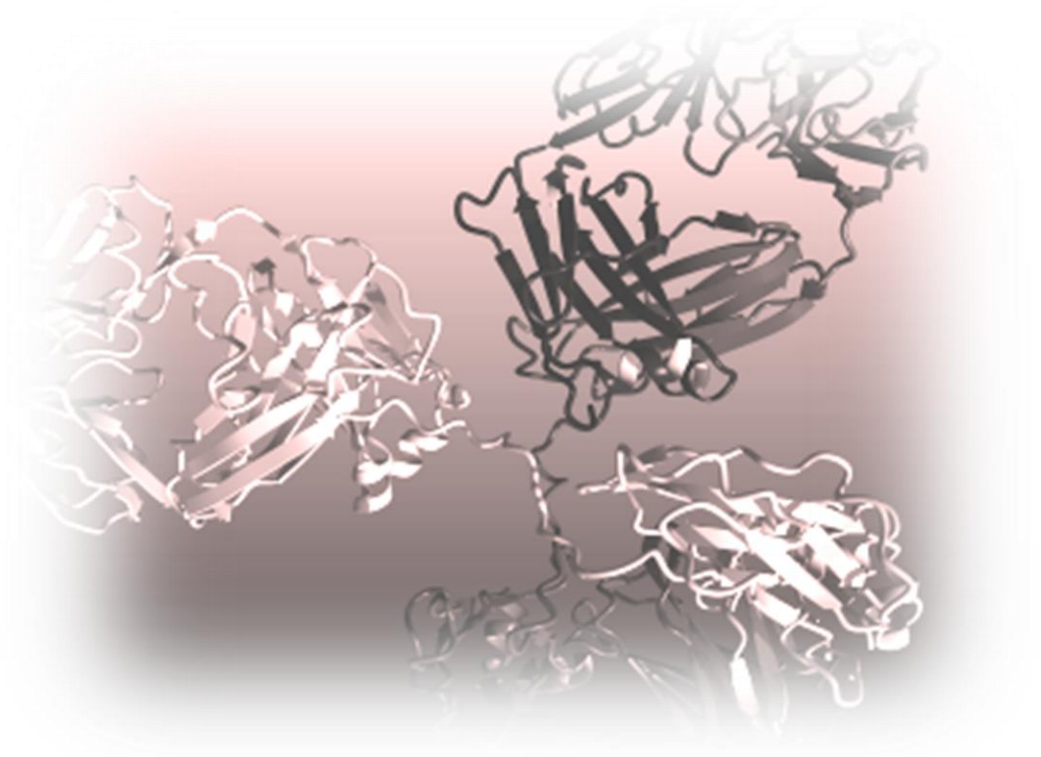
Amplification and sequencing of genes encoding monoclonal antibodies targeting amyloid beta in Alzheimer's disease.

-

A new solution to an old problem



Andreas Møller Stehr

Master Thesis

Medicine with Industrial Specialisation

Aalborg University, 2012

Department of Health Science and Technology
Medicine with Industrial Specialisation
Frederik Bajers Vej 7
9220 Aalborg

Title: Amplification and sequencing of genes encoding monoclonal antibodies targeting amyloid beta in Alzheimer's disease.

Project Period: Sep. 1st 2011 to june 1nd 2012

Project Group: Group 962, MedIS 10th semester

Participants:

Andreas Møller Stehr

Supervisor: Svend Birkelund, Torben Moos

Prints: 5

Number of Pages: 38

Finished: june. 1nd 2012

Synopsis:

Alzheimers disease is known to be the main cause of dementia, and in Europe alone 1.9% over the age of 65 is diagnosed with Alzheimer's.

Monoclonal antibodies are used as a potential new treatment against Alzheimer's disease, and several studies have already been performed. However, one main obstacle remains; the blood brain barrier (BBB). Transfecting cells within the BBB with genes encoding such antibodies might overcome this obstacle.

In order to transfect cells within the BBB, one needs to know the sequence of the antibodies for amplification and cloning. However the antibodies have highly variable areas making it difficult to amplify. Usually, this is performed by polymerase chain reaction using degenerated primers; however the accuracy and specificity can be discussed using this method.

In this study I present a new method for sequencing antibodies using mRNA sequencing technique and tandem mass spectrometry. This new approach gives rise to results of high validity and specificity not previously seen in the literature.

Table of Contents

Abstract

Alzheimers disease is known to be the main cause of dementia, and in Europe alone 1.9% over the age of 65 is diagnosed with Alzheimer's.

Monoclonal antibodies are used as a potential new treatment against Alzheimer's disease, and several studies have already been performed. However, one main obstacle remains; the blood brain barrier (BBB). Transfecting cells within the BBB with genes encoding such antibodies might overcome this obstacle.

In order to transfect cells within the BBB, one needs to know the sequence of the antibodies for amplification and cloning. However the antibodies have highly variable areas making it difficult to amplify. Usually, this is performed by polymerase chain reaction using degenerated primers; however the accuracy and specificity can be discussed using this method.

In this study I present a new method for sequencing antibodies using mRNA sequencing technique and tandem mass spectrometry. This new approach gives rise to results of high validity and specificity not previously seen in the literature.

Preface

The Original idea of this study was first introduced to me by Torben Moos who was inspired by late findings that monoclonal antibodies might be used as a therapeutic drug in Alzheimer disease. This idea was further developed in cooperation with Svend Birkelund, and during the process of amplification we developed a new method for sequencing problematic genes due to high variable areas.

In the current study I would like to thank people at the institute for biotechnology; Anne Rusborg Nyborg for helping with the mRNA extraction, PhD Student Tue Bennike for helping with digestion and MS/MS of antibodies, Assistant professor Mads Sønderkær for data preparation of the extracted mRNA reads and Professor Allan Stensballe for ideas and guidance. Furthermore I like to thank Professor Torben Moos and Svend Birkelund for their guidance and help, Trine Axelsen for Immunohistochemistry, and finally Professor Gunna Christensen for helping with discussion of the manuscript.

Introduction

Alzheimer's disease

Alzheimer's disease (AD) is a neurological disorder known as the main cause of dementia. On a word scale 3.9% of all people over 60 years is diagnosed with dementia and of these cases 75% is due to AD(Qiu & Kivipelto 2009). Dementia is characterized by decreased cognitive functions beyond that of normal aging, and the symptoms include impaired; memory, orientation, language, calculation, comprehension, learning capacity and behavior alterations (WHO 2012). In Europe alone 1.94% of all people over the age of 65 years are diagnosed with AD, and the risk rises almost exponentially up to the age of 85 (Qiu & Kivipelto 2009).

The Neuropathological signs of AD are loss of brain weight and volume due to degeneration of neurons especially seen by shrinking of the hippocampus, enlarged ventricles, narrowed gyri and wide sulci in the temporal and frontal lobe. Further signs include dysfunction and loss of neural-synapses, accumulation of neurofibrilary tangles (NFT) in neurons and extracellular aggregates called amyloid plagues. The cause of AD is not fully known, however the main hypothesis describes it as alterations in generation and degradation of amyloid (A)β leading to formation of the amyloid plagues. These plagues are manly composed of Aβ fibrils which are fragments originating from trans-membrane molecules called the amyloid precursor protein (APP) that undergo protelytic cleavage via the activity of different secretases. The function of APP is not fully understood but is considered to be part of neural plasticity, iron export and synaptic formation. Exactly how the Aβ induces toxicity it not known but it might include changes in activity of neurotransmitters as well as the receptors and the distribution of these, furthermore a disruption of the intracellular calcium homeostasis of cells, and impairment of mitochondrial function and axonal transport. Soluble Aβ is also considered to have a toxic effect on the intracellular environment. here It might damage the cell membrane, organelles, alter biochemical pathways, induce inflammation, oxidative stress, and possibly also increase the activity of the Tau-phosphorylating kinase hereby playing a role in the formation of the NFTs. NFTs are hyperphosphylated tau proteins that under normal circumstances function by stabilizing the microtubules in the neurons. After phosphylation the tau proteins are replaced by tangles leading to the death of the cell. Studies have shown the appearance of tangles before the emergence of Aβ fibrils and as a result some believe that NFT's and not Aβ is the primary source of AD. However, since the mechanism behind AD, and the correlation between NFT's and Aβ is still unclear, and that most studies point at Aβ fibrils as the primary source, Aβ is still seen as the main cause leading to the pathology of AD (Taipa et al. 2012)(Huang & Mucke 2012)(Di Carlo et al. 2012).

Treatment

Today approved therapies targeting AD focus on symptomatic treatment either by inhibiting acetylcholine esterase increasing the levels of the neurotransmitter acetylcholine, which are depleted in the AD brain, or by targeting the glutamate receptors where a drug such as Memantine works as an uncompetitive NMDA receptor antagonist blocking the activity of glutamate. The effect of these drugs are at best described as modest, however a combination therapy might increase the therapeutic benefits (Kathryn L. MacCane 2006)(Huang & Mucke 2012).

This has called for development of new strategies in the therapeutic field of AD, with focus on disease modifying drugs (Bruno vellas 2007). There are five obvious targets in AD; secretase's cleaving the APP, NFT, and Aβ. Drugs targeting these areas are being developed with various success (Samadi & Sultzer 2011)(Huang & Mucke 2012), and especially immunotherapy seems to be the current strongest approach. Immunotherapy can be divided in passive immunization and active immunization. Trials of active immunization is likely to be more cost effective and long lasting compared to passive immunization that requires regular injections, however one of the main concerns by active immunization is stopping the treatment if adverse effects occur such as meningoencephalitis which has been observed in a study, probably due to T helper cell 1 activation. Passive immunization has the advantage of more specific targeting, hereby ensuring that for example not all of the Aβ are removed from the brain, however some cons exist such as; off target-reactivity, microhaemorrhage, antibodies might have trouble crossing the BBB and the patient might produce autoantibodies (Panza et al. 2011). The most promising immunotherapeutic drugs which targets Aβ are called Solanezumab and Bapineuzumab. They are both passive immunization therapies, are in phase 3 trials, are administrated intravenously, are humanized anti-Aβ monoclonal antibody of the IgG class, and work by passive transfer. Solanezumab targets the Aβ at the central part (amino acid 13-28), it functions by binding soluble Aβ, has little to no affinity for amyloid plaques, and is further designed for clearance from the brain. Bapineuzumab targets the N terminal of the Aβ peptide (aminoacid 1-5), however where Solanezumab exclusively binds the central region of soluble Aβ, Bapineuzumab binds both with deposited Aβ as well as the N-terminal of the Aβ peptide, and might decrease the total tau levels in the CSF (Panza et al. 2011)(Panza et al. 2011; Reichert 2011; Samadi & Sultzer 2011).

Clearance of Aβ from the Brain

It is not clear how Aβ is removed from the brain, however there are some theories. 1; Fc-mediated phagocytosis, where microglia cells are attracted to the amyloid plaques resulting in Aβ degradation. 2; Inhibition of fibrillar aggregation by anti Aβ antibodies which binds to the fibrils, thus preventing them from aggregating (Geylis & Steinitz 2006), 3; Peripheral sink mechanism, whereby anti-Aβ antibodies in plasma causes a shift in the equilibrium of the peptide from the CSF to the plasma, reducing the quantity of Aβ in the brain(Geylis & Steinitz 2006)(Kerchner & Boxer 2010).

There is some *inconsistency* in the literature and the clearance is probably due to a combination of the before mentioned mechanisms, or by unknown mechanisms. Furthermore, it is not clear to what extent the antibodies are capable of crossing the BBB.

Barriers of the brain an major obstacle

The majority of all drug candidates targeting the central nervous system never become clinically available due to problems arising when trying to circumvent the barriers of the brain(Hulliger et al. 1989). The barriers are; the blood brain barrier (BBB), the blood-cerebral spinal fluid barrier (B-CSFB) and the arachnoid barrier (Cardoso et al. 2010). Of these barriers the one of most interest is the BBB, because it comprises approximately 95% of the total area of the barriers with an exchange area of 12-18m2 (Pottiez et al. 2009). Furthermore, the cells of the brain lie close to the capillaries giving a short diffusion distance (Abbott et al. 2010). The tight junctions of the BBB are known to be 50-100 times tighter than those of normal endothelium making diffusion difficult (Cardoso et al. 2010).

To overcome the problems arising by passive immunization, and still get the advantages such as specificity in targeting of the antibodies one could therapeutically change the endothelial cells compromising the BBB to synthesize and secrete the molecule of choice (Jiang et al. 2003).

Hypothesis

The overall goal of this study is to design a delivery system for monoclonal antibodies (MAb), cloning the genes encoding the monoclonal antibodies targeting Aβ. The antibodies will be produced in the BBB endothelium cells by the mean of viral transfection leading to a secretion of the antibodies primarily to the basolateral surface as reported in (Jiang et al. 2003). After entering the brain, the antibodies will target and join with Aβ creating immunocomplexes. These complexes will be dragged out of the brain via transcystosis mediated by the Fc-neonatal receptor.

In order to successfully transfect the BBB, amplification and cloning of the genes is essential, however the amplification of the variable areas of the immunoglobulins (Ig) are known to be a major obstacle. This study addresses the problems arising from amplifying the variable regions, using known polymerase chain reaction (PCR) technologies, and presents an alternative solution to an old problem.

Monoclonal antibodies targeting Alzheimer

Our therapeutic drugs of choice are antibodies targeting Aβ. These antibodies are MAb1.1, MAb19.2.33 and MAb23.3.22 supplied by Loke diagnostics. MAb1.1 targets the C-terminal end of the Aβ peptide (Axelsen et al. 2009), while MAb19.2.33 and MAb23.3.22 are targeting the N-terminal part of the Aβ peptide showed by Trine Axelsen phd non published data), for further detail see table (1) and figure (1) . Like all other monoclonal antibodies they consist of two identical heavy chains (HC) and two identical light chains (LC), in these cases kappa LC. Each chain contains a variable area located in the N terminal end of the chain which together defines the Fab fragments of the IgG molecule involved in the antigen binding (Jung et al. 2006), and a constant area defining the Fc fragment, class and effector function. LC contains one constant domain, where HC have between 3 – 4 constant domains. Each variable or constant domain consists of approximately 110-130 amino acids (Schroeder & Cavacini 2010).

Antibody	Target	Amino acid sequence
MAb IgG1.1	C-terminal part of Aβ peptide 37-42	GGVVIA
MAb IgG19.2.33	N-terminal part of Aβ peptide 3-8	EFRHDS
MAb IgG23.3.22	N-terminal part of Aβ peptide 3-8	EFRHDS

Tabel 1, From left; Name of antibody, Which part of the Aβ the antibody is targeting and the area, the specific targeting amino acid sequence.

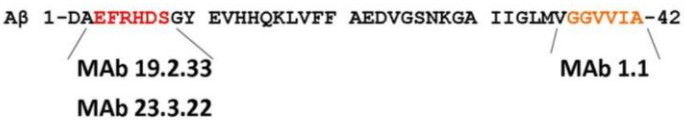


Figure 1; the figure shown the amino acid sequence of Aβ, the red marking are the target sites of the different MAbs.

Known methods and arising problems

The Ig molecules are of interest due to their therapeutic possibilities, however the before described variability often gives rise to problems designing a general method for amplification, especially amplification of the HC. One study that addresses this problem was performed by Z. Wang et al., (2000), who after trying commerce

al primers for PCR, designed highly degenerated primers also known as universal primers, by calculating the frequency for each dominant nucleotide in each position based on the Kabat database 1991. This resulted in a success rate of 70% for V_H and by further modifying the primers giving an overall amplifying success rate of 80%. These primers have since been verified by other studies such as Morrow et al. (2009). Others such as Y. Wang et al. (2006), tries to optimize the amplification by further calculating the most fitting primers by the help of a new algorithm, increasing the coverage. Even so Rohatgi et al. 2008, addresses the problem that the degenerated primers have higher propensity of cross-family amplification, and even if this is countered by designing a panel of primers with minimum degeneracy, the primers might not be sensitive enough for an amplification of genes from a single or small number of B cells. Rohatgi et al. (2009) presents a solution using nested primers which seems to cover all known V_H, V_κ, V_λ, J_H, J_κ, J_λ, C_H, C_κ and C_λ genes. Even so, this technique is still depending upon PCR and even so newer studies such as Tiller et al. (2009), still only get a coverage of 70% of the V_H genes making room for further adjustments and method development.

Understanding the high antibody variability

It is well known that the immune system is capable of initiating a very specific immune response and that the specificity of this immune response is due to the high diversity of the Igs. This diversity is mainly achieved by the mechanism known as the V(D)J recombination. The V (variable), D (diversity) and the J (joint) segments are on germline level, separated by exons and introns over a large area of the chromosome and joined together by random recombination generating the variable areas of the Igs defining the Fab fragments, the specificity and diversity (figure 2). The variable area of the LC is sampled from V and J segments, whereas the HC is a combination of V, D and J segments, strongly increasing the variability of the HC compared to that of LC. Furthermore the constant (C) domain of the Ig is encoded by independent exons located downstream from the V(D)J segments. For B lymphocytes there are 3 different loci located on chromosome 12; Igs heavy chain locus (IgH), light chain Locus (IgL) IgLκ and IgLλ. The IgH are

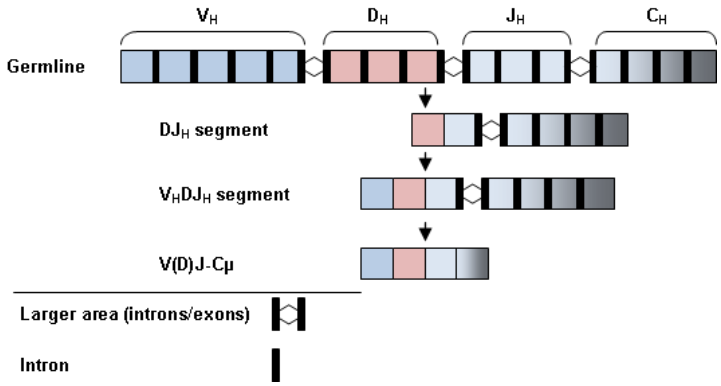


Figure 2; Heavy chain recombination. On germline level the genes (the V_H ,D_H, J_H and C_H) for recombination are placed in clusters along the chromosome separated by lager areas of introns and exons. The different gene coding segments in the clusters are further separated by introns. The recombination starts with combination of the D_H and J_H gene segments, in next step the V_H gene segment is added. the C_H genes is the last one to be added, and as long as no class switch have occurred the constant segment added will be the Cμ segment. Coding for a transcriptional unit called V(D)J-Cμ.

composed of >150 or more V_H genes located upstream of the 12-13 D_H genes (2.8 Mb). The 4-5 J_H genes (60Kb) are located downstream of the D_H genes (1.4kb) followed by The C genes, starting with C_μ See figure. The light chain, in our case IgLκ are composed of 140 Vκ genes (3.2Mb) and 4 Jκ genes (1.4) placed downstream (numbers taken from David G S, 2011).

The recombination process for B lymphocytes starts when the B cell is in its pro stage synthesizing the HC, combining first D_H and J_H forming the DJ_H segment, and then adding the V_H gene, creating the V_HDJ_H segments. The transcript is finished by adding the C_μ region (figure 2). The mature B cell only expresses one allele. This selection, called allelic exclusion, is thought to occur under the recombination as it ensures against autoimmunity. If the recombination of the genes leads to an non-functional Igs, the other allele might become available functioning as an substitute giving the cell a second chance for recombination (Bergman & Cedar 2004).

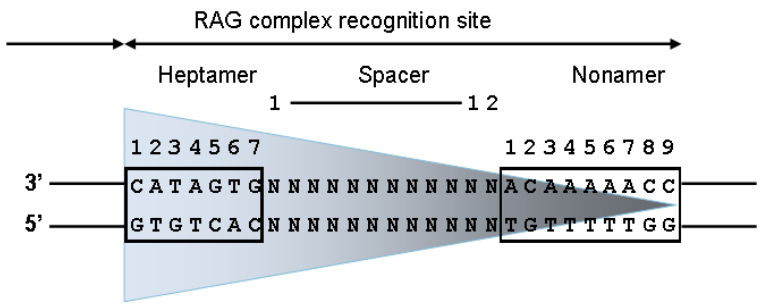


Figure 3, Recombination starts with RAG 1 recognizing the 12 spacer, RAG 2 then associates with RAG-1 creating a complex with RAG-1 and the heptamer group. The complex is illustrated as a blue triangle.

The mechanism is initiated by the recombination activating genes (RAG) encoding for the RAG endonucleases (RAG-1 and RAG-2) which is expressed at high levels in the early stages of the lymphocyte development. The RAG proteins binds and cleaves the DNA at specific sites between two coding segments introducing DNA doubled stranded breaks (DSBs). These sites are flanked by the recombination signal sequences (RSSs). The RSSs are composed of highly conserved heptamer and nonamer sequences separated by either 12 or 23 base pairs (figure 3), The RAG can only combine segments if one is separated with a 12-mer spacer and the other by a 23-mer spacer known as the 12/23 rule. V_H and J_H genes are flanked by 12-bp spacer whereas D_H genes are flanked only by 23-bp spacer herby dictating a VDJ combination, the constant areas are later added in order to finish the sequence (figure 2).(Schatz & Ji, 2011; Schatz & Swanson, 2011;Chaudhuri et al., 2007; Jung et al., 2006; Schroeder & Cavacini, 2010).

Recombination in details

RAG-1 initiates the process by recognition of the 12-bp spacer or 23-bp spacer, RAG-2 associates with RAG-1 and the heptamer, forming a complex (figure 4 A). The complex now captures the other RSS forming a synaptic complex also called paired complex see figure 4 B . DNA binding proteins known as high mobility group proteins (HMG), in this case HMG1 and HMG2 facilitates a creation of a heptamer – to heptamer ligation of the signal sequences (figure 4 B). DSB's occurs when the RAG single strand cuts at the heptamer sequence leaving the coding sequence with a free 3'OH group that ligates to the 5' phosphate group creating a hairpin. The signal sequences are left with blunt ends and are still captured in the RAG complex (figure 4 B). The hairpin junction is often resolved 4- 5 nucleotides from the end of the hairpin leaving an overhang, opening up the possibility

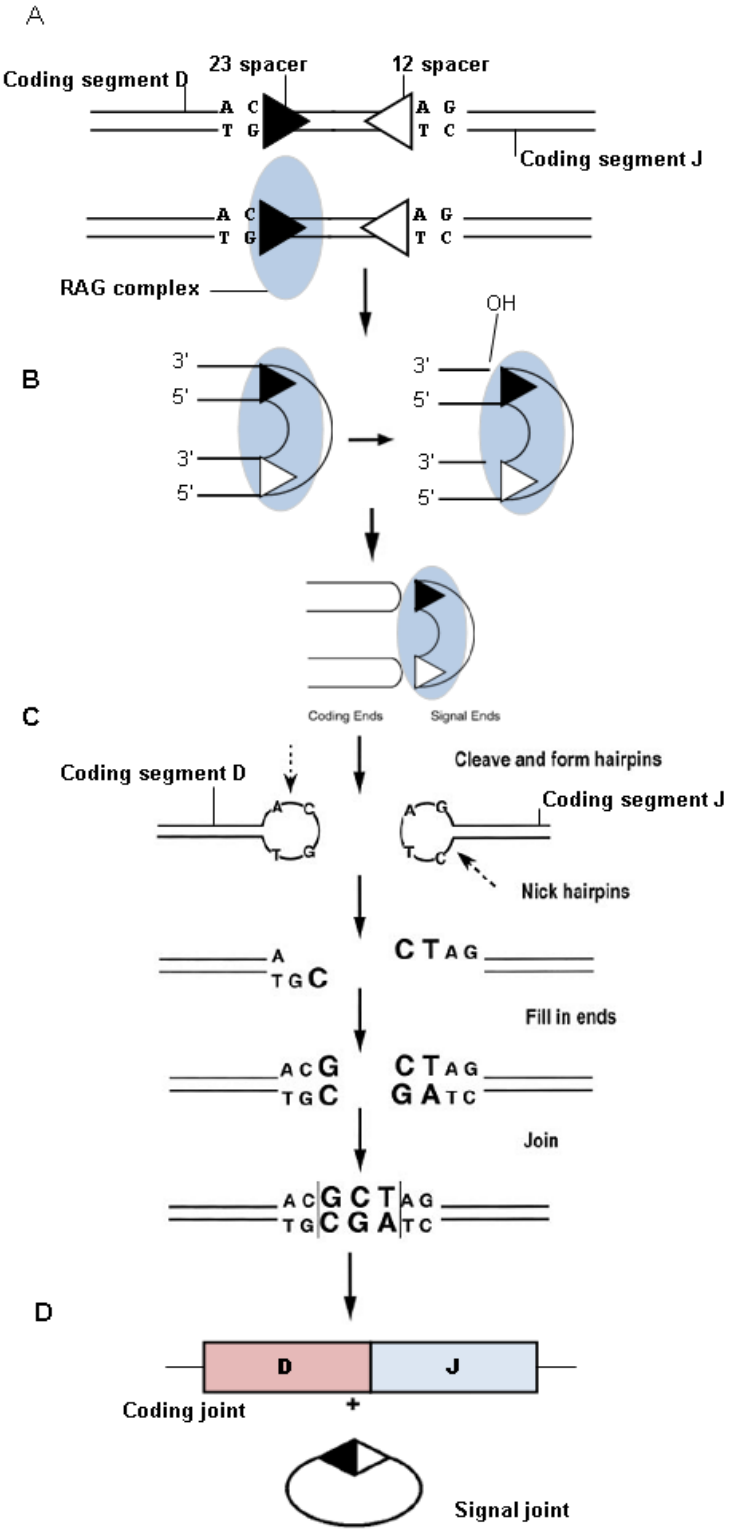


Figure 4; modified from; Martin G el al, 2002; Recombination process in detail. A; The process starts with RAG-1 recognizing a RSS separated by either a 12 or a 23 spacer as in the example. RAG-2 associates with RAG-1 forming a complex with the heptamer group illustrated as a blue circle covering the black triangle. B; The RAG complex now captures another RSS defined with a 12 spacer creating a synaptic complex. HMG facilitates a heptamer – heptamer ligation. RAG now single cuts the 3' DNA strand at the heptamer, leaving a free OH group. The OH group now joins the 5' DNA strand, creating a hairpin loop and a DSB. The RSSs are left with blunt ends, and still captured in the RAG complex. C; the hairpin loops defining the ends of the coding sequences are resolved leaving an overhang, making room for TdT or polymerase to add random nucleotides to the joints, and are repaired by NHEJ and DNA repair factors assembling the joints. D; this assembling and repairing of the joints usually results in a coding joint and a signal joint.

for enzymes such as terminal deoxynucleotidyl transferase (TdT) or polymerase to implant non-germline-encoded nucleotides, further adding to the variability (figure 4 C). The joints are then repaired by by nonhomologous end-joining (NHEJ) DNA repair factors usually resulting in a coding joint, and a signal joint(Schatz & Swanson 2011; Schatz & Ji 2011; Jung et al. 2006; Chaudhuri et al. 2007; Schroeder & Cavacini 2010).

A successful assembling shuts off the RAG mRNA and protein expression leading to a maturation of the pro B cell to the pre B cell by production of a HC signaling protein constituting the main part of the pre B cell receptor, expressed on the cell surface. The HC signaling protein is called μHC and is a product of a transcription unit called V(D)J-Cμ. The unit is composed of the mature mRNA sequence encoding the V(D)J and the - Cμ exon positioned upstream to the V(D)J segment (figure 2). In the pre B-cell a reexpression of RAG occurs, assembling IgL. The IgLκ, which constitutes 95% of the IgL chains produced, is the first to be assembled, and if this should fail the IgLλ is assembled. The assembling happens in the same way as HC with the one difference that LC only composes 2 segments V and J. Assembly of the IgL results in a production of a LC protein, joining with the HC protein, creating an IgM molecule leading to surface IgM also called the B Cell Receptor (BCR) hereby defining the immature B cell. The immature B cell travels from its place of origin into the blood, and increases its transcription to include the constant segment C_δ positioned further downstream. Due to alternative splicing the B cell is now capable of co-producing both the IgM and the IgD molecules sharing the same variable domains and the B cell is now defined as mature (Jung et al. 2006; Chaudhuri et al. 2007).

Class switch and Hypermutation

The B cell enters secondary lymphoid organs such as lymph nodes or spleen and travels to the germinal centers, where IgH class switch recombination (CSR) and somatic hypermutation (SHM) occurs when presented for a cognate antigen. Somatic hypermutations (SHM) can take place in the lgL, and the lgH chain. The SHM can only happen by the help of T cell activation and activation-induced cytidine deaminase (AID) which is up regulated after antigen activation. SHM induces insertions, deletions and point mutations in the variable exons at a rate up to 10^{-4} per bp, per generation. This increases the variability and diversity by allowing creations of B cells with mutated variable regions and higher affinity for specific antigens. The majority of the mutations happen in so called mutation hot spots. (Rich et al, 2008;Chaudhuri et al., 2007; Goodman, Scharff, & Romesberg, 2007)

In extension to SHM, CSR is considered co-initiated by AID, and can happen in the lgH, a process where the constant region of Ig is changed by replacing the Cμ with one of the other C_H genes and keeping the variable region, herby defining a new isotype (figure 5). The C_H genes are organized in units, including a transcriptional promoter, a switch (S) region and the C_h exons. The switch region is an intronic 1-12 kb long sequence in which DSB occurs during the CSR. The DSBs occurs in two S regions removing the unwanted C_H segment, the broken S segments are then joined and repaired. The transcript undergoes further splicing in order to remove the S segments before a functional transcription sequence is ready. There are 8 different mouse C_h genes positioned in the following order on the chromosome; μ, δ, γ3, γ1, γ2b, γ2a, ε and α. Every C_h is capable of undergoing

alternative splicing, producing either a membrane terminus as part of the B cell receptor, or a secretory terminus as part of the Ig (Chaudhuri et al. 2007; Kracker & Durandy 2011; Rich 2008).

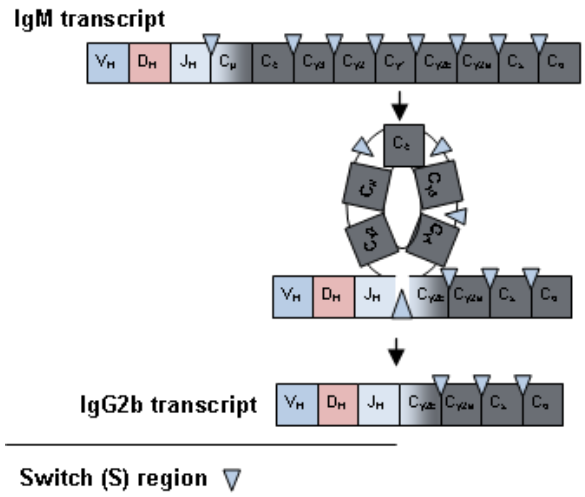


Figure 5 the figure shows the mechanims of CSR. The mature B-cell producing IgM changes to produce Ig2B by replacing the gene for the constant region. From the top, we have a transcript composed of the VD_HJ_H segment and the genes coding for the constant region. Each of the constant genes are separated by switch regions except between Cμ and C_δ. When splicing occurs, in starts with DSBs in the Switch regions followed by a rejoining of the segments surrounding the deleted segment, hereby defining a new transcript with the same variable areas, but with a new constant area, altering the isotype of the Ig molecule.

Variability in numbers

Several factors define the variability; first the diversity of the genes; there are >150 genes encoding the V_H gene in mouse, 5 V_H genes and 12 D_H genes. Each D_H gene can be rearranged by either inversion or deletion and undergo splicing and translation in each of the 3 reading frames, giving each D_H gene the potential of 6 different peptides giving a total of 60 possibilities. $HC_{variability} > 300 * 5 * 12 * 6 = 108{,}000.$

TdT can add up to 9 N nucleotides on either side of the D segment, replacing or adding to the original germline at random, giving 64,000,000 different junctions.

The LC variable are defined by 150 V_κ gene segments, and 2 V_λ gene segments making a total Ig variation number of; $108{,}000 * 64{,}000{,}000 * 250 * 2 = 3{,}456{,}000{,}000{,}000{,}000$ The numbers will vary according to the strain. (these numbers are taken from (Agger 2005).

The number is further increased by the former mentioned nicking of the Hairpin structure leaving an overhang for potential germ line-encoded nucleotides. Secondly, each terminal end of the coding gene segments might lose one or more nucleotides during the recombination.(Chaudhuri et al. 2007; Agger 2005; Schroeder & Cavacini 2010)

In the present study RNA form three hybridoma cell lines was purified. The purified RNA was first used for RT-PCR with degenerated primers. These PCR products were cloned and sequenced. In addition, mRNA was extracted from the cells and sequenced. The sequences were mapped to reference in order to obtain the specific gene sequences that are coding for HC and LC of the hybridoma cells. The result was further verified by a comparitative analysis with results from tandem mass spectrometry from purified antibodies produced by the hybridoma cells. The capability of the produced antibodies to bind Aβ was controlled by enzyme-linked immunosorbent assay (ELISA)

Materials and Methods

Antibody preparations

The antibodies; MAb1.1, 19.2.33 and MAb23.3.22 were produced by hybridroma clones grown in CELLine CL 1000,(Integra) using serum Prodoma 1 JAB-4F (Lonza) medium supplemented with 10 mM hypoxanthine, 1.6 mM thymidine containing, 1 mM pyruvate, and 10 mg/l gentamycin.. The medium in the cell free chamber contained additional 1% foetal calf serum. The hybridoma cells were grown to 10% vol/vol. The antibodies were purified from the supernatant on a Chromatography column XK 26/20 (GE Healthcare), loaded with GamMAbindTM plus sepharose (GE healthcare). Eluded fractions were measured on a Helios Omega Uv-Vis (Thermo). As quality control SDS-PAGE, ELISA and Lowry protein assay was performed.

Cloning

Cloning was done both with non purified and purified PCR products, in our case purification Gene Jet PCR purification kit #K0701 was used according to protocol. *Escherichia coli* Nova blue was used for transfection by Micropusler from Bio-Rad, program Ec2, with two different vectors according to the primers used in PCR. Novogen Ek/Lic cloning Kit protocol tb163 rev Ko908 was used in case of modified primers, while products amplified by non modified primers were transfected using the Clonejet PCR Cloning kit #k1231 according to protocol. The recombinant *E. coli* was streaked on LB Ager plates containing either ampicillin 10mg/ml (the Clonejet PCR Cloning) or kanamycin 5mg/ml(Ek/Lic cloning Kit) and incubated at 37 degrees. After 24-48 hours colonies were streaked and checked by PCR using pjet1.2 primers. DNA was extracted from the positive clones by phenol/chloroform extraction according to (Mygind et al. 2003), and controlled with a restriction enzyme *Bgl*II #FD0083 Fermentas. Successfully extracted samples were sent to DNA technology for sequencing.

RT-PCR

RNA was purified from the hybridoma cells using GeneJet PCR purification kit (#0701 Fermentas), and controlled by RNA gel electrophoresis. The RNA was loaded with 6xDNA loading dye (#R0611, Fermentas) and run for 20 minutes on a 1% agarose gel 10 mg/mL ethidium bromide (#39H8930, SIGMA) at 100 V in order to control the RNA purification, and only positive samples were used for

further investigation. The concentration of the extractions was then controlled by nanophotometer (IMPLEN).

Synthesis of cDNA was done using revertAid H Minus First Strand cDNA Synthesis Kit (#1631, Fermentas). For each sample 1 µg total RNA and random hexamer primers were used, and run according to protocol.

PCR was done using High-Fidelity DNA polymerase (Phusion, F-530S). 1 µL of each cDNA sample and 10 pmol of each of the primers were used. The samples were first run with GAPDH housekeeping gene in order to control the product before running with specific primers. Also, samples without reverse transcriptase (RT minus samples) were run to test for contamination. The samples were run at; 1 cycle 98 °C (1 minute), 35 cycles 98 °C (30 seconds), annealing temperature (30 seconds), and 72 °C (30 seconds), and 1 cycle 72 °C (5 minutes). The melting temperature (***T_m)*** of each primer was calculated according to the following formula:

T_m = 64.9°C + 41°C x (number of G's and C's in the primer – 16.4)/N

T_a was set 5-8 degrees lower than T_m, and further modified.

Primers were chosen from the list made by Want et al. (2000), and specific primer pairs were chosen due to their verification in the literature (Morrow et al. 2009), reporting a 100% coverage. PCR was run both with the original primers and with primers modified for cloning by a ligation-independent cloning (LIC) sequence 5'-primer 5'-GACGACGACAAGATg-3' and 3'primer 5'-taACCGGGCTTCTCCTC-3' , and restriction sites *Bam*HI(GGATCC) and *Sal*I (GTCGAC) were introduced in the forward and reveres primers respectively. All primers were purchased from DNA technology**.** The different primers, their nucleotide sequence and the annealing

Primer	Nucleotide Sequence	T_a
GAPDH Forward primer	5'-CAAGGTCATCCATGACAACTTTG - 3'	55 °C
GAPDH Reverse primer	5'-GTCCACCACCCTGTTGCTGTAG - 3'	55 °C
pJET1.2 Forward sequencing primer	5'-CGACTCACTATAGGGAGAGCGGC-3'	58°C
pJET1.2 Reverse sequencing primer	5'-AAGAACATCGATTTTCCATGGCAG-3'	58 °C
light chains v domain Forward primer	50-GAYATTGTGMTSACMCARWCTMCA- 3'	45°C
light chains v domain Reverse primer	5'-GGATACAGTTGGTGCAGCATC- 3'	45°C
light chains v domain Forward modif. primer	5'-GACGACGACAAGATgGGATCCGTCGAC GAYATTGTGMTSACMCARWCTMCA- 3'	45°C
light chains v domain Reverse modif. primer	5'-GAGGAGAAGCCCGGTTAGGATCC GGATACAGTTGGTGCAGCATC- 3'	45°C
Heavy chains v domain Forward primer	5'-SARGTNMAGCTGSAGSAGTCWGG- 3'	45°C
Heavy chains v domain Forward primer 1.0	5'-GAAGTGAAGCTGGTGGAGTCTGG- 3'	45°C
Heavy chains v domain Reverse primer	5'-CTTGACGCATCCTAGAGTCA- 3'	45°C
Heavy chains v domain Forward modif. primer	5'-GACGACGACAAGATgGGATCCGTCGAC SARGTNMAGCTGSAGSAGTCWGG- 3'	45°C
Heavy chains v domain Reverse modif. primer	5'-GAGGAGAAGCCCGGTTAGGATCC CTTGACGCATCCTAGAGTCA- 3'	45°C

Tabel 2; The degenerated nucleotides are defined according to IUPAC Nomenclature (Y- T/C, M-A/C, R-A/G, W-A/T, S-G/C). The modified primers are market with "modif." in the end, and the modification areas are underlined.

temperatures can be seen in Table 2.

The PCR products were loaded in 1% agarose gels containing 0.5 µg/mL ethidium bromide at 100V for 20 min, and visualized on the Kodak Image Station 4000MM Pro (Carestreamhealth). As ladder 1 kb (#00068368, Fermentas) was used.

Immunohistochemistry

Immunohistichemistry was carried out on paraffin sections of hippocampal brain tissue from human Alzheimer brains as described in (Axelsen et al. 2009; Axelsen et al. 2011)As negative control we used MAb32.3 targeting chlamydia trachomatis described in (Birkelund et al. 1988; Hobolt-Pedersen et al. 2009). Images were taken with Leica DMRBE microscope equipped with a Qimaging cooled color 12 bit camera (QICAM 12-bit).

ELISA

$A\beta^{1-42}$ was diluted in phosphate buffered saline ((PBS) 2µg/mL) and used for coating in ELISA assays as described in (Axelsen et al. 2011). Supernantant containing the anti amyloid antibodies was diluted 1:5, 1:10, 1:20, 1:40, 1:80, 1:160 and 1:320 and added in duplicate, for primary antibody description see introduction section Monoclonal antibodies. As negative control CRL-2434 Fc-neonatal receptor from ATCC was used. As secondary antibody we used Peroxidase – conjugated affinity purified goat-anti-mouse IgG (H+C) (Jackson USA) (Axelsen et al. 2011) diluted 1:40,000.

mRNA analysis

mRNA was purified and cDNA clones were synthesized using TruSeg RNA Sample Preparation v2 Guide from ilumina verified by agerose gel, and run on ilumina HiSeg2000.

The data were imported in CLC bio genomic workbench, trimmed (2<N, reads >100bs.) and mapped against reference genes in order to create a consensus sequence defining the targeted gene. The reference genes were chosen due to their resemblance to the expected gene. For heavy chain MUSH21C accession number: L35037, the gene was further modified by adding the first 77 nucleotides from Mus Musculus immunoglobulin heavy chain complex, mRNA (cDNA clone) accession number: BC003435, to the start of the MUSH21C gene. As Reference for Light chain Mus Musculus mRNA for immunoglobulin, light chain kappa (IGK gene) Accession number X87231 was used.

The results from mapping to reference was further modified manually, in cases of low coverage, by blasting against a database containing the illumine sequences. The sequences were added to consensus and the process was repeated with the new 30 5` end nucleotide sequence. After 5 rounds the 5`end of the complete mRNA was obtained. Finally the new consensus was used as a reference in a new mapping to reference.

Bioinformatics software for nucleotide analysis

Genomics Workbench version 5.1(CLCbio, Aarhus, Denmark) was used. The *blastn* algorithm (Altschul et al. 1997) was used in the programs Blast version 2.2.25+, wwwblast version 2.2.26 (National Center for Biotechnologi Information, U.S. National Library of Medicine, Rockville Pike, Bethesda MD). The *getorf* command was used in the Emboss package version 2.10.0 ((Rice et al. 2000). Hardware used: Sun Fire X2200 M2 server (8 cores, 48 GB RAM) and Sun Fire X4600 server (16 cores, 32 GB RAM) (Sun Microsystem – Oracle, La Jolla, CA).

Constructing, searching and extracting sequences from a blast database

Sequence reads were trimmed in Genomics Workbench (length 120-150, max two N's) and exported in fasta format. A blast database was created with *makeblastdb* (makeblastdb -in [database fasta format] -parse_seqids -dbtype nucl).

A file for database query with sequence for search was made in fasta format with *emacs* and named [query].fasta. A list of matching sequences was obtained with *blastn* search (blastn -num_threads 8 -db [database] -query [query].fasta -outfmt 6 –out [query].out). To obtain the accession numbers from the file *awk* was used (awk '{print $2}' [query].out > [query].fetch. For retrieval of the sequences from the database *blastdbcmd* command was used (blastdbcmd –entry_batch [query].fetch –db [database] –out [query].fasta). The files [query].fasta was imported into Genomics Workbench and assembled.

Protein analysis

Protein Digestion

100 µg expressed purified antibodies were subjected to in-solution enzymatic digestion with sequencing grade Trypsin (*Promega*). The proteins were diluted in 1% sodium deoxycholate, 50mM triethylammonium bicarbonate (TAEB) and heat-denatured at 99 °C for 5 min, then cooled to ambient temperature. Proteins were reduced by tris(2-carboxyethyl)phosphine (*Sigma-Aldrich*) 1:25 (w/w) to sample protein for 20 min at 37 °C. Following reduction the samples were alkylated by iodoacetic acid (*Fluka*) 1:10 (w/w) for 30 min at 37 °C. Finally, 1:50 (w/w) trypsin (*Promega*) was added and the proteins were digested overnight at 37 °C.

LC-MS/MS Analysis

The peptides were analyzed by a nanoflow UPLC (*ThermoFisher Scientific; Dionex Ultimate3000/RSLC*) system coupled online by a nanospray ion source (*ThermoFisher Scientific Proxeon*) to an Orbitrap Q-Exactive mass spectrometer (*ThermoFisher Scientific, Bremen, Germany*). The peptides were loaded onto a 2 cm reversed phase Acclaim PepMap100 C18 Nano-Trap Column (*ThermoFisher Scientific*) with 4 µL/min in 2% solvent B and 98% solvent A. The peptides were then separated using a 15 cm reversed phase Acclaim PepMap300 C18 column (*ThermoFisher Scientific*), and eluted with a linear gradient of 4% solvent B which was increased to 40% solvent B over 35 min at a constant flow rate of 300 nL/min.

The mass spectrometer was operated in a data-dependent mode to switch between full MS scans and tandem MS/MS. A top 12 mode was applied that acquired one full MS scan at a range of m/z

325-2000 at a constant resolution of 70,000 (@ m/z 200), and up top 12 MS/MS scans at a constant resolution of 17,500 (@ m/z 200). Fragmentation was performed using higher-energy collision-induced dissociation (HCD) and sequenced precursor ions were dynamically excluded for 30 s.

The raw mass spectrometry files were analyzed by Thermo Proteome Discover (version 1.3.0.339). For search of MS/MS spectra in house Mascot server version 2.2 (Matrix sceinece Ltd., London, UK) was used, on a Sun Fire X2200 M2 server. BioTools version 2.1 (Bruker Daltronics, Bremen, Germany) was used for visualizing data.

Results

Immunochemistry

In order to determine the ability of the antibodies to target Aβ, MAb1.1, MAb 19.2.33 and MAb23.3.22 were first controlled by ELISA., data is not shown for MAb 23.3.22 Non purified supernantant from the antibody producing hybridoma cell lines, was added wells coated with Aβ. The antibodies were diluted in double series, halving the concentration per dilution (figure 6). Figure 6 presents the dilution series when secondary antibody was added in the concentration of 1:40,000. As expected, the MAbs bind the Aβ antigen, and the binding is decreased according to the dilution concentration. Furthermore the control antibody targeting the FC receptor shows no sign of binding.

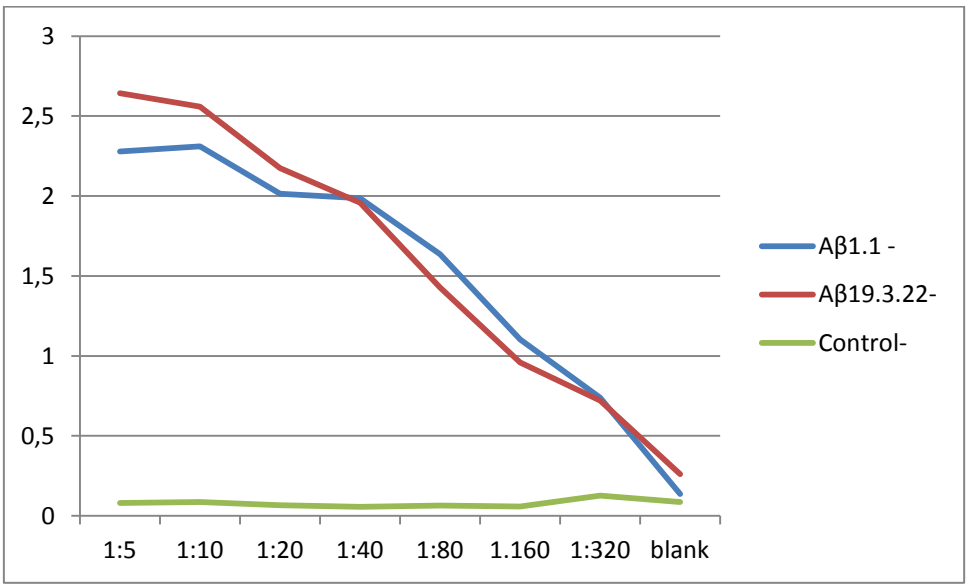


Figure 6; the x-axis presents the delution of supernantant containing antibodies, the Y-axis presents the antibodies binding ability at different concentrations.

To further investigate the properties of the MAb1.1, MAb 19.2.33 and MAb23.3.22, paraffin embedded sections of human AD brain tissue was exposed to supernatant from the Hybridoma cells,

containing the antibodies, ensuring their ability to bind Aβ in human tissue. Pictures A-C are of MAb1.1, D-F; is MAb19.2.33, G-I; MAb23.3.22 and J-L; is negative control. From the pictures A-I it is clear that all three of the antibodies react with Aβ plagues in human brain tissue and the negative control is negative (figure 7).

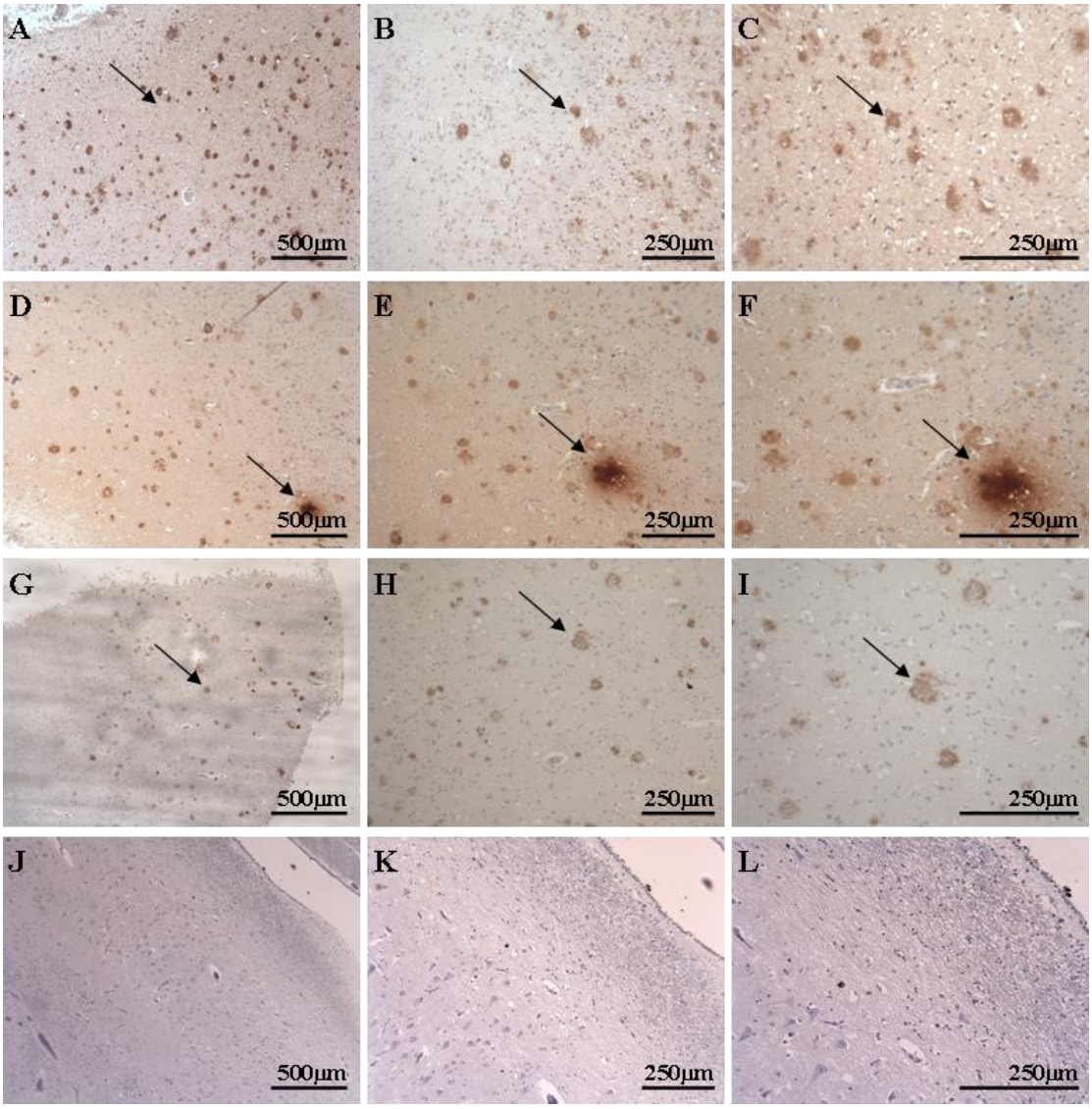


Figure 7; Immunohistochemestry on human AD tissue from Hipocampus, by second antibody peroxidase goat-anti-mouse. A-C; MAb1.1, D-F; MAb19.33, G-I; MAb23.3.22 J-L; is negative control (targeting Chlamydia trachomatis). In each picture stained Aß plaques are highlighted by a black arrow and the scale is in µm.

Conventional method for amplification

RT-PCR

First we tried to amplify the genes encoding the MAbs by the generally known method using RT-PCR and degenerated primers. The RNA was purified from the antibody producing cell, and the RNA quality was controlled by gel electrophoresis. In Figure 8 ribosomal RNA bands of 18s, 28s is seen. Common for each of the cell lines are that 28s seems to be at higher concentration than 18s. Furthermore MAb23.3.22 has clearer bands than that of MAb1.1 and MAb19.2.33 indicating a higher concentration, which is also seen in the concentration measurement by nanodrop (table figure 8).

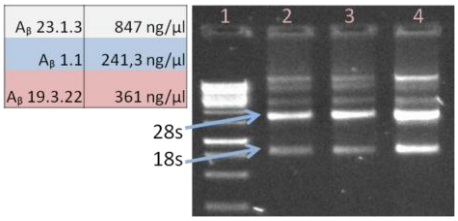


Figure 8, The table to the left displays the concentration of RNA ng/µl, for each of the antibodies. Lane 1; 1 kb ladder, Lane 2; MAb1.1, lane 3; MAb 19.3.22, lane 4; MAb 23.1.33. The ribosomal RNA are indicated by arrows, 18s (1.9kb) and 28s (4.7 kb)

After the verification of the RNA, cDNA synthesis was initiated. cDNA from each cell line was synthesized using random hexamer primers, data not shown.

After cDNA synthesis LC and HC mRNA sequences were amplified using degenerated primers described by (Wang et al. 2000). Both original primers from (Wang et al. 2000), and modified primers for easier cloning were used with different melting temperatures for amplification optimization. We succeeded in amplifying the cDNA for the LC with modified as well as non modified primers for MAb1.1, MAb 19.2.33 and MAb 23.3.22, for non modified primers (figure 9, lane 2-4), modified primers data not shown. We did not succeed in amplifying the HC with either modified or non modified primers, for non modified primers (figure 9 lane 5-7), data for modified primers not shown. A second try for amplification of HC was done by cDNA synthesis with specific primers instead of random primers, modified, and non modified, followed by RT-PCR with specific primers. Even so no positive amplification for HC was achieved, data not shown.

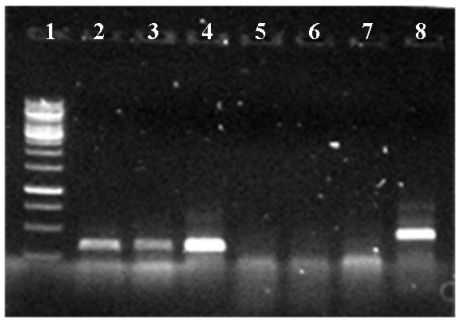


Figure 9, 1% agerose gel, showing the amplification product of heavy and light chain with non modified primers. Lane 1; ladder 1kb, lane 2; MAb1.1, lane 3; LC, MAb19.3.22 LC, Lane 4; MAb23.1.33 LC, lane 5; MAb1.1, lane HC; LC, lane 6; MAb19.3.22 HC, Lane 7; MAb23.1.33 HC, lane 8 Control Gapdh.

The products amplified by RT-PCR were further used for transfection of Nova blue cell using micropulser for cloning. The vectors used for transfection depended upon the use of primers modified/non modified see method and materials for cloning We only succeeded in cloning LC amplified by non modified primers. The clones were controlled by RT-PCR with pjet1.2 primers, and positive clones were further sub-cloned and amplified see example (figure 10). The positive sub-clones were amplified and the plasmid DNA was purified by miniprep. Finally the purified plasmids were controlled by restriction enzyme digestion, and send to DNA Technology for sequencing. Only two positive results were obtained, one MAb 19.2.33, and one MAb23.3.22 data not shown.

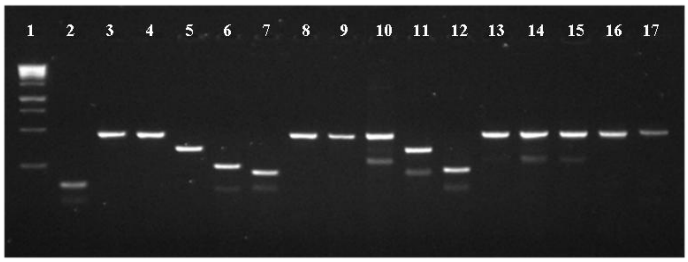


Figure 10, control of clones transfected with RT-PCR product amplified by non modified primers by RT-PCR using pjet1.2 primers. Lane 1; 1kb ladder, lane 2-8 ecoli clones with Aβ1.1 LC RT-PCR product insert, lane 9-16 ecoli clones with Aβ23.1.33 LC RT-PCR product insert. Lane 3-4, 8-10 and 13-17 is positive for LC.

In Summary, we succeeded in purifying the RNA from each of the antibody producing cells, amplifications only succeeded for LH, and after cloning, sub-cloning and miniprep, only DNA sequences from two samples were obtained.

New approach

The lack of results using conventional methods for amplification described both in the literature and shown in our own study calls for a new approach. Our new approach is a comparative controlled analysis ensuring high validity of the results. The approach is divided in two different analyses, one on protein level and one on mRNA level creating of a common consensus sequence.

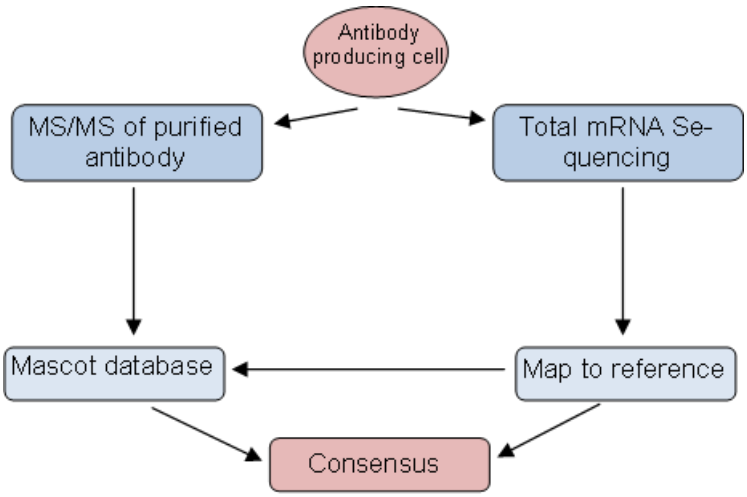


Figure 11, The strategy used, starting with the antibody producing cell at the top. To the left we have the protein analysis starting with tandem mass spectrometry of the antibody – next the formation of a mascot database containing the non immunoglobulin sequences in addition to translated sequences found by mRNA sequencing of total mRNA – and mapped to reference creating a consensus. Finally the experimentally found spectra are searched against the Mascot database – validating the result and the consensus.

The mRNA analysis starts with sequencing of the total mRNA of a specific cell line, in our case, hybridoma cells producing our anti Aβ antibodies. The total mRNA sequence is mapped to a reference gene believed to resemble your desired gene, in our case a gene for IgG heavy and light chain. After mapping to reference, a consensus sequence generated. This consensus sequence is translated to protein and appended to the Mascot database. The experimentally found MS/MS spectra are then used for searching the Mascot database. The mascot database contains among others the mass spectra's from the MS/MS analysis. The map to reference and the result from the search against the mascot database is assembled and compared, thus validating the result.

RNA analysis

For the RNA analyses, the mRNA was first extracted from the antibody producing cells using the before mentioned RNA purification see figure 8 The mRNA extraction was done using oligo dT magnetic beads capturing the poly A tail rich mRNA, while discarding most of the ribosomal and other non-messenger RNA by multiple wash processes.

A_β 1.1	83,1 ng/µl
A_β 19.3.22	55,4 ng/µl
A_β 23.1.3	62,6 ng/µl

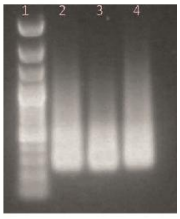


Figure 12, The table to the left displays the concentration of RNA ng/µl, for each of the antibodies. Lane 1; 1 kb ladder, Lane 2; mRMA from MAb1.1 producing cells, lane 3; mRNA from MAb 19.2.33 producing cells, lane 4; mRNA fromMAb 23.3.22 producing cells.

In order to control a successful mRNA extraction, the product was run on a 1% agerose gel, see figure 12. From figure 12 it is clear that the ribosomal RNA is removed from the samples. After mRNA extraction DS cDNA was synthesizes.

The DS cDNA was then run on the HiSeq Systems - *Illumina*2000. In the data presented, the adopters wereere removed, herby not interfering with the analysis. The data from sequencing of cDNA from MAb1.1, MAb19.2.33 and MAb23.3.22 mRMA were imported and trimmed in CLC genomic work bench.

In the only result following the MAb1.1 analysis is shown.

	Count	Average length	Total bases
Reads	68,681,010	148.17	10,176,765,509
Matched	1,906,653	148.61	283,341,550
Not matched	66,774,357	148.16	9,893,423,959
References	3	1,317	3,951
Reads in pairs	1,814,300	182.27	
Broken paired reads	92,353	147.41	

Table 3, summery repport from clc after mapping to reference HC.

The MAb1.1 data were mapped to reference, of the 68,681,010 reads, 1,906,653 matched our reference gene and a consensus sequence was obtained see table 3. Due to the highly variable areas of the LC and HC, large gabs and a low coverage was observed in parts of the mapping figure 13A.

To obtain the variable 5` end of the HC mRNA, a blast database was made with the Illumina sequences. The 30 nucleotides sequence from the 5 `end with high coverage was used for a blast search against the database (max 500 reads). The 500 reads were assembled with sequence assembling in CLC-Bio. The sequences were added to consensus and the process was repeated with the new 30 5` end nucleotide sequence. After 5 rounds the 5`end of the complete mRNA was obtained.

To control the semi manual assembly, the new consensus was used as reference in the “mapping to reference” thus a high coverage was obtained of the full sequence ranging between 50,000-100,000 reads see figure 13B.

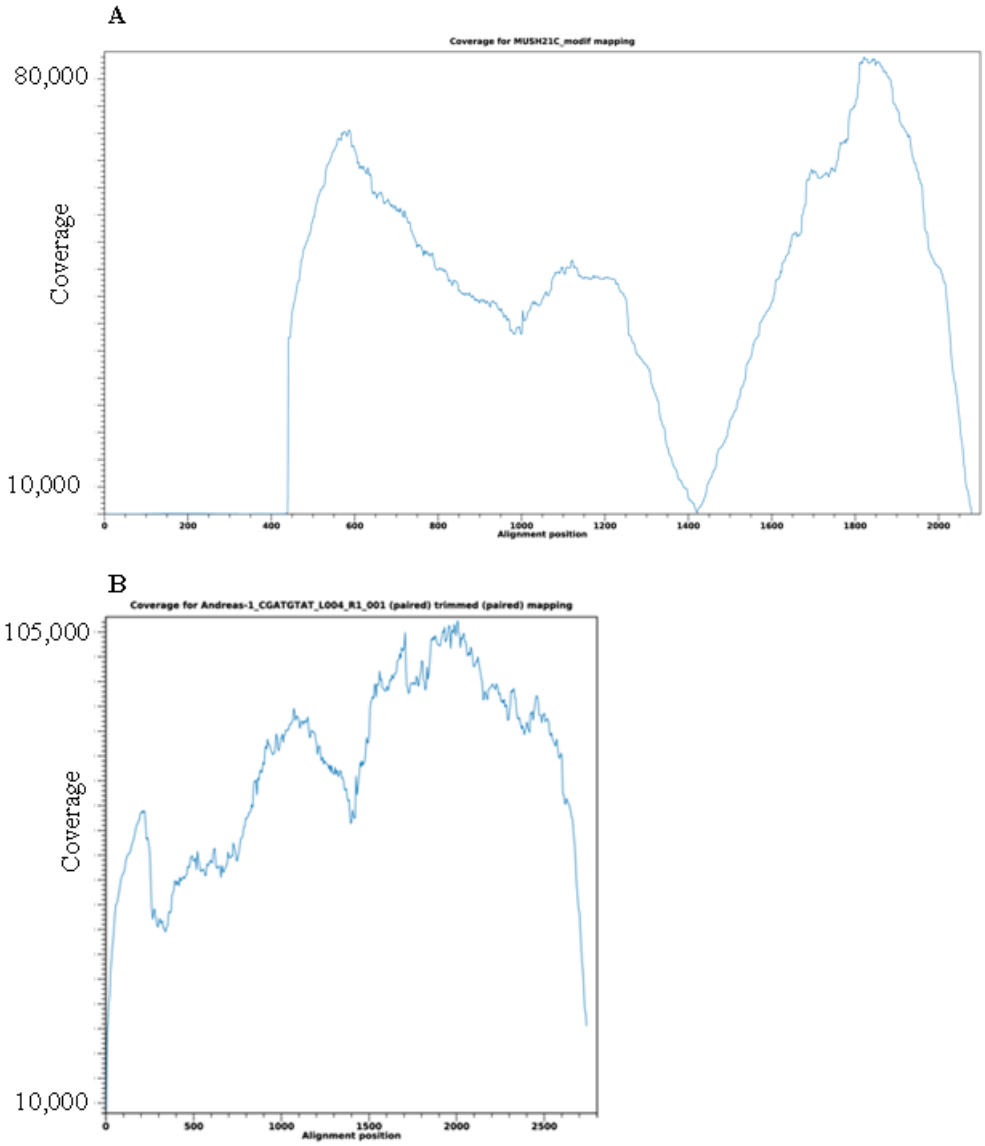


Figure 13, A and B illustrations from CLC bio tools. A illustrates the Coverage of mapping to reference HC, low coverage is seen before 420 and at 1400. B illustrates the Coverage of mapping to reference with new modified consensus, overall full coverage with the lowest coverage at approximately 50,000 reads.

Protein analyses

Before analysis, the purified MAb1.1, MAb 19.2.33 and MAb23.3.22 were digested with trypsin, cutting the protein at Lys and Arg residues, dividing the antibodies in smaller peptides for easier analysis.

The protein analysis was done using tandem in space MS/MS instrument the Orbitrap Q-Exactive mass spectrometer (*ThermoFisher Scientific, Bremen, Germany*). In one part of the instrument, the peptides are selected according to mass-to-charge ratio (m/z) by ionscore, the peptides are then exposed to further fragmentation in a HCD collision cell and finally analyzed in Orbitrap giving a high sensitive result (m/z). Figure 15 is an example of a MS/MS result each spike represents a specific peptide fragment with a specific m/z.

The common peptides after fragmentation and their nomenclature are seen in figure 14. The peptides are named according to their ions preserved at either the N or C terminal: a, b and c if the charge is preserved at the N-terminus, and x, y and z if the charge is preserved at the C terminus. The most common cleavage happens between the carbonyl oxygen and the amide nitrogen, giving a fragment b- and y- ion.

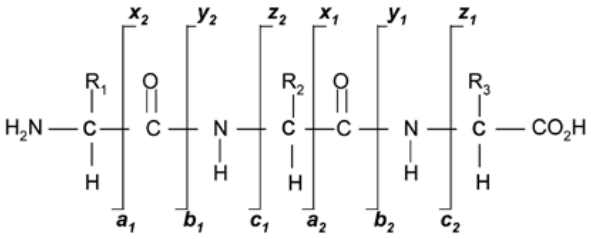


Figure 14; (Lim & Elenitoba-Johnson 2004)*. The figure shows the most common cleave sites and fragmentations by collision induced dissociation. Peptides preserving their ions at the N terminal are called a, b and c, while peptides with preservation at the C terminal are called x, y and z.*

Since the fragmentation is random, peptides of different length and m/z score are observed, see figure 15 for example. The sequence of the observed peptide is **VNDYANQAWFPYWGQGTLVTVSAAK.** The sequence is further indicated on figure 17 by underscore.

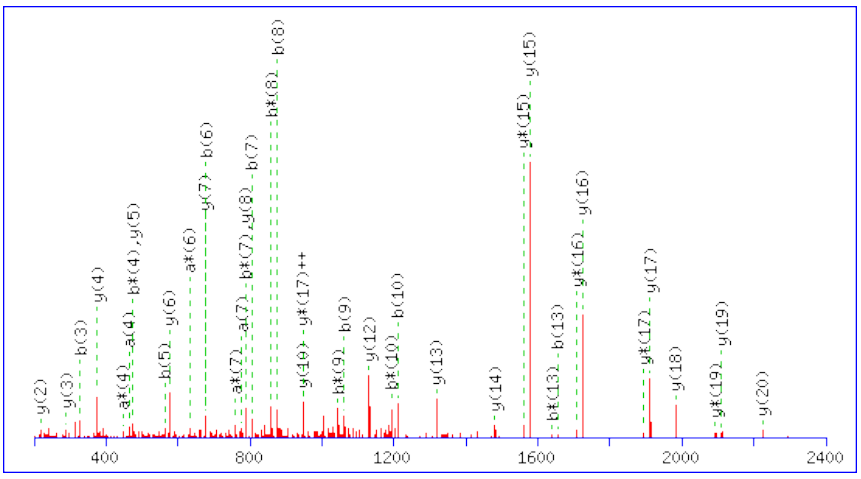


Figure 15; A single peptide found when searched our observed spectra's against our database. Each spike represents a fragment of a peptide, defined by an ion. The collected peptide is; VNDYANQAWFPYWGQGTLVTVSAAK.

The observed ion spectra are then compared with a theoretical database. In our case the database contains, among others, the translated consensuses sequences from the mRNA analysis, translated into peptide sequences. A theoretical digestion of the peptide by trypsin gives rise to the theoretical specters that would be expected from the MS/MS analysis. The observed spectra's are run against the theoretical spectra's comparing the theoretical spectra's with the observed. The proteins are then matched according to their mass matches, and ranked. The higher the matches are, the more likeliness that the discovered protein is valid.

In our case, the database contains among others our translated consensuses from our mRNA analysis, both LC and HC from each cell line and the experimental spectra of MAb1.1 has been searched against these.

Figure 16 shows the search of MAb1.1 experimental spectra's against the database. For each spectrum that can be matched against the peptide, mascot calculates an overall protein score, in this case 12127 see figure 16. Furthermore the observed values and the expected values are given. It is also seen that some of the peptides found, have missing trypsin sites. There are 275 peptides matching our protein, however some of these are duplicates.

MAb1.1_heavy Mass: 51862 Score: 12127 Queries matched: 275 emPAI: 44.64

MAb1.1_heavy

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
70	418.7407	835.4668	835.4664	0.48	1	34	0.00039	1	K.GRFTISR.D
520	550.8325	1099.6505	1099.6489	1.46	0	43	4.8e-05	1	K.DVLTITLTPK.V 518 519 521
549	560.3085	1118.6025	1118.6006	1.73	0	48	1.4e-05	1	K.VTCVVVDISK.D 548
627	579.2653	1156.5160	1156.5149	0.93	0	51	7.5e-06	1	R.EEQFNSTFR.S 628
761	605.8460	1209.6775	1209.6758	1.42	0	63	5.1e-07	1	K.APQVYTIPPPK.E 757 758 759 760 762 763
823	415.2279	1242.6619	1242.6608	0.88	0	69	1.2e-07	1	R.VNSAAFPAPIEK.T 820 821 822 824 825 826 827
1108	662.3587	1322.7028	1322.7016	0.92	0	(67)	1.8e-07	1	R.NIVYLQMSSLR.S 1105 1106 1107 1109 1110 1111 1112
1171	670.3568	1338.6991	1338.6965	1.88	0	76	2.5e-08	1	R.NIVYLQMSSLR.S 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1172
1266	683.7748	1365.5350	1365.5329	1.53	0	(69)	1.3e-07	1	R.SEDTAMYYCAR.V 1264 1265 1267
1328	691.7716	1381.5285	1381.5278	0.52	0	97	1.9e-10	1	R.SEDTAMYYCAR.V 1326 1327 1329 1330 1331 1332 1333
1642	749.4357	1496.8568	1496.8562	0.36	0	95	3.4e-10	1	K.LVESGGGLVKPGGSLK.V 1640 1641 1643 1644 1645 1646 1647 1648 1649 1650 1651 165
2675	899.4840	1796.9535	1796.9495	2.23	1	56	2.7e-06	1	K.APQVYTIPPPKEQMAK.D 2672 2673 2674 2676 2677
2737	605.3230	1812.9472	1812.9444	1.54	1	(39)	0.00012	1	K.APQVYTIPPPKEQMAK.D 2732 2733 2734 2735 2736 2738 2739 2740
2864	927.4676	1852.9206	1852.9142	3.49	0	66	2.3e-07	1	R.SVSELPIMHQDWLNGK.E 2863
2918	935.4622	1868.9098	1868.9091	0.37	0	(36)	0.00025	1	R.SVSELPIMHQDWLNGK.E 2915 2916 2917 2919 2920 2921 2922 2923 2924
3286	655.9684	1964.8833	1964.8826	0.36	0	(83)	4.9e-09	1	K.NTQPIMDTDGSYFVYSK.L 3287 3288 3289 3290 3291 3292 3293 3294 3295
3362	991.4487	1980.8828	1980.8775	2.66	0	93	4.7e-10	1	K.NTQPIMDTDGSYFVYSK.L 3356 3357 3358 3359 3360 3361 3363 3364 3365 3366 3367 33
3805	1057.4790	2112.9434	2112.9398	1.75	0	110	1.1e-11	1	K.VSCVASGFTFSSYAMTWGR.Q 3802 3803 3804 3806
3977	1108.0509	2214.0872	2214.0845	1.26	0	124	4.1e-13	1	R.LEWVASISVSGNTYYPDSVK.G 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 397
4172	753.3820	2257.1241	2257.1201	1.74	1	62	7e-07	1	R.SVSELPIMHQDWLNGKEFK.C
4211	758.7136	2273.1190	2273.1150	1.76	1	(27)	0.002	1	R.SVSELPIMHQDWLNGKEFK.C
4412	1186.1022	2370.1898	2370.1856	1.78	1	110	9.5e-12	1	K.RLEWVASISVSGNTYYPDSVK.G 4385 4386 4387 4388 4389 4390 4391 4392 4393 4394 43
4864	1274.6252	2547.2359	2547.2315	1.72	1	104	4.4e-11	1	K.NTQPIMDTDGSYFVYSKLNVQK.S 4865
4901	855.4154	2563.2244	2563.2265	-0.80	1	(2)	0.58	1	K.NTQPIMDTDGSYFVYSKLNVQK.S
5285	899.7544	2696.2413	2696.2363	1.86	1	12	0.067	1	K.VSCVASGFTFSSYAMTWGRQTPEK.R
5477	1393.6847	2785.3548	2785.3500	1.73	0	117	1.9e-12	1	R.VNDYANQAWFFYWGQGTLVTVSAAK.T 5475 5476 5478 5479 5480
5730	1423.1584	2844.3023	2844.2991	1.12	0	109	1.3e-11	1	K.DDPEVQFSWFVDDVEVHTAQTQPR.E 5729 5731 5732 5733 5734 5735 5736
5821	954.1626	2859.4660	2859.4511	5.21	0	80	1e-08	1	K.TTPPSVYPLAPGSAAQTNSMVTLGCLVK.G 5819 5820
5886	959.4890	2875.4450	2875.4460	-0.33	0	(41)	8e-05	1	K.TTPPSVYPLAPGSAAQTNSMVTLGCLVK.G 5887
5986	969.1107	2904.3103	2904.2998	3.63	0	51	8e-06	1	K.SNWEAGNTFTCSVLHEGLHNHHTEK.S 5983 5984 5985
6040	974.8113	2921.4120	2921.3948	5.88	0	95	3.4e-10	1	R.DCGCKPCICTVPEVSSVFIFPPKPK.D 6018 6019 6020 6021 6022 6023 6024 6025 6026 602
7726	1315.9739	3944.8998	3944.8892	2.70	1	57	2e-06	1	K.VTCVVVDISKDDPEVQFSWFVDDVEVHTAQTQPR.E 7727 7728 7729
9648	1639.5512	6554.1755	6554.1541	3.26	0	48	1.7e-05	1	K.GYFPEPVTVTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVPSSTWPSETVTCNVAHPASSTK.V

Figure 16; A Mascot search of our experimental spectra's against our database. Observed is the mass of the spectra observed, Mr (expt) is the mass of the predicted peptide, Mr (calc) is the calculated mass, miss is the number of trypsin residues missed, score is the ion score, Expect; the frequency that the peptide would occur by chance, and the peptide amino acid sequence.

In total, the MS/MS analysis of the MAb1.1 antibody gives a coverage of 81% for HC see figure 17 and 64% for light chain (figure 18). When searched against our database, the observed peptides are marked with red. Furthermore the search finding the LC with coverage of 64% had a protein score of 7286 and 156 peptides matched the protein.

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **81%**

Matched peptides shown in **Bold Red**

```
  1 MNFGFSLIFL VLVLKGFQCE VKLVESGGGL VKPGGSLKVS CVASGFTFSS
 51 YAMTWGRQTP EKRLEWVASI SVSGNTYYPD SVKGRFTISR DNARNIVYLQ
101 MSSLRSEDTA MYYCARVNDY ANQAWFPYWG QGTLVTVSAA KTTPPSVYPL
151 APGSAAQTNS MVTLGCLVKG YFPEPVTVTW NSGSLSSGVH TFPAVLQSDL
201 YTLSSSVTVP SSTWPSETVT CNVAHPASST KVDKKIVPRD CGCKPCICTV
251 PEVSSVFIFP PKPKDVLTIT LTPKVTCVVV DISKDDPEVQ FSWFVDDVEV
301 HTAQTQPREE QFNSTFRSVS ELPIMHQDWL NGKEFKCRVN SAAFPAPIEK
351 TISKTKGRPK APQVYTIPPP KEQMAKDKVS LTCMITDFFP EDITVEWQWN
401 GQPAENYKNT QPIMDTDGSY FVYSKLNVQK SNWEAGNTFT CSVLHEGLHN
451 HHTEKSLSHS PGK
```

Figure 17 Coverage mark in red, of the consensus translated MAb1.1 HC.

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **64%**

Matched peptides shown in **Bold Red**

```
  1 MVFTPQILGL MLFWISASRG DVVLTQSPAT LSVMPGDSVS LSCRASQSIS
 51 DNLHWYQQQS HASPRLLIKY SSQSVSGIPS RFSGSGSGTV FTLSISSVET
101 EDFGLYFCQQ SDSWPLTFGP GTKLELKRAD AAPTVSIFPP SSEQLTSGGA
151 SVVCFLNNFY PKDINVKWKI DGSERQNGVL NSWTDQDSKD STYSMSSTLT
201 LTKDEYERHN SYTCEATHKT STSPIVKSFN RNEC
```

Figure 18; Coverage mark in red, of the consensus translated MAb1.1 LC.

The coverage and distribution of the peptides for the HC can be seen in the following figure 19 it is seen that the peptides are distributed over the whole sequence, and that especially the beginning of the protein has a high coverage.

Figure 19; Distribution and coverage of peptides covering the MAb1.1. It is seen that there is a general coverage and that it is particularly high in the beginning of the MAb1.1 protein.

Discussion

We succeeded in determining the amino acid sequence of MAb1.1, HC and LC respectively, using a combination of MS/MS data and mRNA sequencing. This was done by mapping to reference in CLC genomic workbench, and further modifying the areas with low coverage and large gaps. Using the new consensus for mapping to reference resulted in sequence converge depth of more than 50.000 (figure 13 B).

The large gaps and low coverage observed in the initial mapping might be explained by the use of the gene as a reference. The difference between the reference and the expected gene was too great in the variable areas, and the areas are stretched over too large an area, since the program is only capable of mapping smaller gaps and gene differences. The gaps seen in the sequence were larger than the reads obtained (>150) and especially the 5´end of the sequence had a low coverage see (13 A). Taking the construct of the gene into consideration this might have been foreseen, since the variable area for the HC is 400 nucleotides in the 5'end of the coding sequence of the gene. Also the sequence encoding the leader sequence is highly variable, which can be seen, by making a Blast search of the sequence against the NCBI database.

After the second round of mapping the consensuses obtained were translated into peptides and once more blasted at NCBI in order to verify that the genes were of the immunoglobulin superfamily. An amino acid identity of 100 percent was to some IgG1 sequences in the constant part of the HC 140-463 peptide sequence. And the most variable areas were found,on protein level, located to the areas of HC (1-140) and LC (25-125).

In order to validate the mapping, the translated sequences were uploaded in a mascot database that in addition contained non immunoglobulin sequences. The spectra's from the MS/MS analysis were searched against the database. Both HC and LC chains were found during the search, HC with coverage of 81% and LC with a coverage of 61%. The real coverage is even higher, since the translated consensuses uploaded in the database contain a leader sequences that are not present in the antibodies because the leader sequence is cut off the peptide chain when entering the endoplasmic reticulum before assembling the chains to the mature antibody molecule (Agger et al, 2005). Thus, no spectra were identified within the leader sequence (figure 17 and 18). When the MS/MS spectras were searched against the NCBI database, bovine antibodies or other mouse proteins were not identified, showing that the cells cultivated in serum free media resulted in production of pure.

Furthermore, the relatively low coverage of the LC compared to HC, can be explained by the lack of trypsin sites in this region resulting in a peptide with a molecular weight of more than 5,000 Da (Figure 18). The size of the peptide prevents the MS/MS from making a throughout analysis. Using other settings for the mass spectrometer or using other proteolytic enzymes may possibly overcome this problem. .

The peptide coverage of the HC found (figure 19). shows that the overall coverage was high, and that specific areas such as the start of the molecule also had a high coverage. Since the variable regions of the HC are located between amino acid 1-140 the high coverage in this area truly validate my results of the sequenced mRNA translated into protein.

In order to further validate our MS/MS results and to insure that the peptides are not just matched by chance, I investigated the ionscore for each of the peptides found (figure 16). The peptides need to have a score >45 in order to be valid. This is calculated according to a threshold of 0,05 (a 1 in 20 chance of being a false positive). As seen in figure 16, most of the peptides score 45 or higher indicating that the peptides are not found by chance and therefore are correct and valid. This is further supported by the low Expect values defining the frequency that the peptide would occur by chance.

Conventional approach

During the conventional approach we succeeded in amplifying the LC for each of the immunoglobulin mRNA, however no bands were seen for HC. In the original study by Wang et al (2000), they had an overall success rate of 80% using 7 different forward HC primers targeting the FR1 region combined with reverse primers targeting the constant region. The primer we used had 256 fold degeneracy and was verified by Morrow et al (2009). Several reasons may account for the lack of success amplifying the HC.

We may have been able to amplify the HC by using one of the other primers also listed by Wang et al (2000), such as a more degenerated primer of 512 fold. For PCR we used a Phusion polymerase with proofreading, however, such a polymerase may degrade the degenerated primers due to mismatch shown by Wang et al (2000), who used a Vent polymerase with proofreading activity, and showed that in some cases a Tag polymerase may work better with degenerated primers.

The validity of my results amplifying the LC could be questioned by the low annealing temperature used. My best results amplifying the genes were at an annealing temperature of 45 degrees. Such low temperature could lead to nonspecific binding (Rychlik et al. 1990), and the non specific binding could further increase when using degenerated primers. However, the sequencing reveled 100% match of the amplified sequence, but the forward prime had 3 mismatch to the sequenced mRNA obtained by the illumina sequencing.

The lack of success amplifying the HC seems evident in the light of our current results using our new approach. After assembling the consensus for the HC and LC, the primers used in RT-PCR have been searched against the sequenced genes showing no binding site for either the forward degenerated 256 fold primer, or the reverse primer which is designed for targeting the constant part see appendix 1. Furthermore, when searching the primers of the LC no binding site for the forward degenerated primer was found, however the reverse primer was located, and is probably the only reason for a positive amplification see appendix 2. One might even consider us lucky in obtaining any results using those primers on our cDNA.

The problems listed above could have been overcome using the nested primer system proposed by Rohatgi et al (2008), however, this would have demanded pooling of 17 different external forward

primers followed by 17 internal forward primers just to amplify the HC, heavily increasing the expenses of the experiment and one would still not be one-hundred percent sure to get an amplification.

Conclusively, our new approach showed it possible to obtain the complete amino acid sequence of MAb1.1. The method was clearly better compared to methods described in the literature. The only control missing to complete our results, is having the gene amplified this time using specific primers, followed by expression of the antibody and testing its binding specificity compared to the MAb1.1.

Bibliography

Abbott, N.J. et al., 2010. Structure and function of the blood-brain barrier. *Neurobiology of disease*, 37(1), pp.13-25. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19664713 [Accessed March 11, 2012].

Agger, R. et al., 2005. *Immunologi* 4th ed., Biofolia.

Altschul, S.F. et al., 1997. Gapped BLAST and PSI-BLAST : a new generation of protein database search programs. , 25(17), pp.3389-3402.

Axelsen, T.V. et al., 2011. Identification of the shortest Aβ-peptide generating highly specific antibodies against the C-terminal end of amyloid-β42. *Vaccine*, 29(17), pp.3260-9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21371581 [Accessed May 31, 2012].

Axelsen, T.V. et al., 2009. Specific recognition of the C-terminal end of A beta 42 by a high affinity monoclonal antibody. *Molecular immunology*, 46(11-12), pp.2267-73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19447496 [Accessed May 31, 2012].

Bergman, Y. & Cedar, H., 2004. A stepwise epigenetic process controls immunoglobulin allelic exclusion. *Nature reviews. Immunology*, 4(10), pp.753-61. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15459667 [Accessed April 8, 2012].

Birkelund, S., Lundemose, a G. & Christiansen, G., 1988. Chemical cross-linking of Chlamydia trachomatis. *Infection and immunity*, 56(3), pp.654-9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=259341&tool=pmcentrez&rendertype=abstract.

Cardoso, F.L., Brites, D. & Brito, M.A., 2010. Looking at the blood-brain barrier: molecular anatomy and possible investigation approaches. *Brain research reviews*, 64(2), pp.328-63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20685221 [Accessed March 12, 2012].

Di Carlo, M., Giacomazza, D. & San Biagio, P.L., 2012. Alzheimer's disease: biological aspects, therapeutic perspectives and diagnostic tools. *Journal of physics. Condensed matter : an Institute of Physics journal*, 24(24), p.244102. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22595372 [Accessed May 19, 2012].

Chaudhuri, J. et al., 2007. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Advances in immunology*, 94(06), pp.157-214. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17560275 [Accessed March 8, 2012].

Geylis, V. & Steinitz, M., 2006. Immunotherapy of Alzheimer's disease (AD): from murine models to anti-amyloid beta (Abeta) human monoclonal antibodies. *Autoimmunity reviews*, 5(1), pp.33-9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16338209 [Accessed March 5, 2012].

Goodman, M.F., Scharff, M.D. & Romesberg, F.E., 2007. AID-initiated purposeful mutations in immunoglobulin genes. *Advances in immunology*, 94(06), pp.127-55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17560274 [Accessed May 15, 2012].

Hobolt-Pedersen, A.-S. et al., 2009. Identification of Chlamydia trachomatis CT621, a protein delivered through the type III secretion system to the host cell cytoplasm and nucleus. *FEMS immunology and medical microbiology*, 57(1), pp.46-58. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2784215&tool=pmcentrez&rendertype=abstract [Accessed April 12, 2012].

Huang, Y. & Mucke, L., 2012. Alzheimer Mechanisms and Therapeutic Strategies. *Cell*, 148(6), pp.1204-1222. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867412002784 [Accessed March 16, 2012].

Hulliger, M. et al., 1989. Flexible fusimotor control of muscle spindle feedback during a variety of natural movements. *Progress in brain research*, 80(2), pp.87-101; discussion 57-60. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2634288 [Accessed May 31, 2012].

Jiang, C. et al., 2003. In vivo delivery of glial cell-derived neurotrophic factor across the blood-brain barrier by gene transfer into brain capillary endothelial cells. *Human gene therapy*, 14(12), pp.1181-91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12908969.

Jung, D. et al., 2006. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annual review of immunology*, 24(D), pp.541-70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16551259 [Accessed March 16, 2012].

Kathryn L. MacCane, S.E.H., 2006. *Pathophysiology the biologic Basis for Disease in Adult and Children* 5th ed.,

Kerchner, G.A. & Boxer, A.L., 2010. Bapineuzumab. *Expert opinion on biological therapy*, 10(7), pp.1121-30. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3000430&tool=pmcentrez&rendertype=abstract [Accessed May 31, 2012].

Kracker, S. & Durandy, A., 2011. Insights into the B cell specific process of immunoglobulin class switch recombination. *Immunology letters*, 138(2), pp.97-103. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21324342 [Accessed March 9, 2012].

Lim, M.S. & Elenitoba-Johnson, K.S.J., 2004. Proteomics in pathology research. *Laboratory investigation; a journal of technical methods and pathology*, 84(10), pp.1227-44. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15311217 [Accessed May 31, 2012].

Morrow, M.P., Pankhong, P. & Weiner, D.B., 2009. Design and characterization of a plasmid vector system capable of rapid generation of antibodies of multiple isotypes and specificities. *Biotechnology letters*, 31(1), pp.13-22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18777012 [Accessed May 31, 2012].

Mygind, T. et al., 2003. Evaluation of five DNA extraction methods for purification of DNA from atherosclerotic tissue and estimation of prevalence of Chlamydia pneumoniae in tissue from a Danish population undergoing vascular repair. *BMC microbiology*, 3, p.19. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=201026&tool=pmcentrez&rendertype=abstract.

Panza, F. et al., 2011. Monoclonal antibodies against b-amyloid (Ab) for the treatment of Alzheimer ' s disease : the Ab target at a crossroads. , pp.679-686.

Pottiez, G. et al., 2009. Understanding the blood-brain barrier using gene and protein expression profiling technologies. *Brain research reviews*, 62(1), pp.83-98. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19770003 [Accessed March 14, 2012].

Qiu, C. & Kivipelto, M., 2009. State of the art. , pp.111-128.

Reichert, J.M., 2011. Antibody-based therapeutics to watch in 2011. *mabs*, 3(1), pp.76-99. Available at: http://www.landesbioscience.com/journals/mabs/article/13895/ [Accessed March 5, 2012].

Rice, P., Longden, I. & Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends in genetics : TIG*, 16(6), pp.276-7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10827456 [Accessed May 31, 2012].

Rich, R., 2008. *No Title Clinical Immunology principles and Practice* 3rd ed. R. Rich, ed., Elsevier limited.

Rohatgi, S., Ganju, P. & Sehgal, D., 2008. Systematic design and testing of nested (RT-)PCR primers for specific amplification of mouse rearranged/expressed immunoglobulin variable region genes from small number of B cells. *Journal of immunological methods*, 339(2), pp.205-19. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18926828 [Accessed April 4, 2012].

Rychlik, W., Spencer, W.J. & Rhoads, R.E., 1990. Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic acids research*, 18(21), pp.6409-12. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=332522&tool=pmcentrez&rendertype=abstract.

Samadi, H. & Sultzer, D., 2011. Solanezumab for Alzheimer's disease. *Expert opinion on biological therapy*, 11(6), pp.787-98. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21504387 [Accessed May 31, 2012].

Schatz, D.G. & Ji, Y., 2011. Recombination centres and the orchestration of V(D)J recombination. *Nature reviews. Immunology*, 11(4), pp.251-63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21394103 [Accessed March 12, 2012].

Schatz, D.G. & Swanson, P.C., 2011. V(D)J recombination: mechanisms of initiation. *Annual review of genetics*, 45(D), pp.167-202. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21854230 [Accessed April 11, 2012].

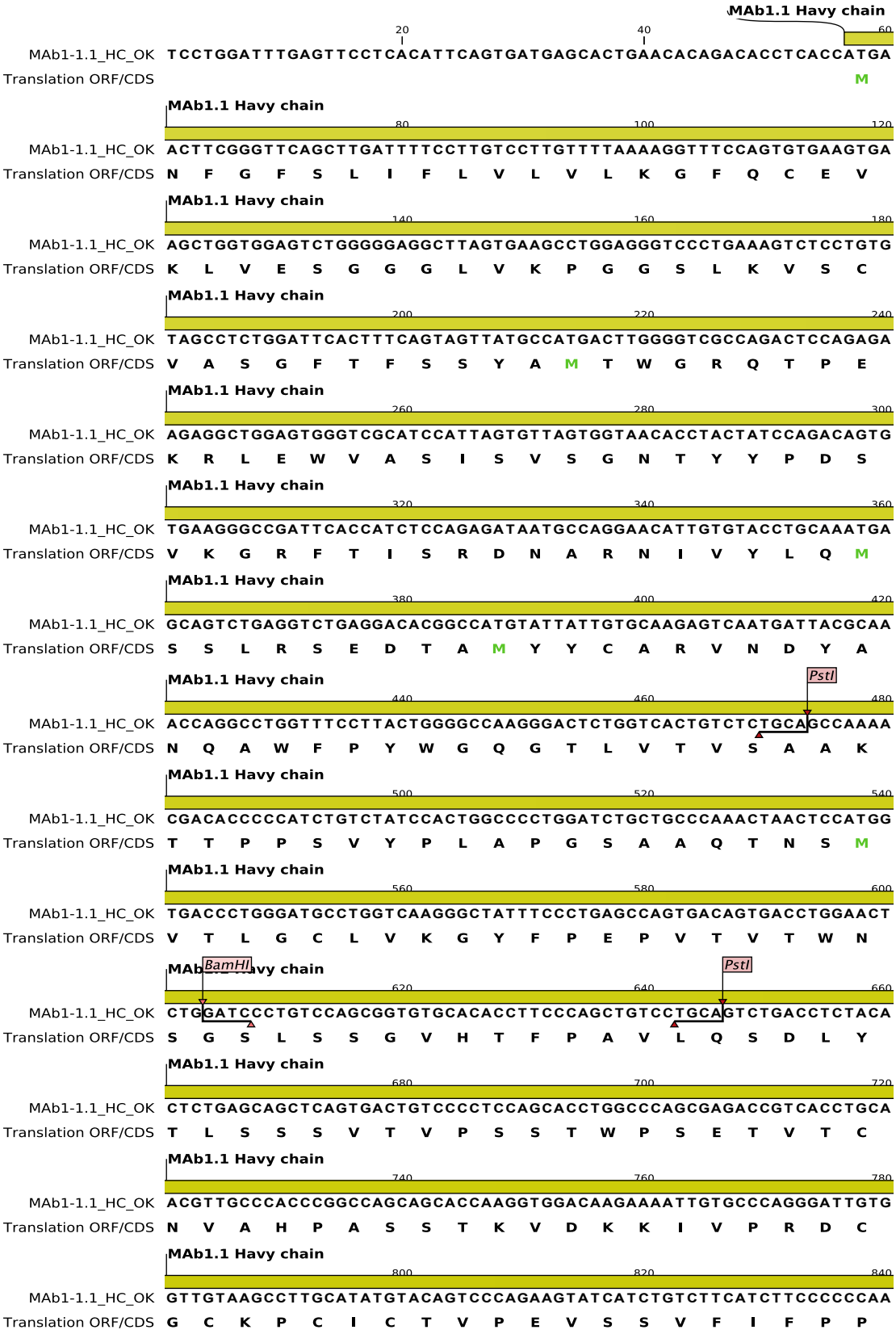
Schroeder, H.W. & Cavacini, L., 2010. Structure and function of immunoglobulins. *The Journal of allergy and clinical immunology*, 125(2 Suppl 2), pp.S41-52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20176268 [Accessed March 1, 2012].

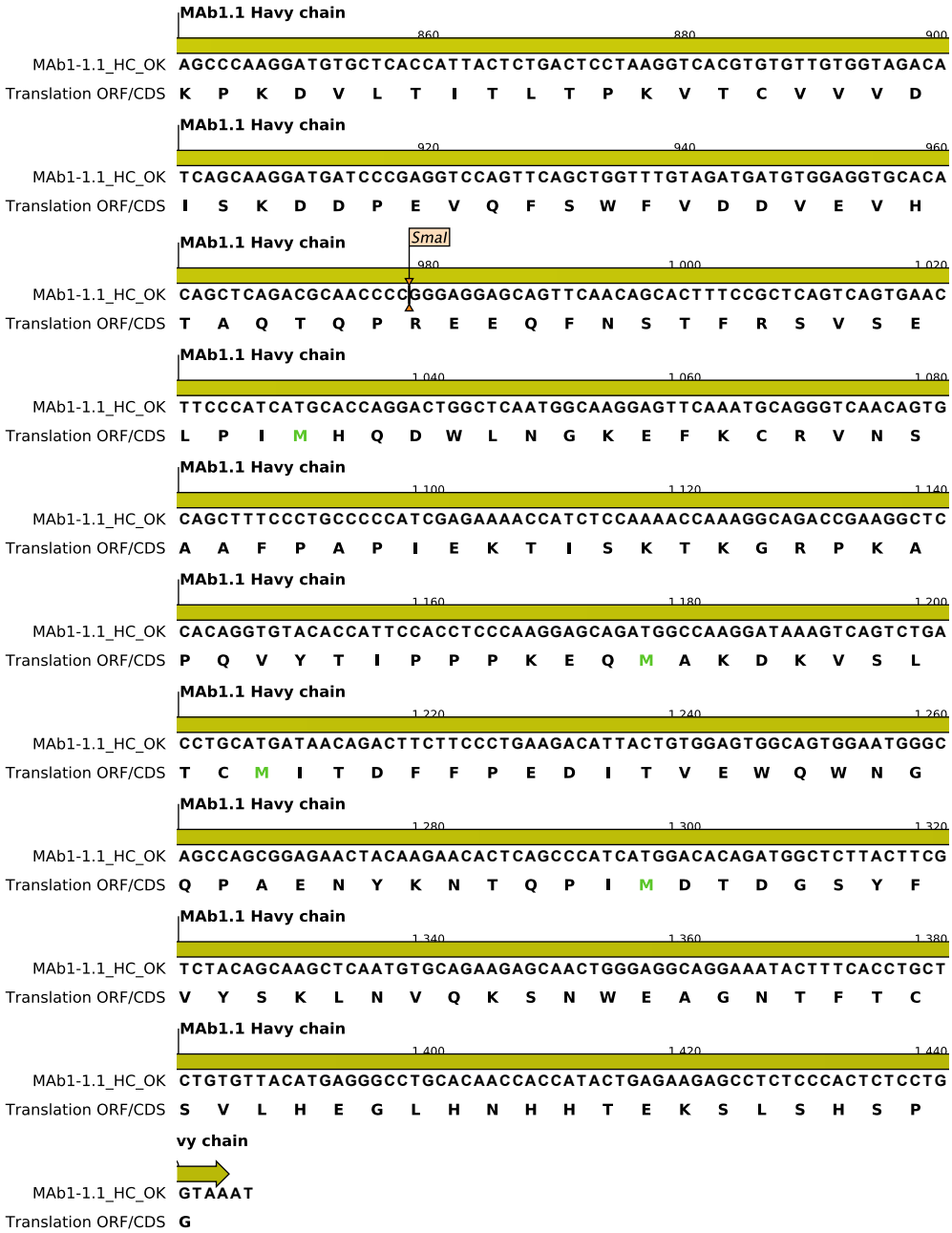
Taipa, R., Pinho, J. & Melo-Pires, M., 2012. Clinico-Pathological Correlations of the Most Common Neurodegenerative Dementias. *Frontiers in Neurology*, 3(May), pp.1-13. Available at: http://www.frontiersin.org/Dementia/10.3389/fneur.2012.00068/abstract [Accessed May 8, 2012].

WHO, 2012. *Dementia a public health priority*, World Health Organisation (WHO).

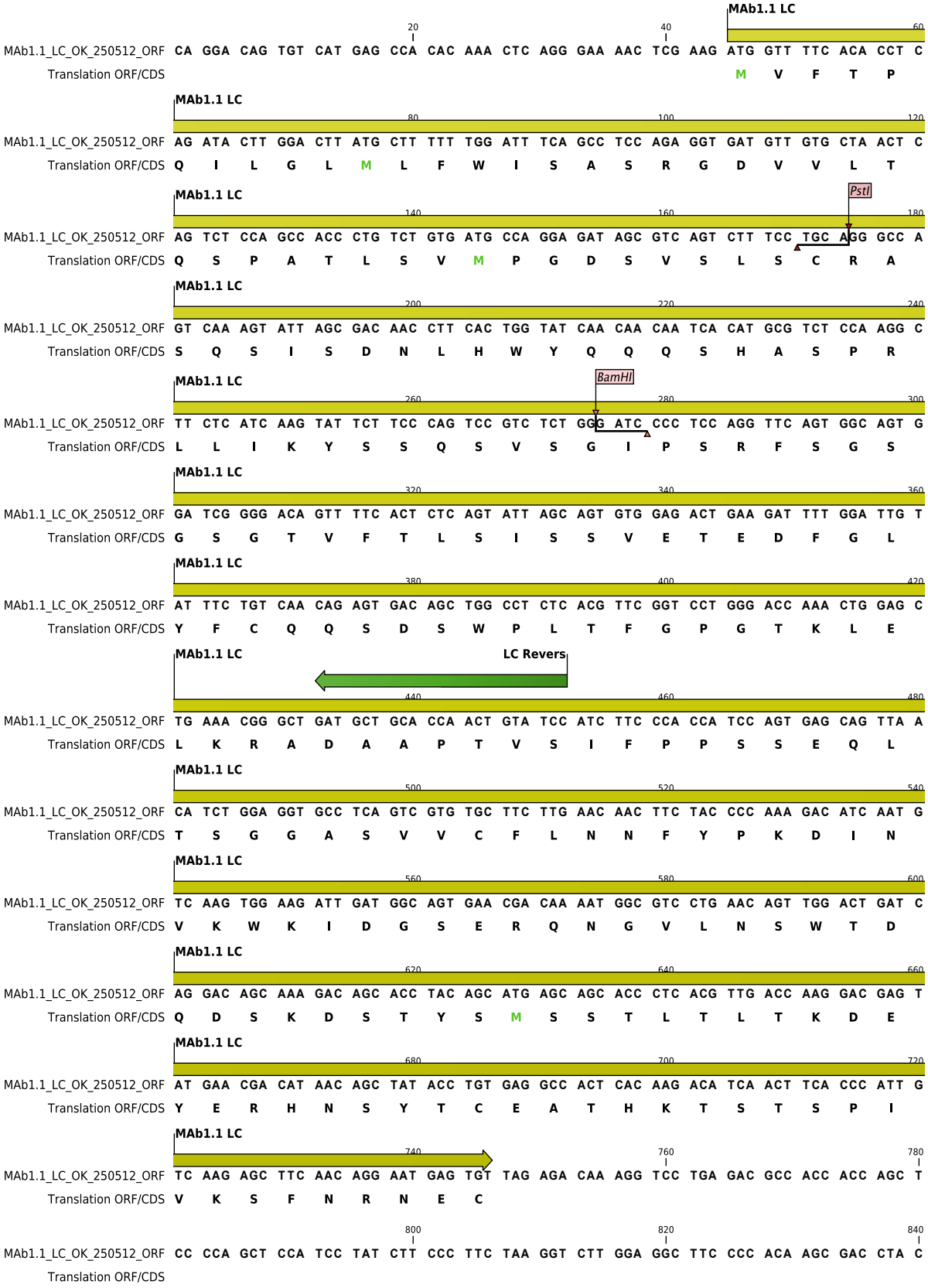
Wang, Z. et al., 2000. Universal PCR amplification of mouse immunoglobulin gene variable regions: the design of degenerate primers and an assessment of the effect of DNA polymerase 3' to 5' exonuclease activity. *Journal of immunological methods*, 233(1-2), pp.167-77. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10648866.

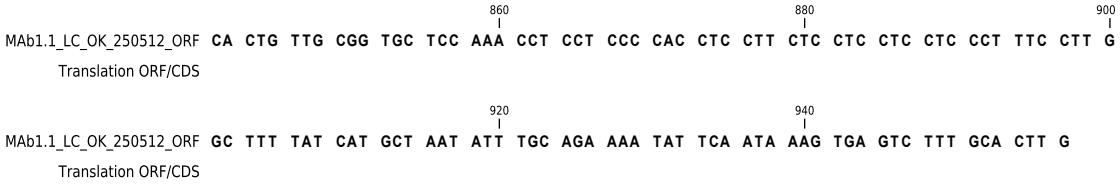
Appendix 1


MAb1.1 Havy chain
MAb1.1_HC_OK TCCTGGATTTGAGTTCCTCACATTCAGTGATGAGCACTGAACACAGACACCTCACCATGA
Translation ORF/CDS M
MAb1.1 Havy chain
MAb1.1_HC_OK ACTTCGGGTTCAGCTTGATTTTCCTTGTCCTTGTTTTAAAAGGTTTCCAGTGTGAAGTGA
Translation ORF/CDS N F G F S L I F L V L V L K G F Q C E V
MAb1.1 Havy chain
MAb1.1_HC_OK AGCTGGTGGAGTCTGGGGGAGGCTTAGTGAAGCCTGGAGGGTCCCTGAAAGTCTCCTGTG
Translation ORF/CDS K L V E S G G G L V K P G G S L K V S C
MAb1.1 Havy chain
MAb1.1_HC_OK TAGCCTCTGGATTCACTTTCAGTAGTTATGCCATGACTTGGGGTCGCCAGACTCCAGAGA
Translation ORF/CDS V A S G F T F S S Y A M T W G R Q T P E
MAb1.1 Havy chain
MAb1.1_HC_OK AGAGGCTGGAGTGGGTCGCATCCATTAGTGTTAGTGGTAACACCTACTATCCAGACAGTG
Translation ORF/CDS K R L E W V A S I S V S G N T Y Y P D S
MAb1.1 Havy chain
MAb1.1_HC_OK TGAAGGGCCGATTCACCATCTCCAGAGATAATGCCAGGAACATTGTGTACCTGCAAATGA
Translation ORF/CDS V K G R F T I S R D N A R N I V Y L Q M
MAb1.1 Havy chain
MAb1.1_HC_OK GCAGTCTGAGGTCTGAGGACACGGCCATGTATTATTGTGCAAGAGTCAATGATTACGCAA
Translation ORF/CDS S S L R S E D T A M Y Y C A R V N D Y A
MAb1.1 Havy chain
PstI
MAb1.1_HC_OK ACCAGGCCTGGTTTCCTTACTGGGGCCAAGGGACTCTGGTCACTGTCTCTGCAGCCAAAA
Translation ORF/CDS N Q A W F P Y W G Q G T L V T V S A A K
MAb1.1 Havy chain
MAb1.1_HC_OK CGACACCCCCATCTGTCTATCCACTGGCCCCTGGATCTGCTGCCCAAACTAACTCCATGG
Translation ORF/CDS T T P P S V Y P L A P G S A A Q T N S M
MAb1.1 Havy chain
MAb1.1_HC_OK TGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACT
Translation ORF/CDS V T L G C L V K G Y F P E P V T V T W N
BamHI
MAb1.1 Havy chain
PstI
MAb1.1_HC_OK CTGGATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCCTGCAGTCTGACCTCTACA
Translation ORF/CDS S G S L S S G V H T F P A V L Q S D L Y
MAb1.1 Havy chain
MAb1.1_HC_OK CTCTGAGCAGCTCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCA
Translation ORF/CDS T L S S S V T V P S S T W P S E T V T C
MAb1.1 Havy chain
MAb1.1_HC_OK ACGTTGCCCACCCGGCCAGCAGCACCAAGGTGGACAAGAAAATTGTGCCCAGGGATTGTG
Translation ORF/CDS N V A H P A S S T K V D K K I V P R D C
MAb1.1 Havy chain
MAb1.1_HC_OK GTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCCCAA
Translation ORF/CDS G C K P C I C T V P E V S S V F I F P P


MAb1.1 Havy chain
860
880
900
MAb1-1.1_HC_OK AGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACA
Translation ORF/CDS K P K D V L T I T L T P K V T C V V V D
MAb1.1 Havy chain
920
940
960
MAb1-1.1_HC_OK TCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACA
Translation ORF/CDS I S K D D P E V Q F S W F V D D V E V H
MAb1.1 Havy chain
SmaI
980
1,000
1,020
MAb1-1.1_HC_OK CAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAAC
Translation ORF/CDS T A Q T Q P R E E Q F N S T F R S V S E
MAb1.1 Havy chain
1,040
1,060
1,080
MAb1-1.1_HC_OK TTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGTTCAAATGCAGGGTCAACAGTG
Translation ORF/CDS L P I M H Q D W L N G K E F K C R V N S
MAb1.1 Havy chain
1,100
1,120
1,140
MAb1-1.1_HC_OK CAGCTTTCCCTGCCCCCATCGAGAAAACCATCTCCAAAACCAAAGGCAGACCGAAGGCTC
Translation ORF/CDS A A F P A P I E K T I S K T K G R P K A
MAb1.1 Havy chain
1,160
1,180
1,200
MAb1-1.1_HC_OK CACAGGTGTACACCATTCCACCTCCCAAGGAGCAGATGGCCAAGGATAAAGTCAGTCTGA
Translation ORF/CDS P Q V Y T I P P P K E Q M A K D K V S L
MAb1.1 Havy chain
1,220
1,240
1,260
MAb1-1.1_HC_OK CCTGCATGATAACAGACTTCTTCCCTGAAGACATTACTGTGGAGTGGCAGTGGAATGGGC
Translation ORF/CDS T C M I T D F F P E D I T V E W Q W N G
MAb1.1 Havy chain
1,280
1,300
1,320
MAb1-1.1_HC_OK AGCCAGCGGAGAACTACAAGAACACTCAGCCCATCATGGACACAGATGGCTCTTACTTCG
Translation ORF/CDS Q P A E N Y K N T Q P I M D T D G S Y F
MAb1.1 Havy chain
1,340
1,360
1,380
MAb1-1.1_HC_OK TCTACAGCAAGCTCAATGTGCAGAAGAGCAACTGGGAGGCAGGAAATACTTTCACCTGCT
Translation ORF/CDS V Y S K L N V Q K S N W E A G N T F T C
MAb1.1 Havy chain
1,400
1,420
1,440
MAb1-1.1_HC_OK CTGTGTTACATGAGGGCCTGCACAACCACCATACTGAGAAGAGCCTCTCCCACTCTCCTG
Translation ORF/CDS S V L H E G L H N H H T E K S L S H S P
vy chain
MAb1-1.1_HC_OK GTAAAT
Translation ORF/CDS G

Appendix 2


MAb1.1 LC
MAb1.1_LC_OK_250512_ORF CA GGA CAG TGT CAT GAG CCA CAC AAA CTC AGG GAA AAC TCG AAG ATG GTT TTC ACA CCT C
Translation ORF/CDS M V F T P
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF AG ATA CTT GGA CTT ATG CTT TTT TGG ATT TCA GCC TCC AGA GGT GAT GTT GTG CTA ACT C
Translation ORF/CDS Q I L G L M L F W I S A S R G D V V L T
MAb1.1 LC
PstI
MAb1.1_LC_OK_250512_ORF AG TCT CCA GCC ACC CTG TCT GTG ATG CCA GGA GAT AGC GTC AGT CTT TCC TGC AGG GCC A
Translation ORF/CDS Q S P A T L S V M P G D S V S L S C R A
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF GT CAA AGT ATT AGC GAC AAC CTT CAC TGG TAT CAA CAA CAA TCA CAT GCG TCT CCA AGG C
Translation ORF/CDS S Q S I S D N L H W Y Q Q Q S H A S P R
MAb1.1 LC
BamHI
MAb1.1_LC_OK_250512_ORF TT CTC ATC AAG TAT TCT TCC CAG TCC GTC TCT GGG ATC CCC TCC AGG TTC AGT GGC AGT G
Translation ORF/CDS L L I K Y S S Q S V S G I P S R F S G S
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF GA TCG GGG ACA GTT TTC ACT CTC AGT ATT AGC AGT GTG GAG ACT GAA GAT TTT GGA TTG T
Translation ORF/CDS G S G T V F T L S I S S V E T E D F G L
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF AT TTC TGT CAA CAG AGT GAC AGC TGG CCT CTC ACG TTC GGT CCT GGG ACC AAA CTG GAG C
Translation ORF/CDS Y F C Q Q S D S W P L T F G P G T K L E
MAb1.1 LC
LC Revers
MAb1.1_LC_OK_250512_ORF TG AAA CGG GCT GAT GCT GCA CCA ACT GTA TCC ATC TTC CCA CCA TCC AGT GAG CAG TTA A
Translation ORF/CDS L K R A D A A P T V S I F P P S S E Q L
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF CA TCT GGA GGT GCC TCA GTC GTG TGC TTC TTG AAC AAC TTC TAC CCC AAA GAC ATC AAT G
Translation ORF/CDS T S G G A S V V C F L N N F Y P K D I N
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF TC AAG TGG AAG ATT GAT GGC AGT GAA CGA CAA AAT GGC GTC CTG AAC AGT TGG ACT GAT C
Translation ORF/CDS V K W K I D G S E R Q N G V L N S W T D
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF AG GAC AGC AAA GAC AGC ACC TAC AGC ATG AGC AGC ACC CTC ACG TTG ACC AAG GAC GAG T
Translation ORF/CDS Q D S K D S T Y S M S S T L T L T K D E
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF AT GAA CGA CAT AAC AGC TAT ACC TGT GAG GCC ACT CAC AAG ACA TCA ACT TCA CCC ATT G
Translation ORF/CDS Y E R H N S Y T C E A T H K T S T S P I
MAb1.1 LC
MAb1.1_LC_OK_250512_ORF TC AAG AGC TTC AAC AGG AAT GAG TGT TAG AGA CAA AGG TCC TGA GAC GCC ACC ACC AGC T
Translation ORF/CDS V K S F N R N E C
MAb1.1_LC_OK_250512_ORF CC CCA GCT CCA TCC TAT CTT CCC TTC TAA GGT CTT GGA GGC TTC CCC ACA AGC GAC CTA C
Translation ORF/CDS


860
880
900
MAb1.1_LC_OK_250512_ORF CA CTG TTG CGG TGC TCC AAA CCT CCT CCC CAC CTC CTT CTC CTC CTC CTC CCT TTC CTT G
Translation ORF/CDS
920
940
MAb1.1_LC_OK_250512_ORF GC TTT TAT CAT GCT AAT ATT TGC AGA AAA TAT TCA ATA AAG TGA GTC TTT GCA CTT G
Translation ORF/CDS