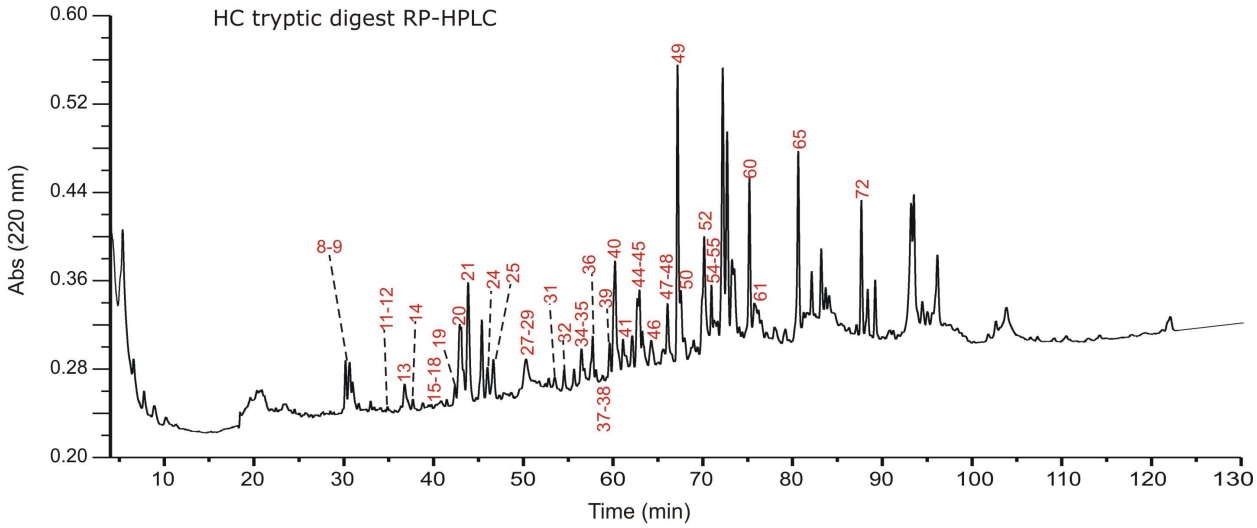


Figure 2.26: RP-HPLC chromatogram of a heavy chain tryptic mixture of the Aß autoantibody. Separation was performed on C4 stationary phase, with a gradient ranging from 0% solvent B to 65% solvent B over 130 minutes.

Table 4: Heavy chain variable region tryptic peptides from the Aß-autoantibody identified by RP-HPLC separation and individual analysis of each fraction by off-line ESI-MS/MS

Peptide Nº	Residue number[a]	Fraction number	Amino acid sequence [b]	MW (exp.)	MW (theor.)	Delta M (Da)
1	[1-19]	44	EVQLVESGGGLVKPGGSLR	1881.40	1881.03	0.37
2	[1-11]	31	pyroXVQLVQSQEVK	1267.60	1267.71	0.11
3	[20-30]	39	LSCAASGFTFR	1215.55	1215.57	0.02
4	[44-50]	39	GLVWVSR	815.75	815.47	0.28
5	[44-50]	34	GLEWVGR	815.62	815.43	0.19
6	[53-65]	35	QFFSG[SP|AL]ATGSVK	1311.61	1311.68	0.07
7	[59-65]	18	YYVDSVK	872.36	872.43	0.07
8	[57-64]	14	TNYNPSLK	935.46	935.47	0.01
9	[66-72]	20	GRPTISR	786.39	785.45	0.94
10	[68-72]	20	FTISR	622.33	622.34	0.01
11	[68-72]	13	FTVSR	608.10	608.33	0.23
12	[68-74]	12	VTITADR	774.22	774.42	0.20
13	[77-82]	16	NQLSLK	701.40	701.41	0.01
14	[77-87]	48	NTLYLQMNSLR	1351.62	1351.69	0.07
15	[83-98]	38	LSSVTAADTAVYYCAR	1746.85	1746.82	0.03
16	[114-125]	14	GTVVTVSSASTK	1135.50	1135.61	0.11
17	[114-125]	15	GTTVIVSSASTK	1149.52	1149.62	0.10
18	[114-125]	21	GTLVTVSSASTK	1149.55	1149.62	0.07

[a] Numbering of the peptides was determined based on homology with the heavy chain sequence AAD30396
[b] Cysteine residues were observed in alkylated form

Compared to DDA analysis, a larger number of peptide isoforms from the heavy chain constant region were identified (Table 5); these groups contained single or double amino acid mutations, such as P136S in the fragment C_H(126-137) (peptides 3 and 4, from IgG_1 subclass), E141G and S142G in the fragment C_H(138-151) (peptides 5 and 6, from IgG_1) and V313L, respectively, in the fragment C_H(306-321) (peptides 13 from IgG_2 and 14 from $IgG_{1,3,4}$). Three different isoforms were determined for the region C_H(361-370) (peptides 20 – 22) and C_H(393-409) (peptides 23 – 25). The extended heterogeneity in the constant region is derived from the fact: *(i)* that peptides containing amino acid substitutions are derived from different subclasses, and *(ii)* that within a specific subclass, nucleobase substitutions in the rearranged heavy chain genes may introduce additional diversity, such as in peptides 20-22 (Table 5). Several peptides were found to contain Cys at the C-terminus in alkylated form (peptides 1, 2 and 28 Table 5). As the peptides are located within the heavy chain region, it appears that in these instances trypsin cleaved at the C-terminus of alkylated Cys. With the alkylation by carbamido methyl moiety, the Cys side chain is rendered a similar length with that of Lys, although the basic character of the amide group is less pronounced than that of the Lys side chain amino group, explaining the lower cleavage rate C-terminal to Cys.

An example of RP-HPLC separation of an Aß-antibody light chain tryptic mixture on C_4 stationary phase is presented in Figure 2.27, and the variable and constant region peptides identified by off-line LC-ESI-MS/MS are summarized in Table 6.

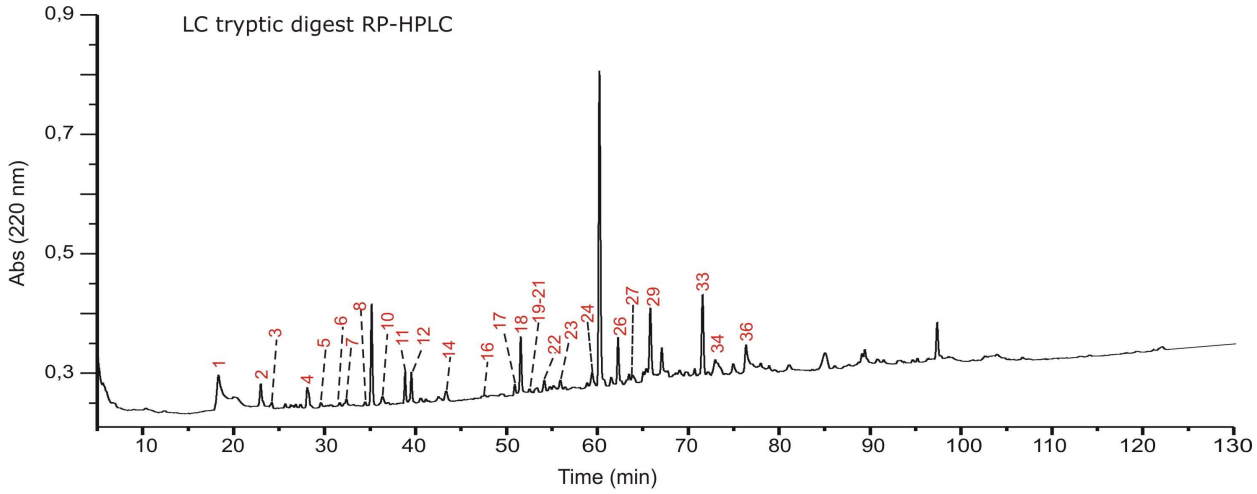


Figure 2.27: RP-HPLC chromatogram of a light chain tryptic mixture of the Aß autoantibody. Separation was performed on C4 stationary phase, with a gradient ranging from 0% solvent B to 65% solvent B over 130 minutes.

Table 5: Heavy chain constant region tryptic peptides from the Aß-autoantibody identified by RP-HPLC separation and individual analysis of each fraction by off-line ESI-MS/MS

Peptide N°	Residue number[a]	IgG Subclass	Fraction number	Amino acid sequence [b]	MW (exp.)	MW (theor.)	Delta M (Da)
1	[126-135]	IgG 2,3,4	47	GPSVFPLAPC	1043.46	1043.51	0.05
2	[126-135]	n/a	47	GPSVFMTAPC	1065.44	1065.46	0.02
3	[126-137]	IgG1	45	GPSVFPLAPSPK	1195.75	1195.66	0.09
4	[126-137]	IgG1	40	GPSVFPLAPSSK	1185.62	1185.64	0.02
5	[138-151]	IgG1	36	STSESTAALGCLVK	1422.82	1422.70	0.12
6	[138-151]	IgG1	34	STSGGTAALGCLVK	1320.64	1320.67	0.03
7	[219-230]	Hinge IgG3	15	TPLGDTTHTCPR	1354.60	1354.63	0.03
8	[226-252]	Hinge IgG2	61	CCVECPPCPAPPVAGPSV FLFPPKPK	2907.15	2907.39	0.24
9	[253-259]	IgG1,2,3 4	19	DLTMISR	850.36	850.42	0.06
10	[260-278]	IgG1,2,3	50	TPEVTCVVVDVSHEDPEVK	2138.37	2138.02	0.35
11	[279-292]	IgG1	49	FNWYVDGVEVHNAK	1676.52	1676.79	0.27
12	[281-292]	IgG3	37	WYVDGVEVHNAK	1415.75	1415.68	0.07
13	[306-321]	IgG2	65	VVSVLTVVHQDWLNGK	1792.72	1792.98	0.26
14	[306-321]	IgG1,3,4	72	VVSVLTVLHQDWLDGK	1808.12	1807.98	0.14
15	[327-334]	IgG1	24	ALPAPIEK	837.44	837.49	0.05
16	[327-334]	IgG2	25	GLPAPIEK	823.53	823.48	0.05
17	[329-334]	n/a	25	PAPIEK	653.45	653.37	0.08
18	[345-355]	IgG1,2,3	29	EPQVYTLPPSR	1285.57	1285.66	0.09
19	[345-360]	IgG1	34	EPQVYTLPPSRDELTK	1871.76	1871.96	0.20
20	[361-370]	IgG1,2,3,4	41	NQVSLTCLVK	1160.47	1160.62	0.15
21	[361-370]	n/a	41	NQVTLTCLVK	1174.56	1174.64	0.08
22	[361-370]	n/a	41	NQVSLTCLLK	1174.74	1174.64	0.10
23	[393-409]	IgG2,3	55	TTPPMLDSDGSFFLYSK	1904.73	1904.88	0.15
24	[393-409]	IgG1	60	TTPPVLDSDGSFFLYSK	1873.05	1872.91	0.14
25	[393-409]	n/a	52	TTPPVLDSVGSFFLYSK	1856.82	1856.96	0.14
26	[410-414]	IgG1,2,3,4	8	LTVDK	574.25	574.33	0.08
27	[410-416]	IgG1,2,3,4	9	LTVDKSR	817.30	817.47	0.17
28	[417-435]	IgG1,2	32	WQQGNVFSC	1124.63	1124.47	0.16
29	[417-439]	IgG4	54	WQEGNVFSCSVMHEALHN HYTQK	2801.45	2801.24	0.21

[a] Residue numbers of the constant regions were determined from reference X

[b] Cysteine residues were observed in alkylated form

From these analyses, a total of 36 different light chain peptides were determined, 18 from the variable region (peptides 1 – 18 in Table 6) and 18 from the constant region (peptides 19 – 36 in Table 6). Within the constant region peptides both κ and λ chains were identified. Based on homology, variable region peptides were fit against the light chain CAE54361 from the NCBInr database. The regions V_L(19-24), V_L(55-61) and V_L(98-104)

were identified exclusively by pre-fractionation of the light chain tryptic digest and off-line MS.

Table 6: Aß-autoantibody light chain variable and constant region tryptic peptides identified by RP-HPLC separation and individual analysis of each fraction by off-line ESI-MS/MS

Peptide Nº	Residue number[a,b]	Fraction number	Amino acid sequence [d]	MW (exp.)	MW (theor.)	Delta M (Da)
1	[19-24][c]	8	VTITCR	748.45	748.36	0.09
2	[46-50][c]	21	LLIYK	648.43	648.41	0.07
3	[46-54][c]	34	LLIYWASTR	1121.78	1121.61	0.17
4	[46-54][c]	23	LLIYAGSTR	992.61	992.55	0.06
5	[46-54][c]	23	LLIYGASAR	962.64	962.54	0.10
6	[46-54][c]	16	LLISDASNR	987.48	987.52	0.04
7	[54-61][c]	7	LQSGVPSR	842.49	842.45	0.04
8	[55-61][c]	6	ATGIPAR	684.43	684.38	0.05
9	[55-61][c]	6	ATDIPAR	742.45	742.38	0.07
10	[55-61][c]	7	ATGIPDR	728.47	728.37	0.10
11	[55-61][c]	10	APGIPDR	724.48	724.37	0.11
12	[62-74][c]	22	FSGSGSGTDFTLK	1302.63	1302.59	0.04
13	[62-70][c]	21	FSGSLLGGK	864.44	864.46	0.02
14	[62-77][c]	27	FSGSGSGTDFTLTISR	1631.71	1631.76	0.05
15	[62-70][c]	19	FSGSILGNK	921.63	921.48	0.23
16	[98-104][c]	5	TFGQGTR	765.47	765.36	0.11
17	[98-104][c]	3	TFGQGTK	737.29	737.36	0.07
18	[98-104][c]	5	VFGGGTK	664.25	664.34	0.07
19	[108-115][a]	14	VTVLGQPK	840.46	840.49	0.03
20	[108-115][a]	16	LTVLGQPK	854.45	854.51	0.06
21	[108-115][a]	17	LTVNGQPK	855.63	855.47	0.16
22	[108-126][b]	33	RTVAAPSVFIFPPSDEQLK	2101.72	2101.11	0.61
23	[109-126][b]	36	TVAAPSVFIFPPSDEQLK	1945.28	1945.01	0.17
24	[111-129][b]	29	AAPSVTLFPPSSEELQANK	1985.06	1985.00	0.06
25	[127-142][b]	33	SGTASVVCLLDNFYPR	1797.95	1797.84	0.11
26	[154-160][a]	2	ADGSPVK	672.25	672.33	0.07
27	[154-160][a]	2	ADSSPVK	702.23	702.34	0.11
28	[150-170][a]	12	VDNALQSGNSQESVTEQDSK	2135.12	2134.95	0.17
29	[170-183][b]	26	DSTYSLSSTLTLSK	1501.88	1501.74	0.14
30	[184-188][b]	1	ADYEK	624.28	624.26	0.02
31	[191-207][b]	18	VYACEVTHQGLSSPVTK	1875.05	1874.89	0.16
32	[191-207][b]	22	LYACEVTHQGLSSPVTK	1888.83	1888.90	0.07
33	[176-190][a]	33	YAASSYLSLTPEQWK	1742.92	1742.84	0.08
34	[194-208][a]	11	SYSCQVTHEGSTVEK	1710.97	1710.72	0.25
35	[209-216][a]	4	TVAPTECS	863.25	863.33	0.08
36	[208-211][b]	2	SFNR	522.24	522.24	0

[a] Peptides derived from lambda light chain constant region, fit against the sequence ABU90549 from NCBInr database;
[b] Peptides derived from kappa light chain constant region fit against the sequence ACF34453 from NCBInr data base;
[c] Variable region peptides were fir against the light chain sequence CAE54361 from NCBInr database;
[d] Cysteine residues were observed in alkylated form.

The primary structures of the Aß-autoantibody heavy and light chains determined with the peptides in Table 2 – Table 6 is summarized in . Collectively, the results obtained for the Aß-autoantibody using *(i)* the data dependent analysis (*Strategy I*), and *(ii)* pre-fractionation followed by off-line LC-ESI-MS/MS (*Strategy II*), demonstrate the applicability

of the second approach to complex mixtures such as the polyclonal Aß-antibody, where duty cycle and dynamic range represent crucial parameters that affect sensitivity. A major disadvantage of *Strategy II* is the time ineffectiveness, whereas data collection using *Strategy I* can be accomplished in several hours.

A.

B.

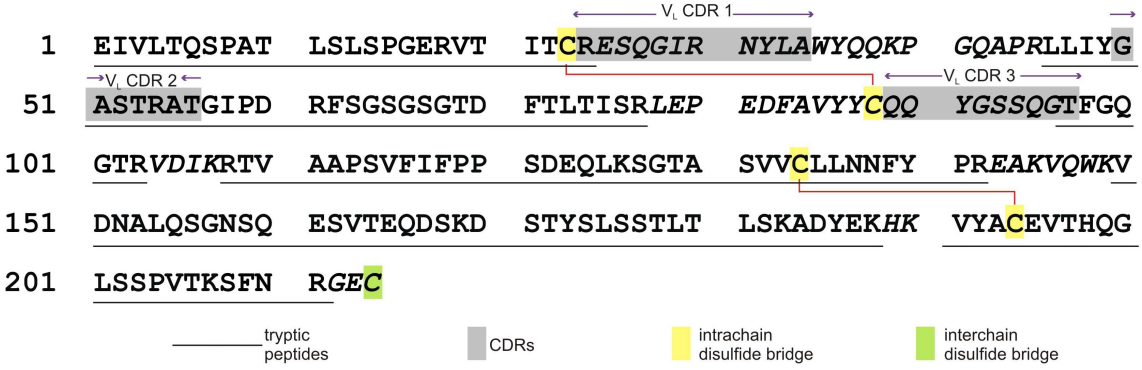


Figure 2.28: (A) Amino acid sequence of the Aß-autoantibody IgG$_1$ heavy chain assembled with the tryptic peptides determined by LC-MS/MS, summarized in Table 2, Table 4 and Table 5, respectively. Variable and constant region peptides were fit against the heavy chain antibody sequence with the NCBI accession number AAV67804 based on homology; (B) Amino acid sequence of the Aß-autoantibody kappa light chain assembled with the tryptic peptides determined by LC-MS/MS summarized in Table 3 and Table 6. Variable and constant region peptides were fit against the light chain antibody sequence with the NCBI accession number AAH16380, based on homology. Peptides resembling the amino acids indicated in italics were not observed. Colour code: grey boxes – complementary determining regions, determined according to Kabat rules [204]; yellow – intrachain disulfide bridges; green – interchain disulfide bridges; purple – consensus of amino acids for N-linked glycosylation.

2.2 Mass spectrometric identification of glycosylated structures in Aß-specific antibodies and Aß-proteins

2.2.1 Post-translational modification by glycosylation

Post-translational modifications (PTMs) represent one of the major sources of protein structure diversity affecting protein conformation, charge and hydrophobicity, either by covalent addition of groups to amino acid side chains or by proteolytic cleavage of one or more peptide bonds in a protein by proteases. There are more than 200 kinds of covalent modifications, among which phosphorylation, glycosylation, ubiquitination, acetylation, oxidation represent few of the most analyzed PTMs, resulting in a large number of publication entries in PubMed every year. An overview of several most common PTMs in given in Appendix 2. Covalent modifications have crucial impact on a protein's three dimensional structure, folding, activity, cellular localization, and are involved in protein-protein interactions and signal transduction pathways. Knowledge of the attachment site of PTMs is of major importance for the understanding of protein function and its regulation in biological pathways. Characterization of modifications represents a formidable analytical challenge, primarily derived from the fact that some PTMs are present at substoechiometric level, while others are highly heterogeneous. In the past years, mass spectrometry proved to be the method of choice for microcharacterization of PTMs, due to the high sensitivity, low sample consumption and its ability to resolve complex mixtures.

Several decades ago, following the discovery of the genetic code, it became clear that carbohydrates represent integral components of a broad range of molecules, collectively termed glycoconjugates. These comprise compounds in which sugar chains (glycans) are covalently attached to either polypeptide or lipid chains to form glycoproteins, proteoglycans and glycolipids. Most cell-surface and secreted proteins are glycosylated – a fact that impacts on efforts to understand the biological relevance of specific protein expression and modification patterns. Several significant biological roles of carbohydrates include location of a protein within the cell, protection of a protein against proteolytic attack, control of the lifetime of circulating cells and glycoproteins, induction and maintenance of the spatial conformation in a biologically active form, facilitation of the extracellular secretion, as well as modulation of the immune response. Surface

carbohydrates serve as the interface between the cell and its environment, and define self versus non-self. Depending on the type of the linkage atom to which carbohydrates are covalently attached, glycans are categorized as *N-, O-, C-* and *S*-glycans. Unlike the core proteins, glycans are expressed as a set of variations on a core structure and are polydisperse in nature. Figure 2.29 shows the most common monosaccharide building blocks found in *N*- and *O*-linked glycans, with the abbreviations and symbols used throughout the following chapters.

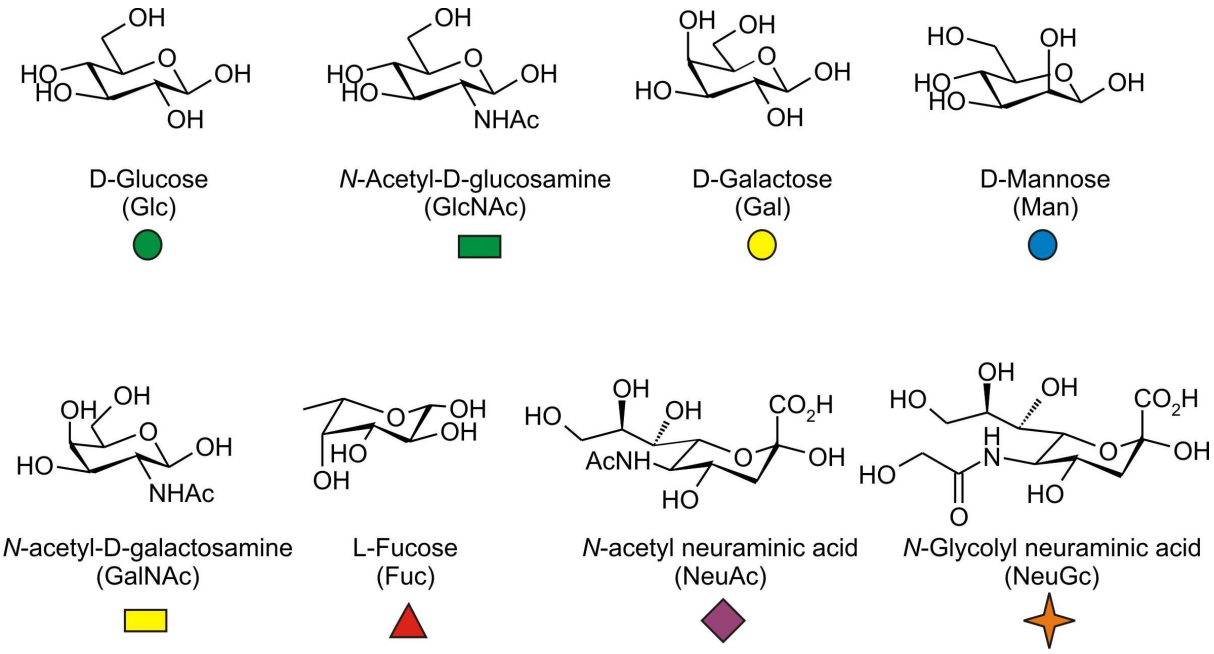


Figure 2.29: Chair conformation of the most common monosaccharide building blocks contained in N- and O-linked glycans.

Prediction of glycoproteins is difficult and challenging, as no consensus sequence has been defined thus far, e.g. for *O*-glycosylation. It has been found that *N*-glycosylation requires the consensus sequence Asn-Xxx-Ser/Thr/Cys, where Xxx can be any amino acid, except for Pro. Protein modification by *N*-glycosylation is initiated in the endoplasmic reticulum (ER) by the attachment of a conserved oligosaccharide precursor containing two *N*-acetyl glucosamine, nine mannose and three glucose residues ($Glc_3Man_9GlcNAc_2$) to the Asn side chain of the nascent polypeptide. A large orchestra of enzymes including glycosidases and glycosyl transferases trim down and elongate the precursor glycan, already in ER and with potential continuation in the Golgi apparatus, such that three major types of mature *N*-glycans may be synthesized, as follows: *(i)* high mannose, consisting primarily of Man with a maximum number of nine mannoses possible unless not fully processed, *(ii)* complex type glycans, mainly composed of GlcNAc and Gal with or without

sialic acid, where a fucose may be added to the first GlcNAc in the core; and *(iii)* hybrid type glycans which are composed of mannose, GlcNAc-Gal and with or without sialic acids (Figure 2.30). High mannose glycans are commonly found in invertebrates, e.g. in glycoproteins forming viral envelopes, whereas only vertebrates appear to be capable of synthesizing hybrid and complex *N*-glycans, in addition to the high mannose.

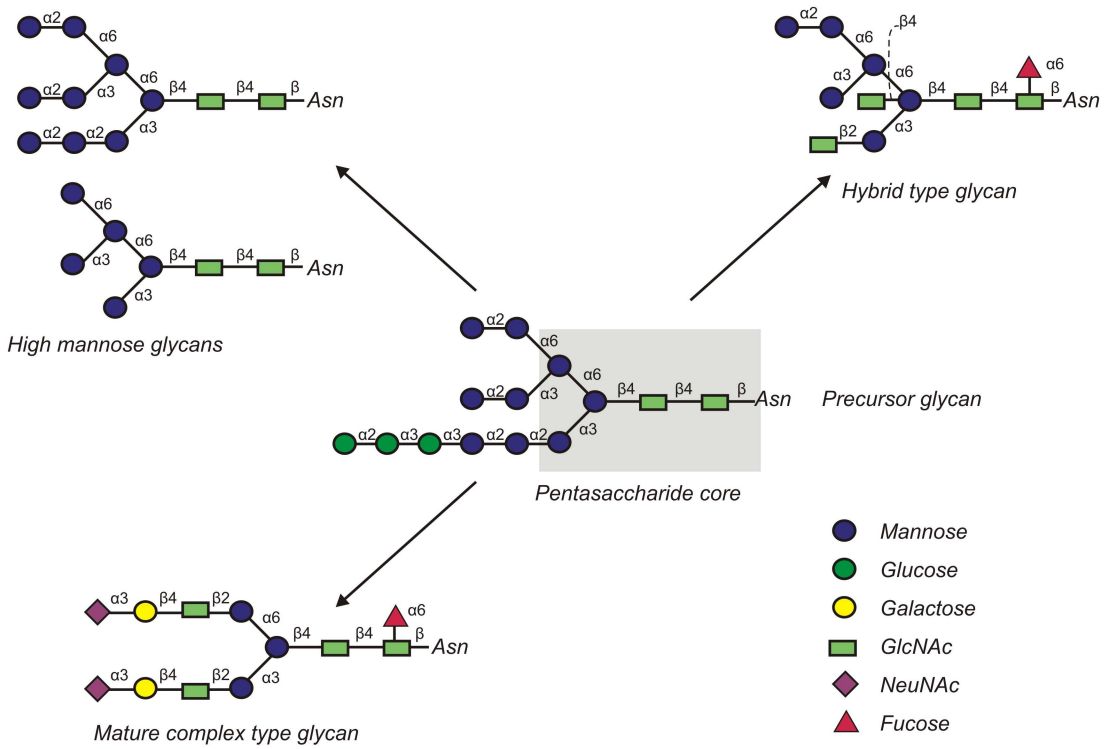


Figure 2.30: Pathways for biosynthesis of N-linked glycans: a precursor glycan with the composition $Glc_3Man_9GlcNAc_2$ (middle) is attached to the asparagine side chain at the consensus sequence (Asn-Xxx-Ser/Thr) during ribosomal protein biosynthesis. Enzymatic processing of the precursor glycan leads to diversification of the N-linked species into high mannose (top left), hybrid (top right) and complex type glycans (bottom left), while a minimal pentasaccharide core of composition $Man_3GlcNAc_2$, highlighted with grey, remains common to all structures. Anomeric forms and linkages between monosaccharides are indicated in black Greek letters and numbers, respectively. Colour code: green rectangle – N-acetyl glucosamine, yellow circle – galactose, blue circle – mannose, green circle – glucose, red triangle – fucose, and purple rhombus – N-acetyl neuraminic acid.

Biosynthesis of *O*-glycans is more complex than that of *N*-glycans and does not require a precursor glycan or a consensus of amino acids. Several hundreds of different *O*-glycan structures have been described, yet little is known about their specific functions. The tremendous structural variation is due to the fact that at least eight different core structures of *O*-glycans exist in mammalian mucins. Sugars are transferred to the Ser and/or Thr side chain and subsequently elongated from specific nucleotide sugar donors by the action of specific membrane-bound glycosyl transferases in the Golgi apparatus. In cancer cells, many of the enzymes involved in *O*-glycan biosynthesis are up- or down-regulated. The

initial step in *O*-glycan biosynthesis is the attachment of *N*-acetyl galactosamine (GalNAc), representing the only core structure common to all *O*-linked oligosaccharides (Figure 2.31).

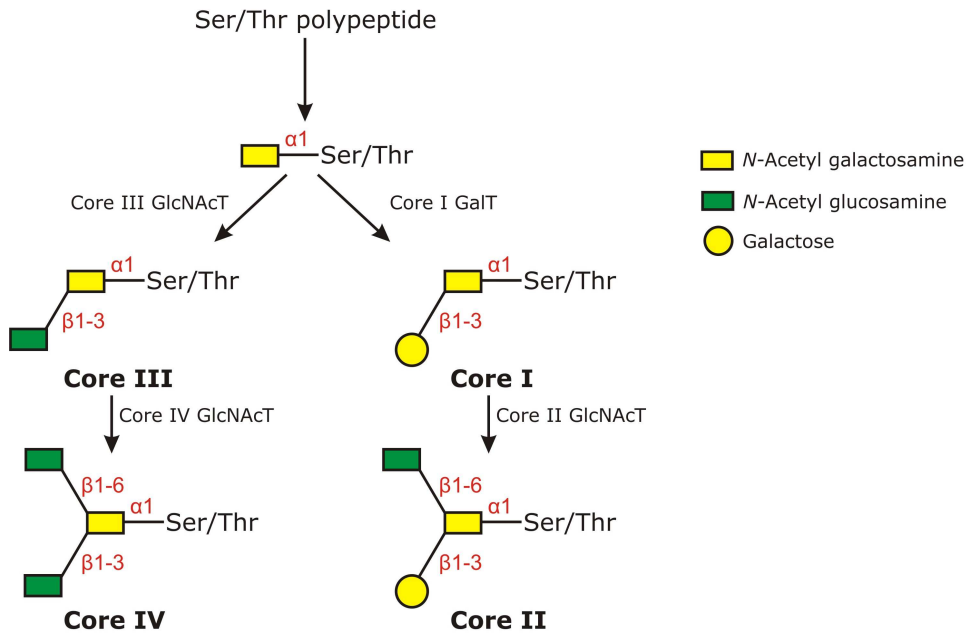


Figure 2.31: Principle pathways of biosynthesis of O-glycans, initiated by the attachment of N-acetyl galactosamine (yellow rectangle) to the side chain of Ser or Thr. Elongation and branching of this substrate by the action of various sugar transferases results in formation of Core I, Core II, Core III and Core IV glycans, which can be further elongated with (GlcNAc-Gal) disaccharide repeat units and terminated by N-acetyl neuraminic acid. Anomeric forms and linkages between monosaccharides are indicated in red Greek letters and numbers, respectively. Colour code: yellow rectangle – N-acetyl galactosamine, green rectangle – N-acetyl glucosamine, and yellow circle – galactose.

Consequently, one glycosylation site may have multiple glycan structures (microheterogeneity), while a protein may contain multiple glycosylation sites (macroheterogeneity). Glycosylation increases the complexity of protein molecules and causes them to migrate as diffuse bands or spots on SDS-PAGE gels to complicate efforts to identify protein expression patterns that correlate with disease state. Whereas unmodified proteins can often be studied by X-ray crystallography or NMR spectroscopy, these methods fail to provide satisfactory results with glycoproteins, as a result of the increased heterogeneity derived from glycosylation.

2.2.2 Glycosylation analysis by mass spectrometry

Glycosylation analysis is recognized as one of the current challenges in proteomics [214] and liquid chromatography – mass spectrometry has become an invaluable technology for the analysis of protein glycosylation. Two distinct analytical strategies were developed for structural characterization of glycoproteins in conjunction with mass spectrometry, *(i)* analysis of carbohydrates cleaved by enzymatic or chemical procedures from a protein [215-218] (*Strategy I* in Figure 2.32), and *(ii)* analysis of glycopeptides released by proteolysis [219] (*Strategy II* in Figure 2.32). Common enzymes used to cleave the *N*-glycan pool from a glycoprotein regardless of their position on the protein backbone are peptide *N*-glycosidases F and A (PNGase F and PNGase A, endo-glycosidases) which hydrolyze the β-aspartylglycosylamine bond between asparagine and the innermost GlcNAc moiety. *O*-glycans are typically removed by alkaline β-elimination [218]. Because oligosaccharides have poor UV signal and ionization efficiency, their derivatization with hydrophobic groups is often necessary to improve their detection [220]. Although LC-MS analysis of released glycans may provide a detailed picture of the structure of the glycans derived from a protein, information of the attachment sites of the glycans is lost.

Identification of glycosylation sites via glycopeptide analysis (*Strategy II*) requires a lengthy process often involving trypsin digestion to produce glycopeptides, separation of the glycopeptides by HPLC, and characterization of each fraction by MS with or without enzymatic release to determine the oligosaccharide and peptide structures of each fraction. Glycopeptides frequently constitute only a minor fraction of a complex peptide mixture, e.g. a tryptic digest of a glycoprotein preparation, so that differentiation between glycosylated and non-glycosylated peptides prior or during LC/MS analysis is essential. This can be achieved either by using lectin affinity enrichment prior to LC separation [221, 222] or by scanning characteristic sugar oxonium ions derived from glycosydic bond cleavage from the non-reducing end (e.g. *m/z* 204.1, protonated *N*-acetylhexosamine, *m/z* 366.1, protonated HexHexNAc or *m/z* 292.1, protonated *N*-acetyl neuraminic acid) which arise during an MS/MS experiment or from in-source decay in the MS mode, after elution from the column [223-226]. Because full scan mass spectra do not always yield unequivocal structural information, tandem mass spectrometry is often necessary to identify unique fragments that correspond to the oligosaccharide moieties [35].

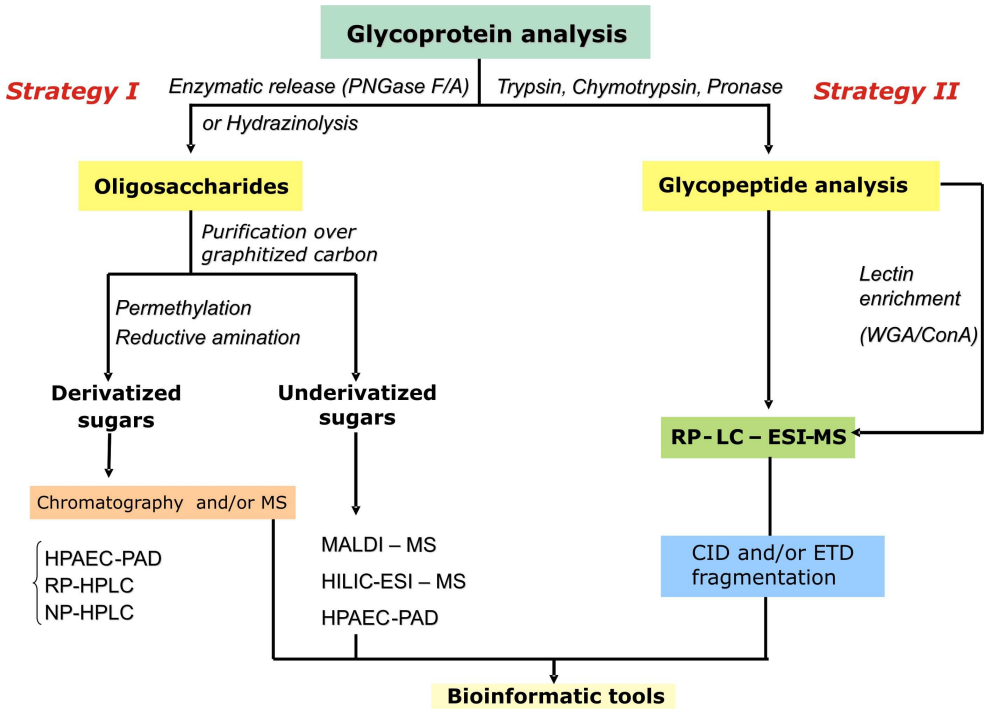


Figure 2.32: Overview of methodologies for glycoprotein analysis by mass spectrometry: (i) Strategy I involves chemical or enzymatic release of sugars from a glycoprotein. The carbohydrate pool can be analyzed in derivatized form (e.g. permethylation) or underivatized. Separation techniques such as normal phase (NP), reversed phase (RP) or anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) may be employed prior to MS; (ii) Strategy II uses enzymatic degradation to produce glycopeptides which can be enriched by lectin affinity, or analyzed directly by RP-LC-ESI-MS. Fragmentation by CID or ETD is used for their structural characterization.

In the present work glycosylation analysis of ß-amyloid antibodies and ß-amyloid precursor protein was performed following *Strategy II*, and employed reduction and alkylation of each protein, followed by SDS-PAGE separation, in-gel digestion with trypsin, LC-MS/MS using ETD and CID fragmentation, and manual interpretation of the data.

Fragmentation of glycopeptide ions under CID and ETD conditions presents several characteristic features that enable their specific identification and characterization. Carbohydrates are more labile than peptides under MS conditions, such that they dissociate by a lower energy pathway than that involved in cleavage of the amide linkage. CID spectra of protonated glycopeptides are dominated by successive loss of sugar moieties from the non-reducing end, with little or no fragmentation of the peptide backbone [227] (see Figure 2.33). Characteristic singly charged sugar oxonium ions (e.g. *m/z* 204.1, protonated *N*-acetylhexosamine, *m/z* 366.1, protonated HexHexNAc or *m/z* 292.1, protonated *N*-acetyl neuraminic acid) populate the low *m/z* range of the spectrum. The retention time of glycopeptides can be determined by extracting the trace of these reporter

ions from the total ion current in the full scan MS or in MS/MS [162, 228-230]. The structure of the glycan is directly attainable from the observation of successive oxonium ions, or from the observation of successive glycopeptide fragment ions

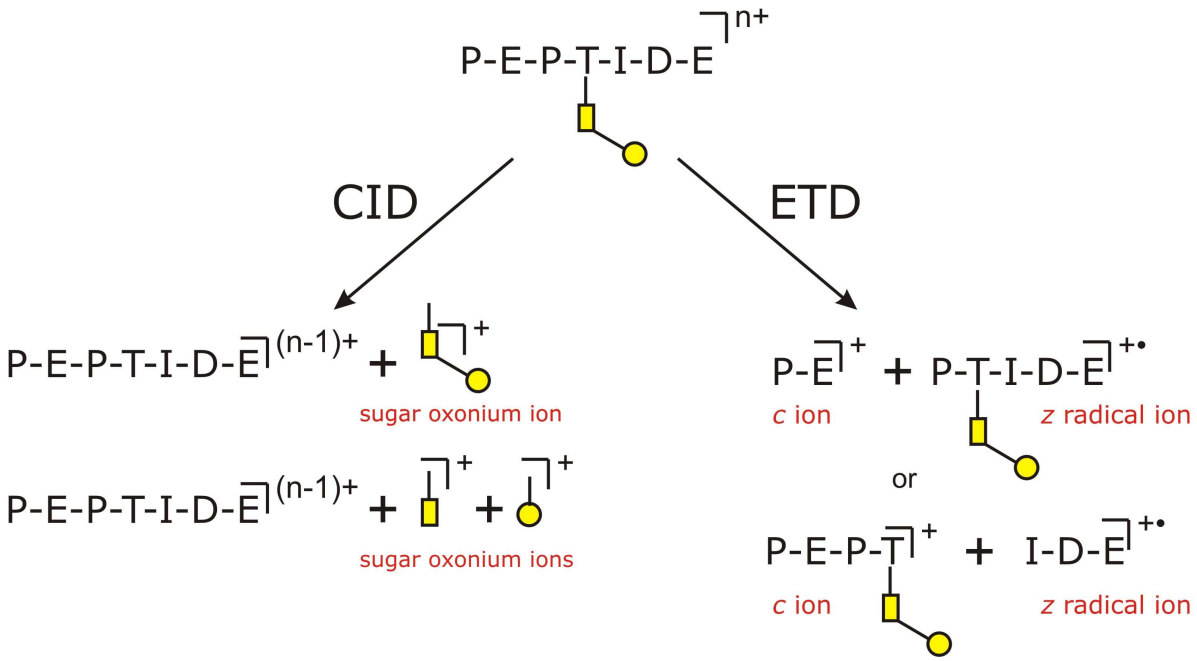


Figure 2.33: Comparison of the CID (left) and ECD/ETD (right) fragmentation of an O-glycopeptide. In CID sugar oxonium ions arise from glycosydic bond cleavages, while the peptide remains unaffected. In ETD, fragmentation occurs predominantly at the peptide backbone, with retention of the sugar modification at the glycosylation site.

In the ETD process, an electron from the fluoranthene radical anion, produced by negative ion chemical ionization is transferred to a multiply charged peptide/glycopeptide ion. The resulting charge reduced odd electron species undergo, in a manner similar to ECD, predominant cleavage of the N – C_α bonds generating c' and z• fragment ions in a sequence independent manner (nomenclature of Zubarev and co-workers [231, 232]). Unlike CID, ETD is independent of amide bond protonation and occurs on a time scale that is short (few milliseconds) compared to the internal energy distribution. Derived from this, the most important feature of ETD is that intact oligosaccharide moieties are retained in the fragment ions containing the site of glycosylation (see Figure 2.33). For this reason, ECD and ETD represent excellent tools for characterization of sites of modification in glycopeptides and other post-translationally modified proteins [166, 229, 233-236].

2.2.3 Glycosylation structure of a ß-amyloid plaque specific, anti-Aß(1-17) monoclonal antibody

One of the characteristics of immunoglobulins is glycosylation of a conserved Asn residue in the C_H2 domain of the heavy chain constant region - one of the sources of molecular heterogeneity in antibodies. Each heavy chain generally contains one glycan moiety. The carbohydrate attached at this conserved Asn residue from the F_c region is an essential component required for high affinity receptor binding, representing one of the pathways developed during the immune response [9, 10]. In addition, glycans help stabilize the immunoglobulin fold by making contact with residues on the protein backbone and with each other within the same molecule [6, 9, 11, 12]. Complete deglycosylation of IgG results in the loss of binding to specific cell surface receptors [237-239] and, consequently, to a failure in the initiation of the corresponding effector functions [240]. Aberrant glycosylation of an antibody heavy chain may be related to diseases, such as myeloma [241, 242] or rheumatoid arthritis (RA), which is thought to be caused by elevated levels of agalactosyl glycoforms. The aberrant glycoforms in RA may become antigenic and lead to formation of antibody – antibody immune complexes [243].

The *N*-linked glycans at the conserved constant region Asn residue on immunoglobulin heavy chains have been previously characterized and found to be of biantennary, complex type attached to the trimannosyl chitobiose core [13]. It has been shown that human IgGs contain high amounts of α(1,6) core fucose (F), with minimal amount of *N*-acetyl neuraminic acid (SA) and bisecting GlcNAc (B) [13], while mouse IgGs contain *N*-glycolyl neuramic acid (NeuGc) instead of *N*-acetyl neuraminic acid. Typical immunoglobulin *N*-glycans from the constant region are shown in Figure 2.34. The physiological IgG glycans have zero to one core fucose units and zero to two galactose residues, G_n (n= 0, 1 or 2). For example, the notation G_1FBSA refers to a complex, core fucosylated biantennary glycan containing bisecting GlcNAc, a single ß(1-4) galactose and a single sialic acid unit, while the notation G_2 indicates a complex type glycan bearing two ß-linked galactose residues (one on each antennae of the trimannosyl chitobiose core) without core fucosylation.

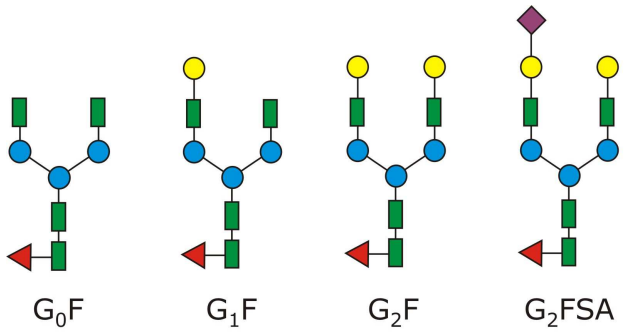


Figure 2.34: Structures of biantennary complex type N-glycans from the constant region of IgG molecules; the notations under each structure refer to agalactosyl, mono- and digalactosyl glycoforms (G_0F, G_1F, and G_2F, respectively), whereas the notation G_1FSA refers to the monogalactosyl glycoform terminated with one N-acetyl neuraminic acid.

The heavy chain of the anti-Aß(1-17) antibody (clone 6E10) was digested with trypsin and analyzed by LC-MS/MS. This antibody belongs to the IgG_1 subclass, such that the corresponding tryptic peptide containing the consensus sequence for *N*-glycosylation is EEQF^{297}NSTFR. The ESI-MS/MS spectrum of the doubly charged precursor ion of *m/z* 1301.53 is shown in Figure 2.35 A, confirming the G_0F glycan structure assigned for this mass and the identity of the tryptic peptide containing the glycan.

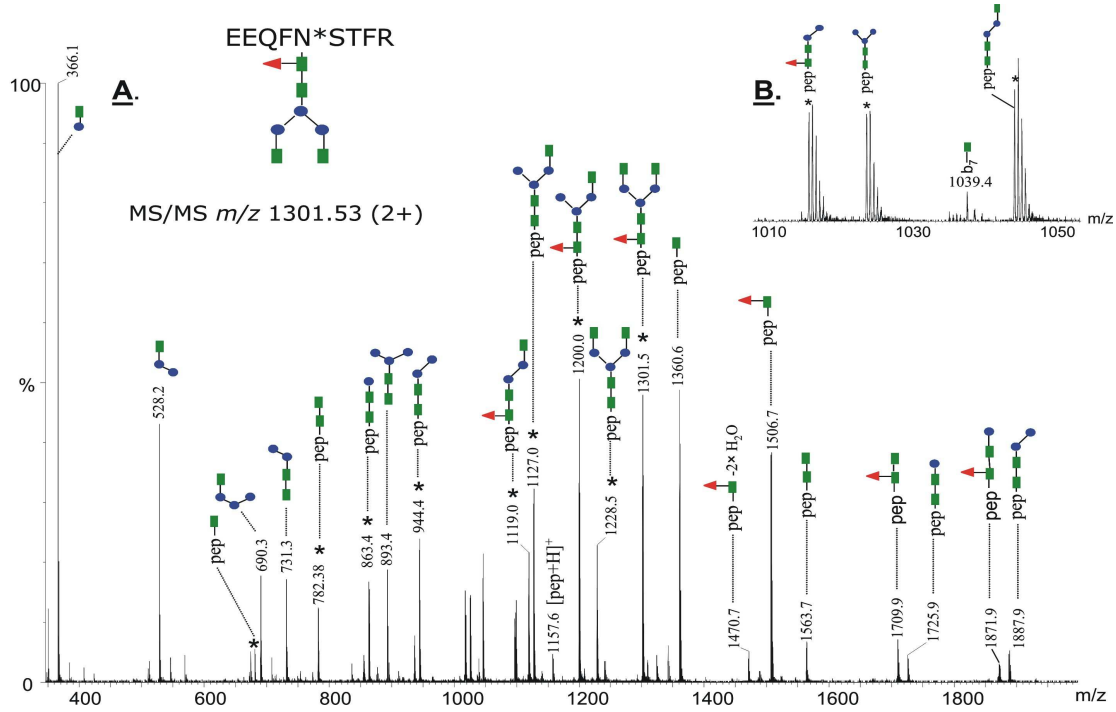


*Figure 2.35: MS/MS of the precursor ion of m/z 1301.53 (2+), corresponding to the glycopeptide EEQFN*STFR containing the glycan indicated at the top (left). Doubly charged ions are indicated with an asterisk (*). The remaining ions are singly charged. (A) MS/MS spectrum obtained in the data dependent mode using a collision energy gradient from 30 to 40 V, showing all the observed fragment ions; (B) The insert in the mass range m/z 1010-1050 indicates the singly charged ion at m/z 1039.4 which arises from backbone cleavage and carries an N-acetyl glucosamine residue. Colour code: green square – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose.*

The MS/MS experiments revealed two characteristic fragmentation pathways: *(i)* neutral loss of the sugar moieties from the non-reducing end of the glycan, which generated doubly protonated fragments (marked with an asterisk in the spectrum) with an intact peptide backbone, and *(ii)*, charge reduction of the precursor, which produced protonated sugar oxonium ions with a single positive charge and singly charged glycopeptide ions containing the remaining sugar residues. The MS/MS spectra of glycopeptides are characterized by abundant fragment ions derived from one of the two fragmentation pathways described above and by low abundance or no backbone fragments. However, a low abundant b_7 fragment formed by peptide backbone cleavage which still has the first GlcNAc unit attached at the Asn residue was observed at *m/z* 1039.42 (see Figure 2.35 B) consistent with the amino acid sequence of the peptide. In addition, complete processing of the glycan from the non-reducing end resulted in a peptide fragment of *m/z* 1157.57, which was assigned as the singly protonated peptide.

The mass spectrum averaged over the chromatographic retention time in which the glycopeptides eluted is presented in Figure 2.36 A. The most abundant species, detected in the positive ion mode as doubly protonated fragments of *m/z* 1301.53, 1382.55 and 1463.60, were assigned to the glycoforms G_0F, G_1F and G_2F respectively. The structural assignment of the *N*-glycans was deduced from the difference between the experimental glycopeptide mass observed and the calculated mass of the tryptic peptide without the sugar moiety (EEQFNSTFR, M_r 1156.6), consistent with the structural information provided by MS/MS (Figure 2.35). The low abundance non-fucosylated structures G_0, G_1 and G_2 were detected as doubly charged fragments of *m/z* 1228.52, 1309.53 and 1390.55. This pattern is consistent with the structures reported for recombinant monoclonal antibodies [224].

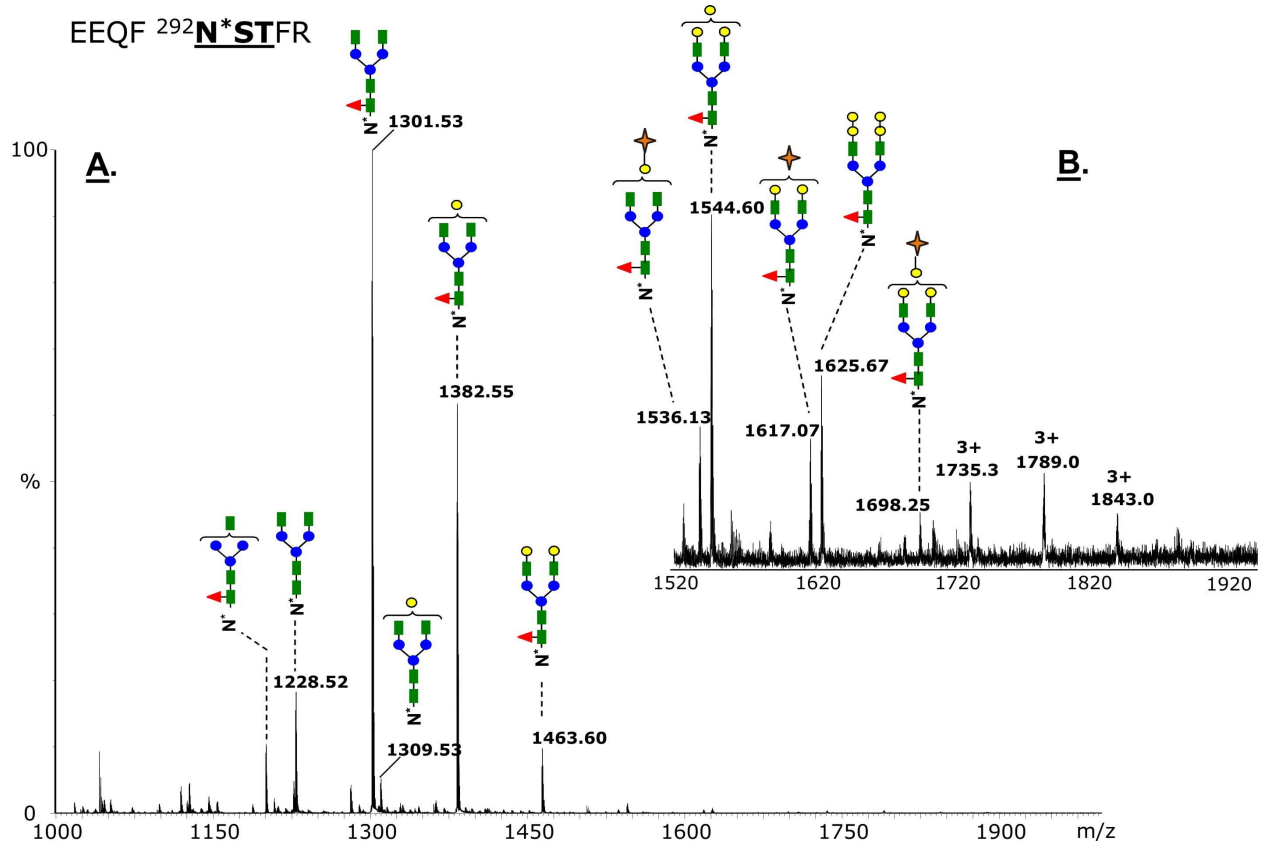


*Figure 2.36: Positive ion nano-HPLC-ESI/MS of the F_c glycopeptides EEQFN*STFR from the mouse monoclonal antibody 6E10; (A) MS spectrum averaged over the chromatographic window where glycopeptides eluted (30.6 minutes, average of 15 full MS scans). The glycan structures of the most abundant glycoforms are indicated above each ion. All ions are doubly charged; (B) Insert in the mass range m/z 1520-1920, showing low abundance di-, tri- and tetragalactosylated glycoforms. The triply charged ions of m/z 1735.0, 1789.0 and 1843.0 show unusual oligosaccharide composition (see discussion in the text). Colour code: green rectangle – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose, yellow circle – galactose, and orange diamond – N-glycolyl neuraminic acid.*

The expanded mass range *m/z* 1500 – 1920 presented in Figure 2.36 B shows two low abundance glycoforms detected as doubly charged fragments of *m/z* 1544.60 and 1625.67, which were assigned to core fucosylated, biantennary, complex type glycans incorporating three and four galactose residues respectively, whereas the third and the fourth galactose units are probably α(1,3) linked to the Gal-ß(1,4)-GlcNAc. Hypergalactosylation of recombinant immunoglobulins was reported previously for antibodies expressed in NS0 cell lines [244]. This feature represents a potential problem if such a monoclonal antibody should be used as a therapeutic agent due to the possible immunogenicity. It has been reported that up to 1 % of the circulating IgG may be specific for binding the α-Gal epitope [245] and that antibodies containing this motif might be highly immunogenic, which may lead to increased proteolytic degradation [246]. Low amounts of *N*-glycolyl neuraminic acid (NeuGc) terminated species were observed at *m/z* 1536.13 (2+), 1617.07 (2+) and 1698.25 (2+), which were assigned to the structures G_1FNeuGc,

G_2FNeuGc and G_3FNeuGc. In addition, small amounts of hybrid glycans were detected. Using ß-galactosidase digestion of the heavy chain tryptic mixture the extent of mannose and galactose in each hybrid structure was determined and their overall structural composition was assigned. However, the exact sugar linkages could not be determined from these data.

Glycopeptides incorporating high molecular weight glycans on the same peptide backbone (EEQFN*STFR) were observed as triply charged species of *m/z* 1735.04, 1789.04 and 1843.03 (Figure 2.36 B); this mass interval corresponds to a hexose unit. The MS/MS of the parent ion of *m/z* 1789.04 is shown in Figure 2.37.

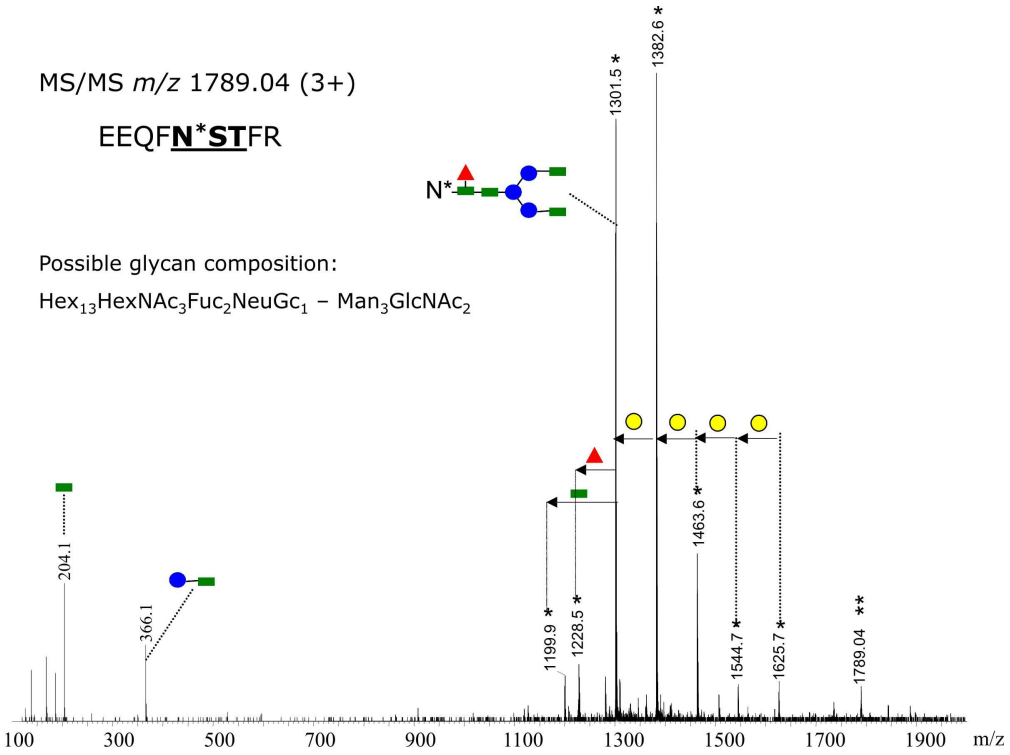


Figure 2.37: CID spectrum of the triply charged precursor ion of m/z 1789.04 obtained with a collision energy ramp of 30 V to 40 V. The most abundant product ions are doubly charged (indicated with a single asterisk (*)) and result from the fragmentation of the glycan moiety. The product ion of m/z 1301.5 (2+) (which provided the fragmentation pattern shown in Figure 2.35) indicates the same amino acid sequence of the peptide backbone. One of the possible glycan compositions is indicated at the left. The successive loss of hexose (either mannose or galactose) is indicated through yellow circles. Colour code: green rectangle – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose, yellow circle galactose.

The fragments detected in the mass range 1100 – 1700 are doubly charged and they have *m/z* values identical intact glycopeptides ions observed in the full scan mass spectrum (Figure 2.36 A and B), confirming the identity of the peptide backbone. As described above, the complete carbohydrate structures could not be derived because various

monosaccharide compositions are possible for this observed mass. These putative structures were obtained using the GlycoMod software, designed to determine possible glycan compositions from experimentally determined glycan/glycopeptide masses (www.expasy.org/tools/glycomod). Based on the fragmentation pattern of the triply charged precursor a hybrid, biantennary structure (or of a higher degree of branching) is suggested from the observed MS/MS and is consistent with a sugar composition $Hex_{13}HexNAc_3Fuc_2NeuGc_1$ – $Man_3GlcNAc_2$. This structure might result from abnormal processing of the precursor glycan of composition $Glc_3Man_9GlcNAc_2$, which is attached to the Asn residue during protein biosynthesis. One antenna may be of the high mannose type and the second one may be elongated by successive addition of GlcNAc and galactose.

2.2.4 Glycosylation structure of the plaque-protective, anti-Aß(17-24) monoclonal antibody

The glycan structures decorating the heavy chain constant region at residue Asn 307 (Figure 2.22, numbering as in full length 4G8 heavy chain) were determined from the LC-MS/MS analysis of glycopeptides derived from in-gel digestion with trypsin of the SDS-PAGE band 1 (Figure 2.18). The *N*-linked glycosylation site was observed in the tryptic peptide ^{304}EDYN*STIR^{311}, characteristic for the constant region of mouse IgG_{2b} isotype. As for the 6E10 antibody, identification of the glycopeptides in the digestion mixture was performed by monitoring the formation of sugar oxonium ions *m/z* 204.1 (protonated GlcNAc) and *m/z* 366.1 (protonated HexGlcNAc). The identity of the peptide backbone and the compositions of the *N*-glycans were determined from the MS/MS analyses of the doubly charged glycopeptide ions using data dependent acquisition, and from the mass differences between the observed glycosylated and non-glycosylated peptides. In order to determine the complete constant region glycosylation, the MS scans over the chromatographic retention time in which the glycopeptides eluted were averaged and the spectrum was deconvoluted over the mass range 2000-3600 (see Figure 2.38). In combination with CID MS/MS analyses, the *N*-linked glycans attached at Asn307 were found to be of complex, biantennary type, containing zero to four galactose residues with completely core-fucosylated structures. The most abundant glycoforms are G_0F, G_1F and G_2F, while those incorporating three and four galactose residues, G_3F and G_4F, respectively, were of lower abundance. From the analysis of constant region glycosylation

of the antibodies 6E10 (IgG_1) and 4G8 (IgG_{2b}) it appears that the presence of tri- and tetragalactosyl species is a common feature in mouse immunoglobulins, irrespective of the heavy chain isotype. The *N*-glycosylation pattern of the 4G8 heavy chain is similar to that determined for the 6E10 antibody; however, the distribution of the glycoforms G_0F, G_1F and G_2F is different, the most abundant being G_1F, followed by G_2F and G_0F, which corresponds to the physiologic glycosylation of human IgGs.

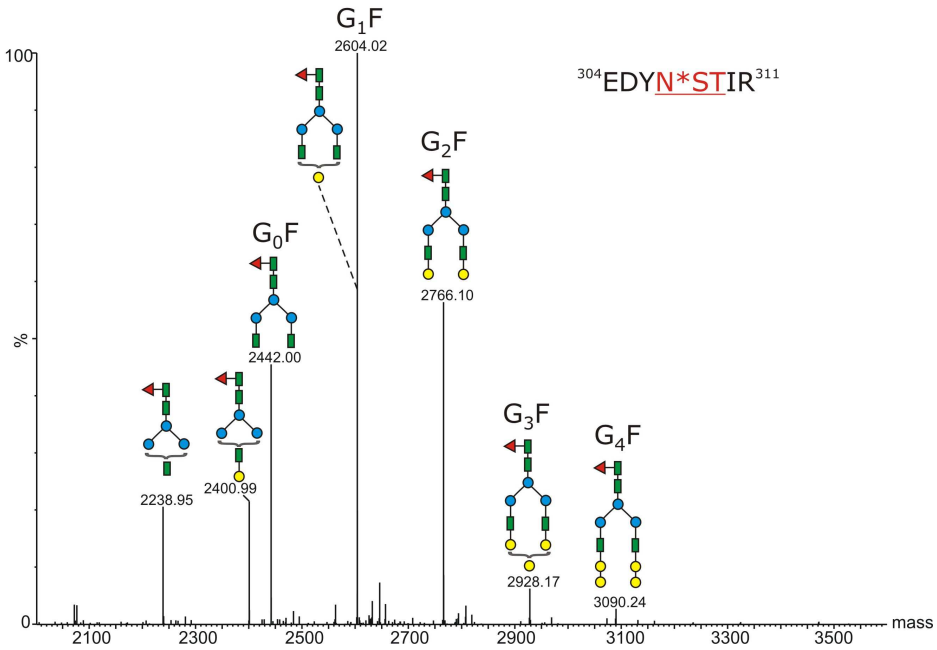


*Figure 2.38: Positive ion nano-HPLC-ESI/MS of the F_c glycopeptides EDYN*STIR from the mouse monoclonal antibody 4G8; The deconvoluted MS spectrum was obtained by averaging the MS scans over the chromatographic window where glycopeptides eluted (32.7 minutes, average of 10 full MS scans). The glycan structures of the most abundant glycoforms are indicated above each ion. Colour code: green rectangle – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose, and yellow circle – galactose.*

In addition to the conserved *N*-linked site from the heavy chain constant region, LC-MS/MS analysis of the mouse monoclonal antibody 4G8 light chain (band 2 in SDS-PAGE, Figure 2.18) revealed the presence of novel consensus sites for *N*-glycosylation, located between the complementary determining regions V_L CDR2 and V_L CDR3 (see Chapter 2.1.3.2). The presence of *N*-linked glycosylation in the antibody variable region is interpreted as result of somatic hypermutation of the V_L and V_H genes during affinity maturation of immunoglobulins which may lead to occurrence of potential of *N*-consensus sites. The amino acids contained in the sequon were deciphered from the ETD of the precursor ion *m/z* 858.6 (4+), corresponding to the tryptic peptide V_L(66-96) (see Figure 2.23); the presence of successive fragment ions c_3 through c_7, confirmed the consensus

sequence Asn-Phe-Ser, between the residues V_L(74-76). In contrast to the heavy chain, which is fully glycosylated at residue V_H Asn307, the light chain of 4G8 is partially glycosylated at residue Asn74. This feature accounts for the differential migration of bands 2 and 3 on SDS-PAGE (Figure 2.18); the glycosylated portion of the light chain (band 2) migrates as a diffuse band at higher molecular weight, whereas the non-glycosylated portion appears as a compact band at lower molecular weight (band 3).

The LC-MS/MS analysis of a tryptic digest of light chain band 2 did not show any glycosylated structure, i.e. no characteristic sugar oxonium ions (*m/z* 204.1, 366.1, 292.1) were observed. This may be due to the length of the region V_L(66-96), which has a molecular weight of 3430.4 Da in non-glycosylated form. The trimannosyl chitobiose core $Man_3GlcNAc_2$ has an average molecular weight of 892 Da, such that peptide V_L(66-96) containing this *N*-linked core pentasaccharide would have a theoretical molecular weight of 4322.4, rendering it difficult to observe by LC-ESI-MS. The *N*-glycans on V_L(66-96) are expected to be considerably larger and highly heterogeneous compared to the minimal core structure, forming a more complex pattern. The reduced ionization efficiency of glycopeptides compared to non-modified peptides, co-elution with unrelated peptides, as well as distribution of the ion signal between different charge states of the various glycoforms represent factors affecting the detection sensitivity.

In order to obtain better sensitivity, the tryptic mixture of band 2 was re-digested over night with chymotrypsin. It was expected that the length of the tryptic peptide V_L(66-96) containing several potential chymotrypsin cleavage sites, could be reduced to smaller chymotryptic peptides containing the *N*-linked site, which should be easier to analyze by LC-MS/MS than full length V_L(66-96). Upon re-digestion, the LC-MS/MS analysis contained the sugar oxonium ions characteristic for the presence of glycosylated structures: *m/z* 204.1 (protonated *N*-acetyl glucosamine, GlcNAc+), *m/z* 292.1 (protonated *N*-acetyl neuraminic acid, $NeuNAc^+$) and *m/z* 308 (protonated N-glycolyl neuraminic acid, $NeuGc^+$), suggesting that the use of the enzymatic cocktail of successive trypsin and chymotrypsin was successful. Furthermore, from the observed singly charged monosaccharide ions, variable region glycosylation was observed with glycans terminated with both NeuNAc and NeuGc, a feature less common for mouse immunoglobulins, that contain minimal amounts of *N*-acetyl neuraminic acid and bisecting GlcNAc. The variable region glycopeptides were observed as both doubly and triply protonated molecules.

Figure 2.39 shows the CID spectrum of the precursor ion *m/z* 1066.780 (3+) obtained with a collision energy ramp from 20 to 30 V.

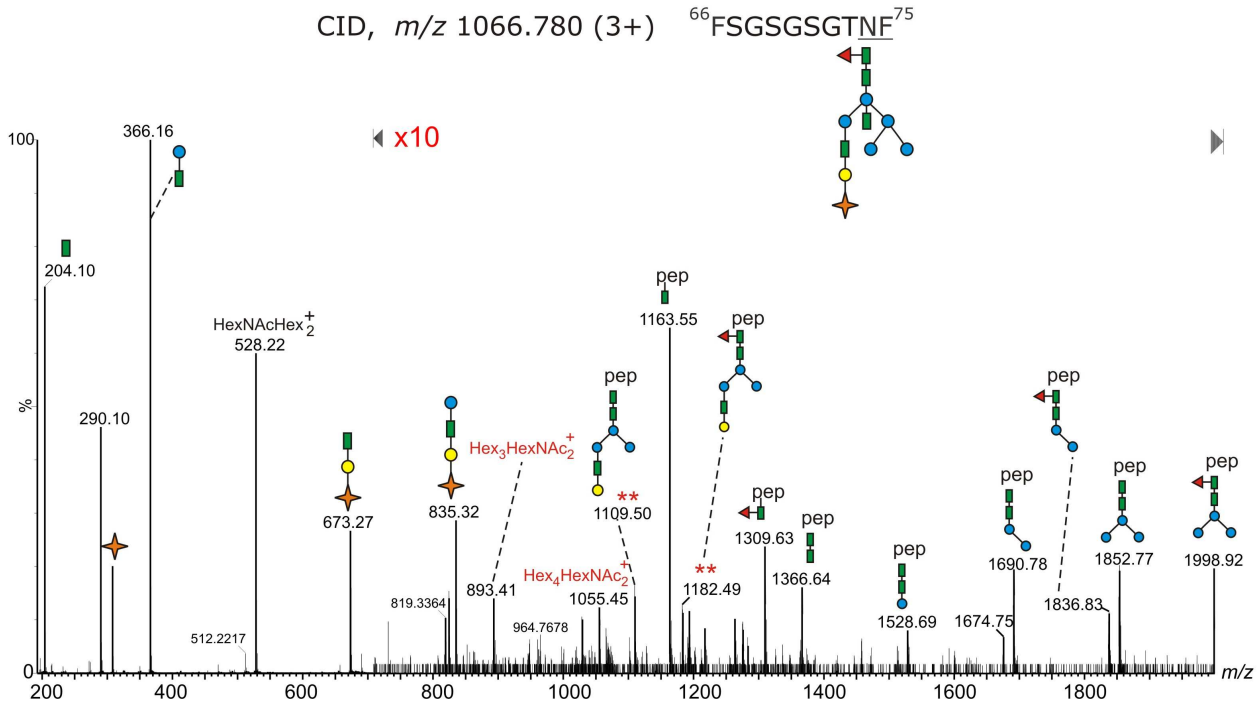


*Figure 2.39: MS/MS of the precursor ion of m/z 1066.780 (3+), corresponding to the glycopeptide FSGSGSGTN*F containing the glycan indicated at the top (right). Doubly charged ions are indicated with double asterisk (**). The remaining ions are singly charged. The CID spectrum was obtained on the Q-Tof Premier in the data dependent mode using a collision energy gradient from 20 to 30 V. Colour code: green square – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose, yellow circle – galactose, orange diamond – N-glycolyl neuraminic acid.*

The dominant decomposition pathway of the precursor ion *m/z* 1066.780 (3+) includes loss of the mono- and oligosaccharide units from the non-reducing end, giving rise to protonated sugar ions observed as *m/z* 204.10 ($GlcNAc^+$), 290.10 (dehydrated $NeuGc^+$), 308.10 ($NeuGc^+$), 366.16 ($GlcNAc_1Hex_1^+$), 528.22 ($GlcNAc_1Hex_2^+$), 673.27 ($GlcNAc_1Gal_1NeuGc_1^+$), 835.32 ($Man_1GlcNAc_1Gal_1NeuGc_1^+$), 893.41 ($GlcNAc_2Hex_3^+$) and 1055.45 ($GlcNAc_2Hex_4^+$), respectively, indicated in Figure 2.39, from which the partial glycan composition was derived. The distinction between the isobaric residues mannose and galactose can not be made based on CID data of glycopeptides; however, in several instances it is possible to assign the hexose residues as one or the other under consideration of the biosynthetic pathways of *N*-glycans. For example, in the fragment ion *m/z* 673.27, the hexose unit may be assigned to galactose, as sialic acids are always terminating *N*-glycans after galactose residues have been incorporated. The sugar ion *m/z* 1055.45, assigned to $GlcNAc_2Hex_4^+$ suggests the presence of a hybrid type glycan, in

which one antenna is elongated with GlcNAc-Gal and terminated with *N*-glycolyl neuraminic acid, whereas the other one is of high mannose type, as indicated in Figure 2.39. Alternatively, this fragment may arise from secondary fragmentation following the loss of NeuGc and of the two terminal mannoses from the depicted structure. The ions observed in the *m/z* range 1100 – 2000 represent singly and doubly protonated glycopeptide fragments comprising partial glycan structures which arise from the loss of the sugar units from the non-reducing end and charge reduction of the precursor. The doubly protonated species are indicated with double asterisk (**) in Figure 2.39.

Based on the observed ion of *m/z* 1163.55, the identity of the peptide backbone was assigned to the fragment ^{66}FSGSGSGTNF^{75} which contains a GlcNAc monosaccharide attached at the next-to-last Asn74 residue. Hence, one cleavage pathway of chymotrypsin in the substrate V_L(66-96) was determined within the consensus sequence for *N*-glycosylation, C-terminal to Phe75. Furthermore, it appears that the enzyme did not cleave C-terminal to Phe66, such that the minimal peptide containing the *N*-linked site was observed as V_L(66-75). The remaining fragment ions were consistent with the glycan compositions assigned in Figure 2.39. From the mass difference between the intact glycopeptide (MW 3197.317 Da) and the non-glycosylated peptide V_L(66-75) (MW 959.387 Da), the mass of the glycan fragment was calculated as 2237.930, and the corresponding sugar composition was determined as $GlcNAc_4Fuc_1Hex_6NeuGc_1$. This corresponds to a hybrid type glycan, containing structural features from both complex and high mannose glycans, as follows: the trimannosyl chitobiose core has a fucose attached at the Asn-linked GlcNAc; one antenna is of complex type, bearing GlcNAc-Gal-NeuGc and bisecting GlcNAc, as in mature complex glycans, while the remaining core mannose contains two additional mannoses (see the structure depicted in Figure 2.39). Obviously, there are several possibilities in which the sugar units could be interconnected to form the glycan with the determined composition, in addition to structure described here. Considering the biosynthetic pathways of hybrid glycans, the structure assigned to this glycan represents the most probable variant, although the exact linkage can not be derived from these MS/MS data.

The results obtained by using the combination of trypsin and chymotrypsin on gel band 2 were corroborated by the results obtained from the LC-MS/MS analyses of the same band

digested only with chymotrypsin. Figure 2.40 shows the MS/MS spectrum of the ion *m/z* 1190.024 (2+).

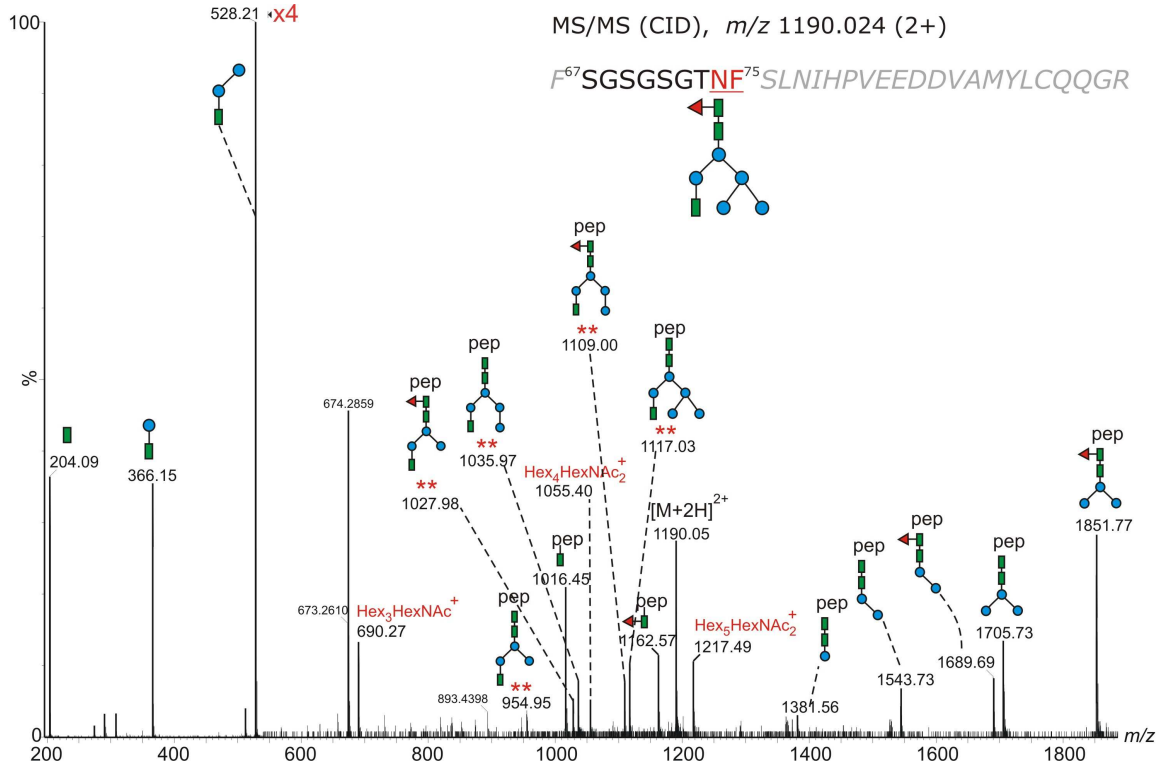


*Figure 2.40: MS/MS of the precursor ion of m/z 1190.024 (2+), corresponding to the glycopeptide ^{67}SGSGSGTN*F^{75} containing the glycan indicated at the top (right). The peptide was derived from chymotryptic digestion of the light chain band 2 (Figure 2.18). Doubly charged ions are indicated with double asterisk (**). The remaining ions are singly charged. The CID spectrum was obtained on the Q-Tof Premier in the data dependent mode using a collision energy gradient from 20 to 30 V. Colour code: green square – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose.*

The fragments observed in the CID of the doubly protonated precursor molecule of *m/z* 1190.024 are consistent with the indicated hybrid glycan of composition $GlcNAc_3Fuc_1Hex_5$ attached to the chymotryptic peptide ^{67}SGSGSGTNF^{75}. The ion *m/z* 1016.45 was assigned to peptide V_L(67-75) containing an *N*-acetyl glucosamine residue attached to Asn74, indicating that chymotrypsin cleaves both Phe66 and Phe75 residues in the full length 4G8 light chain. These two independent analyses ascertain the structural assignment of the glycans and the identity of the peptide.

The deconvoluted mass spectrum showing the glycoforms attached at the residue Asn74 on the light chain, observed in glycopeptides V_L(67-75) is presented in Figure 2.41.

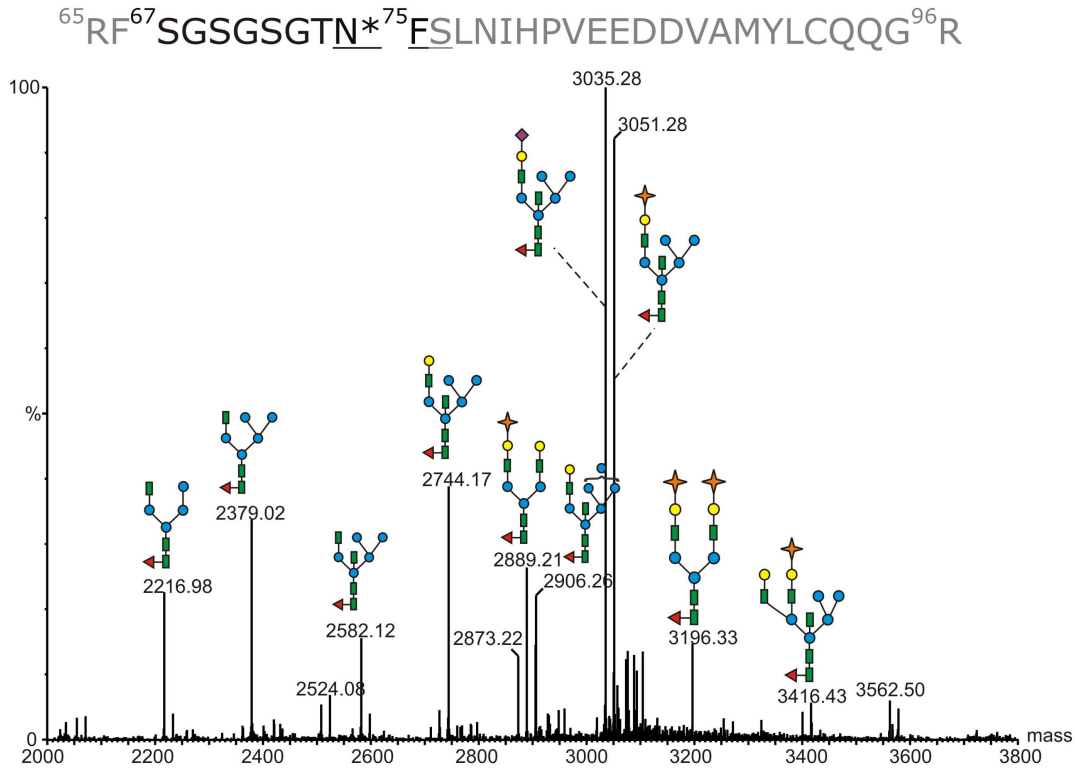


Figure 2.41: Positive ion nano-HPLC-ESI/MS of the V_L glycopeptides ^{67}SGSGSGTN^{75}F from the light chain of the mouse monoclonal antibody 4G8; The deconvoluted MS spectrum was obtained by averaging the MS scans over the chromatographic window where glycopeptides eluted. The glycan structures of the most abundant glycoforms are indicated above each ion. Colour code: green rectangle – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose, and yellow circle – galactose, orange diamond – N-glycolyl neuraminic acid, purple diamond – N-acetyl neuraminic acid.*

The sugar population decorating the variable region is more heterogeneous compared to that of the constant region. The most abundant glycans were of hybrid type, whereas the complex type oligosaccharides, terminated with one or two NeuGc residues, and assigned to the deconvoluted masses 2889.21 and 3196.33, respectively, represent only a minor portion of the total glycan pool. The structures of the major types of *N*-glycans found at Asn74 are shown in Figure 2.42.

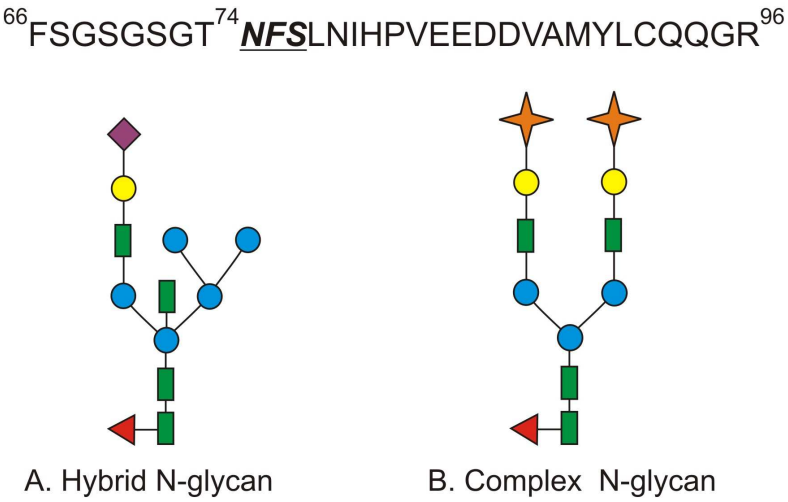


Figure 2.42: Structures of the major type of glycans found at Asn74 on the light chain variable region of the anti(17-24) monoclonal antibody: (A) Core fucosylated hybrid N-glycan containing one antenna of high mannose type, bisecting GlcNAc, and the second antenna of complex type terminated with N-acetyl neuraminic acid; (B) Core fucosylated complex type glycan, containing two ß-galactose residues and N-glycolyl neuraminic acid on each antenna. Colour code: green rectangle – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose, yellow circle – galactose, purple rhombus – N-acetyl neuraminic acid, and orange star – N-glycolyl neuraminic acid.

Under physiological conditions, variable region glycans from human plasma and recombinant IgG sources were found to be of complex type, containing elevated amounts of sialic acids on entirely core-fucosylated structures [247, 248]. Hence, the pattern determined for the mouse antibody 4G8 light chain differs substantially, in particular through the presence of abundant and heterogeneous hybrid structures. At present, little is known about the variable region glycosylation in other species. Another interesting structural feature is derived from the termination of the hybrid oligosaccharides by both *N*-acetyl and *N*-glycolyl neuraminic acid, as mouse immunoglobulins typically contain NeuGc instead of NeuAc. Variable region glycans are known to affect the antigen recognition, such that an interesting future study would be to determine the effects of light chain glycosylation of the 4G8 antibody on antigen binding.

2.2.5 Subclass specific glycosylation profiling of Aß-autoantibodies

Intravenous immunoglobulin (IVIg), a purified IgG fraction from the blood of healthy individuals is an FDA approved therapeutic agent for immune and inflammatory diseases. IVIg contains Aß autoantibodies which have been shown to exert a positive effect on AD patients. A phase 3 study evaluating safety and effectiveness of IVIg for the treatment of mild to moderate AD was initiated in December, 2008 [249]. For glycosylation studies,

affinity isolation of Aß-autoantibodies was performed using the Cys-Aß(12-40) polypeptide, which contains the specific C-terminal epitope recognized by the autoantibodies.

As described for the 6E10 antibody, glycopeptides were selectively detected in this mixture by monitoring the formation of the $GlcNAc^+$ oxonium ion of *m/z* 204.1 in the parent ion detection mode (Figure 2.43). In contrast to the mouse monoclonal antibodies, a complex pattern is observed for the extracted ion current (EIC) of *m/z* 204.1 as a result of the polyclonal nature of the Aß-autoantibody.

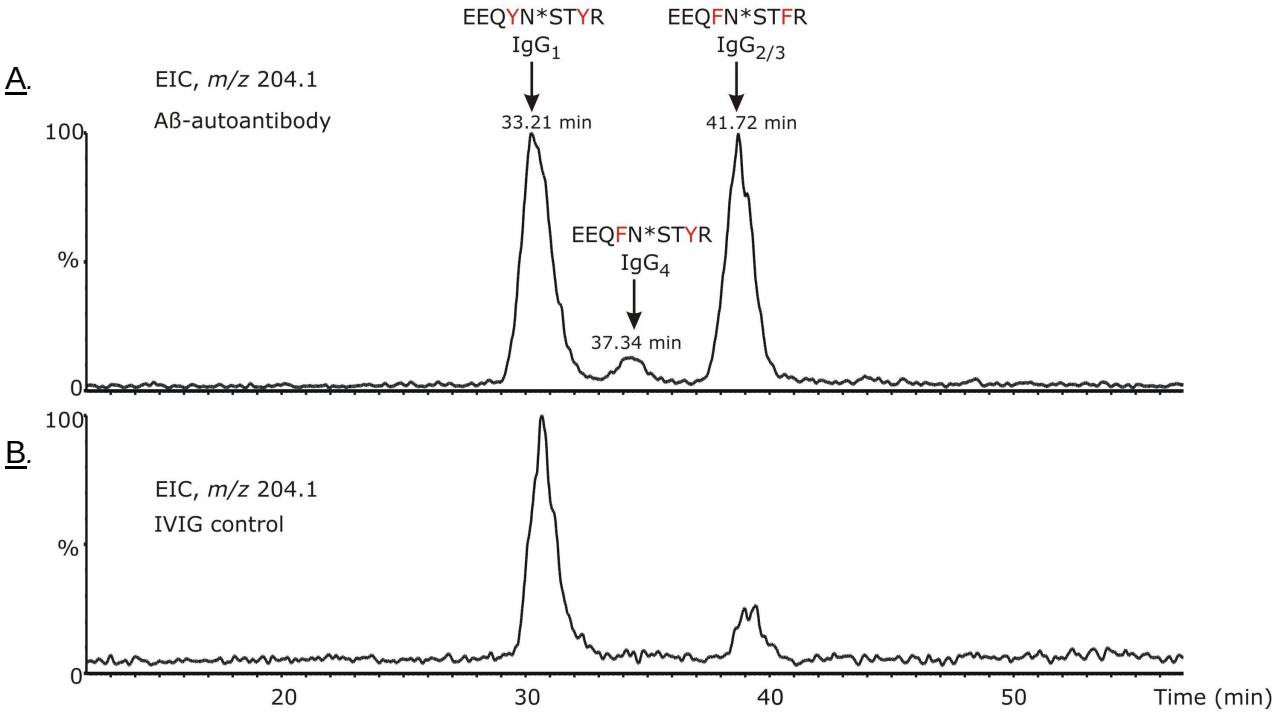


Figure 2.43: Extracted ion chromatogram (EIC) for m/z 204.1, corresponding to protonated GlcNAc+, over the chromatographic retention time 10-60 min for (A) ß-amyloid autoantibody heavy chain tryptic digest, and (B) IVIg heavy chain tryptic digest. The data were obtained using MS-only acquisition. The peaks observed in the chromatogram were assigned to distinct N-glycosylated peptide isoforms, corresponding to individual IgG subclasses found in the Aß-autoantibody and IVIg, as indicated above each peak.

The glycopeptides typically elute early in the chromatogram (15-20% acetonitrile), due to the polar character of the attached glycans. All four IgG subclasses were detected in the heavy chain tryptic mixture (Figure 2.43 A) with IgG_1 and IgG_2 subclasses being observed with the highest abundance. Glycopeptides derived from IgG_1 elute earlier than those of IgG_4 and IgG_2/IgG_3 and within each subclass the neutral glycopeptides elute slightly earlier than the sialylated ones. Human IgG subclasses show more than 95% constant region sequence homology, but characteristic differences are found in the length of the hinge region, in the number of disulfide bridges, and also in the C_H2 domain around the region of

N-linked glycosylation. The tryptic glycopeptides of the Aß autoantibody IgG_1 contained the amino acid sequence $EEQ^{296}Y\underline{NST}^{300}YR$, while, for IgG_2/IgG_3, two simultaneous amino acid substitutions were found, Y296F and Y300F, respectively. The glycopeptides derived from IgG_4 contain (compared to IgG_1) a single amino acid replacement, Y296F [250]. These were observed as both doubly and triply protonated molecules in the full scan MS. In addition, these glycopeptides were found to contain uncleaved arginine and lysine residues, such as the amino acid sequence TKPREEQX<u>NST</u>XR, where X denotes the amino acid mutations characteristic for each IgG subclass. For a rigorous, subclass specific glycosylation analysis of the Aß autoantibody, it was essential to ensure that the tryptic digestion of the antibody heavy chain proceeded to completion. This was based on the observation that miss-cleaved and fully processed glycopeptides from a specific subclass have distinct chromatographic elution times and may co-elute with glycopeptides from other subclasses, thus complicating their overall analysis.

The concentration of each immunoglobulin in serum of healthy individuals depends on several factors, e.g. on the number of plasma cells producing that antibody type. Adults exhibit highest concentrations of IgG_1 (10-12 mg/mL), followed by IgG_2 (2-6 mg/mL), IgA_1, IgM, IgG_3 (0.5-1 mg/mL), IgG_4 (0.2-1 mg/mL), IgA_2, IgD and IgE [251, 252]. The extracted ion chromatogram, EIC, of *m/z* 204.1, obtained for the heavy chain tryptic digest of IVIg, is shown in Figure 2.43 B, and indicates that IgG_1 is the most abundant subclass in this commercial product. For glycosylation analysis of the Aß autoantibody, IVIg was chosen as a control, as this represented the starting material for the epitope-specific isolation of the Aß autoantibody. Furthermore, it has been shown that autoantibodies may exhibit different constant region glycosylation profiles compared to total serum IgG [253]. Because IgG_2 and IgG_3 share identical amino acid sequences around the *N*-glycosylation site, it was not possible to separately analyze their glycosylation profile. However, the amount of IgG_3 is considerably lower than IgG_2 in human plasma. The EIC of *m/z* 204.1 (Figure 2.43) indicate that the Aß autoantibody contains elevated levels of $IgG_{2/3}$ compared to IVIg. From the ion abundances of all glycopeptides observed in each individual subclass, the ratio $IgG_{2/3}$/IgG_1 for the Aß autoantibody was determined to be approximately 1, while, for IVIg, the ratio $IgG_{2/3}$/IgG_1 was determined as ¼. This semiquantitative estimation did not take into account differences in ionization efficiencies of distinct glycoforms and peptide isoforms. Interestingly, the levels of IgG_4, although low, were found to be higher than those

in total serum IgG. The observed levels of IgG_4 in IVIg were found to be close to the limit of detection, and, therefore, no semiquantitative analysis of its abundance was performed.

The *N*-glycosylation profiles for each individual IgG subclass of the Aß autoantibody are shown in Figure 2.44. Each mass spectrum was averaged and deconvoluted over the chromatographic elution times of glycopeptides with the amino acid sequence EEQXNSTXR (where X = F or Y). The identities of the peptide isoforms derived from individual IgG subclasses and of their attached glycans were determined from MS/MS acquired in the data dependent mode, and from the experimental glycopeptide masses and theoretical mass values of the peptides without the sugar.

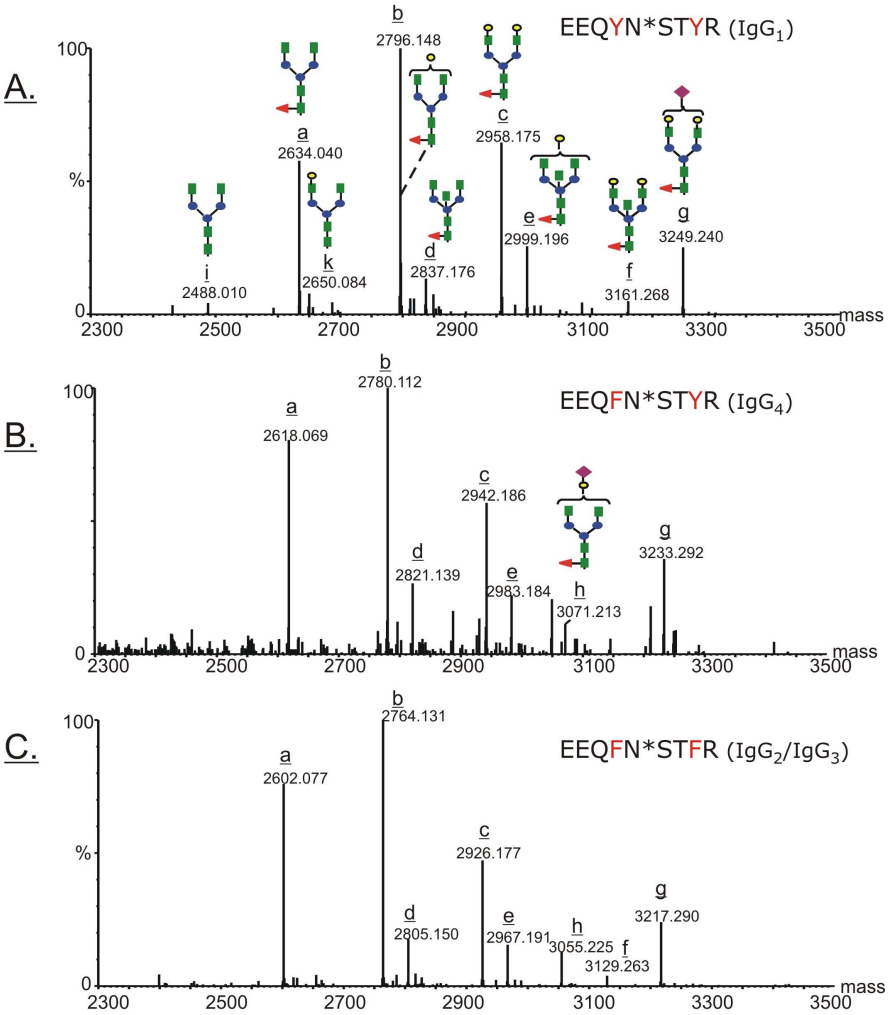


Figure 2.44: Deconvoluted mass spectra over the mass range 2300 – 3500, indicating the glycan populations determined for each IgG subclass: ***(A)*** *EEQYNSTYR (*IgG_1*),* ***(B)*** *EEQFNSTYR (*IgG_4*) and* ***(C)*** *EEQFNSTFR (*$IgG_{2/3}$*). The complex type glycans are represented with the following color code: green square – N-acetyl glucosamine; red triangle – fucose; blue circle – mannose; yellow circle – galactose; purple rhombus – N-acetyl neuraminic acid. The glycoforms are indicated with latin letters from* ***a*** *to* ***k*** *and the structures corresponding to each glycan are represented in* ***(A)*** *and* ***(B)****.*

However, MS/MS was essential to establish the correct glycan compositions, as accurate mass determination alone was not sufficient to discriminate between isobaric structures. For example, the amino acid substitution Y296F in IgG_4 compared to IgG_1 has a mass difference of 16, which is identical with the mass difference between fucose and hexose. Consequently, the deconvoluted mass of 3715.66, calculated for the observed ion of *m/z* 929.69 (4+), could have been assigned to either the "missed cleavage" glycopeptide from IgG_4 (TKPREEQFNSTYR) containing the glycan G_2FSA, or to the glycopeptide from IgG_1 with a glycan G_1F_2SA. The MS/MS of this precursor ion contains fragment ions: *(i)* 929.99 (2+), assigned to the peptide backbone derived from IgG_4 which still has the first GlcNAc residue attached at Asn 297, and *(ii)* 869.96 (2+), which corresponds to the same peptide backbone with a cross-ring cleavage in the first GlcNAc unit (denoted as $^{0,2}X$).

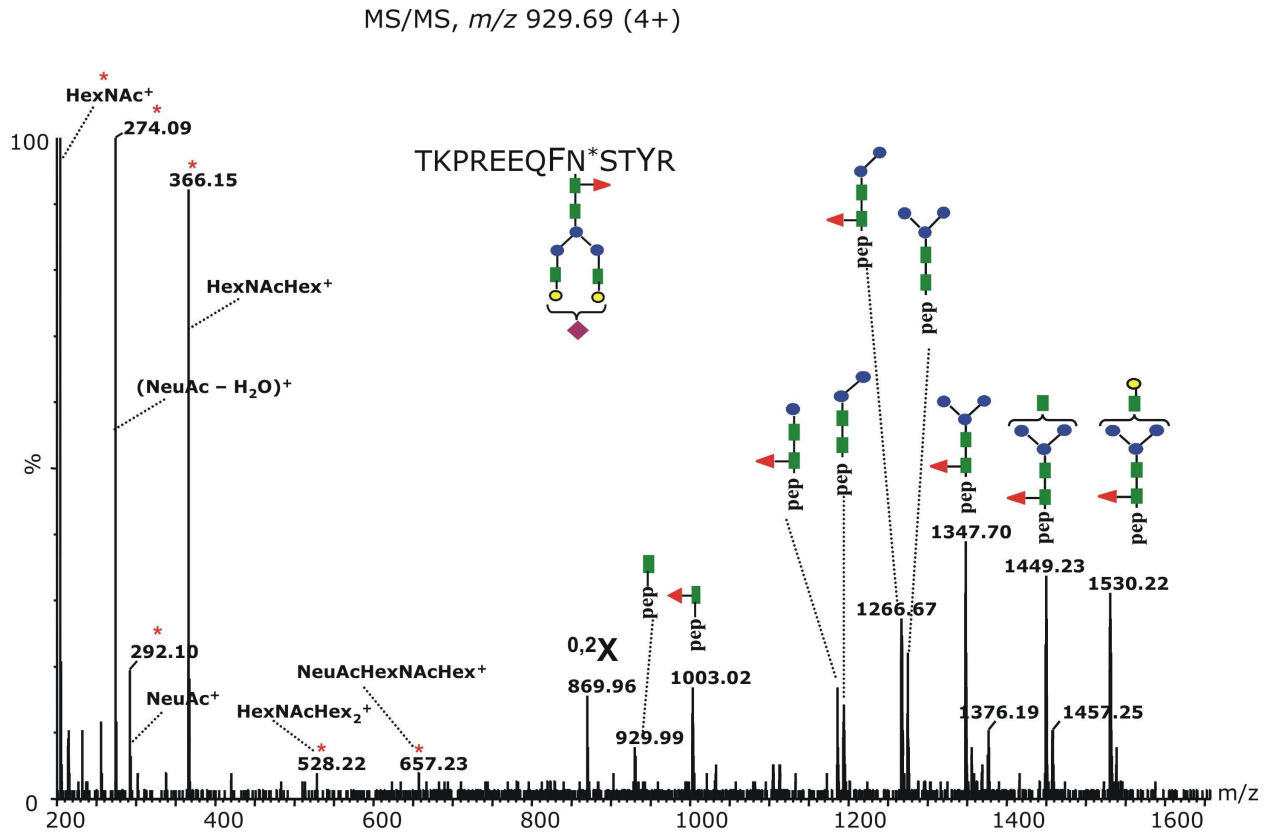


*Figure 2.45: MS/MS spectrum of the precursor ion m/z 929.69 (4+) obtained in data dependent mode with a collision energy ramp from 30 to 40 V, and assigned to the miss-cleaved glycopeptide TKPREEQFN*STYR containing the glycan G_2FSA. The ions marked with an asterisk (*) are singly charged. All the other ions observed in CID are doubly charged and the indicated masses are monoisotopic.*

The glycans decorating the Aß autoantibody constant region are almost entirely core fucosylated and the most abundant glycoform in each IgG subclass is G_1F, followed by G_0F and G_2F. In the case of IgG_1, the digalactosyl and the agalactosyl structures have similar abundances (see Figure 2.44 A), while for $IgG_{2/3}$ and IgG_4 the G_0F population

appears to be higher than the G_2F (Figure 2.44 B and C). The glycoforms containing bisecting GlcNAc (G_0FB, G_1FB and G_2FB) and sialic acid (G_1FSA and G_2FSA) were observed with lower abundance, while the glycoforms lacking the core fucose (G_0, G_1 and G_2) were barely detectable. The subclass specific glycosylation of the Aß autoantibody compared to that of IVIg is presented in Figure 2.46 for the eleven most abundant glycoforms.

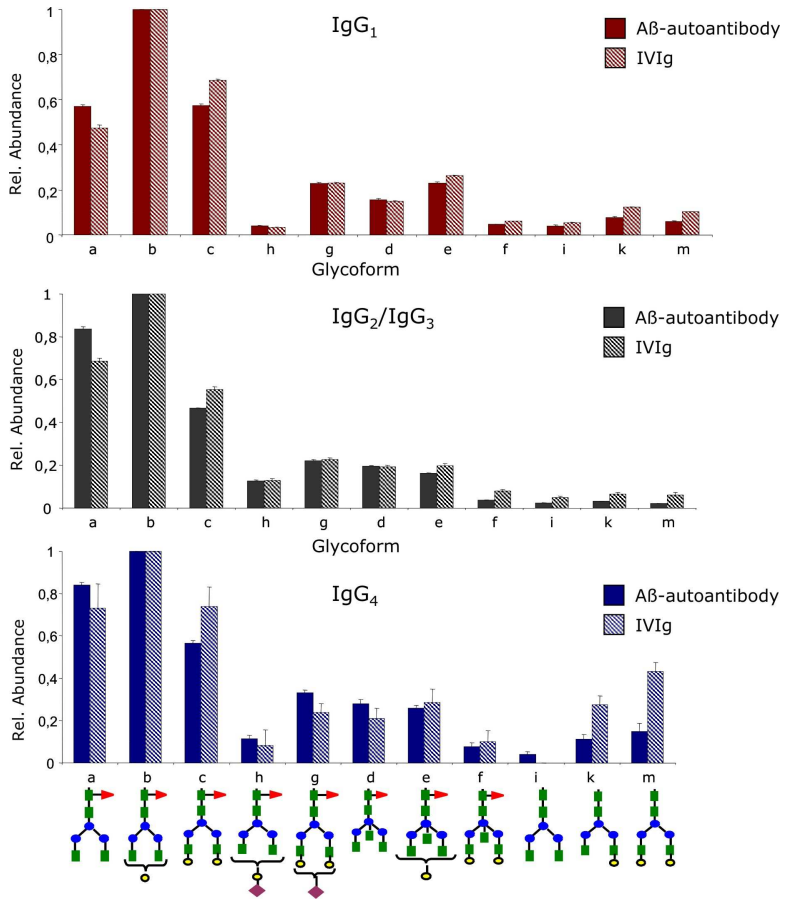


Figure 2.46: Subclass specific glycosylation profiles of the Aß autoantibody constant region compared to IVIg: top – IgG_1; middle – IgG_2/IgG_3; bottom – IgG_4. The profile for each individual subclass was determined for the 11 most abundant glycoforms relative to the abundance of the G_0F glycoform within each subclass. The structure of each glycoform is depicted at the bottom. The one letter annotation is identical with that used in Figure 2.44. Bar code: full bars – Aß autoantibody, stripped bars – IVIg.

In summary, the Aß autoantibody contains lower levels of galactosylation, as G_0F glycoform is elevated and G_2F is decreased within each subclass compared to IVIg. No significant differences were observed among the remaining glycoforms for IgG_1 and $IgG_{2/3}$, respectively. Within the IgG_4 subclass, the slightly elevated levels of G_1 and G_2 in IVIg compared to Aß autoantibody may represent an artifact derived from the isobaric nature of the structures G_2 and G_1 in IgG_4 with the structures G_1F and G_0F, respectively, of IgG_1; because the glycopeptides from this subclass were hardly detectable in IVIg, one can not

exclude the possibility that the values determined for G_1 and G_2 may contain a contribution from the isobaric glycoforms from IgG_1.

2.2.6 Concluding discussion of the glycosylation structure of Aß-specific antibodies

Glycosylation analysis of epitope – specific anti-Aß antibodies provided a detailed picture of the glycans attached at the conserved *N*-linked position on the heavy chain. Moreover, analysis of glycopeptides has advantages over conventional carbohydrate analysis, revealing the specific glycoform microheterogeneities of the individual IgG subclasses and a semiquantitative estimation of their distribution in the Aß-autoantibody. This finding may be important because of the specific effector functions of each IgG subclass. IgG_1 represents the primary secretory product of the adaptive immune system and it is specific for protein antigens, while IgG_2 is secreted in response to stimulation with carbohydrate antigens, e.g. the polysaccharides of the bacterial cell walls [254]. A deficit or increase in selected IgG subclasses may have relevance for the activity of the Aß-antibody. The results derived from the analysis of the Aß-autoantibody glycosylation indicate that the Aß-autoantibody contains approximately four times more IgG_2 compared to IVIg. IgG_4 is the less abundant IgG subclass in human plasma [251, 252, 255] and these antibodies become prominent only after prolonged immunization with protein antigens [256]. It would be interesting to probe whether increased levels of IgG_2 and/or IgG_4 correlate with a possible pathophysiologic role of amyloid-ß autoantibodies.

It is widely accepted that the antibody effector functions are dependent on appropriate glycosylation of the constant region. In the mouse IgG the non-galactosylated species represents the most abundant glycan population, followed by mono- and digalactosyl glycoforms. This pattern is common for other recombinant antibodies [224]. In humans, a high level of non-galactosylated species is characteristic of autoimmune disorders such as rheumatoid arthritis. A possible explanation for this is that the uncovered GlcNAc residues attached to the core pentasaccharide in combination with backbone protein motifs could reveal novel antigenic determinants which are normally masked by galactose [243]. In vitro studies demonstrated two-fold reduced levels of complement lysis activity of recombinant antibodies having reduced levels of galactosylation [257, 258]. The Aß-autoantibody

shows slightly decreased galactosylation compared to IVIg; a feature which is in contrast to its protective nature.

The Aß-autoantibody contains significant amounts of bisecting GlcNAc and terminal *N*-acetyl neuraminic acid on fully core fucosylated structures. It has been reported that fully fucosylated IgG_1 shows a 50-fold decrease in receptor binding affinity compared to the non-fucosylated antibody, and a 100-fold decrease in antibody dependent cellular cytotoxicity [259-261]. A recent study demonstrated that highly sialylated antibodies exhibit anti-inflammatory properties derived from reduced binding to the FcγRIIIa receptor and altered antigen binding [262, 263], which were explained by the lower flexibility of the hinge region induced by the presence of neuraminic acid.

Unlike the human Aß-autoantibody, both mouse IgGs contain some tri- and tetragalactosylated species in the constant region, with immunogenic potential [245], as well as low abundance hybrid structures in the constant region (6E10 antibody) that enhance the molecular microheterogeneity. Variable region glycosylation of the 4G8 anti-Aß(17-24) antibody light chain shows pronounced microheterogeneity of the glycoforms, the most abundant being of hybrid type terminated either with *N*-acetyl neuraminic acid or with *N*-glycolyl neuraminic acid. Under physiologic conditions, variable region glycans from human plasma and recombinant IgG sources were found to be of complex type, containing elevated amounts of sialic acids on entirely core-fucosylated structures [247, 248]. Hence, the pattern determined for the mouse antibody 4G8 light chain differs substantially, in particular through the presence of abundant and heterogeneous hybrid structures. As for now, however, little is known about the variable region glycosylation is other species, although this is known to affect the antigen binding.

In conclusion, investigation of immunoglobulin glycosylation by mass spectrometry represents a highly sensitive method for elucidation of subclass specific glycan populations, and for probing the structural integrity of potential therapeutic candidates. Because immune therapy has received considerable attention in the last years for both treatment and prevention of AD and the molecular mechanisms of AD and the protective role exhibited by ß-amyloid autoantibodies are poorly understood, the molecular characterization of glycosylation of these antibodies represents a new approach to extend the understanding for their physiological role.

2.2.7 Elucidation of *O*-glycosylated structures of amyloid precursor protein

The amyloid precursor protein (sAPP695), was expressed in CHO cells following the procedure described by Sato *et al.* [264]. No fetal calf serum was supplemented to the cell culture medium, which greatly facilitated the purification procedure. The cell culture medium was subjected to Q-sepharose chromatography using a linear NaCl gradient [264], yielding a single peak containing the APP695. An identical binding efficiency of APP695 to the Q-sepharose column was observed, as previously reported by Sato *et al.* [264]. APP695 was detected by Western blot analysis in the cell culture medium and in the fraction eluted from the Q-sepharose column as a band migrating at approximately 75 to 105 kDa on SDS-PAGE. The identity of the protein was confirmed by LC-MS/MS analysis of the tryptic digest and database search using the NCBInr protein database. Despite the co-elution of multiple proteins from the Q-sepharose column, the SDS-PAGE separation of the protein band of interest was satisfactory, so that no additional purification step was required. The secreted (unmodified) APP695 sequence has an expected molecular weight of approximately 68 kDa, and the difference between the theoretical and the apparent molecular weight observed on SDS-PAGE is largely ascribed to glycosylation.

Specific *O*-glycopeptide structures of the secreted APP695 were identified by nano-LC-ESI-MS/MS analysis of a tryptic digest of the corresponding protein band separated by SDS-PAGE. Using data dependent CID-MS, the glycopeptides containing *O*-linked sugars were identified at elution times between 19.5 and 24 minutes by the presence of characteristic sugar-oxonium ions (*m/z* 292.1, protonated neuraminic acid (SA^+); *m/z* 274.1, protonated dehydrated neuraminic acid ($SA^+ - H_2O$); *m/z* 204.1, protonated *N*-acetyl hexosamine ($HexNAc^+$); *m/z* 366.1, protonated hexose-*N*-acetyl hexosamine ($HexHexNAc^+$)). These ions arise during CID of the protonated glycopeptides by glycosydic bond cleavages of the sugar moiety from its non-reducing end (Y/B-type fragmentation [227]), including complete loss of the oligosaccharide [215]. The appearance of ions corresponding to the non-glycosylated tryptic peptides provided initial information of the peptide's identity. Two independent MS/MS analyses of the tryptic peptides were carried out under identical chromatographic conditions: *(i)*, first, data dependent ETD was performed and, *(ii)* second, data dependent CID spectra were acquired. In this manner, the amino acid sequences of the peptides containing the

glycans, and the location of the *O*-linked sites were elucidated. No consensus sequence for *O*-glycosylations has been defined at present. In addition, a significant challenge for the *"de novo"* identification of the *O*-linked glycopeptides and their glycosylation sites was due to the primary structure of full length human APP695 which contains 30 serine and 45 threonine residues, of which 27 serine and 39 threonine residues are contained in the secreted form, sAPPα (see Figure 2.47). A total of 15 serine and 26 threonine residues were observed in the tryptic and chymotryptic digests of sAPP695 by LC-MS/MS analysis.

The sequence coverage obtained by trypsin and α-chymotrypsin digestion, and the combination of both proteases is shown in Figure 2.47. Two peptides with potential *O*-glycosylation were identified in the LC/MS analyses at the partial sequences, (289-302) and (574-587).

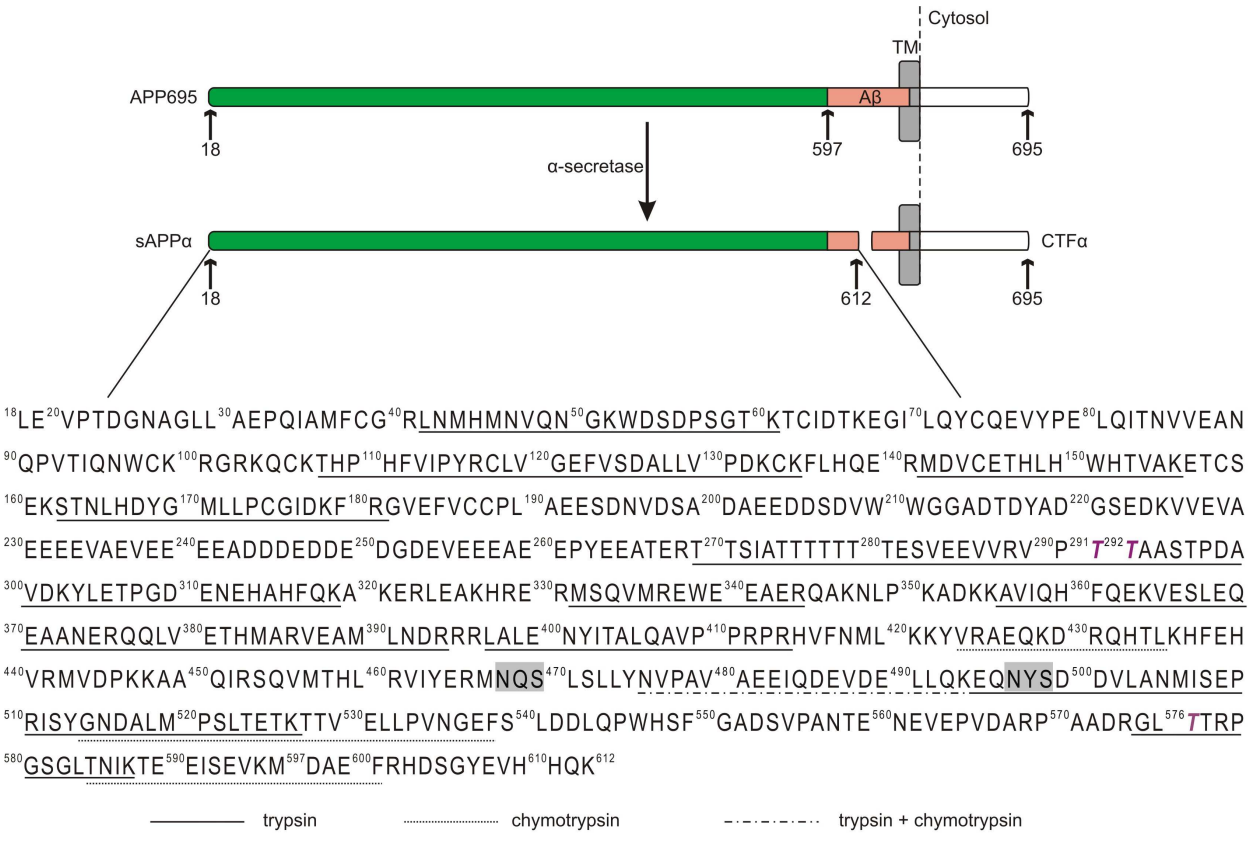


Figure 2.47: Schematic representation of the non-amyloidogenic cleavage pathway of the transmembrane human APP695 generating the secreted fragment sAPPα and the C-terminal fragment CTFα; the amino acid sequence of sAPPα, residues 18-612, is depicted at the bottom. The N-glycosylation sites are highlighted in grey. The identified O-glycosylation sites are indicated in purple. The peptides derived from the use of different enzymes observed by LC-MS/MS are underlined.

Three threonine residues, Thr291, Thr292 and Thr296, and one serine, Ser295, are located in the sequence (289 – 302), while three threonine residues, Thr576, Thr577 and

Thr584, and one serine residue, Ser581, are located in peptide (574 – 587) (numbering as in the full length APP695). The CID and ETD fragmentation of these peptides were interrogated in detail to identify the sites and composition of the glycans.

The averaged ESI-MS spectrum over the chromatographic retention time 19.5 – 20.9 min (Figure 2.48) showed the population of the glycopeptides within the sequence (289 – 302), and the heterogeneity resulting from the attachment of distinct glycans. It can not be excluded that a low extent of in source fragmentation may occur under our MS conditions. However, the extracted ion currents of each individual glycoform are centered at distinct chromatographic elution times, suggesting that the observed heterogeneity is of physiological origine, and not artifactual. The results are presented with the observed relative abundance of each individual glycoform, and were not corrected for differences in ionization efficiencies resulting from the attachment of distinct glycans. In addition to the glycosylated forms, the unmodified peptide (289 – 302) was detected as the doubly-protonated molecule of *m/z* 686.6 and represented the most abundant species.

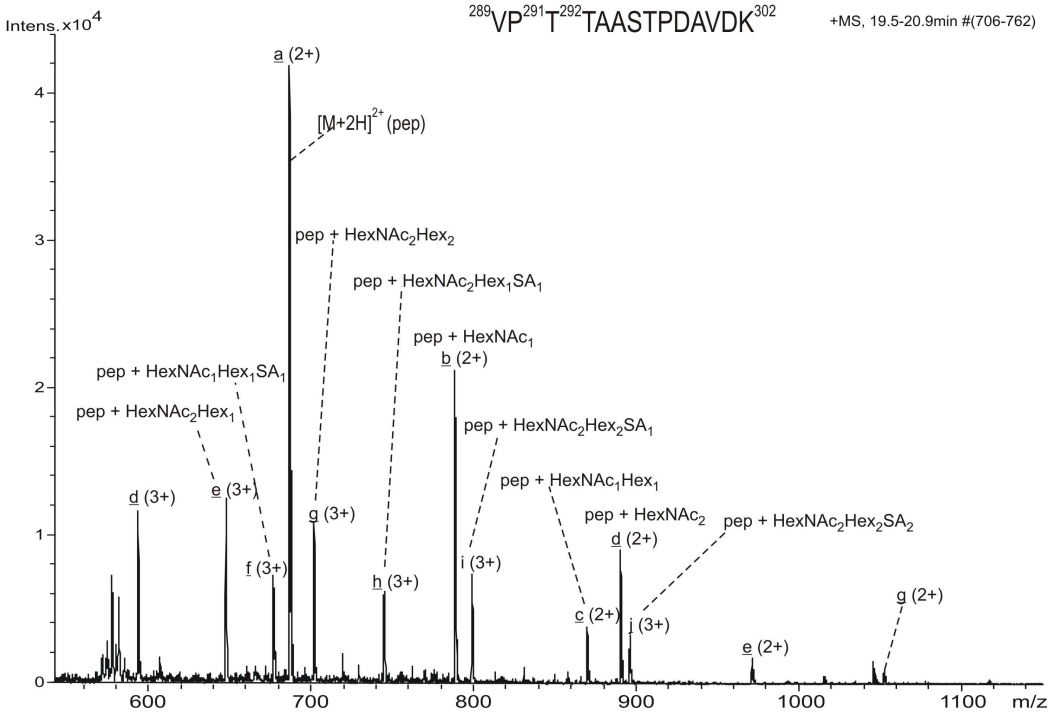


Figure 2.48: Positive ion scan ESI mass spectrum summed over the chromatographic retention time 19.5-20.9 min, indicating the major glycoforms of the peptide 289-302 of the full length APP695. The individual glycopeptides are highlighted with letters from a through j. The charge states and the composition of the glycans determined for every glycopeptide are indicated for each glycoform.

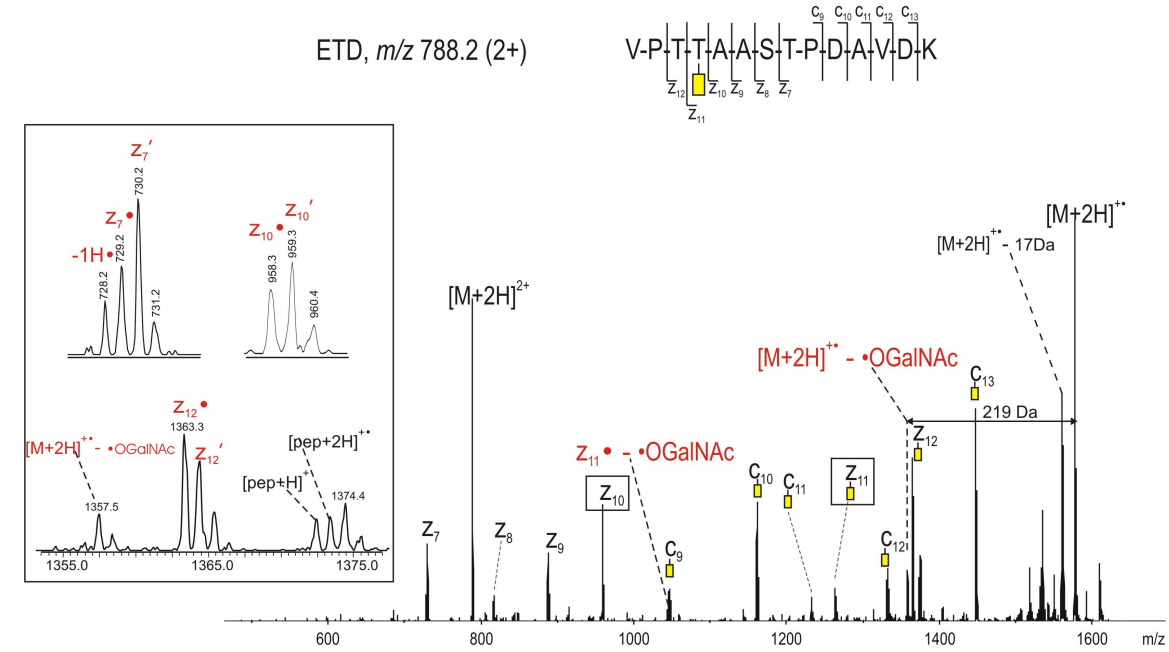
The sugar compositions of the glycopeptides, determined from the molecular weight difference of each individual glycopeptide compared to the non-glycosylated peptide, are indicated in Figure 2.48. The subsequent analysis of the CID and ETD spectra of the ions provided *(i)*, the determination of the composition of the oligosaccharide modification for each glycopeptide from the CID tandem mass spectra, and *(ii)* the identification of the site(s) of modification from the ETD MS/MS.

The most abundant glycoform of peptide (289 – 302), observed as both doubly (*m/z* 788.2) and triply charged (*m/z* 525.7) ions, was identified as containing a single *N*-acetyl hexosamine unit. This monosaccharide is most likely assigned to *N*-acetyl galactosamine (GalNAc), because the biosynthesis of *O*-glycans in vertebrates is initiated by the attachment of α-linked GalNAc to Ser/Thr residues [265]. In the CID spectrum of the ion of *m/z* 788.2 (2+), loss of the GalNAc residue and charge reduction of the precursor give rise to the base peak, *m/z* 1372.3, representing the singly charged unmodified peptide ion. The GalNAc attachment site was unambiguously determined from the ETD of the parent ion *m/z* 788.2 (2+) (Figure 2.49 A). The base peak in this spectrum corresponds to the charge reduced species of *m/z* 1575.4, the singly protonated, even-electron glycopeptide (see Figure 2.49 B). This species may arise either from the transfer of a proton from the 2+ precursor to the fluoranthene (radical) anion or by neutral loss of a hydrogen radical from the charge-reduced radical molecular ion $[M+2H]^{+\bullet}$. The 1 Da heavier isotopomer (peak) (*m/z* 1576.4), having a similar abundance as the ^{12}C isotopomer, represents a mixture of the ^{13}C isotope of the $[M+H]^+$ species and of the ^{12}C isotope of the $[M+2H]^{+\bullet}$ odd-electron species arising from the transfer of a single electron to the precursor ion.

Abundant fragmentation along the peptide backbone resulted in formation of z• and c′ species, which covered almost the complete sequence of the peptide (289 – 302) (Figure 2.49 A). Based on the observed c′ and z• ions, in particular z_{10}• (*m/z* 958.3) and z_{11}• (*m/z* 1262.4), which are separated by the mass increment of the glycosylated threonine residue, it was possible to unambiguously assign Thr292 as the *O*-glycosylation site, and to rule out other possible modification sites. A number of ions was observed to result from hydrogen rearrangement to and from z• fragments, as demonstrated previously in ECD experiments [191, 234]. Specifically, the fragment z_7• (*m/z* 729.2), resulting from N-C_α cleavage between the Ser and Thr residues, showed pronounced hydrogen rearrangement with formation of both species z_7' (1 Da heavier than z_7•) and [z_7• - 1H•]. In contrast, radical

migration to the fragment z_{10}• (*m/z* 958.3) results exclusively in formation of the abundant even-electron species z_{10}' at *m/z* 959.3 (see insert in Figure 2.49 A). Similarly, one hydrogen radical is transferred to the fragment z_{12}• (*m/z* 1363.3) to form the abundant z_{12}'. According to Savitski *et al*., the extent of H• rearrangement is dependent on the nature of the residues adjacent to the N-C_α bond [266], with Thr adjacent to the radical site promoting both H• addition and/or H• loss. In agreement with this observation, pronounced H• transfer to z_7•, z_{10}• and z_{12}• was observed, whereas H• loss was abundant only for the z_7• species.

A.



B.

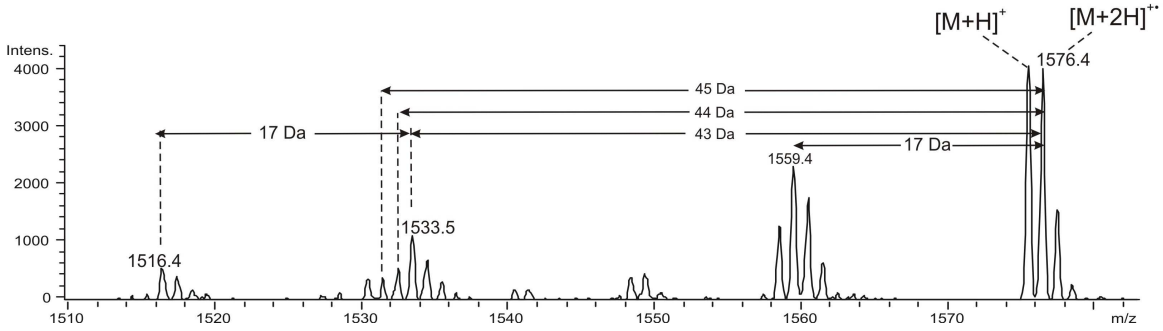


Figure 2.49: (A) ETD of the precursor ion m/z 788.2 (2+), corresponding to the peptide 289-302 of the full length APP695 containing the monosaccharide N-acetyl galactosamine (yellow rectangle) attached at Thr 292. The red label indicates the radical loss of 219 Da (•OGalNAc) from the radical species $[M+2H]^{+\bullet}$ and z_{11}•, respectively. The fragment ions relevant for determination of the glycosylation site are indicated with black boxes. The inset in this spectrum shows the peak distribution of the species z_7, z_{10} and z_{12} (see discussion in the text); (B) Insert over the mass range m/z 1510-1580 of the ETD spectrum obtained for the precursor ion of m/z 788.2 (2+), showing the small neutral losses derived from the molecular radical ion of m/z 1576.4.

The presence of *N*-acetyl hexosamine at Thr292 was found to have a significant effect on H• transfer/abstraction to/from the adjacent z• radical ions. Figure 2.50 shows the comparison of the ETD spectra of the precursor ions of *m/z* 788.2 (2+) and *m/z* 686.7 (2+), corresponding to the peptide modified with HexNAc at Thr292, and to the non-modified peptide (289-302) respectively. The presence of HexNAc at Thr292 inhibits H• transfer to the radical ion z_9• (*m/z* 887.2) (cleavage between Ala293 and Ala294) compared to that observed in the ETD spectrum of the non-modified peptide.

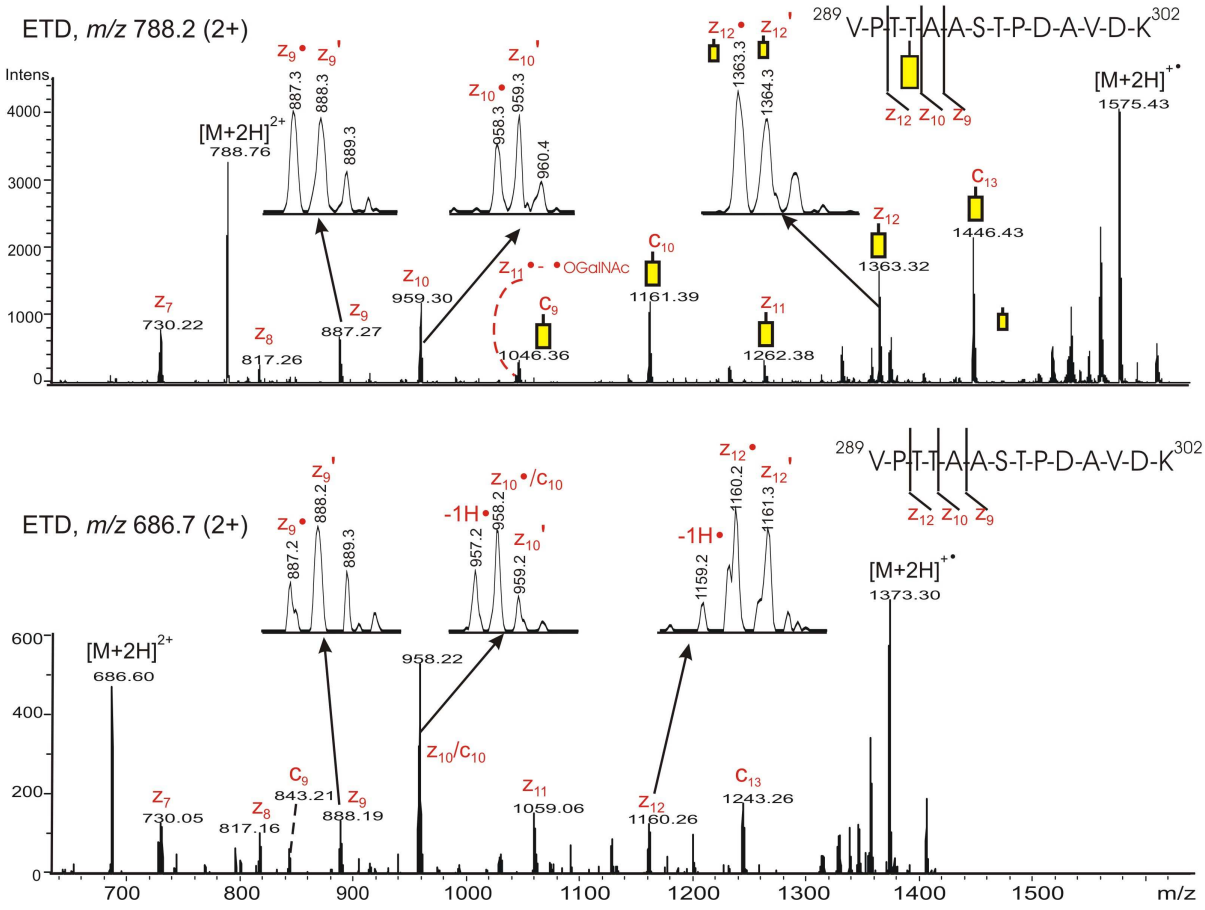


Figure 2.50: ETD of the precursor ions of m/z 788.2 (2+) (top) and 686.7 (2+) (bottom), corresponding to the peptide 289-302 in glycosylated and non-glycosylated form, respectively. The enlarged views of the radical species z_9, z_{10} and z_{12} show the distinct peak distribution of these ions in the glycosylated vs. the non-glycosylated form, derived from the attachment of O-GalNAc at Thr 292.

For the fragmentation leading to formation of the z_{10}• radical ion (*m/z* 958.2), by cleavage C-terminal to the glycosylation site, the H• transfer appears to be promoted when Thr292 is glycosylated. In addition, a species corresponding to the abstraction of one H• was observed at *m/z* 957.2 in the ETD spectrum of the non-glycosylated peptide; however, in the absence of the accurate mass, this species may be assigned as H• abstraction from

z_{10}• and/or c_{10}, as these species are isobaric (*m/z* 958.1). The major cleavage C-terminal to Pro290, in the non-glycosylated peptide ion, is the z_{12}• fragment (*m/z* 1160.2), but enhanced H• subtraction was also observed leading to the ion of *m/z* 1159.2. The latter process was not significant for the z_{12}• species in the ETD spectrum of the glycosylated peptide (*m/z* 1363.3).

Small neutral losses were observed from the diprotonated molecular radical cation $[M+2H]^{+\bullet}$, *m/z* 1576.4 (see Figure 2.49 B), with loss of water (18 Da) and •OH (17 Da) probably from the Ser/Thr side chain representing the most abundant processes. The loss of 43 Da may be explained either by loss of $CH_3HC•CH_3$ from the side chain of valine or by loss of $CH_3C•O$ from the *N*-acetyl galactosamine unit, while the losses of 44 and 45 Da may arise from neutral loss of CO_2 and radical loss $•CHO_2$, respectively, from the Asp side chain. Loss of 60 Da, giving rise to the ion *m/z* 1516.4, may be explained as simultaneous loss of 43 and 17 Da or as loss of CH_3COOH from the aspartic acid side chain. As a result of the limited mass accuracy of the ion trap, it was not possible to unambiguously assign the identity of these neutral losses. Such side chain losses have recently been reported in the ETD spectra of *N*-glycopeptides [235], and previously for the ECD of *O*-glycopeptides [267].

Radical losses of •O-GalNAc (219 Da) from the radical species $[M+2H]^{+\bullet}$ and z_{11}• were observed (see Figure 2.49 A and the bottom spectrum in the insert in Figure 2.49 A). Similar observations have been reported by Catalina, *et al*., who described the loss of the complete glycosylated Asn-side chain from a tryptic *N*-glycosylated peptide under ETD conditions [235], and by Mormann, *et al*., for *O*-glycosylated peptides under ECD conditions [267]. In ECD, side chain losses result from radical transfer from the initially formed C-terminal z• fragment, followed by elimination of radical species from the side chain and subsequent formation of a double bond between the C_α and $C_ß$ atoms of the residue [268]. A similar mechanism based on radical migration or •H transfer explains the loss of •O-GalNAc observed from the charge reduced molecular radical cation. In addition, neutral loss of GalNAc (203 Da) from both $[M+H]^+$ and $[M+2H]^{+\bullet}$ ions was observed (see insert in Figure 2.49 A).

Figure 2.51 shows the ETD spectrum of the precursor ion *m/z* 676.8 (3+), corresponding to the glycopeptide (289 – 302) modified at Thr292 with a Core 1 type glycan terminated by sialic acid (SA), having the structure: GalNAc – Gal – SA. The linear structure of this *O*-

glycan was derived from the CID spectrum of the precursor ion where the sugar oxonium ion, *m/z* 454.1 (Gal – SA^+), indicates that SA and Gal are interconnected. This glycoform was less abundant than that modified with *O*-GalNAc (see Figure 2.48).

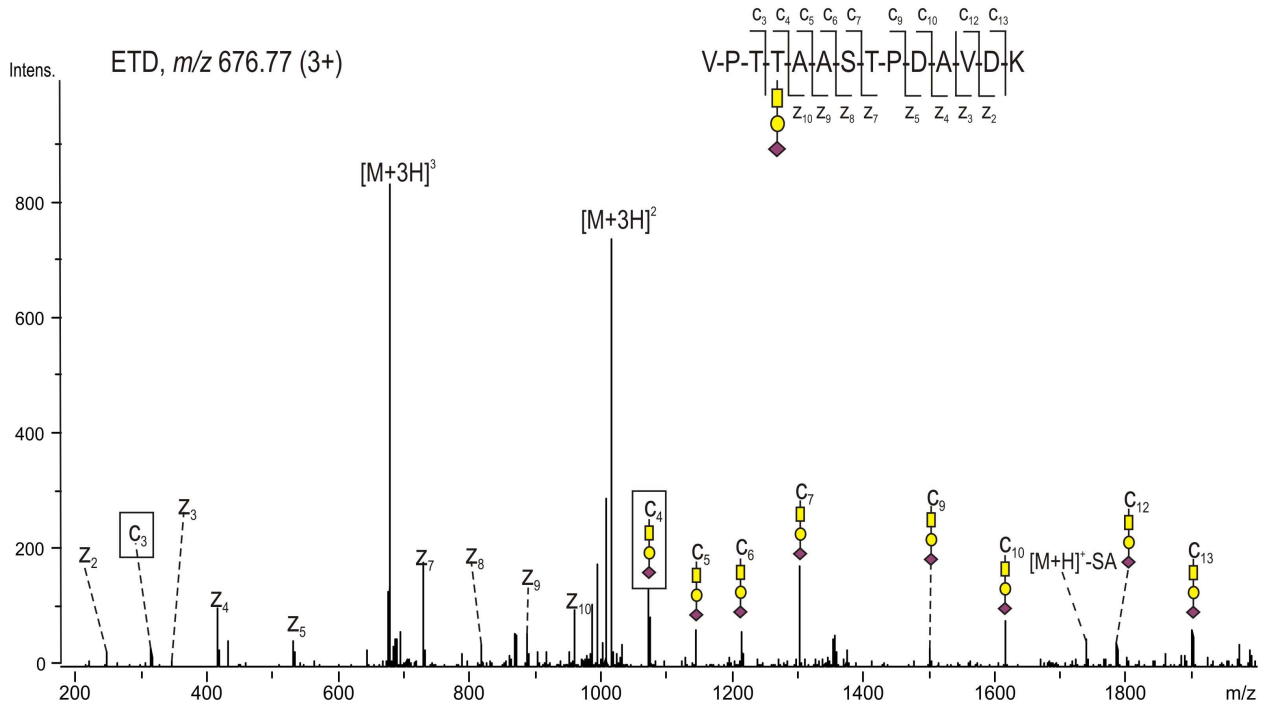


Figure 2.51: ETD of the precursor ion m/z 676.8 (3+), glycopeptide 289-302 of the full length APP695, showing the Core 1 type trisaccharide attached at Thr 292. The spectrum was obtained performing targeted MS/MS of the ion m/z 676.8 (3+), without supplemental ion activation. Colour code: yellow rectangle – N-acetyl galactosamine, yellow circle – galactose, purple – N-acetyl neuraminic acid. The fragment ions relevant for determination of the glycosylation site are indicated with black boxes.

As described above, neutral losses of 17, 43 and 60 Da were observed from the charge reduced odd-electron species *m/z* 1015.3, assigned as $[M+3H]^{2+\bullet}$. The series of c' and z• ions, in particular c_3 (*m/z* 315.0) and c_4 (*m/z* 1072.22), demonstrated that the *O*-glycosylation site is located at Thr292. The fragment ions containing the modified Thr 292 were found to retain the intact glycan, despite the instability of sialic acid containing sugars. Loss of sialic acid by glycosydic bond cleavage as a result of ETD was minimal and was observed as the ion of *m/z* 1737.5, corresponding to $[M+H]^+$-SA. Although the spectrum shown in Figure 2.51 was obtained only by ETD acquisition, pronounced loss of sialic acid was observed upon additional ion activation (typically 0.07 – 0.1 V for these experiments), which induced dissociation of the charge reduced odd-electron species indicating the labile character of this monosaccharide.

An additional glycoform of the peptide (289 – 302) containing GalNAc – Gal at Thr292 was found as a doubly protonated molecule of *m/z* 869.1 (ion c, Figure 2.48). The location of the disaccharide on this peptide was established using ETD. The abundances of this glycoform and of the sialic acid-terminated form were comparable and significantly lower than the *O*-GalNAc modified peptide (ion b, Figure 2.48).

Furthermore, an additional set of six different glycopeptides, with the amino acid sequence (289 – 302), were found glycosylated at both Thr291 and Thr292. Among these, the glycoform containing one GalNAc at each site was the most abundant, albeit its abundance includes both the doubly (*m/z* 889.7) and the triply protonated species (*m/z* 593.5). The composition of the glycan attached to the peptide was derived from the CID data as $GalNAc_2$, in which the base peak ion (*m/z* 788.2, 2+) arises from the loss of a single GalNAc unit. One could argue that the *N*-acetyl hexosamine units are interconnected and attached to a single side chain; however from these data no information upon the connectivity of the sugar moieties was obtained. The ETD spectrum of the ion of *m/z* 593.5 (3+) is shown in Figure 2.52. A nearly complete series of c' and z• ions were observed in addition to the small mass neutral losses from the charged reduced molecule ion. The fragment ions c_2 (*m/z* 214.0), c_3 (*m/z* 518.1) and c_4 (*m/z* 822.2), as well as z_{10}• (*m/z* 958.1) and z_{11}• (*m/z* 1262.1) indicate that Thr291 and Thr292 are each modified by GalNAc, because the mass difference between each of these corresponds to the mass of a threonine residue modified with *N*-acetyl hexosamine. In addition, the ion of *m/z* 1043.1 was assigned as arising from the loss of the •O-GalNAc radical (219 Da) from the z_{11}• radical ion (*m/z* 1262.1). In this spectrum, the loss of •O-GalNAc from the charge reduced radical species was minimal compared to that observed in the ETD spectrum of the precursor ion *m/z* 788.2 (2+) (Figure 2.49 A).

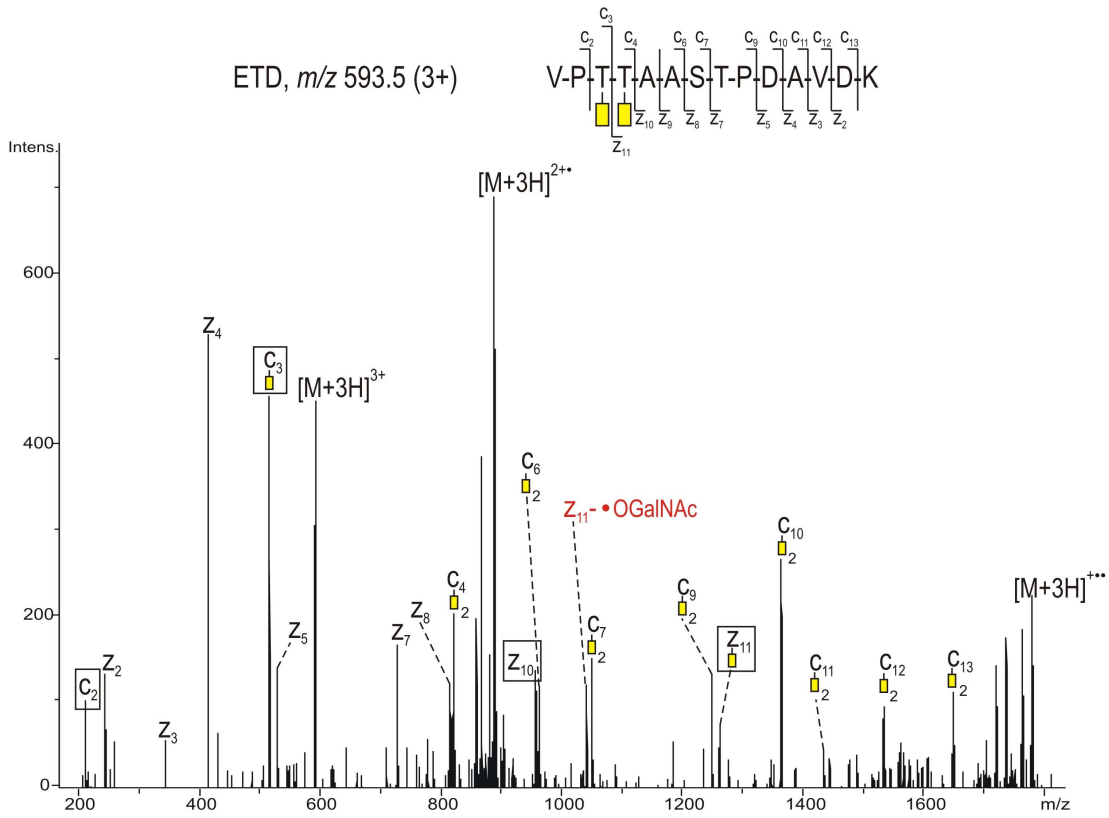


Figure 2.52: ETD spectrum of the precursor ion m/z 593.5 (3+), corresponding to glycopeptide 289-302 of the full length APP695, showing each of the amino acids Thr291 and Thr292 occupied with N-acetyl galactosamine (yellow rectangle). The radical loss of 219 Da (•OGalNAc) from the z_{11} ion is indicated in red. The spectrum was obtained using data dependent acquisition without supplemental ion activation. The fragment ions relevant for determination of the glycosylation site(s) are indicated with black boxes.

The CID and ETD mass spectra of the triply protonated precursor ion *m/z* 744.4, assigned to the glycopeptide (289 – 302) are presented in Figure 2.53. The fragment ions observed in the CID spectrum (Figure 2.53, spectrum at the top) are largely derived from processing of the glycan from the non-reducing end; from this, the composition of the glycan was derived as $HexNAc_2Hex_1SA_1$, which may be assigned to either a Core 2 type glycan, containing GalNAc, GlcNAc, Gal and SA [265], or to two distinct glycans, $GalNAc_1$, and $GalNAc_1Gal_1SA_1$, respectively. The loss of HexNAc from the precursor ion with retention of the remaining monosaccharides indicates that the HexNAc is terminally linked, which may result from decomposition of either a Core 2 type glycan or a glycan containing a single GalNAc.

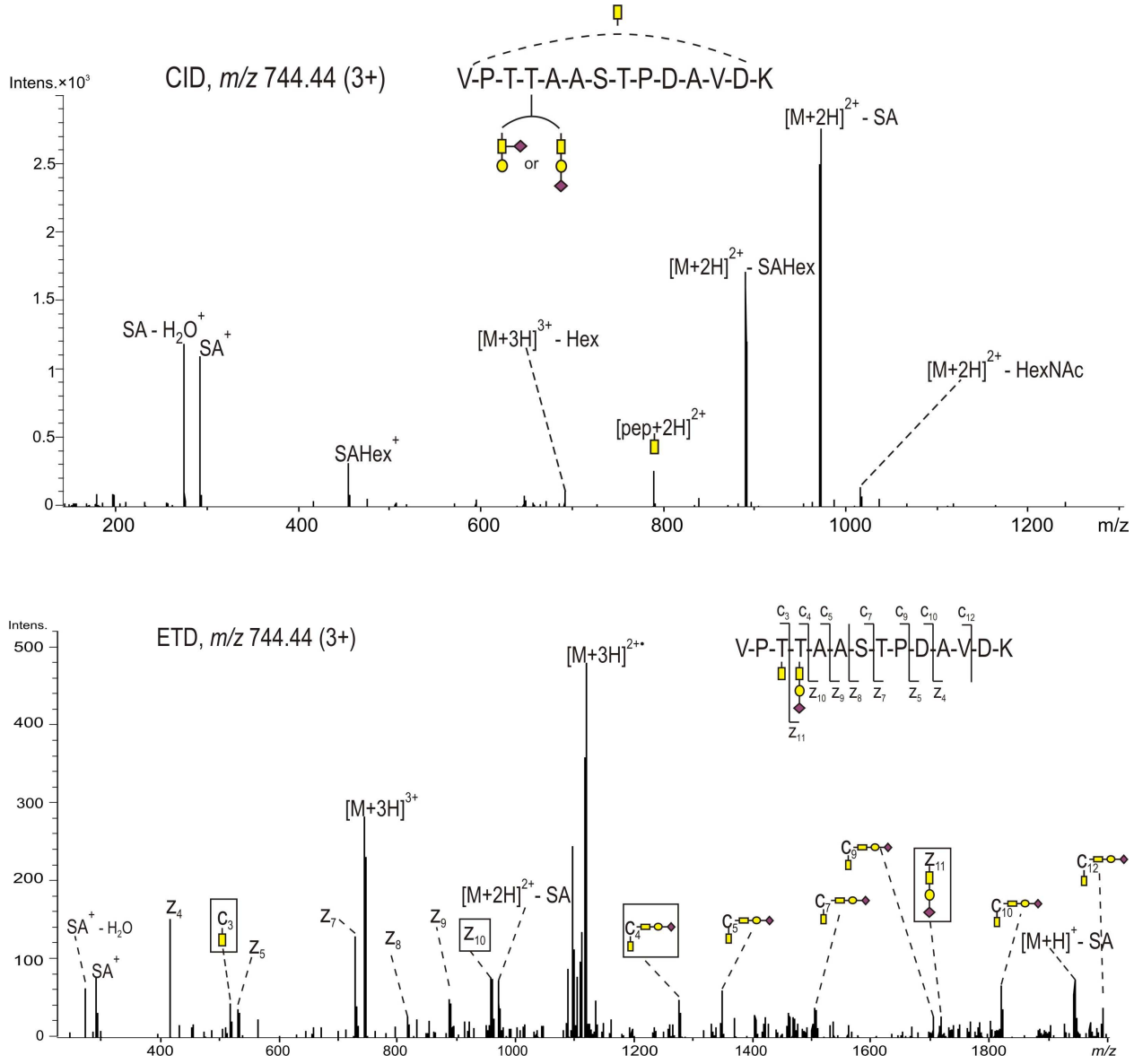


Figure 2.53: CID (top), and ETD spectrum (bottom) of the precursor ion m/z 744.44 (3+), corresponding to glycopeptide 289-302 of the full length APP695, showing each of the amino acids Thr291 and Thr292 occupied with N-acetyl galactosamine and a Core 1 type trisaccharide, respectively. The spectrum was obtained using data dependent acquisition with a supplemental ion activation of 0.07 V. Colour code: yellow rectangle – N-acetyl galactosamine, yellow circle – galactose, and purple diamond – N-acetyl neuraminic acid. The fragment ions relevant for the determination of the glycosylation site(s) are indicated with black boxes.

In addition, the loss of terminal sialic acid and of hexose from the precursor ion indicate that these moieties represent non-reducing ends. This rules out the Core 2 type glycan, because elongation of the GalNAc with three branches has not been reported in vertebrates. This might also indicate a branched Core 1 type glycan, possibly GalNAc (SA) – Gal. However, the oxonium ion *m/z* 454.1 ($SAHex^+$) indicates a linear structure for the glycan GalNAc – Gal – SA. The CID data show that the two residues in the peptide (289 – 302) are specifically modified by the *O*-glycans GalNAc and a mixture of the isomers

GalNAc (SA) – Gal and GalNAc – Gal – SA. The glycosylated residues in this sequence (289 – 302) were unambiguously identified from the ETD spectrum (Figure 2.53, spectrum at the bottom). The ion c_3 (*m/z* 518.1) indicates that the GalNAc unit is attached at Thr291, while c_4 (*m/z* 1275.3) and z_{11}• (*m/z* 1715.3) indicate that Thr292 is occupied with the glycan $GalNAc_1Gal_1SA_1$. The ions z_7• (*m/z* 729.1) and z_8• (*m/z* 816.2) suggest that Ser295 and Thr296 are not modified. The loss of terminal sialic acid was also observed (Figure 2.53, spectrum at the bottom) as a result of the additional ion activation (0.07 V) applied to improve the efficiency of ETD.

In summary, the peptide ^{289}VPTTAASTPDAVDK^{302} was found to contain four potential *O*-glycosylation sites, Thr291, Thr292, Ser295 and Thr296, of which only Thr292 was observed to be *O*-glycosylated in all glycopeptides identified. In addition, Thr291 and 292 were both found modified by multiple short Core 1 type glycans. The minimal *O*-glycan decorating these sites was *N*-acetyl galactosamine (GalNAc), while elongated structures contain the GalNAc – Gal core terminated with sialic acid, attached either in a linear fashion to galactose, or branched with attachment to GalNAc. The observed molecular weights and the retention time of each glycoform are summarized in Table 7.

Table 7: Composition of the Core 1 glycans determined for each glycosylation site in glycopeptide 289-302

Glyco-form	Experimental MW (Da)	Theoretical MW (Da)	Retention time (min)	Occupied sites observed in peptide 289 - 302	
				Thr 291	Thr 292
a	1371.3	1371.69	20.8	-	-
b	1574.3	1574.77	20.6	-	GalNAc
c	1736.2	1736.82	20.4	-	GalNAcGal
d	1777.5	1777.85	20.4	GalNAc	GalNAc
e	1939.4	1939.90	19.8	GalNAc	GalNAcGal
f	2027.3	2027.92	20.7	-	GalNAcGalSA
g	2101.4	2101.95	19.5	GalNAcGal	GalNAcGal
h	2230.3	2230.99	20.3	GalNAc	GalNAcGalSA
i	2392.3	2393.05	20.1	GalNAcGal	GalNAcGalSA
j	2683.3	2684.14	20.2	GalNAcGalSA	GalNAcGalSA

Peptides containing both glycosylated sites Thr291 and Thr292 were found to elute earlier than those containing a single glycosylation (Table 7). Within each glycopeptide form, the sialylated glycopeptides were found with slightly longer retention times than the non-sialylated glycoforms. The distinct chromatographic retention times of each glycoform (see

Table 7) indicates that the observed heterogeneity results primarily from the biochemistry of *O*-glycosylation and to a lesser extent from in source decomposition of the glycan chain.

A further *O*-glycosylation site was identified in the APP sequence (574-587) (^{574}GLTTRPGSGLTNIK^{587}), containing three threonine residues and one serine residue as potential glycosylation sites. The ETD spectrum of the unmodified peptide (Figure 2.54 A) contains a nearly complete z• ion series, identifying the amino acid sequence. The CID and ETD spectra of the corresponding precursor ion of *m/z* 690.8 (3+), eluting at 23.2 min, are shown in Figure 2.54 B and C, respectively. The composition of the glycan was identified as a Core 1 type trisaccharide GalNAc (SA) – Gal, based on the mass difference between the unmodified and the glycosylated peptide, and the glycoforms identified in the CID spectrum (Figure 2.54 B). The presence of sialic acid is clearly shown by the sugar oxonium ions *m/z* 292.1 (SA^+) and *m/z* 274.1 ($SA - H_2O^+$). Primary losses of hexose and sialic acid from the parent ion indicate a branched structure for the glycan: GalNAc (SA) – Gal. Complete decomposition of the glycan resulted in the formation of the doubly protonated peptide of *m/z* 707.7, which is consistent with the amino acid sequence (574-587) for the *O*-glycopeptide.

The amino acid sequence of the peptide was confirmed and the location of the modification site, Thr576, was identified on the basis of the ETD spectrum of the precursor ion of *m/z* 690.8 (Figure 2.54 C). The observed z• ions (z_2 - *m/z* 244.0, z_3 - *m/z* 358.1, z_4 - *m/z* 459.1, z_5 - *m/z* 572.1, z_6 - *m/z* 629.2, z_7 - *m/z* 716.2, z_{11} - *m/z* 1127.3,) are identical to the z• ions observed in the spectrum of the non-modified peptide, indicating that the residues Thr584, Ser581 and Thr577 are not modified.

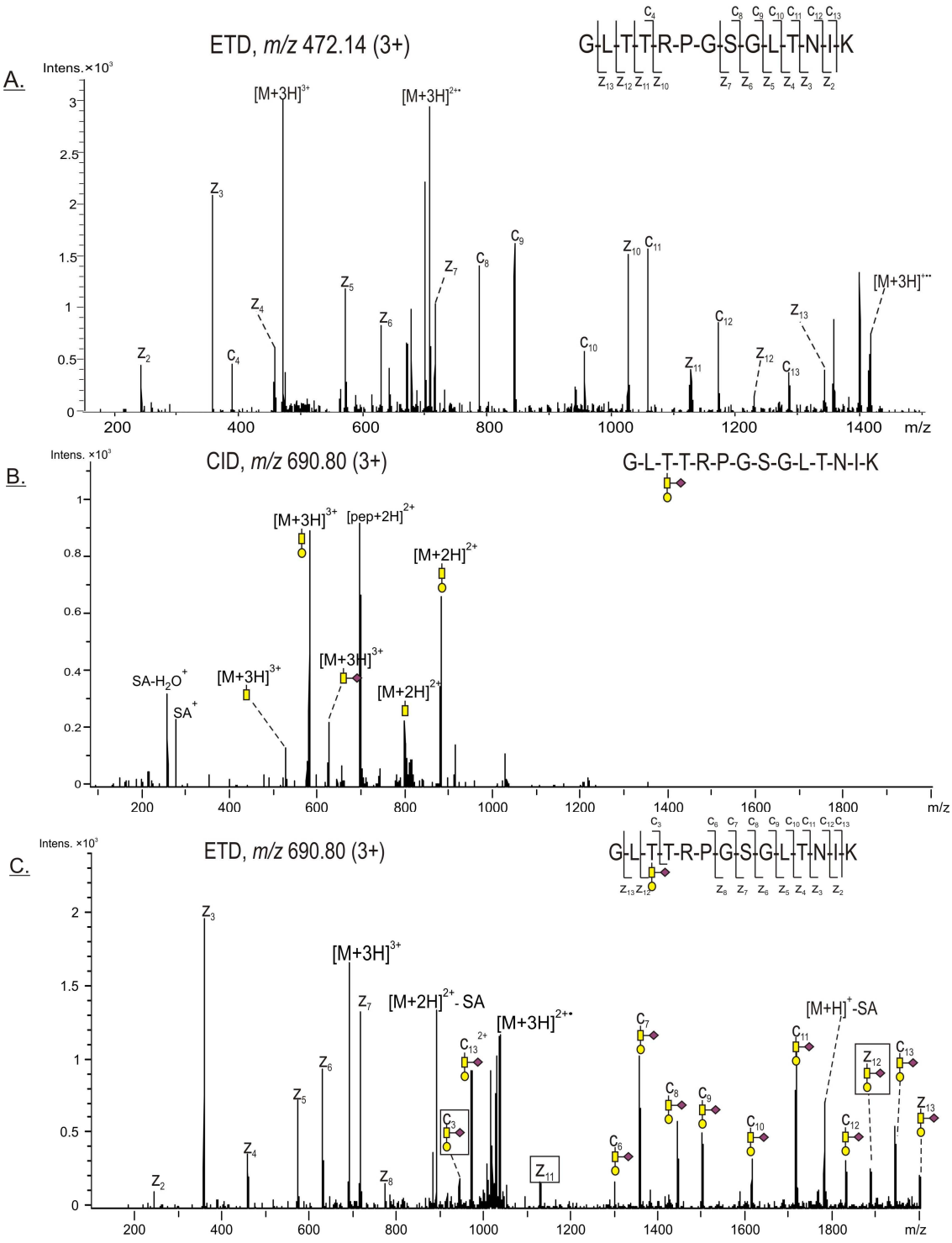


Figure 2.54: (A) ETD spectrum of the precursor ion m/z 472.14 (3+), showing the non-modified peptide 574-587 of the full length APP695, containing four potential O-glycosylation sites; (B) CID, and (C) ETD of the precursor ion m/z 690.80 (3+) corresponding to glycopeptide 574-587 of the full length APP695, showing the amino acid Thr576 occupied with the indicated Core 1 type trisaccharide. The ETD spectrum was obtained using data dependent acquisition and the activation energy was 0.10 V. Colour code: yellow rectangle – N-acetyl galactosamine, yellow circle – galactose, purple diamond – N-acetyl neuraminic acid.

The remaining candidate for *O*-glycosylation, Thr576, was confirmed as glycosylated by the observation of the ions c_3 (*m/z* 945.3), z_{11}• (*m/z* 1127.4) and z_{12}• (*m/z* 1884.6), of which z_{11}• and z_{12}• are separated by the incremental mass corresponding to the

glycosylated threonine residue. Consequently, the c ions indicated in Figure 2.54 C have a mass shift of 656.1 compared to those c ions observed in the ETD spectrum of the non-modified peptide. Pronounced loss of sialic acid from the precursor ion was observed, which may have been induced by the additional CID voltage (0.10 V) employed to increase the fragmentation efficiency of the charge reduced molecule ions. However, no such neutral loss from any of the fragment ions was observed.

The heterogeneity of the oligosaccharides attached at Thr 576 was significantly lower than that observed for the glycopeptide (289 – 302), with only two glycoforms observed at this residue: GalNAc (SA) – Gal (the most abundant glycoform) and GalNAc (SA) – Gal – SA (observed as the triply protonated molecule of *m/z* 787.7). The abundance of the non-glycosylated peptide was considerably larger compared to those of the two glycoforms. Interestingly, the residue Thr576 is only 21 amino acid residues away from the N-terminus of the Aß-sequence (Asp597), suggesting that this glycosylation may play a role in the ß-secretase proteolytic processing of the amyloid precursor protein.

The detailed structures of the two different types of Core I *O*-glycans observed at Thr291, Thr292 and Thr576, respectively, including the linear (Thr291 and Thr292) and the branched attachment of the sialic acid (Thr576), are presented in Figure 2.55.

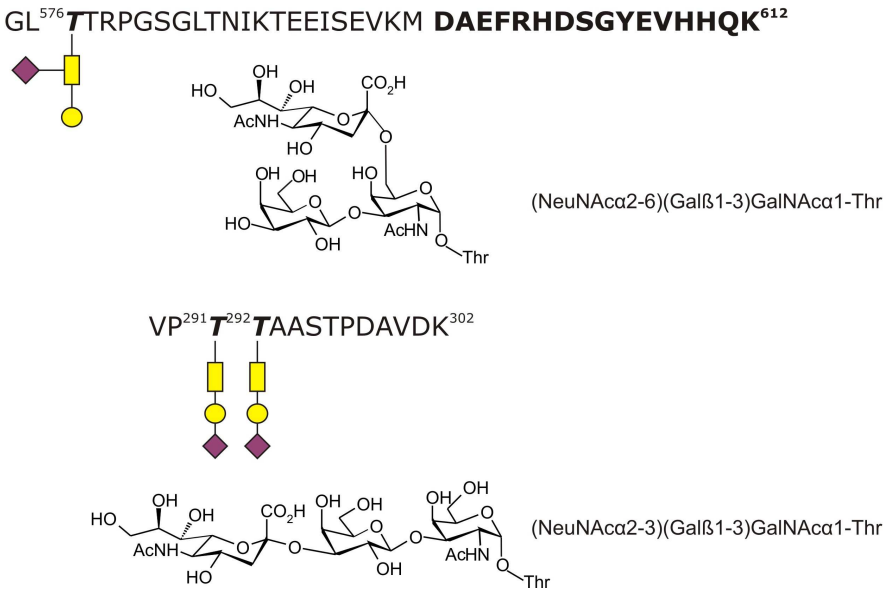


Figure 2.55: Structures of the branched and linear Core I O-glycans attached at Thr576, and Thr291 and Thr292, respectively, including the type of linkage between individual monosaccharides. Colour code: yellow rectangle – N-acetyl galactosamine, yellow circle – galactose, purple rhombus – sialic acid.

2.3 Clinical applications of mass spectrometry to antibody biomarker discovery in myositis patients

2.3.1 Clinical background for investigation of antibody glycosylation in myositis

Current evidence suggests that autoimmune diseases are triggered by chronic immune activation and dysregulation after specific environmental exposures in genetically susceptible individuals [38]. Several systemic rheumatic disorders (SRDs), comprising Rheumatoid Arthritis (RA), Systemic Lupus Erythematosus (SLE), Systemic Sclerosis (SSc) and Idiopathic Inflammatory Myopathies (IIM) share many common clinical manifestations, immune responses, genetic, hormonal and environmental risk factors. A large clinical study initiated at the National Institute of Environmental Health Sciences attempts to explore the pathogenic mechanisms for SRDs by evaluating families with monozygotic or dizygotic twins or other siblings discordant for systemic rheumatic disorders (twin-sibling pairs). This study evaluates children and adults in an attempt to understand possible similarities and differences in pathogeneses of systemic rheumatic disorders based upon the age of onset. Participants in this study are individuals diagnosed with systemic rheumatic disorders (RA, Lupus, Scleroderma, Myositis) within 47 months who have an unaffected twin or a same sex sibling within 47 months of age, as well as sex and age-matched normal volunteers. Initial evaluation and annual follow-up for 5 years is performed, with re-evaluation if new autoimmune diseases develop [269].

The proteins secreted by the humoral immune system, the antibodies, have been found to exhibit characteristic structural changes in diseases such as cancer, tuberculosis, HIV infection, Sjøgren's syndrome and rheumatoid arthritis [243, 270-273]. In general, the distribution of the IgG glycoforms is reproducible among individuals. However, in diseased persons the IgG constant region glycosylation pattern shows decreased levels of galactosylation, with the agalactosyl glycoforms (IgG-G_0F, G_0FB and G_0) representing the dominant carbohydrate structures. In RA, the high incidence of glycoforms lacking terminal galactose is thought to trigger specific anti-IgG immune responses, leading to formation of pathogenic antibody-antibody complexes which are deposited in synovial fluid and joints leading to inflammation [243, 272].

The hypothesis for the existence of a common IgG glycosylation pattern in autoimmune diseases has led us to investigate the IgG glycosylation in patients diagnosed with idiopathic inflammatory myopathies. In an effort to increase the understanding of the pathogenesis of autoimmune disorders, the glycosylation profiles of IgG antibodies isolated from myositis patients were investigated. A total of 30 plasma samples were analyzed, grouped into 10 different sets, each set containing three samples collected from patient, unaffected twin/sibling and non-related age-matched control, respectively. An overview of the plasma samples, the age, race and gender of the subjects and the corresponding disease subtype is presented in Table 8 (Chapter 3.3). Subclass specific glycosylation profiling of the total plasma IgG fraction was performed, followed by statistical analysis of the results.

2.3.2 Isolation of the immunoglobulin fraction from human plasma

The total plasma IgG fraction from adult patients (n=10), their healthy twin/siblings (n=10) and healthy, unrelated age-matched controls (n=10) was isolated from 20 µL plasma using Protein G affinity capture. Protein G and Protein A are cell-surface immunoglobulin-binding proteins that have found application in purifying antibodies from mammalian biological fluids through binding to the IgG Fc region. In addition to two IgG binding domains, Protein G contains sites for albumin and cell surface binding. The affinity of various IgG subclasses, especially from mouse and human, for Protein A, varies more than for Protein G. Unlike Protein A, which selectively binds the human IgG_3 subclass, Protein G captures all four IgG subclasses, $IgG_{1,\ 2,\ 3}$ and $_4$. Hence, Protein G was chosen to isolate all four IgG subclasses from plasma in a single step. The commercially available purification kit supplied by KPL consists of recombinant Protein G expressed in *E. coli*, covalently immobilized onto 4% cross-linked agarose beads. In order to reduce non-specific binding, the albumin and cell surface binding activity was eliminated in the manufacturing process. According to the product data sheet, the binding capacity of 1 mL of drained Protein G – agarose is > 20 mg/mL human IgG.

For IgG purification, a 0.5 mL compact reaction column was packed with approximately 400 µL of slurry and allowed to flow by gravity, resulting in approximately 200 µL drained gel. After equilibration with the washing/binding buffer (pH 7.4) supplied with the kit, 20 µL of plasma were applied on the column and incubated for 45 minutes at room temperature

on a nutator mixer to ensure slow and uniform mixing of the column content. The unbound proteins were removed by passing 5 mL of washing/binding buffer through the column, followed by 5 mL of deionized water. The affinity bound antibody was dissociated with 2 × 0.5 mL glycine buffer (pH 2.85) and the eluting fractions were collected separately. The purity of the IgG was assessed by SDS-PAGE. The washing with deionized water prior to elution was found efficient for removing the salt from the washing/binding buffer, which may cause inefficient dissociation of the Protein G-antibody complex. In order to avoid potential cross-contaminations, separate Protein G columns were prepared for each individual plasma sample.

2.3.3 Subclass specific glycosylation profiling of total plasma IgG in diseased and healthy individuals

Polyclonal IgG purified from human plasma was denatured with DTT at 95° C and separated into heavy and light chains by SDS-PAGE. The heavy chain band was excised and digested with trypsin and the resulting enzymatic mixture was analyzed by LC-MS/MS on the Q-Tof Premier. The structural details of the constant region of each IgG subclass, in particular around the *N*-linked glycosylation site, enabled subclass specific glycosylation analysis by LC-MS. These features were described in Chapter 2.2.5. Selective identification of glycopeptides in the digestion mixture was achieved by monitoring the extracted ion current (EIC) for the ion *m/z* 204.1, corresponding to protonated *N*-acetyl glucosamine, as described previously. A representative EIC of the ion *m/z* 204.1, from the patient sample 016-01 in Set 1, is shown in Figure 2.56. All tryptic glycopeptides were observed in fully processed form, hence all four IgG subclasses were detected in LC-MS. For each detected peptide isoforms, EEQYNSTYR for IgG_1, EEQFNSTYR for IgG_4 and EEQFNSTFR for $IgG_{2/3}$, MS/MS data were acquired to determine the glycan structures and peptide identities. From these experiments, 11 different glycoforms were determined (see the structures at the bottom of Figure 2.57). Additionally, the relative abundances of each glycoform in patient, healthy twin/sibling and age-matched control were analyzed. The most abundant glycoforms were primarily core-fucosylated biantennary complex structures, containing zero to two galactose residues, zero to one bisecting *N*-acetyl glucosamine and zero to one *N*-acetyl neuraminic acid moieties.

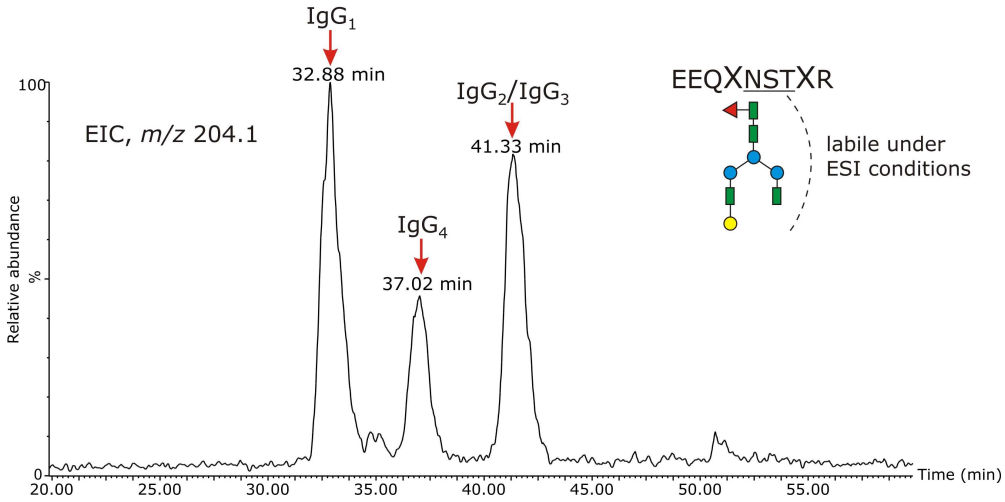


Figure 2.56: Representative extracted ion current (EIC) of the ion m/z 204.1 over the chromatographic time window 20-60 minutes, in the LC-MS analysis of patient sample 016-01 in Set 1. IgG subclasses can be detected based on structural characteristics of the peptide backbone, resulting in distinct chromatographic elution times, as described for the Aß-autoantibodies.

Following the MS/MS analyses, triplicate LC-MS acquisitions were performed for each of the 30 plasma samples and the abundance of each glycoform within each subclass, relative to the abundance of the glycoform G_1F, was determined. For semiquantitative analyses, the EIC of *m/z* 204.1 (e.g. Figure 2.56) indicates the individual chromatographic retention time of the glycopeptide groups from each IgG subclass. The relative abundances (peak heights) in the EIC are proportional to the total ion counts of protonated *N*-acetyl glucosamine, *m/z* 204.1, which in turn is related to the total ion abundances of glycopeptides containing this residue. Because glycans are labile under ESI conditions, a certain degree of in-source decomposition of the oligosaccharide moiety can occur, and therefore, this must be minimized for semiquantitative analysis. However, the sugar oxonium ion aids to the selective detection of glycopeptides in the enzymatic mixture. Another important experimental detail, relevant for relative quantitation, is related to the operation of the instrument in the MS only scanning mode. In contrast to data dependent analysis, employing an MS scan and four subsequent MS/MS cycles, in the MS only mode the quadrupole serves exclusively as a transmission element, allowing all the ions at every point in time to pass and be detected. Hence, the time reserved for MS/MS in the data dependent analysis is efficiently replaced by MS scans in the MS only mode, minimizing the amount of information lost as a result of the duty cycle.

The global picture of the glycoforms present on one specific antibody subclass was obtained by averaging the MS scans over the chromatographic retention time in which

these glycopeptides eluted. For example, the glycosylation pattern of IgG_1 was determined by summing the MS scans corresponding to the peak centered at 32.88 minutes in Figure 2.56. All IgG_1 glycopeptides containing the same peptide (EEQYNSTYR) and various glycans were observed as both doubly and triply protonated molecules. The ion counts for each charge state of each observed glycoform were summed and the numbers were divided by the value of the ion counts of the glycoform G_1F from the IgG_1 subclass. Consequently, the relative abundance value of G_1F is always 1. From the triplicate LC-MS analyses, the average value for each glycoform and the standard deviation were determined. Similarly, the glycosylation pattern of the remaining IgG subclasses was determined.

One example of subclass specific glycosylation profiling from the sample set 4 (see Table 8 in Chapter 3.3) is presented with barographs in Figure 2.57. The yellow bars represent the diseased person (diagnosed with dermatomyositis), the red and the purple bars represent the sibling and the unrelated control, respectively. Dramatic changes of the antibody galactosylation were observed for the IgG_1 subclass (see Figure 2.57A). The levels of IgG_1-G_0F in the patient are elevated compared to G_1F, while the healthy sibling and control have lower relative levels of G_0F. Compared to G_0F of the unrelated control (purple bar in Figure 2.57 A), the patients' IgG_1-G_0F has more than double relative abundance. The relative abundance of IgG_1-G_0F for the unaffected sibling lies between the values determined for the unrelated control and the patient, suggesting that this phenotype may be the result of certain genetic factors and the person may be predisposed to develop an autoimmune condition. The relative abundance of the glycoform IgG_1-G_2F is decreased in patient compared to sibling and control, however this pattern is not as pronounced as for IgG_1-G_0F. A similar trend is observed for the glycoform G_0FB, bearing bisecting GlcNAc. (glycoform 6 in Figure 2.57 A). Within the subclasses $IgG_{2,3}$ and IgG_4, the abundance of G_0F is almost equal (patient IgG_4-G_0F) or higher than that of G_1F for all subjects in this sample set (Figure 2.57 B and C), however, no trend towards increased agalactosylation in patient vs. sibling and control was observed, as highest values $IgG_{2,3}$-G_0F and IgG_4-G_0F were determined for the sibling. These results indicate that the levels of galactosylation may be different within each IgG subclass in any one individual. In this example, the diseased patient shows increased levels of IgG_1-G_0F.

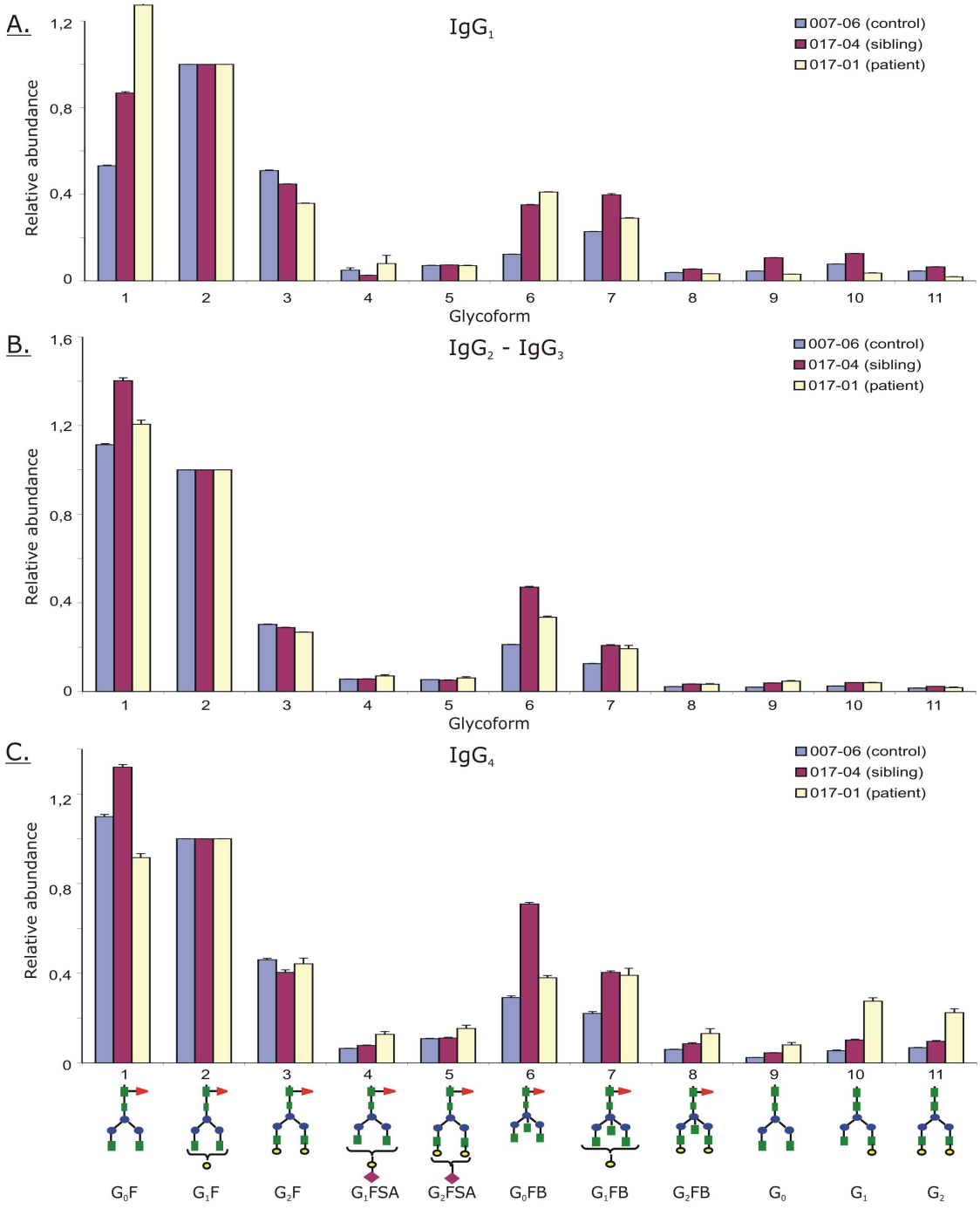


Figure 2.57: Subclass specific glycosylation profiling of IgG fractions isolated from plasma samples contained in set 4: (A.) IgG_1, (B.) $IgG_{2,3}$ and (C.) IgG_4. Colour code: yellow – dermatomyositis patient, red – healthy sibling, purple – unrelated age-matched control. The analyzed glycoforms are indicated at the bottom of the figure. Within individual subclasses, the abundance of each glycoform was normalized to the abundance of the glycoform G_0F. Triplicate LC-MS analyses were acquired; the average values from the three experiments and the standard deviation were calculated and are indicated with the error bars.

Subclass specific glycosylation analysis of sample set 1, collected from monozygotic twins discordant for dermatomyositis, and unrelated control is presented in Figure 2.58.

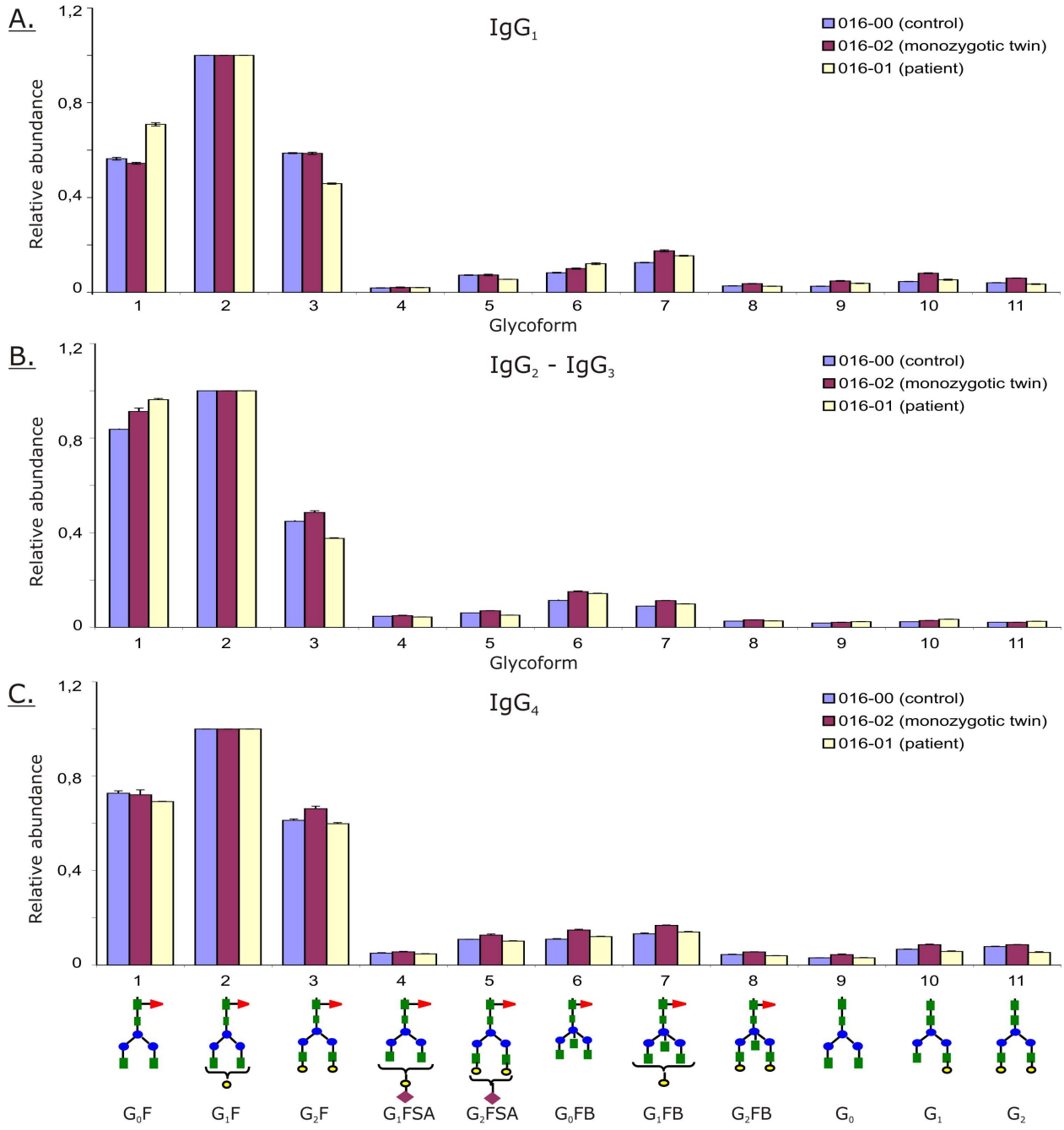


Figure 2.58: Subclass specific glycosylation profiling of IgG fractions isolated from plasma samples contained in set 1: (A.) IgG_1, (B.) $IgG_{2,3}$ and (C.) IgG_4. Colour code: yellow – dermatomyositis patient, red – healthy monozygotic twin, purple – unrelated age-matched control. The analyzed glycoforms are indicated at the bottom of the figure. Within individual subclasses, the abundance of each glycoform was normalized to the abundance of the glycoform G_0F. Triplicate LC-MS analyses were acquired; the average values from the three experiments and the standard deviation were calculated and are indicated with the error bars.

The investigation of monozygotic twins discordant for systemic rheumatic disorders allows the assessment of specific environmental risk factors in the development of autoimmune conditions. The incidence of idiopathic inflammatory myopathies and their low incidence among monozygotic twins complicate the study, i.e. the probability to find a statistically significant number of such twin pairs, is reduced. However, for this one sample set of monozygotic twins, significant changes in IgG_1 galactosylation were determined for patient in comparison to his healthy twin (Figure 2.58 A). This suggests that the increase in patients' IgG_1-G_0F and the decrease of his IgG_1-G_2F are the result of an acquired condition, possibly derived from the action of the environment on his genes. Monozygotic twins have identical genome and hence their phenotype should be identical as well. Presumably, certain environmental exposures of the diseased twin triggered changes in the proteome, which in turn may reflect alterations at genetic level. An identical IgG_1 glycosylation pattern was observed for the healthy twin and the non-related control, reinforcing the idea that the immunologic changes in patients' IgG_1 galactosylation are environmentally triggered.

Glycosylation analysis of the remaining subclasses in the monozygotic twin set shows a slight tendency towards increased agalactosylation in patients' $IgG_{2,3}$ (Figure 2.58 B) and no significant changes among IgG_4 (Figure 2.58 C). The results obtained for sample set 1 indicate that specific galactosylation differences are more pronounced for IgG_1. This can be explained by the fact that IgG_1 has the highest concentration in human plasma compared to IgG_2, IgG_3 and IgG_4; consequently, IgG_1 will represent the substrate affected primarily by the enzymes β-galactosyl transferase and β-galactosydase.

The detailed structure of the G_0F glycoform including the linkages of the individual monosaccharides building blocks is presented in Figure 2.59.

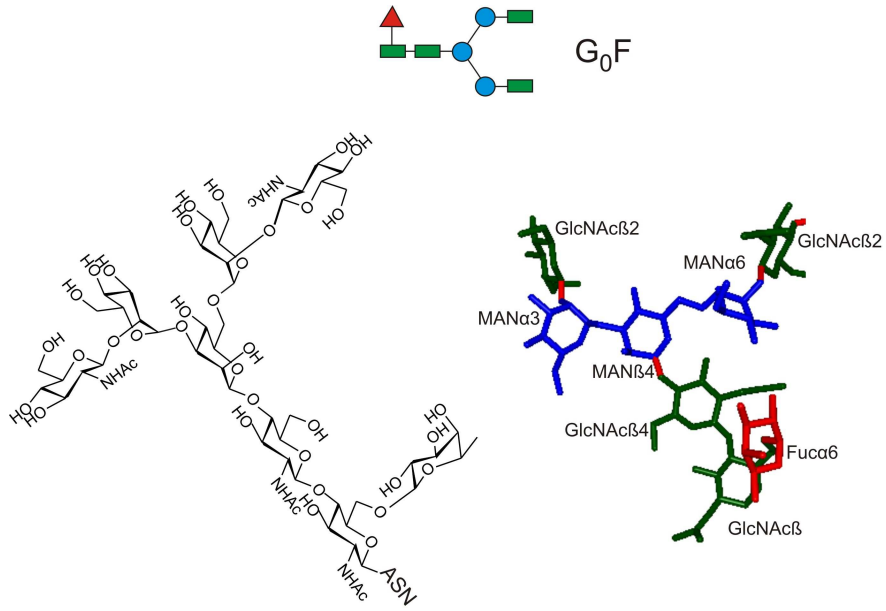
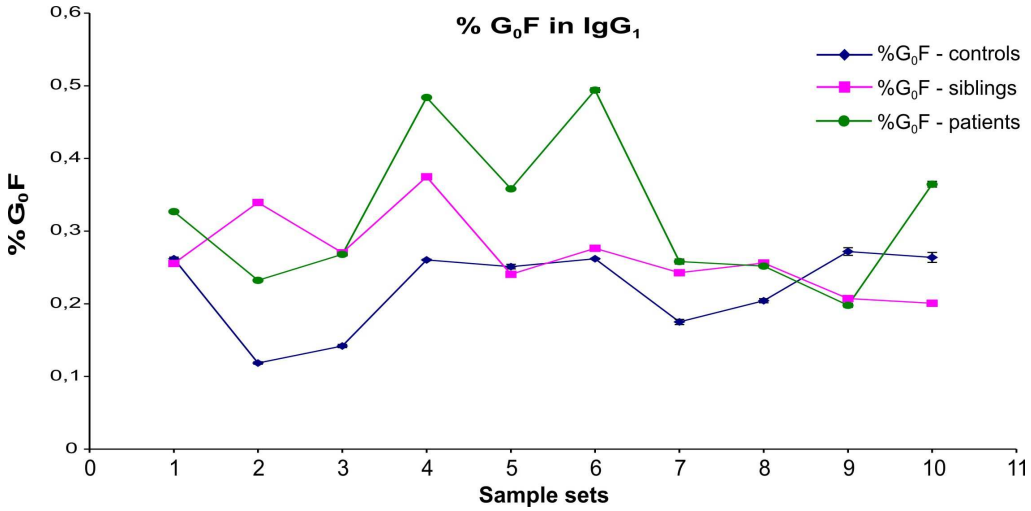


Figure 2.59: Detailed structure of the G_0F glycan attached at the conserved N-linked site of heavy chain constant region of the IgGs isolated from plasma of myositis patients. The linkages of the individual monosaccharides in the chair conformation are shown at the left. The stick model of the G_0F glycoform (right) was prepared with the software BallView 1.1, based on the G_0F glycoform in the antibody with the PDB accession number 1HZH. Colour code: green rectangle – N-acetyl glucosamine, red triangle – fucose, blue circle – mannose.

2.3.4 Statistical analysis of antibody glycosylation

The glycosylation profiling was determined for each IgG subclass of the 30 plasma samples, as described in the previous section. The three most abundant glycoforms observed were G_0F, G_1F and G_2F, such that subclass specific changes in galactosylation were analyzed as ratio of G_0F relative to the sum $G_0F+G_1F+G_2F$. These numbers, representing the percentages of G_0F (% G_0F) for each subclass of an individual, can be obtained by dividing either the ion counts or the relative abundance of G_0F to the sum $G_0F+G_1F+G_2F$ of the corresponding values. Following these calculations, the determined % G_0F were represented for each of the 10 plasma sets, for IgG_1 (Figure 2.60 A) and for $IgG_{2,3}$ (Figure 2.60 B), respectively. The IgG_4 subclass was not analyzed, because in several instances this subclass was barely observed in the LC-MS analyses and hence the ion abundances for the low abundant glycoforms were affected by a greater experimental error than the other subclasses. In Figure 2.60, the green, pink and blue graphs represent the % G_0F in patients, twins/siblings, and unrelated controls, respectively. Statistical analysis of these results was performed in order to determine whether the three groups of % G_0F data – patients, siblings and controls – represent statistically different events.

A.



B.

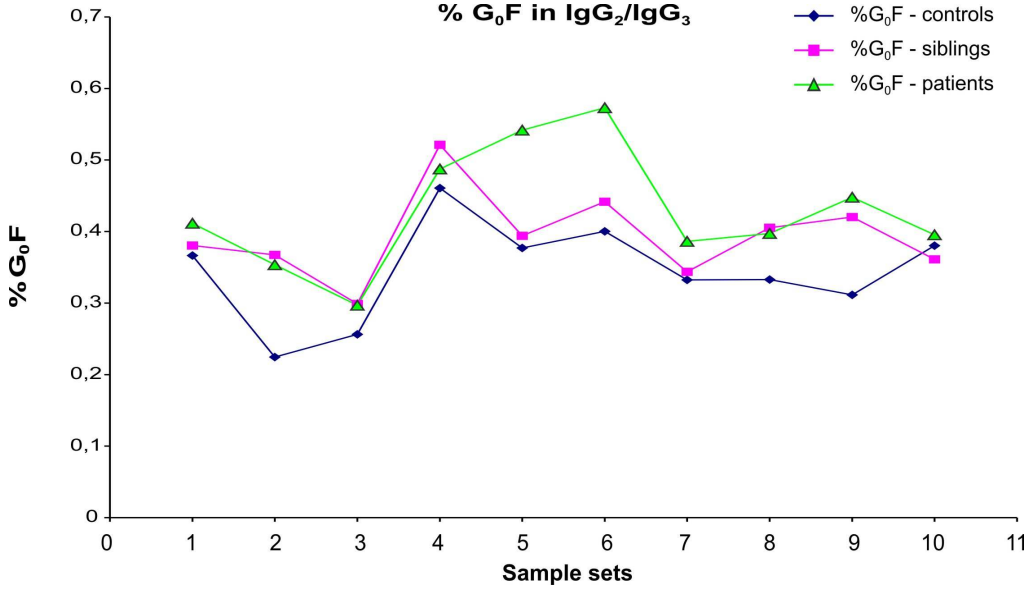


Figure 2.60: Subclass specific representation of % G_0F for each individual in each of the 10 sample sets: (A) for IgG_1, and (B) for $IgG_{2,3}$. The value % G_0F was calculated for each IgG subclass and for each individual as the ratio of the relative abundance of G_0F to the sum (G_0F+G_1F+G_2F). Colour code: green – patients, pink – siblings, and blue – unrelated controls.

The Mann-Whitney U test represents a non-parametric significance test for assessing whether two samples of observables come from the same distribution [274, 275]. The null hypothesis is that the two samples are drawn from a single population, and, therefore, their probability distributions are equal. The prerequisite is that the observables in the two samples are derived from continuous measurements so that the conclusion can be drawn, of any two sets of observations, which is greater. This two-sample test can be thought of as testing the null hypothesis that the probability of the observables from one population exceeding the observables from the second population is equal to 0.5. The test involves calculation of a statistic, termed U, whose distribution under the null hypothesis is known.

In the case of small sample sets, such as the present clinical study (n=10), the distribution is tabulated, while for larger samples a good approximation is the normal distribution. The U statistic is calculated by *ranking* the data contained in the two data sets to be compared, disregarding their raw values. In the case of the twin-sibling study, the % G_0F data sets obtained for IgG_1 and $IgG_{2,3}$ subclasses were assessed two-by-two, i.e. patients vs. siblings, patients vs. unrelated controls, and siblings vs. controls, such that three independent tests were performed. The determined U values were compared to the table of critical values for U, based on the sample size in each group. If U exceeds the critical value for U at a significance level of 0.05, then the null hypothesis can be rejected, i.e. the two sample sets are coming from distinct events. The U test is included in most modern statistical packages. For small sample sets, it can be easily calculated by hand. Statistical analysis of glycosylation in the three data sets was performed using the add-in feature for statistical analysis available in Microsoft Excel 2002.

The results of the U – ranking test for IgG_1 for the three data sets, are shown in Figure 2.61, in which the vertical bars represent the arithmetic mean of % G_0F for all 10 patients, siblings and controls. Glycosylation analysis of the IgG_1 of the 10 myositis patients revealed that these individuals have a statistically significant higher level of the G_0F glycoform, in comparison to healthy unrelated age matched controls (p=0.05), while no significant difference was determined between patients and siblings (p=0.29), or between siblings and controls (p=0.33).

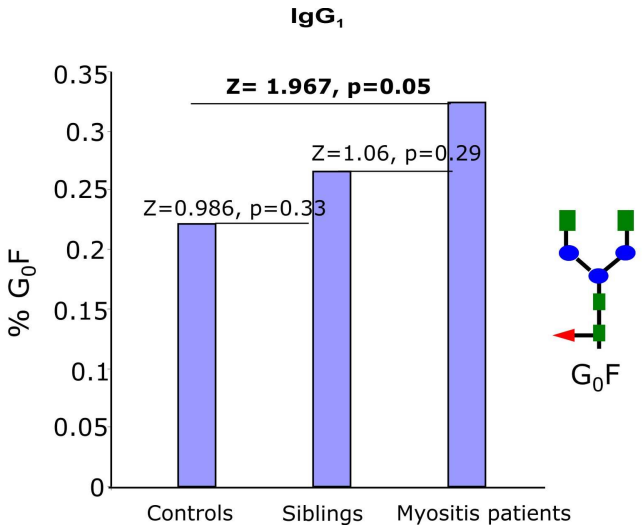


Figure 2.61: Arithmetic mean of % G_0F within IgG_1 subclass for: unrelated controls (left), siblings (middle) and myositis patients (right). As a result of the Mann-Whitney U-test, the levels of galactosylation are statistically different between controls and patients (p=0.05), whereas no statistical difference exists between controls and siblings (p=0.33), and siblings and patients (p=0.29), respectively.

Similar results were obtained for the $IgG_{2,3}$ subclasses, i.e. the levels of agalactosylation in myositis patients are significantly different from those of unrelated controls. These findings are similar to the decreased galactosylation reported for rheumatoid arthritis patients. The observation of increased levels of galactosylation in these autoimmune patients may reflect changes in intracellular processing pathways associated with disease state.

In conclusion, statistical analyses of glycosylation profiles of IgG subclasses in myositis patients, healthy siblings and unrelated controls reveal statistically significant increased levels of agalactosylation in the myositis patients compared to controls, whereas for the healthy twins/siblings intermediate levels of G_0F were determined. We speculate that these findings indicate the existence of a genetic predisposition in the group of unaffected siblings towards the development of autoimmune conditions, which can be observed at the protein level in their total plasma IgG fraction. Furthermore, the phenotype observed in the patient group suggests that the increased agalactosylation is the result of environmental exposures on the genes. We hypothesize that the environment may alter the expression levels of ß-galactosyl-transferases and/or of ß-galactosydases. In order to confirm this hypothesis, future work will employ proteomic approaches for determination of differential expression of total plasma proteins in the three clinical groups.

2.4 Prospectives for mass spectrometry in the analysis of protein structures and modifications

With the development of soft-ionization techniques over two decades ago [125, 127], an explosion of instrumental developments, in the field of ionization sources, mass analyzers and their combination with one- and multidimensional separation, has followed. This has led to significant advances in a number of research fields that have received "omics"-ending names, to indicate the distinct directions of research. These include genomics, proteomics, glycomics, metabolomics, which all were developed with the goals of biomarker discovery, early diagnosis, treatment, understanding the physiological processes involved in the development of diseases, and development of new approaches for disease treatment. The increasing demands in these areas, such as high throughput and automation, have stimulated considerable advances in bioinformatics, with the aim of integrating operations ranging from sample analysis, data acquisition, interpretation, and generation of the end results, in only a few "clicks". For example, if several decades ago it took one up to six months to identify a single protein, hundreds to thousands of proteins are now identified within a short time. New concepts have been introduced, such as "lab-on-a-chip" [276, 277], which refers to the ability to perform laboratory operations on small scale using miniaturized devices, representing the state-of-the-art technology for many applications.

However, successful development of therapeutic agents can only be achieved if the underlying concepts of structure – function relationship of biological molecules and their interactions with binding partners are understood at the molecular level. This requires rigorous structural characterization of proteins and their post-translational modifications, the major players in almost all diseases characterized at the present. Mass spectrometry has become a valuable tool for elucidation of *O*-glycosylation structures, as shown in this thesis for the secreted amyloid precursor protein. This challenging task was mainly derived from our poor understanding of biosynthesis of *O*-glycans, especially with regard to the consensus site required as a substrate for galactosyl transferases. Since it was shown that glycosylation profoundly affects the metabolism of APP, it would be interesting to explore the implications of the identified *O*-linked sites for the interactions of APP. The possible role of these structures for the pathogenic formation of Aß would be interesting to

interrogate by affinity – mass spectrometry, one of the new, just emerging areas of combination of MS within bioanalytical methods. An example for this new field is the molecular characterization of carbohydrate recognition structures using proteolytic excision – MS (carbohydrate recognition domain excision – MS, CREDEX – MS) [278, 279]. Furthermore, identification of the *N*-linked antibody glycosylation structures of Aß-specific antibodies and myositis-associated antibodies from patients constitutes a starting point for many other research directions. These should clarify, e.g. why is IgG_2 the dominant subclass in the Aß-autoantibody, why is galactosylation decreased in autoimmune disorders, are galactosyl-transferases down-regulated in disease, or are other proteins involved in the onset and evolution of a pathogenic phenotype.

In conclusion, mass spectrometry-based structural characterization of molecules of biomedical interest has opened up a myriad of research options ranging from sequence identification, characterization of site-specific glycosylation patterns, tertiary structure determination to characterizing conformational changes and sites of protein:protein interactions. The results of these types of studies may represent key steps in rational drug design and form the basis of various approaches in translational research. Mass spectrometry has and will continue to demonstrate its power and potential in biological and clinical applications.

3 EXPERIMENTAL PART

3.1 Proteins, enzymes and antibodies

Monoclonal antibodies: anti-Aß(1-17) (clone 6E10, Cat.# MAB1560), anti-Aß(17-24) (clone 4G8, Cat.# MAB1561): **Millipore (Billerica, MA)**; mouse ascites, containing an anti-Aß(34-40) monoclonal antibody (clone 11A5-B10, Cat.# 05-799): **Upstate (Lake Placid, NY)**; anti-ß-amyloid precursor protein (695aa) (clone 6A6, Cat.# A2275-87E): **US Biological (Swampscott, MA)**. Intravenous immunoglobulin (IVIg): **Bayer Vital GmbH (Leverkusen, Germany)** and **Calbiochem (San Diego, CA)**

Sequencing grade, TPCK-modified porcine trypsin: **Promega (Madison, WI)**; sequencing grade bovine α-chymotrypsin: **Roche (Penzberg, Germany)**.

Protein G Agarose kit (Cat. # 553-51-00): **KPL (Washington DC)**.

The plasmid pHD-APP695, encoding the full length amyloid precursor protein, isoform 695, was a kind gift from Dr. Bernadette Allinquant (Paul Broca Center, Paris, France).

Plasma samples for the myositis project were kindly provided by Dr. Frederick Miller (Environmental Autoimmunity Group, NIEHS, Bethesda).

3.2 Materials and reagents

The chemicals used in this work were used without further purification and comprised the following: dithiotreitol (DTT), iodoacetamide (IAA), ammonium bicarbonate, 96 % formic acid **(Sigma-Aldrich, St. Louis, MO)**; acetonitrile **(Caledon Laboratories, Ltd., Georgetown, Ontario)**; purified water (17.8 MΩ) was obtained from an in-house (NIEHS) Hydro Picopure 2 system; CHO-cells, WesternBreeze chemiluminescent Western blot immunodetection kit, NuPage 4-12 % Bis-Tris precast gels, sample, running and transfer buffers, and Coomasie SimplyBlue **(Invitrogen, Carlsbad, CA)**.

3.3 Affinity purification of the immunoglobulin fraction (IgG) from human plasma

3.3.1 Experimental design of a Twin-Sibling clinical study

One main focus of the present work was to determine whether the glycosylation pattern of total plasma IgG in patients with myositis follows a similar pattern with that determined for other autoimmune diseases, such as rheumatoid arthritis [243, 272]. Plasma samples were collected in the frame of a Twin-Sibling clinical study initiated at the National Institutes of Environmental Health Sciences, and coordinated by Dr. Frederick Miller (NIEHS, Bethesda, MD). This study is evaluating families with twins or siblings discordant for systemic rheumatic disorders (e.g. Rheumatoid Arthritis, Lupus, Scleroderma, or Myositis), in an attempt to assess potential environmental and genetic risk factors that may contribute to the onset of an autoimmune condition. The participants in this study were selected as follows:

- adults or children with a systemic rheumatic disorder, who have been diagnosed within 47 months and have a healthy twin or sibling of the same sex within 47 months of age.
- their unaffected twins or siblings, their parents.
- other normal, age-matched volunteers.

Ten different sets of plasma were analyzed, each comprising samples from a myositis patient, his/her healthy sibling or twin, and a healthy unrelated sex and age matched control (see Table 8).

3.3.2 Handling and storage of plasma samples

Plasma samples (0.5 – 0.7 mL) were shipped on dry ice and stored at -80℃ upon arrival. To avoid freeze-thaw cycles, sample aliquots (100 µL each) were prepared. No protease inhibitors or other preservatives were added.

3.3.3 Protein G affinity purification of the IgG fraction from human plasma

The isolation of total plasma IgG was performed in compact reaction columns (CRCs), packed with agarose bound Protein G. According to vendor's data sheet, Protein G, expressed in *E. Coli*, was engineered to eliminate the albumin and cell surface binding

domains, while maintaining efficient binding of immunoglobulin Fc region. This enables selective purification of IgGs from complex samples containing albumin, such as plasma.

Table 8: Overview of plasma samples analyzed for the twin-sibling study of myositis

Dataset	Sample ID	Diagnostic	Race	Gender	Age	Keys and abbreviations
1	016-01	DM	B	M	34	
	016-02		B	M	34	
	016-00		B	M	33	
2	014-01	PM	W	F	43	
	014-04		W	F	41	
	004-60		W	F	48	XXX-01 proband
3	003-01	IIM	W	F	33	XXX-02 monozygotic healthy twin
	003-04		W	F	n/a	XXX-04 healthy sibling
	003-00		W	F	30	YYY-00/04/06/60 unrelated age matched controls
4	017-01	DM	W	F	67	
	017-04		W	F	65	
	007-06		W	F	69	DM – dermatomyositis
5	027-01	PM	W	F	66	PM – polymyositis
	027-04		W	F	68	IIM – idiopathic inflammatory myopathies
	126-06		W	F	65	
6	054-01	IIM	W	F	54	B – black
	054-04		W	F	55	W – white
	024-00		W	F	54	M – male
7	053-01	PM	W	F	46	F – female
	053-04		W	F	50	
	042-00		W	F	50	
8	058-01	DM	W	M	62	
	058-04		W	M	65	
	058-00		W	M	64	
9	139-01	DM	W	F	66	
	139-04		W	F	63	
	077-04		W	F	62	
10	140-01	DM	W	F	41	
	140-04		W	F	44	
	004-00		W	F	41	

The purification kit is composed of:

- Agarose bound Protein G, supplied as slurry, consisting of 4 % highly cross linked agarose, with a bead size range of 45-165 µm and a binding capacity > 20 mg/mL human IgG.
- Washing/binding buffer (5X concentrate): 0.5 M sodium phosphate, 0.75 M sodium chloride, pH 7.4.
- Elution buffer (10X concentrate): 2M glycine, pH 2.85.

The CRC used for this procedure has a volume of 0.5 mL; it is equipped with a frit (35 µm pore size) and a screw cap, enabling flushing of solvents and air via a syringe applied through the cap. The column was packed with ~ 200 µL drained agarose, as follows: 400 µL of slurry were mixed with 400 µL washing/binding buffer (supplied in the kit), transferred to the column, and allowed to flow by gravity. The packed affinity resin was equilibrated with 5 mL washing/binding buffer. Plasma samples (20 µL) were diluted to 100 µL with washing/binding buffer and applied on the resin; a volume of 200-300 µL washing/binding buffer was subsequently added, in order to fill the remaining dead volume. A nutator mixer was used to mix the content of the column in a three dimensional, gentle rocking motion, for 45 minutes at room temperature. The non-bound protein fraction was removed by washing the resin with 5 mL washing/binding buffer, followed by 5 mL deionized water. Elution of the IgG fraction was performed with 0.5 mL elution buffer and incubation for 15 minutes at room temperature. The eluted IgG was brought to physiologic pH by the addition of 150 µL of 5X washing/binding buffer. The purity of the IgG fraction was assessed by SDS-PAGE.

3.4 Epitope specific isolation of Aß-autoantibodies

3.4.1 Covalent immobilization of Cys-Aß(12-40) on sepharose

3.7 mg of synthetic Cys-Aß(12-40) peptide were dissolved in 10 mL coupling solution containing 50 mM Tris-HCl, 5 mM EDTA (pH 8.5) to a final concentration of 0.37 mg/mL. This solution was added to 1 mL of Ultralink Iodoacetyl gel dried of liquid and the coupling reaction was allowed to take place for one hour at room temperature with shaking, followed by 30 minutes standing without shaking. The matrix was loaded onto a 2.5 mL column, which allows the liquid to be drained. The column was washed with 3 mL coupling solution. The non-reacted iodoacetyl groups were blocked with 1 mL blocking solution containing 50 mM of L-Cysteine-HCl dissolved in coupling solution, for 45 min at room temperature. This procedure was repeated twice. Subsequently, the column was washed with 5 mL of 1M NaCl, followed by 5 mL of PBS buffer, (pH 7.2) and then stored at 4℃.

3.4.2 Affinity isolation of Aß-autoantibody from IVIg

Affinity isolation of anti-Aß autoantibodies from IVIg was performed as follows: 0.5 mL of Cys-Aß(12-40) containing matrix were loaded onto a 2.5 mL column and washed with 20 mL PBS buffer (pH 7.2). The matrix was then transferred into a 15 mL flask using 5 mL PBS buffer and mixed with 5 mL IVIg. The suspension was spun over night at 4℃, then transferred back onto the column and washed 8 times with each 10 mL PBS and subsequently 2 times with each 10 mL deionized water. The affinity isolated Aß-autoantibody fraction was eluted 10 times with each 0.5 mL 0.1 % trifluoroacetic acid. The quantification of the eluted antibody fractions was performed using the Micro BCA^{TM} kit and the detection was performed at 562 nm using a Wallac ELISA plate reader. The antibody fractions were lyophilized to dryness.

3.5 Expression and purification of secreted amyloid precursor protein (sAPP695)

3.5.1 Cell culture

The CHO-cells (Invitrogen) were grown serum free in suspension culture using CD-CHO (GIBCO) containing HT supplement (Gibco) and 8 mM glutamine.

3.5.2 Protein production

The APP695 protein was produced using the pHD-APP695 plasmid, containing the APP695 gene under the control of the HSV promotor with the SV40 enhancer sequence [280]. CHO-s cells were transiently transfected with the pHD-APP695 plasmid using Fugene HD (Roche). A ratio of 4 uL of Fugene HD per 1 ug of plasmid gave the best results. The APP695 protein was allowed to accumulate in the media for 48 to 72 hrs post-transfection before the media was harvested.

3.5.3 Q-Sepharose chromatography

Ion exchange chromatography relies on charge-charge interactions between the proteins in a mixture and the charges immobilized on a resin. Ion exchange chromatography can be subdivided into cation exchange chromatography, in which positively charged ions bind to a negatively charged resin, and anion exchange chromatography, in which the binding

ions are negative, and the immobilized functional groups are positive. Once the analytes are bound, the column is equilibrated in the starting buffer, of low ionic strength. The bound molecules are eluted using a gradient of a second buffer which steadily increases the ionic strength of the eluent solution. Alternatively, the pH of the elution buffer can be modified in order to provide the protein or the resin with charges such that they no longer interact.

For purification of APP695 anion exchange chromatography was employed. The chemical structure of Q-sepharose, used as a stationary phase, is shown in Figure 1. The quaternary ammonium groups on the surface of the matrix ensure interaction with negatively charged analyte molecules.

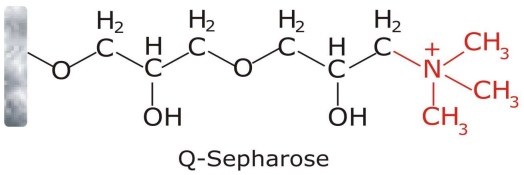


Figure 3.1: Chemical structure of Q-sepharose anion exchange resin

The media containing the APP695 protein was centrifuged at 350 x g for 10 min to remove all of the cells. The media was further clarified by passage through a 0.2 um Acrodisk syringe filter before being flow loaded onto a 10 mm by 20 cm Q sepharose FF (Sigma) column at 1mL/min. The column had been pre-equilibrated with 25 mM HEPES pH 7.5 buffer containing 150 mM NaCl. After the sample was loaded, the column was washed with 3 column volumes of the pre-equilibration buffer. APP695 was eluted from the Q sepharose column using a 10 column volume linear gradient from 150 mM NaCl to 1 M NaCl in 25 mM HEPES pH 7.5. The load flow through and column wash were collected in batch, while the elution was collected in 2 mL fractions. The fractions that contained APP695 were identified by SDS-PAGE and western blot. The chromatography was performed at 4ºC using an AKTA purifier (GE Healthcare).

3.5.4 Western blot analysis of sAPP695

The protein peak eluted from the Q-sepharose column was separated on 4-12 % Bis-Tris pre-cast gels as follows: 30 µL of each fraction were incubated for 10 minutes at 90° C with 10 µL sample buffer containing 100 mM dithiothreitol. Electrophoresis was performed at

200 V and a maximum of 100 mA, for 90 minutes. The bands were stained over night with Coomasie Simply Blue. For Western blot analysis, the proteins separated by SDS-PAGE were transferred onto nitrocellulose membranes for 90 minutes at 25 V and 125 mA. The membrane was rinsed with deionized water and blocked for 30 minutes in the blocking solution prepared from the reagents provided with the WesternBreeze® kit. Subsequently, the membrane was submersed for one hour in a solution containing 1 µg/mL of the mouse anti-ß-APP695 monoclonal antibody (primary antibody). Following washing, the membrane was incubated for 30 minutes with a solution containing the anti-mouse, alkaline phosphatase – conjugated, secondary antibody. The chemiluminiscent substrate was added to the membrane surface and the reaction was allowed to develop for 5 minutes before the membrane was exposed to the X-ray film for 10 seconds.

3.6 One dimensional gel electrophoresis

Electrophoresis represents a well established separation method of protein mixtures, based on differential migration properties of analytes derived from molecular weight. For low complexity protein samples, the one dimensional separation is the method of choice. Upon reduction and alkylation, immunoglobulin samples contain a mixture of heavy (~50 kDa) and light chains (~25 kDa), which can be easily separated on one dimensional SDS-PAGE.

The reduced and alkylated antibody samples were loaded onto a precast 4-12 % Bis-Tris gel (~10 µg/lane) and the electrophoresis was carried our for one hour at 200 V and a maximum of 80 mA.

3.7 Coomasie Simply Blue Staining

The gel was rinsed 3 × 5 minutes in deionized water and stained with ready-to-use Coomasie Simply Blue solution either over night or using microwave heating. For the microwave assisted staining, the gel was submersed in staining solution in a microwave safe Petri dish and heated for 1 minute on high power setting, followed by cooling on the bench for two minutes. The heating procedure was repeated for one additional minute. After cooling, the gel was removed from the staining solution and transferred to a

microwave safe dish containing deionized water and subjected to heating/cooling cycle until the background became crystal clear.

3.8 Chemical modifications and proteolytic degradation of antibodies

3.8.1 Reduction and alkylation of antibodies in solution

Prior to SDS-PAGE, all antibodies used in the present work were treated as follows: antibody stock solutions (~10-20 µg) were incubated with 20 µL sample loading buffer containing 100 mM DTT for one hour at 95℃. Subsequently, iodoacetamide (1M stock solution in deionized water) was added to a final DTT/IAA molar ratio of 1:3, and the reaction mixture was incubated for one additional hour at room temperature before loading onto the SDS-PAGE.

For experiments requiring digestion in-solution, the antibody samples were reduced and alkylated as described, whereas deionized water was substituted for sample loading buffer.

3.8.2 Proteolytic digestion of antibodies in solution with trypsin

Prior to use, TPCK-modified trypsin, supplied as powder, was dissolved in 25 mM ammonium bicarbonate buffer (pH 7.4) to a final concentration of 0.05 µg/µL. The enzyme was added to the reduced and alkylated antibody in a final enzyme-to-substrate ratio of 1:50 (w/w) and the digestion was performed over night at 37℃.

3.8.3 Proteolytic digestion of antibodies in solution with α-chymotrypsin

Alpha-chymotrypsin, supplied as powder, was reconstituted in 1 mM HCl (pH 2-3) to a concentration of 0.25 µg/µL. This stock solution can be stored at 4℃ for up to one week. The use of hydrochloric acid is aimed at preventing the autoproteolysis by changing the enzyme conformation; this conformational change is reversible when the pH is raised to the physiological value. The enzyme was added to reduced and alkylated antibody samples in a final enzyme-to-substrate ratio of 1:30 (w/w) and the digestion was performed over night at 25℃.

3.8.4 In-gel digestion of proteins with trypsin

The protein bands were manually cut and digested with trypsin for 8 hours at 37° C in an automated fashion with a Progest robotic digester (Genomic Solutions). This instrument is capable of performing high-throughput protein digestion in the 96-well plate format. A major advantage over manual digestion is represented by the keratin free environment, as automation minimizes sample handling steps and hence the risk of contamination. A standardized digestion protocol for trypsin comprising steps such as destaining, shrinking, thermal incubation and extraction was used.

3.8.5 In-gel digestion of proteins with α-chymotrypsin

In-gel digestion with α-chymotrypsin was performed manually, as follows. The protein bands were excised, placed in an Eppendorf cup and washed with 100 µL of deionized water for 15 minutes at room temperature. Destaining was performed in 40 % acetonitrile in 25 mM ammonium bicarbonate (2× 20 minutes), followed by drying under vacuum of the shrunk and colourless gel pieces. For digestion, 4 µL of a 0.25 µg/µL stock solution of α-chymotrypsin in 1 mM HCl were diluted to 50 µL with 25 mM ammonium bicarbonate and subsequently added to the destained protein bands. Digestion was performed over night at 25° C. Samples were lyophilized to dryness and resuspended in 0.1 % formic acid. The supernatant was collected in a separate Eppendorf cup; extraction of the remaining peptides from the gel pieces was performed by addition of 60 µL of 10 % formic acid in deionized water (v/v) (1× 20 minutes) followed by addition of 60 µL of 70 % acetonitrile in deionized water containing 0.1 % formic acid (v/v/v) (3× 30 minutes). The extraction steps were performed at room temperature with shaking. The supernatant and the extracted volumes were combined and lyophilized.

3.8.6 Double trypsinization of antibody heavy chain

This experimental approach was employed for semi-quantitative analyses of IgG glycosylation within the twin-sibling study, for which complete enzymatic degradation was essential. The first digestion step using trypsin involved the use of the automatic ProGest robot described above, followed by lyophylization of the extracted digestion products.

Because this mixture contained a significant amount of tryptic glycopeptides with one miscleavage, an additional digestion step with trypsin was employed. To each lyophilized sample, 25 µL ammonium bicarbonate (pH 7.4) containing 0.83 µg trypsin were added. The digestion was performed over night at 37° C. The samples were stored at -80° C until further use.

3.9 Liquid chromatography – mass spectrometry

3.9.1 Nano-LC – ESI Q-Tof mass spectrometry

LC/MS analyses were performed using a Waters Q-Tof Premier mass spectrometer equipped with a nanoAcquity UPLC system (Waters, Milford, MA). Analyses were performed on a 75 µm × 100 mm, Atlantis 3 µm dC18 column (Waters, nanoAcquity), using a flow rate of 300 nL/min. A C18 trapping column (180 µm × 20 mm) with a 5 µm particle size (Waters, nanoAcquity) was positioned in-line with the analytical column and upstream of a micro-tee union used both as a vent for trapping and as a liquid junction. Trapping was performed for 3 minutes at a 5 µL/min flow rate, using the initial solvent composition. Briefly, a 4 µl aliquot of the digest sample was injected onto the column. Peptides were eluted using a linear gradient from 98 % solvent A (0.1 % formic acid in water (v/v)) and 2 % solvent B (0.1 % formic acid in acetonitrile (v/v)) to 40 % solvent B over 60 minutes. For some experiments this gradient was extended over a 90 minutes time window. Mass spectrometer settings for the MS analyses were: capillary voltage, 3.2 kV; cone voltage, 33 V; collision energy, 8.0 V; and source temperature, 80° C. The mass spectra were acquired over the mass range 200 – 2000 Da.

MS/MS data were acquired in the data dependent mode, using collision energies based on mass and charge state of the candidate ions. Alternatively, collision energy ramps from 20 V to 30 V, and from 30 V to 40 V, respectively, were employed in order to obtain optimal fragmentation of the (glyco)peptide ions. Representative examples of MS/MS data with the complete b_x-y_z nomenclature [182-184], obtained for the N-terminal anti-Aß(1-17) antibody heavy chain are shown in Figure 3.2, Figure 3.3, Figure 3.4 and Figure 3.5. A capillary voltage of 3.2 kV and a cone voltage of 20 V were used for glycopeptide analysis, in order to prevent their in-source decomposition. For subclass-specific glycosylation analysis, the instrument was operated in the MS only mode; for each sample, technical triplicates were

acquired. An external lock mass using a separate reference sprayer (LockSpray) and a solution of Glu-Fibrinogen peptide (300 fmol/µL) in water/acetonitrile 80:20 (v/v) and 0.1 % formic acid, with a mass of 785.8496 (2+), was used for calibration. Data analyses were performed using using MassLynx 4.0 software (Waters, Milford, MA).

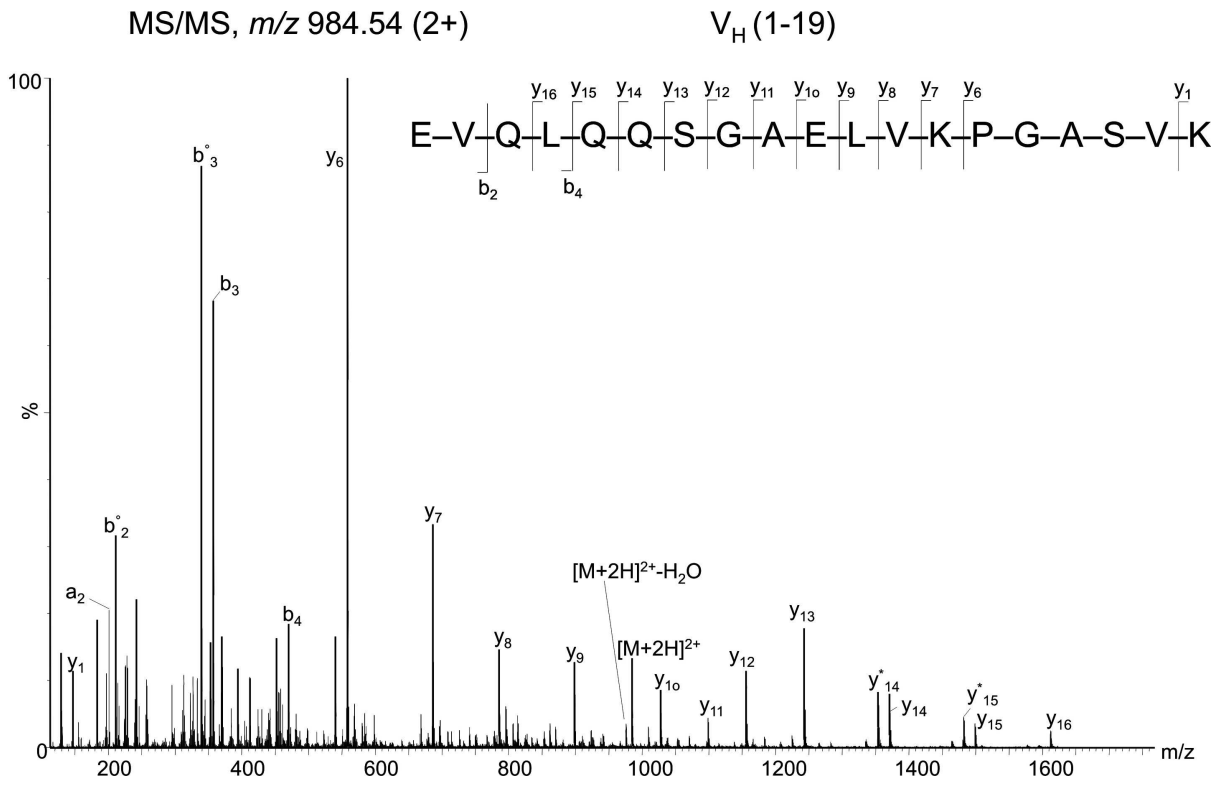


Figure 3.2: Collision induced dissociation spectrum of the precursor ion of m/z 984.54 (2+) corresponding to the N-terminal region V_H(1-19) of the 6E10 antibody heavy chain, indicated at the top. The peptide was assigned by NCBInr protein database search. The observed amide bond cleavages are indicated in the peptide sequence. The asterisk (*) indicates loss of ammonia from the y fragment ions.

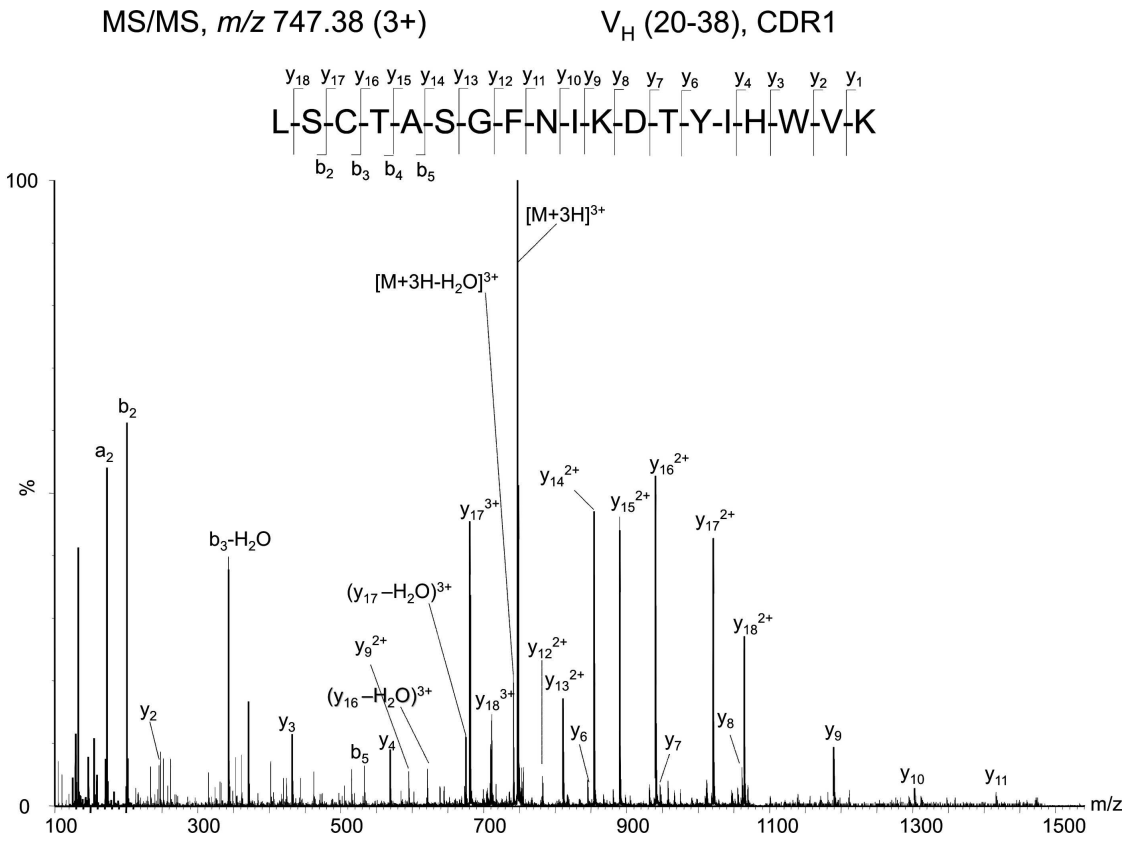


Figure 3.3: Fragment ion spectrum (CID) of the precursor ion of m/z 747.38 (3+) assigned to peptide V_H(20-38) of the 6E10 heavy chain. The sequence of the peptide V_H(1-19) and the observed backbone cleavages are indicated at the top. The spectrum was obtained on the Q-Tof Premier using data dependent acquisition, with a collision energy ramp from 30 V to 40 V.

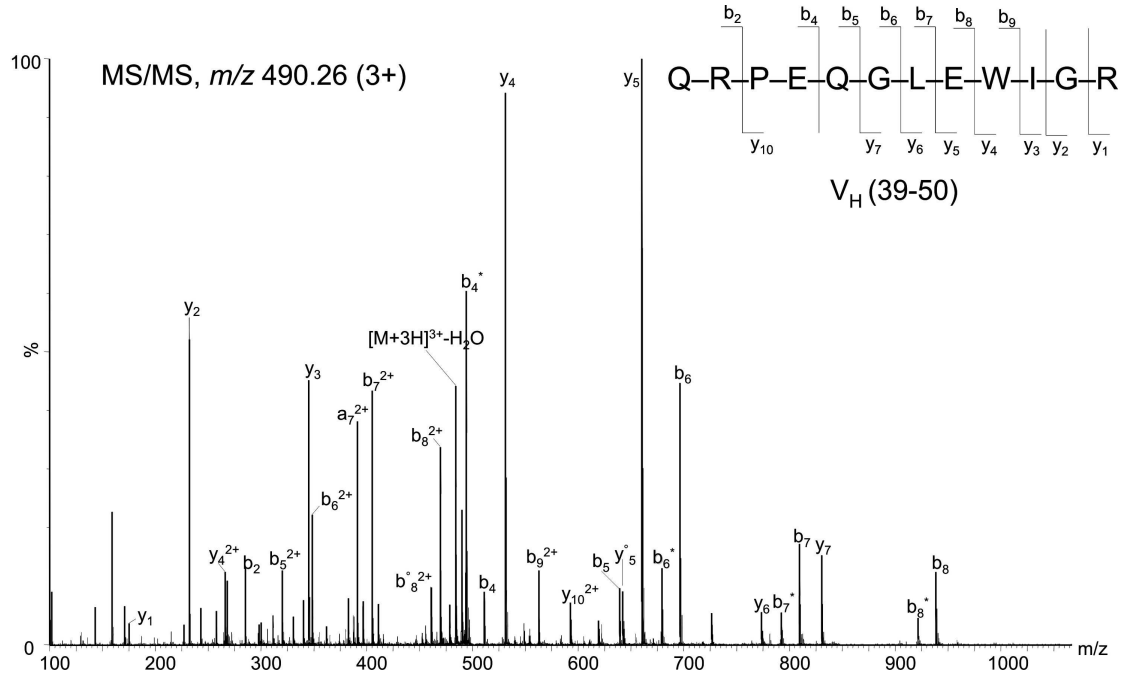


Figure 3.4: Fragment ion spectrum of the precursor ion of m/z 490.26 (3+) assigned by NCBInr database search to the peptide V_H(39-50) of the 6E10 antibody heavy chain. The peptide sequence and the observed amide bond cleavages are indicated at the top (right). The asterisk (*) indicates loss of ammonia from the b and y fragments during collision induced dissociation.

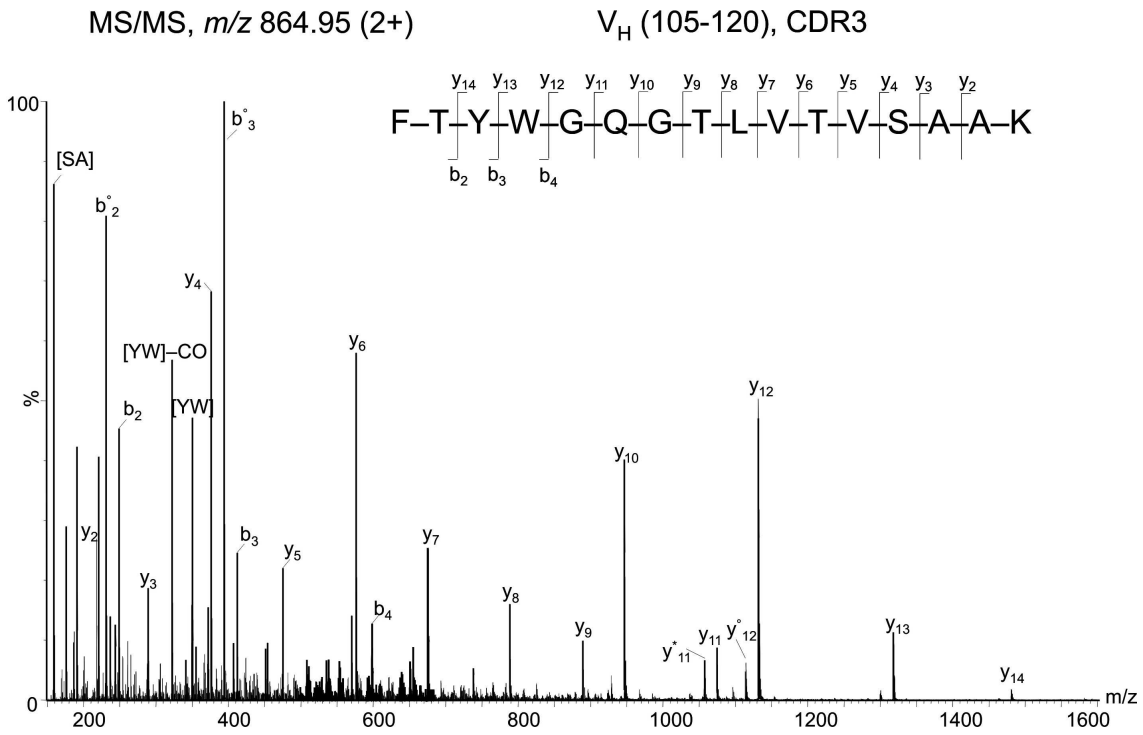


Figure 3.5: Fragment ion spectrum of the precursor ion of m/z 864.96, corresponding to peptide V_H(105-120) from the 6E10 antibody heavy chain. The peptide sequence and observed backbone cleavages are indicated at the top. The loss of water from the b and y fragment ions arising during collision induced dissociation is indicated with circles (○). The spectrum was obtained on the Q-Tof Premier using data dependent acquisition, with a collision energy ramp from 30 V to 40 V.

3.9.2 Nano LC – ESI Ion trap mass spectrometry

Nano-LC/ESI/MS/MS analyses were performed using an Agilent 6340 (Santa Clara, CA) Ion Trap equipped with an HPLC Chip Cube MS interface, an Agilent 1200 nanoHPLC and an electron transfer dissociation module. The solvents used for chromatography were 0.1% formic acid in deionized water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). Ion trap-MS/MS analyses were performed as follows: 20 µL injections of the tryptic or chymotryptic digests dissolved in 0.1% formic acid were loaded onto a 40 nL enrichment column followed by a 43 mm x 75 µm analytical column, packed with ZORBAX 300SB C18 particles. Linear gradients of 3-50% solvent B were performed over 50 min at a flow rate of 500 nL/min. The parameter settings for positive ion ESI-MS were as follows: capillary voltage – 2000 V, end plate offset – 500 V, capillary exit – 180 V, nebulizer – 2 psi, dry gas – 4 L/min, dry gas temperature – 325 °C. For MS/MS (ETD or CID), automated data dependent acquisitions of the four most abundant ions were employed. For CID, the fragmentation amplitude was 0.80 V, which was scanned from 30% to 200% of this preset value (SmartFrag parameter on the instrument tune page). For ETD

analyses, the accumulation time of gaseous fluoranthene anions was 10 – 12 msec, and the reaction time was typically 100 or 150 ms. The ETD/CID voltage was set at 0.07 – 0.10 V when supplemental ion activation of the charged reduced molecule radical ions was employed.

3.9.3 Analytical RP – HPLC combined with "off-line" ESI – Ion trap MS

3.9.3.1 RP-HPLC

Separation of the Aß-autoantibody heavy and light chain tryptic digests was performed on a Bio-Rad analytical HPLC system (Bio-Rad, Richmond, VA) equipped with a 4.6 mm × 250 mm Vydac 5 µm C_4 column, using a flow rate of 0.5 mL/min. The mobile phases were: 0.1 % trifluoroacetic acid (TFA) in deionized water (v/v, solvent A), and 80 % acetonitrile in deionized water (v/v), acidified with 0.1 % TFA (v/v, solvent B). A linear gradient ranging from 0 to 65 % solvent B over 130 minutes was employed for peptide separation. The UV detection of the eluting peptides was performed at 220 nm, representing the absorption wavelength characteristic for the amide bond. The eluting fractions were collected in 1 mL Eppendorf tubes and lyophilized to dryness.

3.9.3.2 "Off-line" ESI – ion trap mass spectrometry

Prior to MS, the lyophilized peptide fractions were resuspended in 12-15 µL 0.1 % formic acid in deionized water (v/v). LC-ESI-Ion trap MS was performed on an Esquire 3000 ion trap instrument (Bruker Daltonik, Bremen, Germany) connected to an Agilent 1100 HPLC system (Santa Clara, CA). LC-analyses of individual fractions were performed on a 1 mm ×10 cm Discovery BIO Wide Pore 3 µm C_{18} column (Sigma-Aldrich, St. Louis, MO), at a flow rate of 50 µL/min. The mobile phases were: 0.1 % formic acid in deionized water (v/v, solvent A) and 80 % acetonitrile in deionized water (v/v), acidified with 0.1 % formic acid (v/v, solvent B). A linear gradient from 20 to 50 % to solvent B over 3 minutes was employed for peptide separation. For each fraction, 4-5 injection cycles (3 µL injection/cycle) were performed, as follows: *(i)* a first analysis was performed using MS only acquisition, in which potential multiply charged peptide precursor ions were identified, and *(ii)* subsequent injection cycles were performed using targeted MS/MS acquisition of the precursor ions observed in the first injection cycle; a typical value of the fragmentation

amplitude employed to achieve dissociation of the precursor ions was 1 V. Additional MS parameters were: capillary voltage – 3000 V, capillary exit – 120 V, skimmer – 40 V, dry gas temperature – 300 V, dry gas – 7 L/ min, nebulizer gas pressure 14-16 psi. The total analysis time for one injection cycle was 25 minutes. Singly charged ions were not considered for targeted MS/MS.

3.10 Bioinformatic tools for mass spectrometry

3.10.1 Database search

Various computer software employed to identify proteins from mass spectrometric data are designed to compare the MS/MS data submitted by the user with the fragmentation spectra of peptides deposited in protein databases. Fragment ions from the acquired data are matched, based on their *m/z* ratio against theoretical spectra of peptides derived from *in-silico* digestion of proteins, taking into account the digestion parameters pre-set by the user. For example, the use of trypsin as a proteolytic enzyme, as well as the number of miss-cleavages can be indicated in the search engine. The user has the option to select amino acid side chain modifications that were derived from chemical procedures (e.g. carbamidomethyl cysteine, if the protein was reduced and alkylated prior to digestion), or that are thought to occur during sample preparation (e.g. methionine oxidation, deamidation, pyro-Glu formation). The type of mass spectrometer can be selected as an input parameter, such as to provide information about the nature of the fragment ions formed using different dissociation regimes (b_x-y_z, or c – z), hence directing the search towards specific match of the fragment ions types. To perform the searches, the non-redundant NCBI protein database (NCBInr) was employed, which contains entries compiled from a variety of sources, such as GenBank (a genome sequence database from the National Institutes of Health), SWISS-Prot, PRF (Protein Research Foundation), PIR (International Protein Sequence Database), and PDB (Brookhaven Protein Databank).

Prior to database searching, LC-MS/MS data were processed (including ions with a S/N ratio greater than 3:1) using Mascot Distiller software (Matrix Science, UK) and saved as an *.mgf file. These data were searched against the NCBInr protein data base by means of the Mascot MS/MS Ion Search engine, using a precursor tolerance of 0.2 Da and an MS/MS tolerance of 0.1 Da. Carbamidomethyl cysteine was defined as a fixed

modification, whereas variable modifications included methionine and tryptophane oxidation, deamidation (Gln and/or Asn), and N-terminal formation of pyro-Glu and pyro-Gln. The sequences determined from the MS/MS data obtained for the identified peptides were validated manually.

3.10.2 Glycopeptide data analysis

The glycopeptide data obtained for the Aß-autoantibody and for the IgG fraction from the subjects of the twin-sibling study were analyzed using the MassLynx 4.1 software. For subclass-specific glycosylation analysis, data from three isolation batches of Aß-autoantibody were averaged, and the mean and standard deviation were calculated for each glycoform of each subclass using nine separate experimentally determined ion abundances (3 separation batches × 3 analyses/antibody batch). For each sample from the clinical study, technical LC-MS triplicates were acquired, and these data were averaged. All observed charge states for a particular glycopeptide were included in these analyses. Subclass-specific glycosylation analyses were performed by averaging the MS scans over the chromatographic retention time in which glycopeptides from a specific subclass eluted. The relative abundance for each glycoform was determined by dividing the determined ion abundance for a particular glycopeptide to the ion abundance of the G_1F glycoform within each IgG subclass.

3.10.3 Glycomod software

Glycomod (http://www.expasy.org/tools/glycomod/) is an open-source software which derives glycan compositions from mass spectrometric data of glycopeptides, and that of free and/or derivatized carbohydrates released by chemical or enzymatic procedures from a glycoprotein. For glycopeptides, the input information consists of the experimentally determined glycopeptides' masses, the protein sequence containing those modifications (or its accession number), and the proteolytic enzyme used to generate glycopeptides. The results were corroborated with the information derived from MS/MS, because in some instances isobaric structures may lead to false-positive results. This software was employed for the analysis of the glycosylation data of the 6E10 and 4G8 monoclonal antibodies.

3.10.4 Mann-Whitney U-test

Statistical analysis of the glycoform levels in IgG antibodies isolated from myositis patients, their siblings and unrelated controls was performed using the add-in feature for non-parametric statistical analysis available in Microsoft Excel 2002; this enables quick and easy statistical evaluation of mass spectrometric data tabulated in an Excel spreadsheet.

3.10.5 BallView 1.1.1

BallView 1.1 is an open-source application employed for molecular modelling and visualization of protein structures, such as antibodies. The input files containing structural information derived from X-ray or NMR data are deposited in Protein Data Bank, can be downloaded and visualized using BallView 1.1. Differential selection and highlighting of protein domains can be employed.

4 SUMMARY

Autoimmune-related diseases have become a major medical concern in recent years due to the increasing number individuals affected, poor understanding of disease etiology, and the lack of an efficient therapy. In the last years, evidence suggests the implication of the immune system in the pathogenesis and progression of autoimmune diseases, and new therapeutic approaches targeting the immune system are being developed. Future advances in the design of effective vaccines will substantially depend on a more complete understanding of the structural basis of immune responses. Mass spectrometry has emerged as a high performance tool for protein structure determination and characterization of protein – protein interactions, because of its sensitivity, high mass accuracy, short analysis time and applicability to mixtures. The successful development of analytical strategies of mass spectrometry, such as differential chemical modification of proteins and proteolytic epitope excision, have improved the level of structural information provided by classical techniques by revealing information about molecular recognition structures in antigen-antibody interactions. In particular, tandem mass spectrometry, producing different types of fragment ions by various dissociation techniques, has been developed and applied as a major tool in this dissertation for determination of protein structures and post-translational modifications .

Alzheimer's disease related neurodementia has become a great medical threat among the aging population, with the histopathological hallmark being the deposition of ß-amyloid plaques in brain. Innovative immunotherapeutic approaches, aimed at reducing the plaques and preventing neurotoxic effects and aggregation of Aß, are based on the epitope structures recognized by Aß-specific antibodies. The first two chapters of this thesis are focused on the determination of complete primary structures, and elucidation of the *N*-glycosylation structures of Aß-specific antibodies. Plaque-specific antibodies produced by active immunization with Aß recognize an epitope located at the Aß N-terminus (^{4}FRHDSGY^{10}) and were shown to reduce the amyloid burden and to improve cognitive functions. Using high performance liquid chromatography – tandem mass spectrometry LC-MSMS, the complete structural characterization of a plaque-specific mouse monoclonal antibody was performed. Using both database search and *de novo* sequence analysis, near-complete sequence determinations were obtained for both heavy and light chains, and the amino acid sequences of five of the six complementary

determining regions were completely elucidated. Several sequence microheterogeneities were identified in the region spanning the CDR2 of the light chain, which may provide key information for the antigen recognition arising during the B-cell affinity maturation. The major glycan structure attached at the heavy chain constant region was identified to be of complex type with zero to four terminal galactose residues and minimal amounts of *N*-glycolyl neuraminic acid.

Human physiological Aß-antibodies (“plaque-protective" antibodies) were identified in plasma of healthy individuals and recognize an epitope located in the C-terminal region of Aß (residues (21-37)), and this feature is suggested to prevent accumulation and deposition of Aß. The complete primary structure and glycosylation of a plaque-protective mouse monoclonal antibody, raised against an epitope located at residues (17-24) of Aß, was obtained by LC-MS/MS, in combination with collision induced dissociation and electron-transfer dissociation (CID and ETD) fragmentation. In this study, the structures of all six CDRs, and an unexpected consensus amino acid sequence of *N*-linked glycosylation were elucidated in the light chain. This glycosylation site was identified to carry partially hybrid glycans having one antenna of high mannose type, and a second one of complex type, terminated with *N*-acetyl and *N*-glycolyl neuraminic acid. Furthermore, small amounts of complex type structures were identified in the variable region, containing up to two galactose residues and terminated with *N*-glycolyl neuraminic acid. Glycosylation of the heavy chain constant region was found less heterogeneous than that of the variable region, consisting of complex type structures with zero to four galactose residues.

Epitope-specific affinity purification of plaque-protective, human Aß-autoantibodies from commercial immunoglobulins was performed using immobilized Cys-Aß(12-40), and the determination of their primary structures indicated the presence of all four IgG subclasses. Subclass- specific glycosylation analysis, performed at the glycopeptide level, revealed elevated amounts of IgG_2/IgG_3 and IgG_4, respectively. The glycan structure of each subclass indicated comparable amounts of the individual glycoforms in the Aß-autoantibody and pooled immunoglobulins, with a slight tendency towards increased agalactosylation of the autoantibody.

The ß-amyloid peptide is derived by proteolytic processing of the amyloid precursor protein APP, a transmembrane protein of hitherto largely unknown function. Although the

importance of *N*- and *O*-glycosylation in the secretory pathway of APP was previously suggested, no sites of *O*-glycosylation have been reported. The elucidation of the *O*-glycosylation structures of APP at three specific sites, Thr291, Thr292 and Thr576, was obtained using LC-MS and a combination of ETD and CID fragmentation. The *O*-glycosylation sites, Thr291, Thr292 and Thr576, were found to comprise multiple short Core-1 type glycans. The minimal *O*-glycan structure was *N*-acetyl galactosamine (GalNAc), while elongated structures were identified to contain the GalNAc – Gal core terminated with sialic acid, attached either in a linear fashion to galactose, or branched with attachment to GalNAc.

In the last chapter of this thesis, the determination of subclass specific glycosylation structures of total plasma IgG from myositis patients was performed as part of a clinical study aimed at delineating genetic and environmental risk factors for pathogenesis of systemic rheumatic disorders. Total plasma IgG from patients, healthy twins/siblings and unrelated age matched controls was isolated using protein G affinity chromatography. Analysis of glycopeptides from each subclass revealed elevated amounts of core-fucosylated agalaytosyl glycoform (G_0F) in patients compared to unrelated controls, and these two groups were found to be statistically different. In contrast, statistical analysis of patients – siblings and siblings – controls groups suggested a random structure distribution of antibody glycosylation. This indicated the existence of a genetic predisposition of healthy siblings towards development of an autoimmune condition, possibly as a result of various environmental exposures.

5 ZUSAMMENFASSUNG

Autoimmun-Erkrankungen haben in den letzten Jahren zunehmende medizinische Bedeutung aufgrund der zunehmeden Zahl individueller Erkrankungen, der geringen Kenntnisse über Krankheitsursache(n), und der bisher weitgehend fehlenden Therapiemöglichkeiten erlangt. In jüngster Zeit wurden eindeutige Hinweise auf die Bedeutung des Immunsystems für Pathogenese und Verlauf von Autoimmun-Erkrankungen erhalten, die derzeit zur Entwicklung von neuen therapeutischen Ansätzen führen; künftige Fortschritte in der Entwicklung und Herstellung von wirksamen Vaccinen werden jedoch wesentlich von dem besseren Verständnis der strukturellen Grundlagen des Immunsystems abhängen. Die Massenspektrometrie hat sich in den letzten Jahren aufgrund ihrer hohen Nachweisempfindlichkeit, Genauigkeit der Massenbestimmung, schnellen Analysenzeiten und Anwendbarkeit zur Analyse komplexer Gemische als Hochleistungsmethode für Strukturaufklärung von Proteinen und Charakterisierung von Protein-Interaktionen erwiesen. Die erfolgreiche Entwicklung von analytischen Verfahren mittels Massenspektrometrie, insbesondere der selektiven chemischen Modifizierung von Proteinen sowie von proteolytischen Methoden zur Epitop-Identifizierung, haben Strukturinformationen mit klassischen Methoden erheblich verbessert und zur Aufklärung von molekularen Erkennungsstrukturen in Antigen-Antikörperkomplexen beigetragen. In der vorliegenden Dissertation wurde vor allem die Tandem- Massenspektrometrie, mit verschiedenen Verfahren der Herstellung und Analytik von Fragment-ionen, als eine zentrale Methode zur Identifizierung von Protein-Strukturen und –Strukturmodifikationen entwickelt und angewendet.

Neurodegenerative Erkrankungen vom Typ der Alzheimer-Krankheit sind eine der wichtigsten medizinischen Herausforderungen einer zunehmend älter werdenden Bevölkerung; ein zentrales Merkmal der Alzheimer-Krankheit ist die Bildung und Ablagerung von Aggregaten (Plaques) des ß-Amyloid (Aß) im Gehirn. Neue Immun-therapeutische Ansätze zur Verringerung von Aß-Aggregaten und Verhinderung/Verringerung ihrer neurotoxischen Effekte basieren auf der Entwicklung von Aß-Epitop-spezifischen Antikörpern. Die beiden ersten Abschnitte der vorliegenden Dissertation behandeln die Aufklärung der vollständigen Primärstukturen sowie Identifizierung der *N*-Glykosilierung von Aß-spezifischen Antikörpern. Durch aktive

Immunisierung mit Aß hergestellte, „Plaque-spezifische“ Antikörper erkennen ein N-terminales Aß-Epitop (^{4}FRHDSGY^{10}), sie verringern die Aß-Plaques und verbessern kognitive Funktionen von Alzheimer-Patienten. Mittels HPLC-Tandem-Massenspektrometrie (LC-MS/MS) konnte die vollständige Struktur eines Plaque-spezifischen monoklonalen Maus-Antikörpers aufgeklärt werden. Nahezu vollständige Sequenzbestimungen wurden mittels Datenbank-Suchverfahren sowie „*De novo*“-Sequenzierung für die schweren und leichten Ketten erhalten, insbesondere die vollständige Aufklärung von 5 der 6 CDR-Sequenzen des Antikörpers. Ferner wurden in der CDR2- Sequenz der leichten Kette eine Reihe von Sequenzheterogenitäten aufgeklärt, die wesentliche Informationen über Antigen-Erkennung bei der B-Zellprozessierung liefern. Die Glycosilierung im F_c-Teil des Antikörpers wurde als Komplexstruktur-Typ mit bis zu 4 terminalen Galactosylresten und einem geringen Anteil an *N*-Glycolyl-Neuraminsäure identifiziert.

Im Gegensatz zu Antikörpern nach Aß-Immunisierung wurden physiologische humane Aß-Antikörper („Plaque-protektive“ Antikörper) im Plasma von gesunden Personen identifiziert, die ein Epitop im C-terminalen Bereich (Aß-(21-37)) erkennen und die Bildung von Aß-Aggregaten inhibieren. Die vollständige Primärstruktur und Glykosilierung eines Plaque-protektiven monoklonalen Antikörpers der Maus gegen ein Aß(17-24) – Epitop wurde mittels LC-MS/MS, in Kombination mit Stoss-induzierter Dissoziation (CID) und Elektronentransfer- Dissoziation (ETD) aufgeklärt. Die Strukturen aller 6 CDR-Domänen sowie eine bisher unbekannte Konsensus-Sequenz der Glycosilierung in den leichten Ketten konnten bestimmt werden. Die Glycosilierungsstruktur enthält teilweise hybridisierte Glycane der Mehrantennenstruktur mit einem „High-mannose“-Typ und einem Komplex-Typ, und terminalen *N*-Acetyl- und *N*-Glycolyl-Neuraminyl-Gruppen. Darüber hinaus wurden geringe Mengen von Komplextyp-Glykosilierung in der variablen Region mit Galaktose- und terminalen *N*-Glycolyl-Neuraminyl-Resten identifiziert. Die Glycosilierung im F_c-Teil der schweren Kette wurde mit geringer Heterogenität als Komplextyp-Struktur mit 2-4 Galactosylresten identifiziert.

Weitere Strukturuntersuchungen wurden von humanen, Plaque-protektiven Aß-Autoantikörpern aus Immunglobulinpräparaten nach Affinitätsisolierung mit immobilisiertem Aß(12-40) durchgeführt. Durch Bestimmung der Primärstrukturen konnten alle 4 IgG-Subklassen auf der Ebene der entsprechenden Glycopeptide bestimmt werden.

Die Subklassen-spezifische Glykosilierung ergab relativ erhöhte Anteile an IgG_2/IgG_3 sowie IgG_4. Die Bestimmung der Glykanstruktur der einzelnen IgG-Subklassen ergab ähnliche Glycosilierungen mit leicht erhöhtem Agalactosylanteil.

ß-Amyloid wird durch proteolytische Prozessierung des Transmembran-Amyloid-Vorläuferproteins APP gebildet, dessen genaue Struktur und Funktionen bisher nicht bekannt ist. Obwohl die mögliche Bedeutung von *N*- und/oder *O*-Glykosilierung für die Sezernierung von APP in früheren Arbeiten vermutet wurde, lagen bisher keine Strukturinformationen über die Glycosilierung vor. In Strukturuntersuchungen im Rahmen der vorliegenden Arbeit konnten 3 spezifische *O*-Glycosilierungen an den Threoninresten T291, T292 und T576 mittels LC-MS/MS in Kombination mit ETD- und CID-Fragmentierung aufgeklärt werden. Die *O*-Glycosilierungen wurden als kurze Core-1 Typ-Glycane mit *N*-Acetylgalactosamin und terminalen Silinsäureresten, sowie linearen (T291, T292) und verzweigten Strukturen (T576) identifiziert.

Im letzten Abschnitt der Arbeit wurde die Subklassen-spezifische Glycosilierung von Plasma-Immunglobulin bei Patienten mit Myositose, im Rahmen einer klinisch-diagnostischen Studie zur Ermittlung von Risikofaktoren bei Patienten mit systemisch-rheumatischen Erkrankungen untersucht. Plasma IgG wurde von Patienten, gesunden Zwillingen sowie Kontrollgruppen gleichen Alters mittels Protein G-Affinitätschromatographie isoliert. Die Analyse der Glycosilierungsstruktur ergab statistisch signikante Erhöhungen von fucosylierten Agalactosyl-Glycoformen bei Patienten im Vergleich zur Kontrollgruppe. Diese Ergebnisse deuten auf eine genetische Prädisposition für die Entwicklung von Autoimmun-Erkrankungen, möglicherweise aufgrund von Umweleinflüssen.

6 BIBLIOGRAPHY

1. Woods Schindler, L., *Understanding the immune system*, ed. U.S.D.o.H.a.H. Services. 1991: DIANE Publishing, ISBN 0788115197, 9780788115196. 40.

2. Voet, D., Voet, J.G., *Biochemistry*. 3rd ed. 2004, New York: Wiley.

3. Parham, P., *The immune system*. 2nd ed. 2004: Garland Science. 431.

4. Goldsby, R.A., Kindt, T.J., Osborne, B.A., Kuby, J., *Immunology*. 5th ed. 2003: W.H. Freeman. 603.

5. Amit, A.G., R.A. Mariuzza, S.E. Phillips, and R.J. Poljak, *Three-dimensional structure of an antigen-antibody complex at 2.8 A resolution.* Science, 1986. **233**(4765): p. 747-53.

6. Huber, R., J. Deisenhofer, P.M. Colman, M. Matsushima, and W. Palm, *Crystallographic structure studies of an IgG molecule and an Fc fragment.* Nature, 1976. **264**(5585): p. 415-20.

7. Tonegawa, S., *Somatic generation of antibody diversity.* Nature, 1983. **302**(5909): p. 575-81.

8. Sundberg, E.J. and R.A. Mariuzza, *Molecular recognition in antibody-antigen complexes.* Adv Protein Chem, 2002. **61**: p. 119-60.

9. Jefferis, R. and J. Lund, *Interaction sites on human IgG-Fc for FcgammaR: current models.* Immunol Lett, 2002. **82**(1-2): p. 57-65.

10. Ravetch, J.V. and J.P. Kinet, *Fc receptors.* Annu Rev Immunol, 1991. **9**: p. 457-92.

11. Deisenhofer, J., P.M. Colman, O. Epp, and R. Huber, *Crystallographic structural studies of a human Fc fragment. II. A complete model based on a Fourier map at 3.5 A resolution*. Hoppe Seylers Z Physiol Chem, 1976. **357**(10): p. 1421-34.

12. Deisenhofer, J., *Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution*. Biochemistry, 1981. **20**(9): p. 2361-70.

13. Dwek, R.A., A.C. Lellouch, and M.R. Wormald, *Glycobiology: 'the function of sugar in the IgG molecule'.* J Anat, 1995. **187 (Pt 2)**: p. 279-92.

14. Mizuochi, T., J. Hamako, and K. Titani, *Structures of the sugar chains of mouse immunoglobulin G.* Arch Biochem Biophys, 1987. **257**(2): p. 387-94.

15. Buckley, R.H., *Primary immunodeficiency diseases due to defects in lymphocytes*. N Engl J Med, 2000. **343**(18): p. 1313-24.

16. Buckley, R.H., R.I. Schiff, S.E. Schiff, M.L. Markert, L.W. Williams, T.O. Harville, J.L. Roberts, and J.M. Puck, *Human severe combined immunodeficiency: genetic, phenotypic, and functional diversity in one hundred eight infants.* J Pediatr, 1997. **130**(3): p. 378-87.

17. Candotti, F., L. Notarangelo, R. Visconti, and J. O'Shea, *Molecular aspects of primary immunodeficiencies: lessons from cytokine and other signaling pathways.* J Clin Invest, 2002. **109**(10): p. 1261-9.

18. Hong, R., *The DiGeorge anomaly.* Clin Rev Allergy Immunol, 2001. **20**(1): p. 43-60.

19. Conley, M.E., J.B. Beckwith, J.F.K. Mancer, and L. Tenckhoff, *The spectrum of the DiGeorge syndrome.* The Journal of Pediatrics, 1979. **94**(6): p. 883-890.

20. Conley, M.E., A.K. Dobbs, D.M. Farmer, S. Kilic, K. Paris, S. Grigoriadou, E. Coustan-Smith, V. Howard, and D. Campana, *Primary B Cell Immunodeficiencies: Comparisons and Contrasts.* Annual Review of Immunology, 2009. **27**(1): p. 199.

21. Durandy, A., N. Taubenheim, S. Peron, A. Fischer, and W.A.a.T.H. Frederick, *Pathophysiology of B-Cell Intrinsic Immunoglobulin Class Switch Recombination Deficiencies*, in *Advances in Immunology*. 2007, Academic Press. p. 275-306.

22. Levy, J.A., *HIV pathogenesis and long-term survival.* Aids, 1993. **7**(11): p. 1401-10.

23. Robinson, W.E., D.C. Montefiori, and W.M. Mitchell, *Complement-mediated antibody-dependent enhancement of HIV-1 infection requires CD4 and complement receptors.* Virology, 1990. **175**(2): p. 600-604.

24. Ephgrave, K., *Extra-Intestinal Manifestations of Crohn's Disease.* Surgical Clinics of North America, 2007. **87**(3): p. 673-680.

25. C Scully, T.H.H.L., *Auto-inflammatory syndromes and oral health.* Oral Diseases, 2008. **14**(8): p. 690-699.

26. Akdis, C.A. and M. Akdis, *Mechanisms and treatment of allergic disease in the big picture of regulatory T cells.* Journal of Allergy and Clinical Immunology, 2009. **123**(4): p. 735-746.

27. Pollard, K.M., *Autoantibodies and Autoimmunity: Molecular Mechanisms in Health and Disease.* 2006: Wiley-VCH. 633.

28. Whitacre, C.C., *Sex differences in autoimmune disease.* Nat Immunol, 2001. **2**(9): p. 777-80.

29. Jawaheer, D., M.F. Seldin, C.I. Amos, W.V. Chen, R. Shigeta, J. Monteiro, M. Kern, L.A. Criswell, S. Albani, J.L. Nelson, D.O. Clegg, R. Pope, H.W. Schroeder, Jr., S.L. Bridges, Jr., D.S. Pisetsky, R. Ward, D.L. Kastner, R.L. Wilder, T. Pincus, L.F. Callahan,

D. Flemming, M.H. Wener, and P.K. Gregersen, *A genomewide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases.* Am J Hum Genet, 2001. **68**(4): p. 927-36.

30. Fairweather, D., S. Frisancho-Kiss, and N.R. Rose, *Sex differences in autoimmune disease from a pathological perspective.* Am J Pathol, 2008. **173**(3): p. 600-9.

31. Ostensen, M. and P.M. Villiger, *The remission of rheumatoid arthritis during pregnancy.* Semin Immunopathol, 2007. **29**(2): p. 185-91.

32. Ostensen, M. and P.M. Villiger, *Immunology of pregnancy-pregnancy as a remission inducing agent in rheumatoid arthritis.* Transpl Immunol, 2002. **9**(2-4): p. 155-60.

33. Russell, A.S., C. Johnston, C. Chew, and W.P. Maksymowych, *Evidence for reduced Th1 function in normal pregnancy: a hypothesis for the remission of rheumatoid arthritis.* J Rheumatol, 1997. **24**(6): p. 1045-50.

34. Nelson, J.L., K.A. Hughes, A.G. Smith, B.B. Nisperos, A.M. Branchaud, and J.A. Hansen, *Remission of rheumatoid arthritis during pregnancy and maternal-fetal class II alloantigen disparity.* Am J Reprod Immunol, 1992. **28**(3-4): p. 226-7.

35. Zhu, W.H., C.Z. Lu, Y.M. Huang, H. Link, and B.G. Xiao, *A putative mechanism on remission of multiple sclerosis during pregnancy: estrogen-induced indoleamine 2,3-dioxygenase by dendritic cells.* Mult Scler, 2007. **13**(1): p. 33-40.

36. Amadori, A., R. Zamarchi, G. De Silvestro, G. Forza, G. Cavatton, G.A. Danieli, M. Clementi, and L. Chieco-Bianchi, *Genetic control of the CD4/CD8 T-cell ratio in humans.* Nat Med, 1995. **1**(12): p. 1279-83.

37. Huygen, K. and K. Palfliet, *Strain variation in interferon gamma production of BCG-sensitized mice challenged with PPD II. Importance of one major autosomal locus and additional sexual influences*. Cell Immunol, 1984. **85**(1): p. 75-81.

38. Atassi, M.Z. and P. Casali, *Molecular mechanisms of autoimmunity.* Autoimmunity, 2008. **41**(2): p. 123 - 132.

39. Alzheimer's, A., *2009 Alzheimer's Disease Facts and Figures*, A.s. Association, Editor. 2009. p. 80.

40. Citron, M., *Alzheimer's disease: treatments in discovery and development. Nat Neurosci*, 2002. **5 Suppl**: p. 1055-7.

41. Li, Y. and A. Grupe, *Genetics of late-onset Alzheimer's disease: progress and prospect.* Pharmacogenomics, 2007. **8**(12): p. 1747.

42. Christensen, H., P. Batterham, A. Mackinnon, A. Jorm, H. Mack, K. Mather, K. Anstey, P. Sachdev, and S. Easteal, *The association of APOE genotype and cognitive*

decline in interaction with risk factors in a 65-69 year old community sample. BMC Geriatrics, 2008. **8**(1): p. 14.

43. Birks, J., *Cholinesterase inhibitors for Alzheimer's disease.* Cochrane Database Syst Rev, 2006(1): p. CD005593.

44. Areosa, S.A. and F. Sherriff, *Memantine for dementia.* Cochrane Database Syst Rev, 2003(3): p. CD003154.

45. Lindwall, G. and R.D. Cole, *Phosphorylation affects the ability of tau protein to promote microtubule assembly*. J Biol Chem, 1984. **259**(8): p. 5301-5.

46. Grundke-Iqbal, I., K. Iqbal, Y.C. Tung, M. Quinlan, H.M. Wisniewski, and L.I. Binder, *Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology.* Proc Natl Acad Sci U S A, 1986. **83**(13): p. 4913-7.

47. Hardy, J. and D.J. Selkoe, *The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics*. Science, 2002. **297**(5580): p. 353-6.

48. Selkoe, D.J., *Cell biology of the amyloid beta-protein precursor and the mechanism of Alzheimer's disease.* Annu Rev Cell Biol, 1994. **10**: p. 373-403.

49. Gralle, M. and S.T. Ferreira, *Structure and functions of the human amyloid precursor protein: the whole is more than the sum of its parts. Prog Neurobiol*, 2007. **82**(1): p. 11-32.

50. Goldgaber, D., M.I. Lerman, O.W. McBride, U. Saffiotti, and D.C. Gajdusek, *Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease.* Science, 1987. **235**(4791): p. 877-80.

51. Palmert, M.R., M.B. Podlisny, D.S. Witker, T. Oltersdorf, L.H. Younkin, D.J. Selkoe, and S.G. Younkin, *The beta-amyloid protein precursor of Alzheimer disease has soluble derivatives found in human brain and cerebrospinal fluid.* Proc Natl Acad Sci U S A, 1989. **86**(16): p. 6338-42.

52. Selkoe, D.J., *Alzheimer's disease: genes, proteins, and therapy.* Physiol Rev, 2001. **81**(2): p. 741-66.

53. Selkoe, D.J., *Normal and abnormal biology of the beta-amyloid precursor protein*. Annu Rev Neurosci, 1994. **17**: p. 489-517.

54. Lammich, S., E. Kojro, R. Postina, S. Gilbert, R. Pfeiffer, M. Jasionowski, C. Haass, and F. Fahrenholz, *Constitutive and regulated alpha-secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3922-7.

55. Allinson, T.M., E.T. Parkin, A.J. Turner, and N.M. Hooper, *ADAMs family members as amyloid precursor protein alpha-secretases.* J Neurosci Res, 2003. **74**(3): p. 342-52.

56. Meziane, H., J.C. Dodart, C. Mathis, S. Little, J. Clemens, S.M. Paul, and A. Ungerer, *Memory-enhancing effects of secreted forms of the beta-amyloid precursor protein in normal and amnestic mice.* Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12683-8.

57. Ninomiya, H., J.M. Roch, L.W. Jin, and T. Saitoh, *Secreted form of amyloid beta/A4 protein precursor (APP) binds to two distinct APP binding sites on rat B103 neuron-like cells through two different domains, but only one site is involved in neuritotropic activity*. J Neurochem, 1994. **63**(2): p. 495-500.

58. Smith-Swintosky, V.L., L.C. Pettigrew, S.D. Craddock, A.R. Culwell, R.E. Rydel, and M.P. Mattson, *Secreted forms of beta-amyloid precursor protein protect against ischemic brain injury.* J Neurochem, 1994. **63**(2): p. 781-4.

59. Mullan, M., F. Crawford, K. Axelman, H. Houlden, L. Lilius, B. Winblad, and L. Lannfelt, *A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid.* Nat Genet, 1992. **1**(5): p. 345-7.

60. Chartier-Harlin, M.C., F. Crawford, H. Houlden, A. Warren, D. Hughes, L. Fidani, A. Goate, M. Rossor, P. Roques, J. Hardy, and et al., *Early-onset Alzheimer's disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene.* Nature, 1991. **353**(6347): p. 844-6.

61. Levy, E., M.D. Carman, I.J. Fernandez-Madrid, M.D. Power, I. Lieberburg, S.G. van Duinen, G.T. Bots, W. Luyendijk, and B. Frangione, *Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type.* Science, 1990. **248**(4959): p. 1124-6.

62. Weidemann, A., G. Konig, D. Bunke, P. Fischer, J.M. Salbaum, C.L. Masters, and K. Beyreuther, *Identification, biogenesis, and localization of precursors of Alzheimer's disease A4 amyloid protein.* Cell, 1989. **57**(1): p. 115-26.

63. Pahlsson, P. and S.L. Spitalnik, *The role of glycosylation in synthesis and secretion of beta-amyloid precursor protein by Chinese hamster ovary cells.* Arch Biochem Biophys, 1996. **331**(2): p. 177-86.

64. Sodhi, C.P., R.G. Perez, and N.R. Gottardi-Littell, *Phosphorylation of beta-amyloid precursor protein (APP) cytoplasmic tail facilitates amyloidogenic processing during apoptosis.* Brain Res, 2008. **1198**: p. 204-12.

65. Ando, K., M. Oishi, S. Takeda, K. Iijima, T. Isohara, A.C. Nairn, Y. Kirino, P. Greengard, and T. Suzuki, *Role of phosphorylation of Alzheimer's amyloid precursor protein during neuronal differentiation.* J Neurosci, 1999. **19**(11): p. 4421-7.

66. Tomita, S., Y. Kirino, and T. Suzuki, *Cleavage of Alzheimer's amyloid precursor protein (APP) by secretases occurs after O-glycosylation of APP in the protein secretory pathway. Identification of intracellular compartments in which APP cleavage occurs without using toxic agents that interfere with protein metabolism.* J Biol Chem, 1998. **273**(11): p. 6277-84.

67. Goate, A., M.C. Chartier-Harlin, M. Mullan, J. Brown, F. Crawford, L. Fidani, L. Giuffra, A. Haynes, N. Irving, L. James, and et al., *Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease.* Nature, 1991. **349**(6311): p. 704-6.

68. Levy-Lahad, E., E.M. Wijsman, E. Nemens, L. Anderson, K.A. Goddard, J.L. Weber, T.D. Bird, and G.D. Schellenberg, *A familial Alzheimer's disease locus on chromosome 1.* Science, 1995. **269**(5226): p. 970-3.

69. Sherrington, R., E.I. Rogaev, Y. Liang, E.A. Rogaeva, G. Levesque, M. Ikeda, H. Chi, C. Lin, G. Li, K. Holman, and et al., *Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.* Nature, 1995. **375**(6534): p. 754-60.

70. Mann, D.M., P.O. Yates, and B. Marcyniuk, *Alzheimer's presenile dementia, senile dementia of Alzheimer type and Down's syndrome in middle age form an age related continuum of pathological changes.* Neuropathol Appl Neurobiol, 1984. **10**(3): p. 185-207.

71. Mann, D.M. and M.M. Esiri, *The pattern of acquisition of plaques and tangles in the brains of patients under 50 years of age with Down's syndrome*. J Neurol Sci, 1989. **89**(2-3): p. 169-79.

72. Wisniewski, K.E., A.J. Dalton, C. McLachlan, G.Y. Wen, and H.M. Wisniewski, *Alzheimer's disease in Down's syndrome: clinicopathologic studies.* Neurology, 1985. **35**(7): p. 957-61.

73. Schenk, D., *Amyloid-beta immunotherapy for Alzheimer's disease: the end of the beginning*. Nat Rev Neurosci, 2002. **3**(10): p. 824-8.

74. Weksler, M.E., G. Gouras, N.R. Relkin, and P. Szabo, *The immune system, amyloid-beta peptide, and Alzheimer's disease.* Immunol Rev, 2005. **205**: p. 244-56.

75. Schenk, D.B., P. Seubert, M. Grundman, and R. Black, *A beta immunotherapy: Lessons learned for potential treatment of Alzheimer's disease.* Neurodegener Dis, 2005. **2**(5): p. 255-60.

76. Vasilevko, V. and D.H. Cribbs, *Novel approaches for immunotherapeutic intervention in Alzheimer's disease*. Neurochem Int, 2006. **49**(2): p. 113-26.

77. Gelinas, D.S., K. DaSilva, D. Fenili, P. St George-Hyslop, and J. McLaurin, *Immunotherapy for Alzheimer's disease.* Proc Natl Acad Sci U S A, 2004. **101 Suppl 2**: p. 14657-62.

78. Monsonego, A. and H.L. Weiner, *Immunotherapeutic approaches to Alzheimer's disease.* Science, 2003. **302**(5646): p. 834-8.

79. McLaurin, J., R. Cecal, M.E. Kierstead, X. Tian, A.L. Phinney, M. Manea, J.E. French, M.H. Lambermon, A.A. Darabie, M.E. Brown, C. Janus, M.A. Chishti, P. Horne, D. Westaway, P.E. Fraser, H.T. Mount, M. Przybylski, and P. St George-Hyslop, *Therapeutically effective antibodies against amyloid-beta peptide target amyloid-beta residues 4-10 and inhibit cytotoxicity and fibrillogenesis.* Nat Med, 2002. **8**(11): p. 1263-9.

80. Check, E., *Nerve inflammation halts trial for Alzheimer's drug*. Nature, 2002. **415**(6871): p. 462.

81. Orgogozo, J.M., S. Gilman, J.F. Dartigues, B. Laurent, M. Puel, L.C. Kirby, P. Jouanny, B. Dubois, L. Eisner, S. Flitman, B.F. Michel, M. Boada, A. Frank, and C. Hock, *Subacute meningoencephalitis in a subset of patients with AD after Abeta42 immunization.* Neurology, 2003. **61**(1): p. 46-54.

82. Brody, D.L. and D.M. Holtzman, *Active and passive immunotherapy for neurodegenerative disorders.* Annu Rev Neurosci, 2008. **31**: p. 175-93.

83. Du, Y., R. Dodel, H. Hampel, K. Buerger, S. Lin, B. Eastwood, K. Bales, F. Gao, H.J. Moeller, W. Oertel, M. Farlow, and S. Paul, *Reduced levels of amyloid beta-peptide antibody in Alzheimer disease.* Neurology, 2001. **57**(5): p. 801-5.

84. Gaskin, F., J. Finley, Q. Fang, S. Xu, and S.M. Fu, *Human antibodies reactive with beta-amyloid protein in Alzheimer's disease.* J Exp Med, 1993. **177**(4): p. 1181-6.

85. Xu, S. and F. Gaskin, *Increased incidence of anti-beta-amyloid autoantibodies secreted by Epstein-Barr virus transformed B cell lines from patients with Alzheimer's disease.* Mech Ageing Dev, 1997. **94**(1-3): p. 213-22.

86. Dodel, R., H. Hampel, C. Depboylu, S. Lin, F. Gao, S. Schock, S. Jackel, X. Wei, K. Buerger, C. Hoft, B. Hemmer, H.J. Moller, M. Farlow, W.H. Oertel, N. Sommer, and Y. Du, *Human antibodies against amyloid beta peptide: a potential treatment for Alzheimer's disease.* Ann Neurol, 2002. **52**(2): p. 253-6.

87. Dodel, R.C., H. Hampel, and Y. Du, *Immunotherapy for Alzheimer's disease.* Lancet Neurol, 2003. **2**(4): p. 215-20.

88. Przybylski, M., Stefanescu, R., Manea, M., Perdivara, I., Cozma, C., Moise, A., Paraschiv, G., Juszczyk, P., Marquardt, M., *New molecular approaches for immunotherapy and diagnosis of Alzheimer's disease based on epitope-specific serum beta-amyloid antibodies.* 7th Austral. Pept. Symposium, Cairns, abstr. p.32, 2007.

89. Przybylski, M., Stefanescu, R., Manea, M., Bacher, M., Dodel, R., *Diagnosis and treatment of Alzheimer's and other neurodementing diseases.* EPA&US Patents, UBK 003/004, University of Konstanz, University of Marburg, PCT Patent application 029860-0183, 2008.

90. Przybylski, M., Stefanescu, R., Manea, M., Perdivara, I., Cozma, C., Moise, A., Paraschiv, G., Juszczyk, P., Marquardt, M., Nature (submitted for publication).

91. Stefanescu, R., *Molecular Identification of Antigen Recognition Structures in Immune Complexes for Immunotherapeutic Applications by Proteolytic and Mass Spectrometric Methods*, in Department of Chemistry. 2007, University of Konstanz Konstanz. p. 175.

92. Weksler, M.E., N. Relkin, R. Turkenich, S. LaRusse, L. Zhou, and P. Szabo, *Patients with Alzheimer disease have lower levels of serum anti-amyloid peptide antibodies than healthy elderly individuals.* Exp Gerontol, 2002. **37**(7): p. 943-8.

93. Dodel, R.C., Y. Du, C. Depboylu, H. Hampel, L. Frolich, A. Haag, U. Hemmeter, S. Paulsen, S.J. Teipel, S. Brettschneider, A. Spottke, C. Nolker, H.J. Moller, X. Wei, M. Farlow, N. Sommer, and W.H. Oertel, *Intravenous immunoglobulins containing antibodies against beta-amyloid for the treatment of Alzheimer's disease.* J Neurol Neurosurg Psychiatry, 2004. **75**(10): p. 1472-4.

94. Limaye, V.S., P. Blumbergs, and P.J. Roberts-Thomson, *Idiopathic Inflammatory Myopathies.* Intern Med J, 2008.

95. Wiendl, H., *Idiopathic inflammatory myopathies: current and future therapeutic options*. Neurotherapeutics, 2008. **5**(4): p. 548-57.

96. Plotz, P.H., L.G. Rider, I.N. Targoff, N. Raben, T.P. O'Hanlon, and F.W. Miller, *Myositis: Immunologic Contributions to Understanding Cause, Pathogenesis, and Therapy.* Ann Intern Med, 1995. **122**(9): p. 715-724.

97. Miller, F.W., L.G. Rider, P.H. Plotz, D.A. Isenberg, and C.V. Oddis, *Diagnostic criteria for polymyositis and dermatomyositis.* The Lancet, 2003. **362**(9397): p. 1762-1763.

98. Ginn, L.R., J.P. Lin, P.H. Plotz, S.J. Bale, R.L. Wilder, A. Mbauya, and F.W. Miller, *Familial autoimmunity in pedigrees of idiopathic inflammatory myopathy patients suggests*

common genetic risk factors for many autoimmune diseases. Arthritis Rheum, 1998. **41**(3): p. 400-5.

99. Love, L.A., Miller, F.W., *Understanding the idiopathic inflammatory myopathies.* Contemp Int Med, 1995. **7**: p. 29-43.

100. Miller, F.W., Myositis-specific autoantibodies. *Touchstones for understanding the inflammatory myopathies*. Jama, 1993. **270**(15): p. 1846-9.

101. Targoff, I.N., *Immune manifestations of inflammatory muscle disease.* Rheum Dis Clin North Am, 1994. **20**(4): p. 857-80.

102. Miller, F.W., K.A. Waite, T. Biswas, and P.H. Plotz, *The role of an autoantigen, histidyl-tRNA synthetase, in the induction and maintenance of autoimmunity.* Proc Natl Acad Sci U S A, 1990. **87**(24): p. 9933-7.

103. Miller, F.W., S.A. Twitty, T. Biswas, and P.H. Plotz, *Origin and regulation of a disease-specific autoantibody response. Antigenic epitopes, spectrotype stability, and isotype restriction of anti-Jo-1 autoantibodies.* J Clin Invest, 1990. **85**(2): p. 468-75.

104. Love, L.A., R.L. Leff, D.D. Fraser, I.N. Targoff, M. Dalakas, P.H. Plotz, and F.W. Miller, *A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups.* Medicine (Baltimore), 1991. **70**(6): p. 360-74.

105. Mathews, M.B. and R.M. Bernstein, *Myositis autoantibody inhibits histidyl-tRNA synthetase: a model for autoimmunity.* Nature, 1983. **304**(5922): p. 177-9.

106. Nagaraju, K., N. Raben, L. Loeffler, T. Parker, P.J. Rochon, E. Lee, C. Danning, R. Wada, C. Thompson, G. Bahtiyar, J. Craft, R. Hooft Van Huijsduijnen, and P. Plotz, *Conditional up-regulation of MHC class I in skeletal muscle leads to self-sustaining autoimmune myositis and myositis-specific autoantibodies.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 9209-14.

107. Targoff, I.N., A.E. Johnson, and F.W. Miller, *Antibody to signal recognition particle in polymyositis.* Arthritis Rheum, 1990. **33**(9): p. 1361-70.

108. Hirakata, M., Matsuura, Y., Suwa, A., et al., *Immunological and histopathological features of patients with anti-SRP autoantibodies.* Arthritis Rheum, 1997. **40**: p. S146.

109. Mierau, R., T. Dick, P. Bartz-Bazzanella, E. Keller, E.D. Albert, and E. Genth, *Strong association of dermatomyositis-specific Mi-2 autoantibodies with a tryptophan at position 9 of the HLA-DR beta chain.* Arthritis Rheum, 1996. **39**(5): p. 868-76.

110. Targoff, I.N., *Idiopathic inflammatory myopathy: autoantibody update.* Curr Rheumatol Rep, 2002. **4**(5): p. 434-41.

111. Casciola-Rosen, L., F. Andrade, D. Ulanet, W.B. Wong, and A. Rosen, *Cleavage by granzyme B is strongly predictive of autoantigen status: implications for initiation of autoimmunity.* J Exp Med, 1999. **190**(6): p. 815-26.

112. Rosen, A. and L. Casciola-Rosen, *Autoantigens as substrates for apoptotic proteases: implications for the pathogenesis of systemic autoimmune disease*. Cell Death Differ, 1999. **6**(1): p. 6-12.

113. Kalovidouris, A.E., *Immune aspects of myositis.* Curr Opin Rheumatol, 1992. **4**(6): p. 809-14.

114. Engel, A.G., K. Arahata, and A. Emslie-Smith, *Immune effector mechanisms in inflammatory myopathies.* Res Publ Assoc Res Nerv Ment Dis, 1990. **68**: p. 141-57.

115. Mantegazza, R., F. Andreetta, P. Bernasconi, F. Baggi, J.R. Oksenberg, O. Simoncini, M. Mora, F. Cornelio, and L. Steinman, *Analysis of T cell receptor repertoire of muscle-infiltrating T lymphocytes in polymyositis. Restricted V alpha/beta rearrangements may indicate antigen-driven selection.* J Clin Invest, 1993. **91**(6): p. 2880-6.

116. Oddis, C.V., *Therapy for myositis.* Curr Opin Rheumatol, 1991. **3**(6): p. 919-24.

117. Edman, K., P. Nollert, A. Royant, H. Belrhali, E. Pebay-Peyroula, J. Hajdu, R. Neutze, and E.M. Landau, *High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle.* Nature, 1999. **401**(6755): p. 822-6.

118. Burzlaff, N.I., P.J. Rutledge, I.J. Clifton, C.M. Hensgens, M. Pickford, R.M. Adlington, P.L. Roach, and J.E. Baldwin, *The reaction cycle of isopenicillin N synthase observed by X-ray diffraction.* Nature, 1999. **401**(6754): p. 721-4.

119. Wimberly, B.T., D.E. Brodersen, W.M. Clemons, Jr., R.J. Morgan-Warren, A.P. Carter, C. Vonrhein, T. Hartsch, and V. Ramakrishnan, *Structure of the 30S ribosomal subunit.* Nature, 2000. **407**(6802): p. 327-39.

120. Simpson, A.A., Y. Tao, P.G. Leiman, M.O. Badasso, Y. He, P.J. Jardine, N.H. Olson, M.C. Morais, S. Grimes, D.L. Anderson, T.S. Baker, and M.G. Rossmann, *Structure of the bacteriophage phi29 DNA packaging motor.* Nature, 2000. **408**(6813): p. 745-50.

121. Ban, N., P. Nissen, J. Hansen, P.B. Moore, and T.A. Steitz, *The complete atomic structure of the large ribosomal subunit at 2.4 A resolution.* Science, 2000. **289**(5481): p. 905-20.

122. Kumar, A., R.R. Ernst, and K. Wuthrich, *A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules.* Biochem Biophys Res Commun, 1980. **95**(1): p. 1-6.

123. Williamson, M.P., T.F. Havel, and K. Wuthrich, *Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry.* J Mol Biol, 1985. **182**(2): p. 295-315.

124. Wagner, G. and K. Wuthrich, *Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor.* J Mol Biol, 1982. **155**(3): p. 347-66.

125. Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong, and C.M. Whitehouse, *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.

126. John B. Fenn, M.M.C.K.M.S.F.W.C.M.W., *Electrospray ionization-principles and practice.* Mass Spectrometry Reviews, 1990. **9**(1): p. 37-70.

127. Hillenkamp, F., M. Karas, R.C. Beavis, and B.T. Chait, *Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers.* Anal Chem, 1991. **63**(24): p. 1193A-1203A.

128. Joseph, A.L., *Studying noncovalent protein complexes by electrospray ionization mass spectrometry.* Mass Spectrometry Reviews, 1997. **16**(1): p. 1-23.

129. Albert J. R. Heck, R.H.H.v.d.H., *Investigation of intact protein complexes by mass spectrometry.* Mass Spectrometry Reviews, 2004. **23**(5): p. 368-389.

130. Benesch, J.L.P., B.T. Ruotolo, D.A. Simmons, and C.V. Robinson, *Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics.* Chemical Reviews, 2007. **107**(8): p. 3544-3567.

131. Uetrecht, C., C. Versluis, N.R. Watts, W.H. Roos, G.J.L. Wuite, P.T. Wingfield, A.C. Steven, and A.J.R. Heck, *High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids. Proceedings of the National Academy of Sciences*, 2008. **105**(27): p. 9216-9220.

132. Stephen D. Fuerstenau, W.H.B.J.J.T.C.B.B.B.G.S., *Mass Spectrometry of an Intact Virus13.* Angewandte Chemie International Edition, 2001. **40**(3): p. 541-544.

133. Tito, M.A., K. Tars, K. Valegard, J. Hajdu, and C.V. Robinson, *Electrospray Time-of-Flight Mass Spectrometry of the Intact MS2 Virus Capsid.* Journal of the American Chemical Society, 2000. **122**(14): p. 3550-3551.

134. Suckau, D., M. Mak, and M. Przybylski, *Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping.* Proc Natl Acad Sci U S A, 1992. **89**(12): p. 5630-4.

135. Kalkum, M., M. Przybylski, and M.O. Glocker, *Structure Characterization of Functional Histidine Residues and Carbethoxylated Derivatives in Peptides and Proteins by Mass Spectrometry.* Bioconjugate Chemistry, 1998. **9**(2): p. 226-235.

136. Fiedler, W., C. Borchers, M. Macht, S.-O. Deininger, and M. Przybylski, *Molecular Characterization of a Conformational Epitope of Hen Egg White Lysozyme by Differential Chemical Modification of Immune Complexes and Mass Spectrometric Peptide Mapping*. Bioconjugate Chemistry, 1998. **9**(2): p. 236-241.

137. Happersberger, H.P., M. Przybylski, and M.O. Glocker, *Selective bridging of bis-cysteinyl residues by arsonous acid derivatives as an approach to the characterization of protein tertiary structures and folding pathways by mass spectrometry.* Anal Biochem, 1998. **264**(2): p. 237-50.

138. Fligge, T.A., K. Bruns, and M. Przybylski, *Analytical development of electrospray and nanoelectrospray mass spectrometry in combination with liquid chromatography for the characterization of proteins.* J Chromatogr B Biomed Sci Appl, 1998. **706**(1): p. 91-100.

139. Smedley, J.G., 3rd, J.S. Sharp, J.F. Kuhn, and K.B. Tomer, *Probing the pH-dependent prepore to pore transition of Bacillus anthracis protective antigen with differential oxidative protein footprinting.* Biochemistry, 2008. **47**(40): p. 10694-704.

140. Sanjay Venkatesh, K.B.T.J.S.S., *Rapid identification of oxidation-induced conformational changes by kinetic analysis.* Rapid Communications in Mass Spectrometry, 2007. **21**(23): p. 3927-3936.

141. Sharp, J.S. and K.B. Tomer, *Analysis of the oxidative damage-induced conformational changes of apo- and holocalmodulin by dose-dependent protein oxidative surface mapping.* Biophys J, 2007. **92**(5): p. 1682-92.

142. Sharp, J.S., D.M. Sullivan, J. Cavanagh, and K.B. Tomer, *Measurement of multisite oxidation kinetics reveals an active site conformational change in Spo0F as a result of protein oxidation*. Biochemistry, 2006. **45**(20): p. 6260-6.

143. Sharp, J.S., J.M. Becker, and R.L. Hettich, *Protein surface mapping by chemical oxidation: structural analysis by mass spectrometry.* Anal Biochem, 2003. **313**(2): p. 216-25.

144. Kalkhof, S. and A. Sinz, *Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters*. Anal Bioanal Chem, 2008. **392**(1-2): p. 305-12.

145. Mouradov, D., G. King, I.L. Ross, J.K. Forwood, D.A. Hume, A. Sinz, J.L. Martin, B. Kobe, and T. Huber, *Protein structure determination using a combination of cross-linking, mass spectrometry, and molecular modeling.* Methods Mol Biol, 2008. **426**: p. 459-74.

146. Ihling, C., A. Schmidt, S. Kalkhof, D.M. Schulz, C. Stingl, K. Mechtler, M. Haack, A.G. Beck-Sickinger, D.M. Cooper, and A. Sinz, *Isotope-labeled cross-linkers and Fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex.* J Am Soc Mass Spectrom, 2006. **17**(8): p. 1100-13.

147. Andrea, S., *Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions.* Mass Spectrometry Reviews, 2006. **25**(4): p. 663-682.

148. Suckau, D., J. Kohl, G. Karwath, K. Schneider, M. Casaretto, D. Bitter-Suermann, and M. Przybylski, *Molecular epitope identification by limited proteolysis of an immobilized antigen-antibody complex and mass spectrometric peptide mapping*. Proc Natl Acad Sci U S A, 1990. **87**(24): p. 9848-52.

149. Macht, M., W. Fiedler, K. Kurzinger, and M. Przybylski, *Mass spectrometric mapping of protein epitope structures of myocardial infarct markers myoglobin and troponin T*. Biochemistry, 1996. **35**(49): p. 15633-9.

150. Iacob, R.E., Z. Keck, O. Olson, S.K. Foung, and K.B. Tomer, *Structural elucidation of critical residues involved in binding of human monoclonal antibodies to hepatitis C virus E2 envelope glycoprotein*. Biochim Biophys Acta, 2008. **1784**(3): p. 530-42.

151. Williams, J.G., K.B. Tomer, C.E. Hioe, S. Zolla-Pazner, and P.J. Norris, *The antigenic determinants on HIV p24 for CD4+ T cell inhibiting antibodies as determined by limited proteolysis, chemical modification, and mass spectrometry.* J Am Soc Mass Spectrom, 2006. **17**(11): p. 1560-9.

152. Parker, C.E., L.J. Deterding, C. Hager-Braun, J.M. Binley, N. Schulke, H. Katinger, J.P. Moore, and K.B. Tomer, *Fine Definition of the Epitope on the gp41 Glycoprotein of Human Immunodeficiency Virus Type 1 for the Neutralizing Monoclonal Antibody 2F5*. J. Virol., 2001. **75**(22): p. 10906-10911.

153. Hager-Braun, C. and K.B. Tomer, *Determination of protein-derived epitopes by mass spectrometry.* Expert Rev Proteomics, 2005. **2**(5): p. 745-56.

154. Manea, M., A. Kalaszi, G. Mezo, K. Horvati, A. Bodor, A. Horvath, O. Farkas, A. Perczel, M. Przybylski, and F. Hudecz, *Antibody recognition and conformational flexibility of a plaque-specific beta-amyloid epitope modulated by non-native peptide flanking regions. J* Med Chem, 2008. **51**(5): p. 1150-61.

155. Manea, M., M. Przybylski, F. Hudecz, and G. Mezo, *Design, structural, and immuno-analytical properties of antigenic bioconjugates comprising a beta-amyloid-plaque specific epitope.* Biopolymers, 2008. **90**(2): p. 94-104.

156. Marquardt, A., S. Muyldermans, and M. Przybylski, *A synthetic camel anti-lysozyme peptide antibody (peptibody) with flexible loop structure identified by high-resolution affinity mass spectrometry.* Chemistry, 2006. **12**(7): p. 1915-23.

157. Hamers-Casterman, C., T. Atarhouch, S. Muyldermans, G. Robinson, C. Hamers, E.B. Songa, N. Bendahman, and R. Hamers, *Naturally occurring antibodies devoid of light chains.* Nature, 1993. **363**(6428): p. 446-8.

158. Han, X., M. Jin, K. Breuker, and F.W. McLafferty, *Extending top-down mass spectrometry to proteins with masses greater than 200 kilodaltons.* Science, 2006. **314**(5796): p. 109-12.

159. Breuker, K., M. Jin, X. Han, H. Jiang, and F.W. McLafferty, *Top-down identification and characterization of biomolecules by mass spectrometry.* J Am Soc Mass Spectrom, 2008. **19**(8): p. 1045-53.

160. McLafferty, F.W., K. Breuker, M. Jin, X. Han, G. Infusini, H. Jiang, X. Kong, and T.P. Begley, *Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics.* Febs J, 2007. **274**(24): p. 6256-68.

161. Zhang, Z., H. Pan, and X. Chen, *Mass spectrometry for structural characterization of therapeutic antibodies.* Mass Spectrom Rev, 2009. **28**(1): p. 147-76.

162. Perdivara, I., L. Deterding, A. Moise, K.B. Tomer, and M. Przybylski, *Determination of primary structure and microheterogeneity of a beta-amyloid plaque-specific antibody using high-performance LC-tandem mass spectrometry*. Anal Bioanal Chem, 2008. **391**(1): p. 325-36.

163. Nitanai, Y., Satow, Y., Adachi, H., Tsujimoto, M., *Crystallization and preliminary X-ray investigation of a glycoprotein, human renal dipeptidase* Journal of Crystal Growth, 1996. **168**(1-4): p. 280-283.

164. Zubarev, R.A., Kelleher N. L., McLafferty , F. W., *Electron Capture Dissociation of Multiply Charged Protein Cations - a Nonergodic Process.* J Am Chem Soc, 1998. **120**: p. 3265-3266.

165. Beu, S.C., Senko, M. W., Quinn, J. P., Wampler, F. M., McLafferty, F. W. , *Fourier-transform electrospray instrumentation for tandem high-resolution mass spectrometry of large molecules.* J Am Soc Mass Spectrom, 1993. **4**: p. 557-565.

166. Syka, J.E., J.J. Coon, M.J. Schroeder, J. Shabanowitz, and D.F. Hunt, *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9528-33.

167. Pitteri, S.J., P.A. Chrisman, J.M. Hogan, and S.A. McLuckey, *Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: reactions of doubly and triply protonated peptides with SO2*.* Anal Chem, 2005. **77**(6): p. 1831-9.

168. Nilsson, C.L. and P. Davidsson, *New separation tools for comprehensive studies of protein expression by mass spectrometry.* Mass Spectrom Rev, 2000. **19**(6): p. 390-7.

169. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis*. Anal Chem, 2000. **72**(6): p. 1156-62.

170. Qizhi, H., J.N. Robert, L. Hongyan, M. Alexander, H. Mark, and R.G. Cooks, *The Orbitrap: a new mass spectrometer.* Journal of Mass Spectrometry, 2005. **40**(4): p. 430-443.

171. de Hoffmann, E., Stroobant, V., *Mass spectrometry: principles and applications* 2nd ed. 2003, Toronto: John Wiely & Sons, Ltd. p65.

172. Wiley, W.C., McLaren, I.H., *Time-of-flight mass spectrometer with improved resolution.* The Review of Scientific Instruments, 1955. **26**: p. p1150.

173. Cotter, R.J., *Time-of-flight mass spectrometry for the structural analysis of biological molecules.* Anal Chem, 1992. **64**(21): p. 1027A-1039A.

174. Mamyrin, B.A., *Laser assisted reflectron time-of-flight mass spectrometry.* International Journal of Mass Spectrometry and Ion Processes, 1994. **131**: p. 1-19.

175. Yost, R.A. and C.G. Enke, *Selected ion fragmentation with a tandem quadrupole mass spectrometer.* Journal of the American Chemical Society, 1978. **100**(7): p. 2274-2275.

176. Morris, H.R., T. Paxton, A. Dell, J. Langhorne, M. Berg, R.S. Bordoli, J. Hoyes, and R.H. Bateman, *High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer.* Rapid Commun Mass Spectrom, 1996. **10**(8): p. 889-96.

177. Giles, K., Pringle, S., Worthington, K., Bateman, R. *Travelling wave ion propulsion in collision cells.* in 51st Annual Meeting of the American Society for Mass Spectrometry. 2003. Montreal, Canada.

178. De Hoffmann, v.E., Stroobant, V., *Mass Spectrometry: principles and applications*. 3rd ed. 2007: Wiley-Interscience. 489.

179. Standing, K.G., *Peptide and protein de novo sequencing by mass spectrometry.* Curr Opin Struct Biol, 2003. **13**(5): p. 595-601.

180. Bandeira, N., V. Pham, P. Pevzner, D. Arnott, and J.R. Lill, *Automated de novo protein sequencing of monoclonal antibodies.* Nat Biotechnol, 2008. **26**(12): p. 1336-8.

181. Johnson, R.S. and J.A. Taylor, *Searching sequence databases via de novo peptide sequencing by tandem mass spectrometry*. Methods Mol Biol, 2000. **146**: p. 41-61.

182. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides.* Biomed Mass Spectrom, 1984. **11**(11): p. 601.

183. Biemann, K., *Contributions of mass spectrometry to peptide and protein structure.* Biomed Environ Mass Spectrom, 1988. **16**(1-12): p. 99-111.

184. Johnson, R.S., S.A. Martin, K. Biemann, J.T. Stults, and J.T. Watson, *Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine*. Anal Chem, 1987. **59**(21): p. 2621-5.

185. Paizs, B. and S. Suhai, *Fragmentation pathways of protonated peptides.* Mass Spectrom Rev, 2005. **24**(4): p. 508-48.

186. Summerfield, S.G., Gaskell, S.J., *Fragmentation efficiencies of peptide ions following low energy collisional activation.* International Journal of Mass Spectrometry and Ion Processes, 1997. **165-166**: p. 509-521.

187. Jones, J.L., A.R. Dongre, A. Somogyi, and V.H. Wysocki, *Sequence Dependence of Peptide Fragmentation Efficiency Curves Determined by Electrospray Ionization/Surface-Induced Dissociation Mass Spectrometry.* Journal of the American Chemical Society, 1994. **116**(18): p. 8368-8369.

188. Summerfield, S.G., Whiting, A., Gaskell, S.J., *Intra-ionic interactions in electrosprayed peptide ions.* International Journal of Mass Spectrometry and Ion Processes, 1997. **162**: p. 149-161.

189. Summerfield, S.G., Cox, K.A., Gaskell, S.J., *The promotion of d-type ions during the low energy collision-induced dissociation of some cysteic acid-containing peptides.* Journal of the American Society for Mass Spectrometry, 1997. **8**(1): p. 25-31.

190. McLafferty, F.W., D.M. Horn, K. Breuker, Y. Ge, M.A. Lewis, B. Cerda, R.A. Zubarev, and B.K. Carpenter, *Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance.* J Am Soc Mass Spectrom, 2001. **12**(3): p. 245-9.

191. Zubarev, R.A., D.M. Horn, E.K. Fridriksson, N.L. Kelleher, N.A. Kruger, M.A. Lewis, B.K. Carpenter, and F.W. McLafferty, *Electron capture dissociation for structural characterization of multiply charged protein cations.* Anal Chem, 2000. **72**(3): p. 563-73.

192. Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty, *Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process.* Journal of the American Chemical Society, 1998. **120**(13): p. 3265-3266.

193. Hunt, D.F., G.C. Stafford, F.W. Crow, and J.W. Russell, *Pulsed positive negative ion chemical ionization mass spectrometry.* Analytical Chemistry, 1976. **48**(14): p. 2098-2104.

194. Hunt, D.F. and F.W. Crow, *Electron capture negative ion chemical ionization mass spectrometry.* Analytical Chemistry, 1978. **50**(13): p. 1781-1784.

195. Kohler, G. and C. Milstein, *Continuous cultures of fused cells secreting antibody of predefined specificity*. Nature, 1975. **256**(5517): p. 495-7.

196. Chadd, H.E. and S.M. Chamow, *Therapeutic antibody expression technology.* Curr Opin Biotechnol, 2001. **12**(2): p. 188-94.

197. Vaughan, T.J., J.K. Osbourn, and P.R. Tempest, *Human antibodies by design.* Nat Biotechnol, 1998. **16**(6): p. 535-9.

198. Brekke, O.H. and I. Sandlie, *Therapeutic antibodies for human diseases at the dawn of the twenty-first century.* Nat Rev Drug Discov, 2003. **2**(1): p. 52-62.

199. Schrama, D., R.A. Reisfeld, and J.C. Becker, *Antibody targeted drugs as cancer therapeutics.* Nat Rev Drug Discov, 2006. **5**(2): p. 147-59.

200. Terai, K., A. Iwai, S. Kawabata, Y. Tasaki, T. Watanabe, K. Miyata, and T. Yamaguchi, *beta-amyloid deposits in transgenic mice expressing human beta-amyloid precursor protein have the same characteristics as those in Alzheimer's disease.* Neuroscience, 2001. **104**(2): p. 299-310.

201. Maddalena, A.S., A. Papassotiropoulos, C. Gonzalez-Agosti, A. Signorell, T. Hegi, T. Pasch, R.M. Nitsch, and C. Hock, *Cerebrospinal fluid profile of amyloid beta peptides in patients with Alzheimer's disease determined by protein biochip technology.* Neurodegener Dis, 2004. **1**(4-5): p. 231-5.

202. Tian, X., R. Cecal, J. McLaurin, M. Manea, R. Stefanescu, S. Grau, M. Harnasch, S. Amir, M. Ehrmann, P. St George-Hyslop, M. Kohlmann, and M. Przybylski, *Identification and structural characterisation of carboxy-terminal polypeptides and antibody epitopes of Alzheimer's amyloid precursor protein using high-resolution mass spectrometry.* Eur J Mass Spectrom (Chichester, Eng), 2005. **11**(5): p. 547-56.

203. Manea, M., F. Hudecz, M. Przybylski, and G. Mezo, *Synthesis, solution conformation, and antibody recognition of oligotuftsin-based conjugates containing a beta-amyloid(4-10) plaque-specific epitope.* Bioconjug Chem, 2005. **16**(4): p. 921-8.

204. Kabat EA, W.T., Reid-Miller M, Perry HM, Gottesman KS, *Sequences of proteins of immunological interest*. 4th ed: Bethesda: Department of Health and Human Services.

205. Kabat, E.A. and T.T. Wu, *Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites.* J Immunol, 1991. **147**(5): p. 1709-19.

206. Johnson, A., *Molecular Biology of the Cell*. 4th ed. 2002: Garland Science.

207. Harris, R.J., *Processing of C-terminal lysine and arginine residues of proteins isolated from mammalian cell culture.* J Chromatogr A, 1995. **705**(1): p. 129-34.

208. Beck, A., M.C. Bussat, N. Zorn, V. Robillard, C. Klinguer-Hamour, S. Chenu, L. Goetsch, N. Corvaia, A. Van Dorsselaer, and J.F. Haeuw, *Characterization by liquid chromatography combined with mass spectrometry of monoclonal anti-IGF-1 receptor antibodies produced in CHO and NS0 cells.* J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **819**(2): p. 203-18.

209. Chelius, D., K. Jing, A. Lueras, D.S. Rehder, T.M. Dillon, A. Vizel, R.S. Rajan, T. Li, M.J. Treuheit, and P.V. Bondarenko, *Formation of pyroglutamic acid from N-terminal glutamic acid in immunoglobulin gamma antibodies.* Anal Chem, 2006. **78**(7): p. 2370-6.

210. Yu, L., R.L. Remmele, Jr., and B. He, *Identification of N-terminal modification for recombinant monoclonal antibody light chain using partial reduction and quadrupole time-of-flight mass spectrometry.* Rapid Commun Mass Spectrom, 2006. **20**(24): p. 3674-80.

211. Bolukbasi Hatip, F.F., Y. Matsunaga, and T. Yamada, *Specific reactivity of mild/severe Alzheimer's disease patient's sera to antibody against Abeta1-40 epitope 17-21.* Acta Neurol Scand, 2008. **117**(6): p. 404-8.

212. O'Hair, R.A.J., Reid, G.E.,, *Neighboring groups versus cis-elimination mechanisms for side chain loss from protonated methionine, methionine sulfoxide and their peptides.* Eur J Mass Spectrom 1999. **5**: p. 325-334.

213. Clark, M.R., *IgG effector mechanisms.* Chem Immunol, 1997. **65**: p. 88-110.

214. Reinders, J., U. Lewandrowski, J. Moebius, Y. Wagner, and A. Sickmann, *Challenges in mass spectrometry-based proteomics.* Proteomics, 2004. **4**(12): p. 3686-703.

215. Zaia, J., *Mass spectrometry of oligosaccharides.* Mass Spectrom Rev, 2004. **23**(3): p. 161-227.

216. Park, Y. and C.B. Lebrilla, *Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides.* Mass Spectrom Rev, 2005. **24**(2): p. 232-64.

217. Harvey, D.J., *Collision-induced fragmentation of underivatized N-linked carbohydrates ionized by electrospray.* J Mass Spectrom, 2000. **35**(10): p. 1178-90.

218. Mechref, Y. and M.V. Novotny, *Structural investigations of glycoconjugates at high sensitivity.* Chem Rev, 2002. **102**(2): p. 321-69.

219. Wuhrer, M., A.M. Deelder, and C.H. Hokke, *Protein glycosylation analysis by liquid chromatography-mass spectrometry.* J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **825**(2): p. 124-33.

220. Ciucanu, I., *Per-O-methylation reaction for structural analysis of carbohydrates by mass spectrometry.* Anal Chim Acta, 2006. **576**(2): p. 147-55.

221. Hirabayashi, J., *Lectin-based structural glycomics: glycoproteomics and glycan profiling*. Glycoconj J, 2004. **21**(1-2): p. 35-40.

222. Kaji, H., H. Saito, Y. Yamauchi, T. Shinkawa, M. Taoka, J. Hirabayashi, K. Kasai, N. Takahashi, and T. Isobe, *Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins.* Nat Biotechnol, 2003. **21**(6): p. 667-72.

223. Ritchie, M.A., A.C. Gill, M.J. Deery, and K. Lilley, *Precursor ion scanning for detection and structural characterization of heterogeneous glycopeptide mixtures*. J Am Soc Mass Spectrom, 2002. **13**(9): p. 1065-77.

224. Sullivan, B., T.A. Addona, and S.A. Carr, *Selective detection of glycopeptides on ion trap mass spectrometers.* Anal Chem, 2004. **76**(11): p. 3112-8.

225. Itoh, S., N. Kawasaki, M. Ohta, and T. Hayakawa, *Structural analysis of a glycoprotein by liquid chromatography-mass spectrometry and liquid chromatography with tandem mass spectrometry. Application to recombinant human thrombomodulin.* J Chromatogr A, 2002. **978**(1-2): p. 141-52.

226. Wang, F., A. Nakouzi, R.H. Angeletti, and A. Casadevall, *Site-specific characterization of the N-linked oligosaccharides of a murine immunoglobulin M by high-performance liquid chromatography/electrospray mass spectrometry.* Anal Biochem, 2003. **314**(2): p. 266-80.

227. Domon, B., Costello, C.E., A Systematic *Nomenclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates.* Glycoconjugate J, 1988. **5**: p. 397-409.

228. Perdivara, I., L.J. Deterding, C. Cozma, K.B. Tomer, and M. Przybylski, *Glycosylation profiles of epitope - specific anti-ss-amyloid antibodies revealed by liquid chromatography - mass spectrometry*. Glycobiology, 2009.

229. Perdivara, I., R. Petrovich, B. Allinquant, L.J. Deterding, K.B. Tomer, and M. Przybylski, *Elucidation of O-glycosylation structures of the beta-amyloid precursor protein by liquid chromatography-mass spectrometry using electron transfer dissociation and collision induced dissociation.* J Proteome Res, 2009. **8**(2): p. 631-42.

230. Iacob, R.E., I. Perdivara, M. Przybylski, and K.B. Tomer, *Mass spectrometric characterization of glycosylation of hepatitis C virus E2 envelope glycoprotein reveals extended microheterogeneity of N-glycans.* J Am Soc Mass Spectrom, 2008. **19**(3): p. 428-44.

231. Kjeldsen, F., Haselmann, K.F., Budnik, B.A., Jensen, F., Zubarev, R.A., *Dissociative capture of hot (3-13 eV) electrons by polypeptide polycations: An efficient process accompanied by secondary fragmentation.* Chem. Phys. Lett., 2002. **356**: p. 201-206.

232. Haselmann, K.F., T.J. Jorgensen, B.A. Budnik, F. Jensen, and R.A. Zubarev, *Electron capture dissociation of weakly bound polypeptide polycationic complexes.* Rapid Commun Mass Spectrom, 2002. **16**(24): p. 2260-5.

233. Mikesh, L.M., B. Ueberheide, A. Chi, J.J. Coon, J.E. Syka, J. Shabanowitz, and D.F. Hunt, *The utility of ETD mass spectrometry in proteomic analysis.* Biochim Biophys Acta, 2006. **1764**(12): p. 1811-22.

234. Hakansson, K., H.J. Cooper, M.R. Emmett, C.E. Costello, A.G. Marshall, and C.L. Nilsson, *Electron capture dissociation and infrared multiphoton dissociation MS/MS of an N-glycosylated tryptic peptic to yield complementary sequence information*. Anal Chem, 2001. **73**(18): p. 4530-6.

235. Catalina, M.I., C.A. Koeleman, A.M. Deelder, and M. Wuhrer, *Electron transfer dissociation of N-glycopeptides: loss of the entire N-glycosylated asparagine side chain.* Rapid Commun Mass Spectrom, 2007. **21**(6): p. 1053-61.

236. Schroeder, M.J., D.J. Webb, J. Shabanowitz, A.F. Horwitz, and D.F. Hunt, *Methods for the detection of paxillin post-translational modifications and interacting proteins by mass spectrometry.* J Proteome Res, 2005. **4**(5): p. 1832-41.

237. Jefferis, R., J. Lund, and J.D. Pound, *IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation.* Immunol Rev, 1998. **163**: p. 59-76.

238. Tao, M.H. and S.L. Morrison, *Studies of aglycosylated chimeric mouse-human IgG. Role of carbohydrate in the structure and effector functions mediated by the human IgG constant region.* J Immunol, 1989. **143**(8): p. 2595-601.

239. Sarmay, G., J. Lund, Z. Rozsnyay, J. Gergely, and R. Jefferis, *Mapping and comparison of the interaction sites on the Fc region of IgG responsible for triggering antibody dependent cellular cytotoxicity (ADCC) through different types of human Fc gamma receptor.* Mol Immunol, 1992. **29**(5): p. 633-9.

240. Burton, D.R. and J.M. Woof, *Human antibody effector function.* Adv Immunol, 1992. **51**: p. 1-84.

241. Takahashi, N., I. Ishii, H. Ishihara, M. Mori, S. Tejima, R. Jefferis, S. Endo, and Y. Arata, *Comparative structural study of the N-linked oligosaccharides of human normal and pathological immunoglobulin G.* Biochemistry, 1987. **26**(4): p. 1137-44.

242. Mizuochi, T., T. Taniguchi, A. Shimizu, and A. Kobata, *Structural and numerical variations of the carbohydrate moiety of immunoglobulin G.* J Immunol, 1982. **129**(5): p. 2016-20.

243. Parekh, R.B., R.A. Dwek, B.J. Sutton, D.L. Fernandes, A. Leung, D. Stanworth, T.W. Rademacher, T. Mizuochi, T. Taniguchi, K. Matsuta, and et al., *Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG.* Nature, 1985. **316**(6027): p. 452-7.

244. Sheeley, D.M., B.M. Merrill, and L.C. Taylor, *Characterization of monoclonal antibody glycosylation: comparison of expression systems and identification of terminal alpha-linked galactose*. Anal Biochem, 1997. **247**(1): p. 102-10.

245. Galili, U., *Interaction of the natural anti-Gal antibody with alpha-galactosyl epitopes: a major obstacle for xenotransplantation in humans*. Immunol Today, 1993. **14**(10): p. 480-2.

246. Borrebaeck, C.K., A.C. Malmborg, and M. Ohlin, *Does endogenous glycosylation prevent the use of mouse monoclonal antibodies as cancer therapeutics?* Immunol Today, 1993. **14**(10): p. 477-9.

247. Lim, A., A. Reed-Bogan, and B.J. Harmon, *Glycosylation profiling of a therapeutic recombinant monoclonal antibody with two N-linked glycosylation sites using liquid chromatography coupled to a hybrid quadrupole time-of-flight mass spectrometer.* Anal Biochem, 2008. **375**(2): p. 163-72.

248. Mimura, Y., P.R. Ashton, N. Takahashi, D.J. Harvey, and R. Jefferis, *Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry*. J Immunol Methods, 2007. **326**(1-2): p. 116-26.

249. ClinicalTrials.gov.[Internet]; Bethesda (MD); National Library of Medicine (US); 2000 Feb 29. Identifier: NCT00818662, *A Phase 3 Study Evaluating Safety and Effectiveness of*

Immune Globulin Intravenous (IGIV 10%) for the Treatment of Mild to Moderate Alzheimer´s Disease; 2009 Jan 7th [cited 2009 June 24th].; Available from: http://www.clinicaltrials.gov/ct2/show/NCT00818662?term=alzheimer%27s+ivig&rank=3.

250. Wuhrer, M., J.C. Stam, F.E. van de Geijn, C.A. Koeleman, C.T. Verrips, R.J. Dolhain, C.H. Hokke, and A.M. Deelder, *Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum.* Proteomics, 2007. **7**(22): p. 4070-81.

251. Shakib, F. and D.R. Stanworth, *Human IgG subclasses in health and disease. (A review). Part I.* Ric Clin Lab, 1980. **10**(3): p. 463-79.

252. French, M., *Serum IgG subclasses in normal adults.* Monogr Allergy, 1986. **19**: p. 100-7.

253. Wuhrer, M., L. Porcelijn, R. Kapur, C.A. Koeleman, A.M. Deelder, M. de Haas, and G. Vidarsson, *Regulated Glycosylation Patterns of IgG during Alloimmune Responses against Human Platelet Antigens.* J Proteome Res, 2008.

254. Jefferis, R., *Antibody therapeutics: isotype and glycoform selection.* Expert Opin Biol Ther, 2007. **7**(9): p. 1401-13.

255. Shakib, F. and D.R. Stanworth, *Human IgG subclasses in health and disease. (A review). Part II.* Ric Clin Lab, 1980. **10**(4): p. 561-80.

256. Aalberse, R.C. and J. Schuurman, *IgG4 breaking the rules.* Immunology, 2002. **105**(1): p. 9-19.

257. Boyd, P.N., A.C. Lines, and A.K. Patel, *The effect of the removal of sialic acid, galactose and total carbohydrate on the functional activity of Campath-1H.* Mol Immunol, 1995. **32**(17-18): p. 1311-8.

258. Hodoniczky, J., Y.Z. Zheng, and D.C. James, *Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro.* Biotechnol Prog, 2005. **21**(6): p. 1644-52.

259. Shields, R.L., J. Lai, R. Keck, L.Y. O'Connell, K. Hong, Y.G. Meng, S.H. Weikert, and L.G. Presta, *Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity.* J Biol Chem, 2002. **277**(30): p. 26733-40.

260. Shinkawa, T., K. Nakamura, N. Yamane, E. Shoji-Hosaka, Y. Kanda, M. Sakurada, K. Uchida, H. Anazawa, M. Satoh, M. Yamasaki, N. Hanai, and K. Shitara, *The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity.* J Biol Chem, 2003. **278**(5): p. 3466-73.

261. Yamane-Ohnuki, N., S. Kinoshita, M. Inoue-Urakubo, M. Kusunoki, S. Iida, R. Nakano, M. Wakitani, R. Niwa, M. Sakurada, K. Uchida, K. Shitara, and M. Satoh, *Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity.* Biotechnol Bioeng, 2004. **87**(5): p. 614-22.

262. Kaneko, Y., F. Nimmerjahn, and J.V. Ravetch, *Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation.* Science, 2006. **313**(5787): p. 670-3.

263. Scallon, B.J., S.H. Tam, S.G. McCarthy, A.N. Cai, and T.S. Raju, *Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality.* Mol Immunol, 2007. **44**(7): p. 1524-34.

264. Sato, Y., C. Liu, B.S. Wojczyk, A. Kobata, S.L. Spitalnik, and T. Endo, *Study of the sugar chains of recombinant human amyloid precursor protein produced by Chinese hamster ovary cells.* Biochim Biophys Acta, 1999. **1472**(1-2): p. 344-58.

265. Varki, A., Cummings, R., Esko, J., Freeze, H., Hart G., Marth, J., *Essentials of Glycobiology.* 1999: p. 101-113.

266. Savitski, M.M., F. Kjeldsen, M.L. Nielsen, and R.A. Zubarev, *Hydrogen rearrangement to and from radical z fragments in electron capture dissociation of peptides.* J Am Soc Mass Spectrom, 2007. **18**(1): p. 113-20.

267. Mormann, M., H. Paulsen, and J. Peter-Katalinic, *Electron capture dissociation of O-glycosylated peptides: radical site-induced fragmentation of glycosidic bonds.* Eur J Mass Spectrom (Chichester, Eng), 2005. **11**(5): p. 497-511.

268. Savitski, M.M., M.L. Nielsen, and R.A. Zubarev, *Side-chain losses in electron capture dissociation to improve peptide identification.* Anal Chem, 2007. **79**(6): p. 2296-302.

269. [cited; Available from: http://www.niehs.nih.gov/research/clinical/join/bethesda/participants/studies/psftsdsrd.cfm.

270. Huhn, C., M.H. Selman, L.R. Ruhaak, A.M. Deelder, and M. Wuhrer, *IgG glycosylation analysis.* Proteomics, 2009. **9**(4): p. 882-913.

271. van Zeben, D., G.A. Rook, J.M. Hazes, A.H. Zwinderman, Y. Zhang, S. Ghelani, T.W. Rademacher, and F.C. Breedveld, *Early agalactosylation of IgG is associated with a more progressive disease course in patients with rheumatoid arthritis: results of a follow-up study.* Br J Rheumatol, 1994. **33**(1): p. 36-43.

272. Parekh, R., D. Isenberg, G. Rook, I. Roitt, R. Dwek, and T. Rademacher, *A comparative analysis of disease-associated changes in the galactosylation of serum IgG.* J Autoimmun, 1989. **2**(2): p. 101-14.

273. Moore, J.S., X. Wu, R. Kulhavy, M. Tomana, J. Novak, Z. Moldoveanu, R. Brown, P.A. Goepfert, and J. Mestecky, *Increased levels of galactose-deficient IgG in sera of HIV-1-infected individuals.* Aids, 2005. **19**(4): p. 381-9.

274. Mann, H., Whitney, D., *On a test of whether one of two random variables is stochastically larger than the othe*r. Annals of Mathematical Statistics, 1947. **18**: p. 50-60.

275. Wilcoxon, F., *Individual Comparisons by Ranking Methods.* Biometrics Bulletin, 1945. **1**: p. 80-83.

276. Craighead, H., *Future lab-on-a-chip technologies for interrogating individual molecules*. Nature, 2006. **442**(7101): p. 387-93.

277. Hogan, J., *Lab on a chip: a little goes a long way.* Nature, 2006. **442**(7101): p. 351-2.

278. Moise, A., Siebert, H. -C., Gabius, H. -J., Przybylski, M., *Molecular determination of carbohydrate recognition structures in proteins by a proteolytic excision - affinity mass spectrometric approach.* Proc Natl Acad Sci U S A, 2009: p. in preparation.

279. Moise, A., Siebert, H. -C., Gabius, H. -J., Przybylski, M., *Determination of carbohydrate recognition sequences of lectins and related carbohydrate binding proteins - CREDEX*, U.o.K. Patent Application, Editor. 2009.

280. Muller, G., S. Ruppert, E. Schmid, and G. Schutz, *Functional analysis of alternatively spliced tyrosinase gene transcripts.* Embo J, 1988. **7**(9): p. 2723-30.

7 APPENDIX

7.1 Appendix 1

ABBREVIATIONS

Aß	amyloid ß-peptide
ACN	acetonitrile
AD	Alzheimer's disease
APP	amyloid precursor protein
sAPPα	secreted APPα fragment
C_L	antibody light chain constant region
C_H	antibody heavy chain constant region
CHO	Chinese hamster ovary
CID	collision induced dissociation
CDR	complementary determining region
CSF	cerebrospinal fluid
DC	direct current
DDA	data dependent acquisition
DM	dermatomyositis
DTT	dithiothreitol
ECD	electron capture dissociation
ETD	electron transfer dissociation
ESI	electrospray ionization
Fc	antibody (heavy chain) crystallisable constant region fragment
FT-ICR	Fourier transform ion cyclotron resonance
Fuc	fucose
Gal	galactose
GalNAc	*N*-acetyl galactosamine
Glc	glucose
GlcNAc	*N*-acetyl glucosamine
Hex	hexose

IAA	iodoacetamide
IgG	immunoglobulin G
IIM	idiopathic inflammatory myopathies
IVIg	intravenous immunoglobulin
LC-MS/MS	liquid chromatography – tandem mass spectrometry
mAb	monoclonal antibody
Man	mannose
MS	mass spectrometry
MS(n)	multistage tandem mass spectrometry
MSc	multiple sclerosis
m/z	mass-to-charge
NeuNAc	*N*-acetyl neuraminic acid
NeuGc	*N*-glycolyl neuraminic acid
PBS	phosphate buffer salin
PM	polymyositis
PTM	post-translational modification
Q-Tof	quadrupole time-of-flight
RA	rheumatoid arthritis
RF	radio frequency
RP-HPLC	reversed phase – high performance liquid chromatography
SA	sialic acid
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SLE	systemic lupus erythematosus
SMH	somatic hypermutation
UPLC	ultra performance liquid chromatography
V_L	antibody light chain variable region
V_H	antibody heavy chain variable region

7.2 Appendix 2

Incremental mass difference for some important post-translational modifications

PTM type	Δ Mass (Da)	Subcellular location	Function and notes
Phosphorylation		Nucleus, cytosol, extracellular fluid, extracellular matrix	Reversible activation/inactivation of enzyme activity, modulation of molecular interactions, signaling
pTyr	+80		
pSer, pThr	+80		
Acetylation	+42	Nucleus, cytosol, extracellular fluid	Protein stability, protection of the N-terminus, regulation of protein-DNA interactions (histones)
Methylation	+14	Cytosol	Regulation of gene expression
Acylation, Fatty acid modification			Cellular localization and targeting signals, membrane tethering, mediator of protein-protein interactions
Farnesyl	+204	Ribosome	
Myristoyl	+210	Golgi	
Palmytoyl	+238		
Glycosylation		Endoplasmic reticulum, Golgi, plasma membrane, extracellular fluid and extracellular matrix	Secreted proteins, cell-cell recognition/signalling O-GlcNAc, , reversible regulatory functions
N-linked	> 800		
O-linked	> 800		
O-GlcNAc	+203	Nucleus, cytosol	
GPI anchor	>1000	Endoplasmic reticulum, plasma membrane	Glycosylphosphatidylinositol (GPI) anchor. Membrane tethering of enzymes and receptors, mainly to outer leaflet of plasma membrane
Hydroxyproline	+16	Extracellular matrix	Protein stability and protein-ligand interactions
Sulfation (sTyr)	+80	Golgi	Modulator of protein-protein and receptor-ligand intereactions
Disulfide bond formation	-2		Intra- and intermolecular crosslink, protein stability
Deamidation	+1		Possible regulator of protein-ligand and protein-protein interactions, also a common chemical artefact
Pyroglutamic acid	-17		Protein stability, blocked N-terminus
Ubiquitination	>1000		Destruction signal. After tryptic digestion, ubiquitination site is modified with the Gly-Gly dipeptide
Nitration of tyrosine	+45		Oxidative damage during inflammation