

**Detection of novel serological biomarkers in two  
neurodegenerative disorders: Morbus Alzheimer and Multiple  
Sclerosis**

**Inaugural-Dissertation**

To obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry and Pharmacy  
of Freie Universität Berlin

by

**Yuliya Georgieva**

born in Orenburg, Russian Federation

Oktober 2013

I hereby declare that all work and writing contained within this thesis was conducted by myself, all used references are cited accordingly and any personal assistance has been acknowledged by name.

All work for this thesis was conducted from June 2009 to June 2013 in the working group of Dr. Zoltán Konthur at the Department of Vertebrate Genomics under the supervision of Prof. Dr. Hans Lehrach in the Max Planck Institute for Molecular Genetics.

1. Reviewer: Prof. Dr. Hans Lehrach

Max Planck Institute for Molecular Genetics

Department of Vertebrate Genomics

Innstraße 63-73

14195 Berlin, Germany

2. Reviewer: Prof. Dr. Burghardt Wittig

Freie Universität Berlin

Foundation Institute Molecular Biology and Bioinformatics

Arnimallee 22

14195 Berlin, Germany

Date of defense: 14.02.2014.

## **Acknowledgment**

I would like to thank Prof. Dr. Hans Lehrach, Director at the Max Planck Institute for Molecular Genetics, for giving me the great opportunity to conduct my dissertation at his department in such an interesting research area. Furthermore, I thank Prof. Dr. Burghardt Wittig from Freie Universität Berlin for his kindness to serve as an academic advisor of my thesis and for his invaluable advice on this manuscript.

I would like to express my sincere gratitude to my direct supervisor Dr. Zoltán Konthur for the support of my dissertation, for his scientific advice, as well as for giving me the opportunity to take part at international conferences and for enabling my voluntary work at btS e.V.. I also would like to thank all the members of AG Konthur, current and previous, for their helpful hands and many inspiring discussions. In particular, my two companion Ph.D. students Dr. Stephan Klatt and Dr. Florian Rubelt, as well as Carola Stoschek, for great working atmosphere, support and amity. I especially want to thank my student assistant Sabrina Grasse for her precious assistance during most of the experimental work and her close friendship.

My sincere thanks go to our project cooperation partner, Dr. Steffen Hennig for his collaboration and advice in the field of bioinformatics and next generation sequencing. I would also like to thank Dr. Matthias Kaup for the critical revision of this manuscript and for his scientific advice.

Last but not least, I would like to thank my parents Svetlana and Ivan Georgievi, my brothers Georgi and Dimo Georgievi and all my dear friends for their invaluable support, patience and encouraging in dire straits.

---

## Table of Contents

Acknowledgement .....	3
List of abbreviations .....	9
1. Introduction.....	14
1.1. Diagnostic aspects in neurodegenerative disorders .....	16
1.1.1. Morbus Alzheimer (AD) diagnostics .....	16
1.1.2. Multiple sclerosis (MS) diagnostics.....	22
1.1.3. Serological autoimmunity-based biomarkers.....	25
1.2. High-throughput screening technologies .....	28
1.2.1. Protein macroarray technology.....	29
1.2.2. Filamentous phage display of full-ORF polypeptides .....	31
1.2.3. Next generation sequencing .....	36
2. Objective .....	40
3. Materials.....	42
3.1. Consumables .....	42
3.2. Chemicals, buffers and solutions.....	43
3.3. Growth culture media and additives.....	46
3.4. <i>E. coli</i> strains.....	47
3.5. Helper phages.....	47
3.6. Plasmids and OC source libraries.....	47
3.7. Protein macroarrays.....	48
3.8. Human blood sera.....	48
3.9. Antibodies .....	50
3.10. Magnetic beads and affinity columns.....	51
3.11. DNA and Protein markers .....	51
3.12. Kits .....	52
3.13. Enzymes .....	52
3.14. Oligonucleotides .....	53

---

3.15. Laboratory equipment.....	54
3.16. Software.....	55
4. Methods.....	57
4.1. Molecular biology based methods .....	57
4.1.1. DNA purification .....	57
4.1.1.1. DNA extraction from agarose gel .....	57
4.1.1.2. DNA precipitation with ethanol .....	57
4.1.1.3. Plasmid purification .....	57
4.1.2. DNA digestion with restriction enzymes .....	58
4.1.3. Dephosphorylation of digested plasmids .....	58
4.1.4. DNA separation in agarose gel electrophoresis .....	58
4.1.5. DNA amplification by polymerase chain reaction (PCR) .....	58
4.1.6. DNA ligation .....	59
4.1.7. Gateway LR recombination .....	60
4.1.8. DNA Sanger sequencing .....	61
4.1.9. Preparation of electro-competent <i>E. coli</i> cells .....	61
4.1.10. Transformation of <i>E. coli</i> cells by electroporation .....	61
4.2. Protein biochemistry based methods.....	62
4.2.1. IPTG-induced protein expression in <i>E.coli</i> cells .....	62
4.2.2. Protein extraction under native and denatured conditions.....	62
4.2.2.1. Cytoplasmic protein extraction .....	62
4.2.2.2. Periplasmic protein extraction .....	63
4.2.3. IMAC purification of recombinant His <sub>6</sub> -tagged proteins .....	63
4.2.3.1. Purification with Ni-NTA-Agarose at gravity flow (batch purification) ...	63
4.2.3.2. Purification with Ni-NTA columns on an FPLC system .....	63
4.2.4. Concentrating proteins.....	64
4.2.5. Protein separation in SDS PAGE .....	64
4.2.6. Coomassie staining .....	64
4.2.7. Silver staining .....	64

---

4.2.8. Western Blot .....	65
4.2.9. Protein ELISA .....	65
4.2.10. Determination of total immunoglobulin titers in blood sera .....	66
4.3. Protein macroarray technology .....	66
4.3.1. Protein macroarrays hybridization with human sera .....	66
4.3.2. Data analysis with AIDA Image Analyzer .....	67
4.4. Filamentous phage display based methods.....	67
4.4.1. Preparation of M13K07 helper phage .....	67
4.4.2. Phage titer determination.....	67
4.4.3. Preparation of recombinant M13 phages .....	68
4.4.4. Loading of tosyl-activated magnetic beads with human autoantibodies.....	69
4.4.5. Bio-panning procedures.....	71
4.4.5.1. Semi-automated selection on a magnetic particle processor.....	71
4.4.5.2. Polyclonal phage ELISA.....	74
4.5. Next generation sequencing on Illumina Genome Analyzer.....	75
4.5.1. Preparation of phage-derived full-ORF gene inserts for sequencing.....	75
4.5.2. Processing of sequencing samples and applied NGS protocols .....	76
4.5.3. Bioinformatical raw data processing .....	76
4.6. Statistical analysis of final ELISA results .....	76
5. Results .....	78
5.1. Determination of total immunoglobulin titers in blood sera .....	78
5.2. High-throughput autoantibody screening on protein macroarrays.....	78
5.3. Semi-automated selection of human autoantigens, presented on M13 phages.....	83
5.3.1. Generation of human full-ORF phagemid libraries .....	83
5.3.1.1. Construction of destination pYG-vector series .....	85
5.3.1.2. LR reactions and phagemid library validation .....	89
5.3.1.3. Phagemid libraries evaluation with EGFP .....	93
5.3.2. Autoantigens selection procedures.....	98

---

5.3.2.1. Loading magnetic beads with human autoantibodies .....	98
5.3.2.2. Polyclonal phage ELISA.....	100
5.3.3. Preparation of DNA samples for sequencing on Illumina Genome Analyzer ...	105
5.3.4. Enrichment analysis of NGS results .....	106
5.4. Validation of identified biomarker candidates .....	107
5.4.1. Selection of biomarker candidates for recombinant bacterial expression.....	107
5.4.2. Recombinant bacterial expression, IMAC purification and autoantigen ELISA of selected biomarker candidates .....	108
5.4.2.1. PRDX1 .....	109
5.4.2.2. TANK.....	113
5.4.2.3. TRAF4 .....	114
5.4.2.4. ANXA2 .....	115
5.4.2.5. DTNBP1 .....	118
5.4.2.6. GDI1 .....	121
5.4.2.7. NDRG4.....	122
5.4.2.8. PAX6 .....	125
5.4.2.9. ANKHD1.....	126
5.4.2.10. DEAF1 .....	129
5.4.3. Cross-reactivity tests of HRP-conjugated anti-human antibodies.....	132
6. Discussion .....	135
6.1. Results from protein macroarray screening .....	135
6.2. Results from phage display screening .....	140
6.3. Results from biomarkers validation in ELISA .....	145
6.4. Could I achieve the aims of my thesis? .....	149
7. Summary .....	151
8. Zusammenfassung.....	153
9. References .....	156
10. Supplementary .....	169
10.1. Lists of autoantigens from protein macroarray screenings .....	169

---

10.2. Lists of autoantigens from phage display screenings .....	175
11. Appendix .....	179
11.1. List of figures.....	179
11.2. List of tables.....	181

---

## List of abbreviations

°C	grade Celsius
α-	anti-
AA	amino-acid
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic) acid
AD	Morbus Alzheimer
ADRDA	Alzheimer's Disease and Related Disorders Association
AG	Arbeitsgruppe
ALZAS	Alzheimer-associated protein
Amp	ampicillin
ANKHD1	Ankyrin repeat and KH containing 1
ANXA2	Annexin 2
AP	alkaline phosphatase
APC	antigen-presenting cell
APP	Amyloid-β peptide precursor
APS	ammonium persulfate
Aβ	Amyloid-β
BBB	blood-brain barrier
bp	base pair
BSA	bovine serum albumin
C	control (sample)
Cam	chloramphenicol
CCD	charge-coupled device
cDNA	complementary DNA
CDS	coding DNA sequence
CDT	Clock-Draw-Test
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
cfu	cell forming units
ChIP-Seq	chromatin immunoprecipitation sequencing
CIS	clinically isolated syndrome
CNS	central nervous system
CSF	cerebrospinal fluid
CV	column volume
DARPin	Designed ankyrin repeat protein
ddH <sub>2</sub> O	double-distilled H <sub>2</sub> O

---

ddNTP	dideoxynucleotide triphosphate
DEAF1	Deformed epidermal autoregulatory factor 1
DKFZ	Deutsches Krebsforschungszentrum
DLB	Dementia with Lewy bodies
DNA	deoxyribonucleic acid
DRG	dorsal root ganglion
dsDNA	double-stranded DNA
DTNBP1	Dystrobrevin binding protein 1
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EAE	Experimental autoimmune encephalomyelitis
EBV	Epstein-Barr virus
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EEG	electroencephalography
EGFP	enhanced green fluorescent protein
ELISA	enzyme-linked immunosorbent assay
EU	European Union
f	female
FDG	fluoro-2-deoxy-glucose
FPLC	fast protein liquid chromatography
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
Gb	giga byte
GDI1	GDP dissociation inhibitor 1
GDP	guanosine diphosphate
GER	Germany
Glu	glucose
GTP	guanosine triphosphate
GUS	$\beta$ -glucuronidase
GW	Gateway
H	healthy
HGNC	HUGO Gene Nomenclature Committee
HHV	human herpes virus
HIP	Harvard Institute of Proteomics
His <sub>6</sub> -tag	hexahistidine-tag
HMFM	Hogness-modified freezing medium

---

HRP	horseradish peroxidase
HUGO	Human Genome Organization
IFN $\gamma$	Interferon $\gamma$
Ig	immunoglobulin
IL	Interleukin
IMAC	immobilized metal-ion affinity chromatography
IPA	Ingenuity pathway analysis
IPTG	isopropyl- $\beta$ -D-thiogalactoside
JNK	c-jun N-terminal kinase
Kan	kanamycin
KH	(protein) K homology
LB	lysogeny broth
LZ	leucine zipper
m	male
mAb	monoclonal antibody
MBP	myelin basic protein
MCI	mild cognitive impairment
MES	2-(N-morpholino)ethanesulfonic acid
MGC	Mammalian Gene Collection
MHC	major histocompatibility complex
MMCE	Mini-Mental State Examination
MOG	myelin oligodendrocyte glycoprotein
MPI-MG	Max-Planck-Institute for Molecular Genetics
MRI	magnetic resonance imaging
mRNA	messenger RNA
MS	Multiple Sclerosis
MTP	micro-titer plate
NDRG4	N-myc downstream regulated gene 4 protein
NEB	New England Biolabs
NFT	neurofibrillary tangle
NF $\kappa$ B	nuclear factor "kappa-light-chain-enhancer" of activated B-cells
NGS	next generation sequencing
NIH	National Institute of Health
NINCDS	National Institute of Neurological and Communicative Disorders and Stroke
Ni-NTA	nickel nitrilotriacetic acid

---

NMSS	National Multiple Sclerosis Society
OC	ORFeome collection
OCB	oligoclonal bands
OD	optical density
ON	overnight
ORF	open reading frame
pAb	polyclonal antibody
PAX6	Paired box 6
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PD	Morbus Parkinson
PEG	polyethylene glycol
PET	positron emission tomography
pfu	plaque forming units
PiB	Pittsburgh compound B
PMSF	phenylmethylsulfonylfluoride
POI	polypeptide of interest
PP	polypropylene
PPMS	Primary Progressive Multiple Sclerosis
PRDX1	Peroxiredoxin 1
PRMS	Progressive relapsing Multiple Sclerosis
PSP	progressive supranuclear palsy
PST	polystyrene
p-tau	phosphorylated tau
PVC	polyvinylchloride
PVDF	polyvinylidene difluoride
RBS	ribosome binding site
Rd	round of selection
RNA	ribonucleic acid
rpm	rounds per minute
RRMS	Relapsing-Remitting Multiple Sclerosis
RT	room temperature
s.	see
SDS PAGE	SDS polyacrylamid gel electrophoresis
SDS	sodium dodecyl sulfate
SLE	Systemic Lupus Erythematosus

---

SMRT	single molecule real-time sequencing
SOB	super optimal broth
SOC	super optimal broth with catabolite repression
SP	senile plaques
SPMS	Secondary Progressive Multiple Sclerosis
SRP	signal recognition particle
ssDNA	single-stranded DNA
TANK	TRAF family member-associated NF $\kappa$ B activator
TAT	twin arginine translocation
TBS	tris-buffered saline
TEMED	tetramethylethylenediamine
TNF- $\alpha$	Tumor necrosis factor $\alpha$
TRAF(4)	TNF receptor associated factor (4)
t-tau	total tau
TUB	Tubulin
UBI	Ubiquitin
UK	United Kingdom
USA	United States of America
UTR	untranslated region
UV	ultra violet
VEP	visual evoked potential
w/o	without
WHO	World Health Organization
WTSI	Welcome Trust Sanger Institute

## 1. Introduction

The advances in biosciences and medical research in the last century led to enormous improvement in healthcare and with this to better life conditions in most countries worldwide. As a consequence, life expectancy at birth is constantly rising. Within the EU, for example, these have increased by ten years in average over the last 50 years [1]. According to Eurostat's latest population project scenario (Europop 2010), this trend will probably persist in the next decades and contribute strongly to another demographical tendency: the inevitable population aging [2, 3]. Although these are beyond doubt positive perspectives, the global demographic aging brings also some serious economical and social challenges. Thus, proliferation of many age-related disorders can be expected in near future, which will unavoidably lead to a considerable increase in global healthcare costs.

Neurodegeneration is a general term for neurological disorders with progressive loss of neuronal structures and functions. While advancing age is a major risk factor for many neurodegenerative diseases, dementia is probably the most frequent clinical symptom. It should be taken into consideration that real pathological dementia, as discussed here, is a serious decline in cognitive capacity, which goes beyond "normal" obliviousness in older individuals. Referring again to the global trend of aging population in combination with ever-growing life expectancy, some researchers forecast a real epidemic towards 2050 – 2060. Hence, WHO designates dementia a "public health priority" in its latest report from 2012 [4, 5]. According to this, cases of dementia worldwide will at least double until 2030, increasing from almost 36 Mio presently, up to 66 Mio. For 2060 even more than 115 Mio cases are being predicted, which is more than triple of what we probably have today. Estimations of overall global financial costs for 2010 are in the range of 600 billion US \$, which is 1% of the worldwide gross domestic product and the costs are constantly rising. Thus, dementia has significant economical and social impact on societies worldwide. Recently, governments in UK, France, USA and other countries have declared dementia for a health- and social-care priority and developed accordant strategies [6]. Others are expected to follow soon.

Beside the considerable social impact, dementia has a profound influence on patients' life and environment. The progressive memory loss is often accompanied by further psychiatric problems like depression, anxiety, psychosis and others, even leading to personality changes [7-9]. The outcome is finally a major impairment of patients' life quality, often supported by social exclusion [10]. The situation is quite dramatic also for family members and caretakers, especially during late disease stages, causing heavy emotional and financial

burdens. The awareness of the irreversibility and incurability of this severe condition contributes further to the suffering of all persons involved.

Dementia itself, however, is not a single disease. It is rather a syndrome that can occur in the course of several different neurodegenerative disorders, such as Vascular Dementia, Dementia with Lewy Bodies, Parkinson Dementia, Frontotemporal Dementia and others. Yet, it mostly proves to be Morbus Alzheimer (AD), accounting for estimated 60% – 80% of all diagnosed dementia cases [11]. According to this, AD prevalence is extremely high in older population. About 13% of the over 65-years-olds is affected, which is every eighth person in this age group. Reaching the age of 85 leads to a drastic escalation of a 45% chance to fall ill [5]. Thus, statistical data and economical impacts, mentioned above in regard to dementia, can be directly assigned to AD. Hence, hundreds of scientific groups worldwide, both in academia and industry, work on AD research to date.

Another prominent neurodegenerative disorder in focus of both scientific and pharmaceutical communities, is Multiple Sclerosis (MS). In contrast to AD, MS is in no way typical for the elderly. On the contrary, in the majority of cases it is being diagnosed in adults between the ages of 20 – 40 and females are affected 2 – 3 times more frequently than men [12]. Besides, although also associated with cognitive impairments (30% – 70% of cases), this symptom does not play such a central role, as in AD [13, 14]. Depending on where the inflammation occurs in the nervous system, it can become manifest in almost every neurological condition: visual disturbance, muscle weakness and spasms, difficulties in coordination, problems in speech, bladder dysfunction and many others [15]. Consequently, this chronic disorder can considerably affect patients' quality of life, leading in many cases to severe physical disabilities and paralysis. Disease prevalence is estimated at 2 – 150 per 100,000 people, depending strongly on geographical region and ethnical background [16]. Thus, to date around 2.5 Mio people worldwide suffer from MS. Because of its usually early onset and life-long duration, it causes substantial economical costs and personal struggling to patients, their families and the society as a whole [17].

Hence, although being both neurodegenerative disorders, AD and MS have quite different characteristics. Yet, they also have some important features in common. Despite many crucial findings and numerous publications (over 4,200 and 3,600 Pub Med entries per year for AD and MS in 2012, respectively), for both is valid: explicit cause is still not entirely understood, no cure is available and diagnosis is difficult.

## **1.1. Diagnostic aspects in neurodegenerative disorders**

Accurate and preferably early diagnosis has obvious advantages and, therefore, attracts particular scientific attention. It would allow appropriate medication, lead to better disease prognosis and even decrease treatment and care costs [18]. Yet, precise disease assignment is in both cases a considerable challenge, especially at early disease stages. Currently there are no established early clinical tests available and no distinct biomarkers are known.

### **1.1.1. Morbus Alzheimer (AD) diagnostics**

A definite diagnosis of AD is only possible postmortem by pathological examination of the brain tissue and determination of histological AD hallmarks in an autopsy. On a microscopic level the two key neuronal lesions are amyloid senile plaques (SPs) and neurofibrillary tangles (NFTs). These are concentrated especially in brain regions such as the temporal and the parietal lobes [19, 20]. The overall appearance of the late stage AD brain also demonstrates the devastating outcome of the disease: massive neuron atrophy and diminution of whole brain regions with most severe shrinking in the cortex and hippocampus areas [21]. Consequently, an AD brain can be up to 15% smaller in size and volume, compared to controls of same age at death point. Moreover, normal aging can lead to a moderate brain tissue reduction as well, further increasing the dramatic extent of degeneration in AD [22]. With this, two big problems concerning AD diagnostic have been revealed. First, a brain biopsy for histological analysis from a living person is simply not feasible. And second, even if possible, differential diagnosis would still remain difficult, since many other conditions share same or at least similar features.

To begin with, although being classical AD hallmarks, SPs and NFTs are not limited to it. They can also be found in cognitively normal older individuals [23], as well as in several other neurodegenerative diseases. Thus, SP's have been associated with other disorders like Parkinson's Disease Dementia (PDD) and Dementia with Lewy Bodies (DLB) too [24, 25]. AD brain plaques are formed by aggregated Amyloid- $\beta$  ( $A\beta$ ) peptides. These are generated in the so-called "amyloidogenic" pathway of the Amyloid- $\beta$  precursor protein (APP) after its abnormal consecutive proteolytic cleavage by the  $\beta$ - and then the  $\gamma$ -Secretase [26]. The  $A\beta$ -plaques are believed to be mainly responsible for neuron toxicity and hence the primary cause for neurodegeneration (Amyloid Cascade Hypothesis) [27, 28]. Although being probably one of the most deeply studied proteins in molecular biology, the normal physiological functions of APP are still not fully understood. It is ubiquitously expressed in the CNS and is believed to be involved in several brain key pathways, such as neuroprotection,

synapse formation, adhesion and nutrition of neurons [29-31]. Despite huge amounts of literature, it is also still unclear, which molecular mechanisms lead to the fatal amyloidogenic APP processing in particular and how exactly the A $\beta$ -peptide acts as a neurotoxic agent on biochemical level. Consequently, some researches challenge the A $\beta$  hypothesis, since no precise evidence exists whether the SPs are causal to AD or just a symptom of the “real” pathological process, which is still unrevealed [32-34]. However, the correlation between SPs and dementia, regardless of causality direction, is generally accepted. Yet again, there are some exceptional cases known, which must be considered. For example, Berlau *et al.* published in 2007 an intriguing case study, where an elderly woman did not suffer from any significant memory loss or cognitive decline until her death with 92 years of age. Nevertheless, the brain autopsy showed SPs in an advanced stadium, typical for AD [35]. Apparently, further unknown factors exist in the complex APP pathology, like the special APOE-Y2 isoform, proposed by the authors in this case. These may play a compensatory and/or a protective role, thus being potentially of therapeutic interest.

The situation with the other primary AD hallmark, the NFTs, is just as unclear and controversial, as it is with SPs. NFTs are intracellular aggregates, mainly consisting of the hyperphosphorylated tau protein. Besides AD, at least 10 other neurodegenerative disorders belong to the so-called tauopathies, such as Pick’s Disease, Progressive Supranuclear Palsy (PSP), Frontotemporal Dementia and others [36]. Tau was the first polypeptide, reported to be a microtubule-associated protein and essential for microtubule assembly and stability in neurons [37]. Further studies led to the assumption that its central function would be axonal stabilization and development [38]. According to the most broadly accepted hypothesis, the abnormal phosphorylation leads to tau self-assembly and formation of highly insoluble NFTs. In this state, tau’s association with the microtubule is inhibited, consequently leading to axonal instability and potentially to neurotoxicity and neurodegeneration [39]. Although the precise causality between NFT-presence in neurons and dementia, respectively AD, remains unproven, indications about their correlation continue accumulating. Thus, a direct connection between tau and A $\beta$ -neurotoxicity could be shown by Rapoport *et al.* in 2002 [40]. More recent studies also suggest possible synergistic effects between tau and A $\beta$ , which could be causal to neurodegeneration [41-43]. Furthermore, some researchers discuss a novel hypothesis and propose that NFPs could be only a side effect in the course of AD pathology. Instead, soluble rather than aggregated tau-forms could be the real disease agents [44].

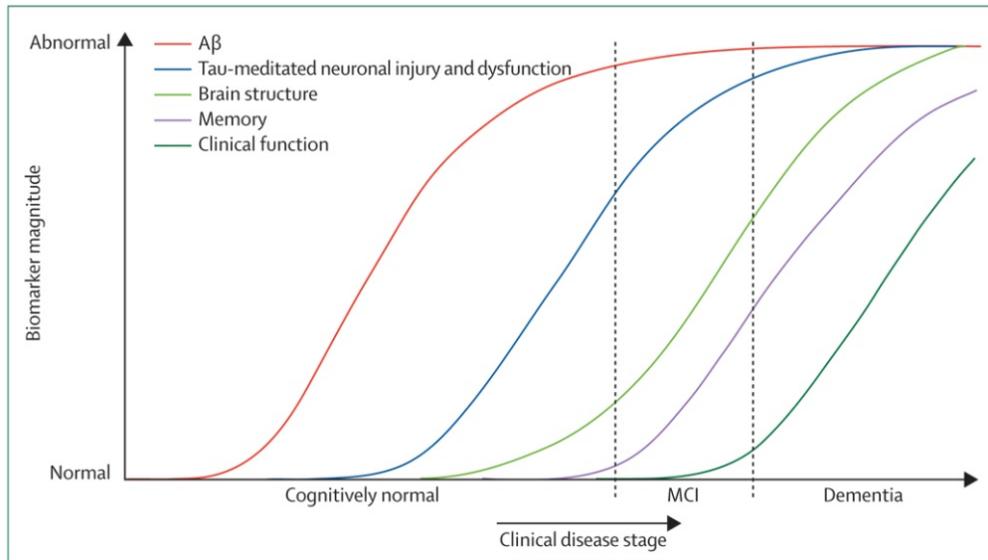
So, after all: How is AD diagnosed, when even its most prominent hallmarks are not limited to it and so many questions about its etiology and pathogenesis are still unclear? The pathophysiological difference between a normally aged and an AD brain seems to be rather

the quantity than the appearance of SPs and NFTs. Thus, the higher density and the bulk of plaques is one criterion. Another one is the relatively specific topography of the plaques, which are mostly concentrated in certain brain regions, as described above. In comparison, the plaques found in non-diseased brains of elderly individuals appear more randomly scattered [45, 46]. Levels of A $\beta$  and tau can be monitored in living individuals. For this, two diagnostic tools are currently available: protein evaluation in the cerebrospinal fluid (CSF) and brain imaging procedures.

The three biomarkers that are usually measured in the CSF are A $\beta_{1-42}$ , total tau (t-tau) and tau phosphorylated at the threonine in position 181 (p-tau<sub>181</sub>). Typical for AD-patients are decreased A $\beta_{1-42}$ -levels and increased levels of tau-proteins [47-49]. Albeit very useful as additional diagnostic parameters, the three CSF biomarkers are not considered to be able to predict AD or any other neurological disorder on their own. On the one hand, protein levels seem to change, depending on disease stage. On the other, currently no validated reference values exist that can precisely assign pathological conditions. Also the large inter-laboratory coefficient of variation of up to 35% in protein level results is a considerable problem [50, 51]. These variations are due to differences in sample collection, storage and transport as well as different assay and laboratory standards. Currently, several projects work on developing standard protocols and operating procedures to overcome this drawback [52, 53]. Finally, taking a CSF sample via a lumbar puncture is not a simple medical intervention. It can only be performed by a specialist in a clinical environment. Taking into account the large amount of dementia cases, it is unlikely that CSF analysis will become a routine diagnostic practice in near future. Yet, CSF biomarkers harbor a big advantage, since they could potentially be able to predict pathological changes in a very early presymptomatic stage. Also the very important, but difficult to characterize conversion from the mild cognitive impairment (MCI) state to full-blown AD may become easier to monitor. It is well known today that AD lesions accumulate slowly over time, first aggregates possibly occurring 20 – 25 years before any signs of cognitive impairment begin to manifest [34, 54]. Thereby, A $\beta$ -plaques seem to form already in a very early stage, leading to a detectable decrease of A $\beta$  in the CSF. Measurable tau alterations are rather associated with later AD stadia, as tau-species are believed to be released from damaged neurons, thus reflecting their stage of degeneration [47, 55, 56]. Good sensitivity and specificity values were achieved when using at least two CSF biomarkers in combination, e.g. t-tau/A $\beta_{1-42}$  [57]. For the discrimination between AD patients and non-diseased of same age sensitivity of 85.7% and specificity of 84.6% could be shown [55]. In a study to distinguish between late stage AD and individuals with MCI even better performance was reached: sensitivity of 95% and specificity of 85% [58]. To sum up, the “CSF AD signature” remains an interesting and potential diagnostic tool. However, it needs

further improvement on specification and validation before it can be applied on a broad scale in practice.

Rapid technological development in the field of medical imaging in the last decades has massively extended our options for diagnosis of disorders in the CNS. Since biopsying in this particular area is fairly critical, powerful *in vivo* imaging techniques, such as positron emission tomography (PET) and magnetic resonance imaging (MRI), are helpful non-invasive alternatives in visualizing histological and morphological changes in brain tissues. Among these, PET is currently the most popular technique for that purpose. It is based on the detection of radiolabeled compounds, called tracers, which accumulate in certain parts of the brain, depending on the targeted molecule or tissue structure [59]. Two kinds of tracers are most widely used, addressing two different targets and with this also two different AD stages. The PiB tracer contains the so-called Pittsburgh compound B, labeled with an  $^{11}\text{C}$ -isotop. It is the most common amyloid-tracer and applied for detection of abnormal A $\beta$  changes in an early presymptomatic AD stage. The information output gained through this analysis overlaps largely with the results from CSF A $\beta$ -level measurements. Therefore, the two procedures are usually not applied in parallel [60]. PET-PiB performs generally very well, as almost all clinically diagnosed AD cases also show distinct pathological A $\beta$  alterations. Yet, it has a grave disadvantage as a sole biomarker, since around 30% of cognitively healthy individuals would deliver positive A $\beta$ -PET-images as well [61]. The other frequently used tracer is  $^{18}\text{F}$ -FDG (Fluoro-2-deoxy-glucose). Unlike PiB, it is rather applied to monitor decreased glucose uptake in neuron cells, which would reflect synaptic activity and level of neurodegeneration. Hence, it is more useful in later dementia, respectively AD stages [62]. Finally, magnetic resonance imaging (MRI) is usually the method of choice to visualize structural brain changes, e.g. atrophied areas. However, these are first detectable only in a relatively late stage of neuron damage and not sufficiently AD-specific. Thus, the three most widespread imaging procedures are rather complementary to each other and to other biomarkers and not likely to become applicable as sole diagnostic criteria in future [63, 64]. According to many researchers, the great potential of both CSF and imaging biomarkers is their ability to address and reflect different stages in the pathogenesis of AD. Hypothetically, each biomarker, or rather a combination of these, would be able not only to confirm the presence of pathological changes in the CNS, but also the actual stage of the condition, as schematically depicted in Fig. 1 below. Such more accurate diagnosis would consequently allow more appropriate medication in future [65].



**Fig. 1 Correlation between the temporal course of AD development and changing biomarker values.** X-axis shows clinical disease stages from the presymptomatic period (Cognitively normal), through the transition period with signs of mild cognitive impairment (MCI), to full-blown dementia. Y-axis indicates chronological alterations in each biomarker according to AD stadium. A $\beta$ -biomarkers are measurements in CSF and PET-PiB (red). Neurodegeneration is detectable by tau-measurements in CSF and PET-FDG (blue). Structural changes in brain can be visualized by MRI (light green). Memory problems (purple) and final cognitive decline (dark green) are assessed by psychiatric tests. Source: Jack *et al.* [66].

Since the five AD biomarkers, described so far, are already known today, the question is: Are they included in routine clinical diagnosis of dementia causing neurological disorders? The answer is, they are not, or only in isolated cases at the utmost. Their major application is rather in the scientific environment, e.g. for characterization and selection of appropriate subjects for clinical trials [65]. One major problem is certainly the lack of standardized procedure protocols and uniform reference values, as discussed above. Additionally, since no real therapeutic options exist anyway and both CSF analysis and medical imaging are invasive and relatively expensive procedures, there is obviously not enough motive force for their clinical application.

Recently the so-called “NINCDS-ADRDA Alzheimer’s Criteria”, which are the official guidelines for AD diagnostic, were thoroughly revised. This was greatly needed, as they have not been changed since their introduction in 1984 by NINCDS (National Institute of Neurological and Communicative Disorders and Stroke) and ADRDA (Alzheimer’s Disease and Related Disorders Association) [67]. While the original comprised solely neuropsychiatric tests to ascertain cognitive impairment, the updated version, published in April 2011, now also includes CSF analysis and neuroimaging as a recommendation to confirm the psychiatric diagnosis [68]. Thus, the general need for novel diagnostic parameters based on

histopathological and biochemical data is clearly distinguishable. Yet, because of the major restrictions of the currently available biomarkers, clinical AD diagnosis remains depending on the established psychiatric criteria. These include mainly neuropsychological tests to assign the level of memory loss and cognitive decline, observation by family members (history by proxy) and exclusion of other neurological pathologies, as far as possible. Hence, they are concentrating basically on dementia determination. The most widely used tests are the so-called Mini-Mental State Examination (MMSE), its extended version CERAD (Consortium to Establish a Registry for Alzheimer's Disease) and the Clock-Draw-Test (CDT). All three perform with a high accuracy, e.g. sensitivity of up to 90% and specificity of 80% [69]. Moreover, it could be shown that MMSE can be helpful to discriminate between AD and other dementia types like Dementia with Lewy bodies for example [70, 71]. However, neuropsychological tests have also a considerable disadvantage, since they are first applicable when clear signs of cognitive impairment begin to manifest, which usually correlates with an advanced disease stadium.

To sum up, diverse diagnostic possibilities are available and applicable to date. Yet, none of them meets all the requirements for an exclusive AD biomarker alone. Best performances are reached when using at least two different parameters in combination and this will probably remain so in future [72]. Table 1 below compares the characteristics of the AD diagnostic tools discussed so far according to the most important features, which a hypothetical "ideal" biomarker should possess.

**Table 1. Comparison of the most widely used AD diagnostic tools** (modified from Borroni *et al.* [72]).

	AD specific	presymptomatic diagnosis	autopsy proven	non-invasive	standardized	simple to perform	specificity > 70% sensitivity > 85%	low costs
<b>Biopsy</b>	yes	unknown	yes	no	no	no	yes	unknown
<b>CSF A<math>\beta</math></b>	no	yes	yes	no	no	no	yes	no
<b>CSF tau</b>	no	no	yes	no	no	no	yes	no
<b>PET-PiB</b>	no	yes	yes	yes	yes	no	no	no
<b>PET-FDL</b>	no	no	yes	yes	yes	no	no	no
<b>MRI</b>	no	no	yes	yes	yes	yes	no	no
<b>Psychological tests</b>	no	no	yes	yes	yes	yes	yes	yes

Table 1 reveals the weak points of the available AD biomarkers, which have to be especially taken into consideration and improved. Thus, the biggest problem is obviously the absence of a parameter that would allow a real differential AD diagnosis. Discovery of an outstanding hallmark, reflecting a strictly specific AD feature, would be certainly a great breakthrough in this research field. Also applicability, including rapid and simple performance, low strains for patients and preferably low costs, is a challenge. For this reason many research groups have been focusing their biomarker discovery efforts on fluids and body regions apart from the

CNS, such as blood, urine, skin and others [72]. Last but not least, early presymptomatic diagnosis is meanwhile of high priority as well. In conclusion, the keen search for more accurate and better applicable biomarkers will undoubtedly proceed in the next years and will hopefully bring novel insights into the complex AD pathophysiology.

### **1.1.2. Multiple Sclerosis (MS) diagnostics**

Diagnosis of MS is, similar to AD, still a difficult issue. This is certainly again due to inaccessibility of the involved tissues and a relatively large overlap of symptoms with many other neurological disorders. However, the situation here is probably even more complicated due to the heterogeneity in regard of disease subtypes, affected physiological areas and symptoms. Thus, four different clinical courses have been ascribed to MS in 1996 by the National Multiple Sclerosis Society (NMSS) and are valid standards since then [73]. These are:

- Relapsing-remitting MS (RRMS)
- Secondary progressive MS (SPMS)
- Primary progressive MS (PPMS)
- Progressive relapsing MS (PRMS)

The RRMS subtype is the most common, accounting for 85% – 90% of all MS cases at onset. It is characterized by distinct disease episodes, in which neurological symptoms occur, followed by recovery periods without disease progression. 65% of the RRMS patients, however, go later over to SPMS. This is the state, which usually leads to the greatest disabilities [15, 74, 75]. 10% of all patients develop PPMS already at onset, having no acute attacks but undergoing a steady decline of neurological functions [76]. Finally, PRMS is characterized by progressive disease development from the beginning, additionally accompanied by intense relapse phases. The differentiation of the four courses is crucial for diagnostic and prognostic purposes, as well as for adequate treatment and medication.

Another difficulty with diagnosing MS is the highly variable presentation of symptoms, depending on the current area of axon demyelization in brain and/or spinal cord. First disease peaks are often being neglected, because of their usually quick decay and full functional recovery. Moreover, each episode can occur in another CNS region and can differ completely in its physiological manifestation from the previous one. Thus, it can take years until first medical examinations would begin. These again can last for months, as no standard laboratory tests exist and a vast number of other neurological causes have to be excluded. Finally, the first neurological attack, called “clinically isolated syndrome” (CIS), earmarks the beginning of MS in only 60% – 80% of cases, while the rest do not develop the disease.

Although numerous risk factors have been described so far, forecasting the outcome of a CIS remains fairly difficult [77].

There are currently few guidelines for MS diagnosis existing, which are being applied in practice. Most popular is the “McDonald Criteria”, introduced in 2001 by NMSS [78] and last revised in 2010 [79]. Yet, no full consensus on differential MS diagnosis has been reached so far. New amendments and improvements are consistently being published [80, 81]. However, in general, diagnostic procedures rely mainly on three clinical tools, which are currently available: electroencephalography (EEG), MRI and analysis of CSF proteins.

EEG measurements have been used for over 30 years to detect irregularities in the electrical brain activity in many neurological disorders. In case of MS, especially measurement of so-called visual evoked potentials (VEP) proved valuable for diagnostic purposes [82]. In contrast to classical EEG, which detects mainly spontaneous potentials, VEP appear first as a response to a distinct stimulus, for example light. Visual dysfunctions, such as optic neuritis, are one of the most common clinical MS manifestations. Moreover, they usually appear very early in the disease course [83]. Distinctive for MS is a delayed VEP latency after stimulation, which can be detected in 50% – 80% of cases [84]. VEP can reveal abnormalities even if the provoking lesions are clinically silent at measurement point. Thus, it is a well-established diagnostic tool for MS, but not specific for it. Also other disorders, such as neurosarcoidosis, SLE, hereditary ataxias, brain compression and even Vitamin B<sub>12</sub> deficiency, can lead to positive VEP measurements [85]. Currently, further diagnostic methods dealing with MS-related visual disturbances are under investigation. Among those, Optical Coherence Tomography and Low-contrast Letter Acuity seem to be the most promising tests [86, 87].

Like in AD, MR imaging is used to visualize dissemination of MS lesions in brain and spinal cord. It is especially helpful for the prognosis whether a CIS has the potential to convert to a clinically definite MS. Thus, number of lesions and specific brain locations were correlated to the risk of conversion [88, 89]. Sensitivity and specificity of MRI as a diagnostic tool for MS were determined as 72% and 87%, respectively [90]. Nevertheless, overlap of indications with other neurological disorders that mimic MS are again the greatest challenge. For example, while relatively effective in differentiating MS from SLE and Sjögren’s syndrome, attempts to distinguish MS from different forms of CNS vasculitis using MRI have been unsuccessful [91].

Major MS biomarker in the CSF is the presence of IgG oligoclonal bands (OCB) and their accordant absence in blood serum. This is usually a clear sign of an ongoing inflammatory reaction in the CNS, which is again not typical for MS only. OCBs can be found in 60% –

98% of MS patients, especially during acute relapse phases. Yet, this ratio is strongly variable in different studies, depending on the applied detection method [77, 92, 93]. Thus, OCB can help to support MS diagnosis, but cannot necessarily confirm MS. Since CSF extraction via lumbar puncture is an additional difficulty, many patients are not even being examined for OCB presence, regardless of their disease stage [94]. Yet, the majority of authors still strongly recommend this procedure to support findings from other tests and eventually to narrow down the ongoing disease subtype and prognosis [95, 96].

Beside OCB, also numerous other CSF biomarkers have been discussed in the literature, which could have the potential to become clinically applicable. Many of them reflect the acute immune response in the relapsing phase. These are, for example, the presence of mature B cells [97], elevated free  $\kappa$  light chains [98], myelin-specific antibody response, as well as antibodies against other lipid antigens [99], increased cytokine and chemokine levels, such as IL-12p40 [100], IL-6 [101], IL-17 [102], TNF- $\alpha$  [103], CXCL13 [104], CCL5 [105]. Other protein biomarkers are associated with the level of neuronal tissue damage and degeneration. Representatives from this group are: increased levels of nitric oxide metabolites [106], chitinase 3-like 1 protein [107] and tau protein [108]. Also neurofilaments, specific structural proteins in neurons, are being released after injury. Thus, increased levels of neurofilament light chains were correlated to relapsing MS phases and functional disease progression [109]. Elevated amounts of neurofilament heavy chains, on the other hand, were associated with a high probability for conversion from CIS to RRMS and a potentially more severe disability [110]. The structural protein of the myelin sheath, the myelin basic protein (MBP), was a research object in several MS studies. Some of them could identify specific MBP antibodies, which could play a significant role in pathogenesis and help predict disease outcome [111]. Much popularity has gained the so-called “molecular mimicry hypothesis”, according to which a peptide of the MBP could equally cross react and subsequently activate T-cells, which would normally recognize a peptide from the HHV-6 (Human herpes virus 6) capsid [112]. Finally, increased MBP levels in the CSF were successfully linked to the level of demyelization [113].

The current state of MS biomarker research was recently reviewed in detail by Graber and Dhib-Jalbut [114]. Unfortunately, none of the markers seem to be able to translate into applicable diagnostic tool in near future, mainly because of their insufficient specificity for MS. Further considerable hurdles are: lack of standardized assays, studies with limited patient numbers and the lack of verification by independent laboratories [77]. Thus, as in case of AD, the search for better MS biomarkers is very likely to persist further.

### 1.1.3. Serological autoimmunity-based biomarkers

Until some 20 years ago, the CNS was believed to be an immunologically “privileged” body compartment, inaccessible for cells and mediators of the immune system. However, this assumption has been entirely refuted by numerous recent findings, showing that although access is in fact restricted, it is not fully intercepted. Thus, brain and CSF are not totally isolated from peripheral circulation, especially not during inflammatory processes [115, 116].

The blood-brain barrier (BBB) is the main gatekeeper between CNS and circulating blood. Its highly dense layer of endothelial cells, equipped with tight junctions that do not appear in any other blood vessel tissues, is the physical limitation that keeps potential hazards away from the fragile, vitally important organs of the CNS [85]. While permeable for water, liposoluble molecules and essential nutrients, such as glucose and vitamins (via special carriers), it is fairly impassable for the majority of polar molecules, e.g. proteins. Therefore, effectors of the immune system like immunoglobulins (Ig) and complements are normally only very scarcely present in the CNS. Also, there is usually no passage through the BBB wall for bacterial and immune cells [117].

However, situation changes, when inflammation begins. Pathological conditions, such as microbial infections, neoplasia and neurodegeneration lead to increased BBB permeability, induced by pro-inflammatory signals. Activated T-cells are the first to pass through an even still intact BBB, probably by secreting specific enzymes to degrade endothelial cell membranes [118, 119]. Shortly after, accessory inflammatory cells, macrophages and further components of the immune system flood the CNS through the “leaky” BBB, which can even become completely broken in severe cases [120]. Earlier it was also believed that the CNS completely lacks antigen-presenting cells (APC), necessary for T-cell interactions and hence regulation of the inflammatory process. However, this argument in favor of the “isolation hypothesis” was proven to be only partially true. Indeed, no professional APCs circulate in the CNS, but there are resident cells, able to conduct this function. These facultative APCs are mainly microglia, expressing Class II MHC molecules [121]. But also astrocytes and certain macrophage-like cells were shown to be inducible to present antigens [122]. In conclusion, communication certainly exists between CNS and the immune system, which is normally muted in a healthy organism, but can be rapidly activated in case of inflammation.

Increased BBB permeability during a real infectious process in the CNS is beyond doubt essential for survival. Yet, if certain control mechanisms in the highly complex immune response cascade fail, undesirable intracerebral immune reactions can occur. Classical examples for such neurological autoimmune disorders are MS and its animal model Experimental Autoimmune Encephalomyelitis (EAE). At this point, following considerations

should be kept in mind: EAE is de facto a T-cell mediated autoimmune disease, initiated by myelin antigens such as MBP and MOG. In case of MS neither MS-specific autoantibodies, nor distinct disease causing autoantigens have been revealed so far. Both MBP and MOG have been under thorough investigation, but sufficient evidence for their initiative immunoreactive role could not be found. Thus, it still remains unproven, if autoimmunity is the factual basis for MS pathogenesis [123]. Nevertheless, its central role is undoubted. Major indications for this assumption are the lack of obvious infectious agents, the positive therapeutic effect of immunosuppressants and the strong analogies to the EAE disease phenotype. However, for the aims of the present study not the causality of MS pathogenesis, but rather its distinct autoimmune aspects were of importance. Thus, MS was chosen as a representative of a neurodegenerative disorder with a clear linkage to autoimmunity in contrast to AD, where it does not seem to play such a pivotal role. Yet, immune response in general is definitely part of the AD pathophysiology as well, since massive neuron degradation is its major morphological hallmark. Intense neuroinflammation with its highly complex course is a key characteristic in AD, which we are only beginning to understand [124, 125]. Furthermore, both neurodegenerative disorders, AD and MS, are characterized by a massive breakdown of the BBB [126-129]. Accordingly, not only components of the immune system gain access to the CNS, but most likely, also a converse seepage of CNS compounds into peripheral circulation occurs. Thus, it is by all means possible that specific disease-relevant polypeptides, resulting from particular inflammation and degradation processes, would accumulate in blood serum of AD and MS patients. And this circumstance could potentially be an essential differential between diseased and healthy individuals, which in turn can be useful for diagnostic purposes. Besides, blood serum assays are excellent diagnostic tools: non-harming, convenient in performance and usually low priced. Hence, many efforts have been made to discover and establish specific serological biomarkers for both AD and MS.

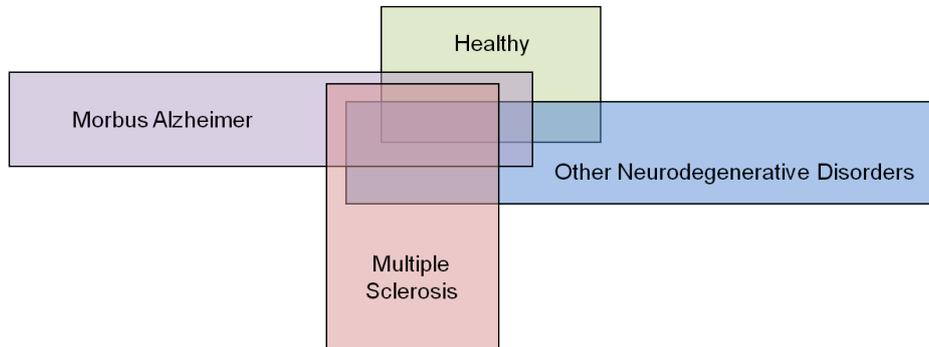
In case of AD, many researchers are again primarily concentrating on the pathological hallmark Amyloid  $\beta$ . In fact,  $A\beta_{1-42}$  could be found in plasma of AD patients, but no agreement exists about its correlation to disease stage and/or dementia severity. Findings remain fairly controversial, since some studies proposed elevated [130] and others reduced levels of serum  $A\beta_{1-42}$  [131] to be indicative. Moreover, significantly high inter- but also intra-person variability was observed, making the establishment of uniform diagnostic criteria difficult [132]. Proteomic analyses of AD blood revealed altered levels of numerous other molecules, such as inflammatory agents [133, 134],  $\alpha$ -1-antitrypsin and apolipoprotein J [135], diverse signaling proteins [136, 137] and many others. Much attention raised the identification of an unknown  $A\beta$  protein in 2002, which occurred to be expressed in elderly individuals, diagnosed with possible AD: The Alzheimer-associated protein (ALZAS) [138]. Interestingly,

not the increased concentration of ALZAS itself, but rather, increased levels of  $\alpha$ -ALZAS IgG antibodies were determined in AD patients in comparison to healthy controls of same age. This finding led to the conclusion that autoimmune reactions are probably also part of the inflammatory process in AD [139]. Also other research groups focus on screening AD patients' sera for disease-specific autoantibodies. For example, Nagele *et al.* were able to identify 10 different autoantigens with specificity and sensitivity for AD of over 90% in a protein microarrays screening [140]. Recently, most well-known AD autoantigens were reviewed by Colasanti *et al.* [141]. These include quite diverse molecules and functional entities, once again demonstrating the complex AD pathophysiology: Amyloid  $\beta$ , lipoproteins, neurotransmitters, microglia, redox-reactive species, aldolase, ATP synthase and numerous others have been discussed.

Attempts to identify serological biomarkers in MS yielded a large number of potential candidates with altered concentration levels compared to healthy individuals. Most of them reflect the acute inflammatory situation and neurodegeneration. Among others, immune mediators such as cytokines and chemokines, as well as apoptotic molecules like TNF- $\alpha$ , could be determined. Findings in this research field were recently reviewed by Hagman *et al.* [142]. Also further blood biomarker candidates with potential diagnostic applicability have been recently described, such as thrombomodulin [143] and the translocator protein 18 kDa TSPO [144]. Further, variable serum autoantibody signatures against common MS antigens in the four different disease subtypes have been described by Quintana *et al.* [145].

To sum up, several promising protein biomarkers have already been isolated from blood sera of both AD and MS patients, which can prove useful for diagnostic purposes in future. But so far, none of them has been sufficiently validated in regard of specificity and sensitivity for the particular disease. Thus, further and larger studies have to be performed, before they can be applied in real diagnostic assays. Currently, the perception seems to become generally accepted that in complex multilateral disorders like AD and MS not a single top biomarker, but rather a combination of many would be a more efficient diagnostic tool. Such multiparametric approach lay at the basis of the present study as well. However, my working hypothesis is based on the assumption that blood sera comprise not only disease-specific protein profiles, but disease-specific autoantibody profiles as well. Major arguments are the massive neuroinflammation and tissue degeneration, accompanied by the breakdown of the blood-brain barrier, allowing CNS polypeptides to enter the peripheral circulation and hence to initiate specific autoimmune reactions. However, also blood sera of non-diseased individuals show high autoimmune reactivity, as very recently demonstrated by Nagele *et al.* [146]. Thus, natural autoimmunity is obviously not implicitly a pathological indication. Its physiological role in healthy organisms is not yet understood, but should be taken into

consideration when screening for disease-specific auto-antigenicity profiles, as illustrated in Fig. 2 below.



**Fig. 2 Hypothetical intersection of auto-antigenicity profiles between AD, MS, healthy individuals and other neurodegenerative disorders.**

To reveal these hypothetical auto-antigenicity profiles, I applied two different high-throughput proteomic-based screening platforms: protein macroarray technology and filamentous phage display of human full-ORF polypeptides.

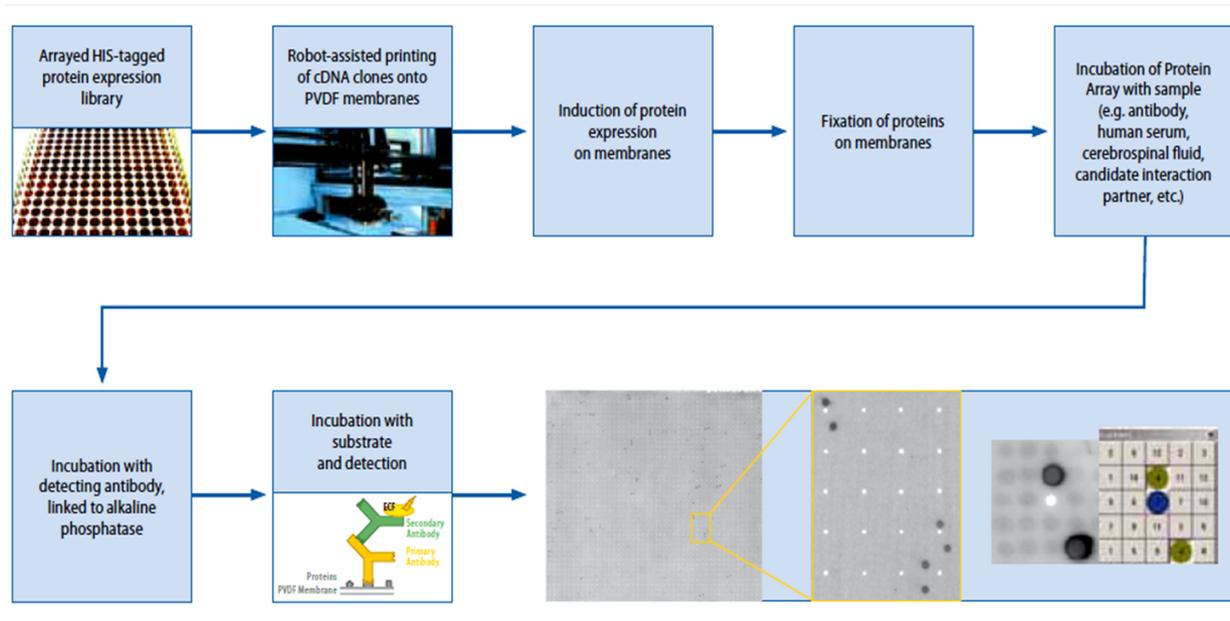
## 1.2. High-throughput screening technologies

Proteomics-based high-throughput screening platforms have become valuable tools in contemporary biochemical and medical research, allowing a rapid and straightforward scanning of large sample sets. They are particularly popular in the field of drug target and biomarker discovery. Thus, in laboratory practice many projects begin with a high-throughput screening in order to quickly prove the initial hypothesis and to narrow down the number of potential candidates, which are then further evaluated in more specific assays. In addition, most platforms are automatable, making high-throughput approaches fairly easy to handle [147]. Automation typically leads to less susceptibility for manual aberrations and to better reproducibility. Especially highly heterogeneous samples such as blood serum are ideal for this kind of screening technologies. In my thesis I applied two proteomics-based high-throughput screening methods that are entirely different in mode of action and their initial source libraries are complementing each other: protein macroarray and phage display of human full-ORF polypeptides. The latter was additionally coupled to another high-throughput technique as a read out tool: Next generation sequencing. Here, all three are described in detail, with a special focus on their mode of application through my thesis.

### 1.2.1. Protein macroarray technology

The first high-density protein arrays were developed at our research institute in the 90's by Büssow *et al.* [148, 149]. They comprise of thousands of arrayed bacteria on filter membranes, harboring human cDNA in expression plasmids. Once bacteria cells are grown and expression of recombinant proteins is induced, cells are lysed and the protein content is cross-linked to the membrane. Since that time, the protein macroarrays were several times modified and refined. Currently, the final version (hEXselect) can be commercially obtained from Source BioScience, Nottingham, UK. The macroarrays are large sheets of PVDF membranes (22 cm x 22 cm), on which 23,806 different cDNA expression clones of the original human hEx1 library, derived from human fetal brain tissue mRNA, are directly spotted. 9,709 clones (~40%) are annotated as expressed in frame, thus representing 3,657 real genes. Clone redundancy, i.e. the number of different clones covering the same gene, varies between 1 and 467 clones per gene. Yet, the median value of the redundancy distribution is 2 clones per gene: over 75% of all genes are represented by only 3 clones and less than 5% of all genes are covered by more than 10 different clones. Approximately 30% of all clones are so called "singletons", being the only clone representing a gene. These have been physically replicated in the library (real duplicates).

All clones have been sequenced and are entirely annotated. They possess an N-terminal His<sub>6</sub>-tag and have been subjected to thorough in-frame analysis. Each clone is spotted in duplicate and in a certain pattern, which allows a reliable assignment of the positive hits. Fig. 3 below shows the workflow of filter production, screening procedures and final analysis of the hit clones.



**Fig. 3 Overview of the protein macroarray production and usage.** Bacterial cDNA expression library is grown in a 384-well microtiter plate and subsequently printed on a PVDF membrane by a spotting robot. Overnight protein expression in *E. coli* is induced directly on the membrane by IPTG. During fixation, bacteria are simultaneously lysed and lyophilized. Screenings are performed by incubating the filter in the fluid of interest, e.g. blood serum, in an appropriate dilution. Detection is achieved by a secondary antibody, conjugated with an alkaline phosphatase (AP) enzyme and a suitable fluorescent substrate (e.g. AttoPhos®, Roche). Since all clones are spotted in duplicate and arranged in a specific 5 x 5 pattern around guiding ink dots (here white spots), it can be easily distinguished between positive hits (here black spots) and residual non-reactive clones (here grey spots). Digitalized pictures can be analyzed by commercially available software packages like AIDA Image Analyzer (Raytest, Berlin, GER)). Source: Protein Macroarrays Manual, Source BioScience Imagenes, November 2010 [150].

Virtually all kinds of screenings, based on protein interactions, can be performed on hEX protein macroarrays. They have been successfully applied in studies to identify and analyze novel protein-protein interactions [151-153] or DNA-binding proteins [154]. However, their largest utilization is probably the screening of protein-antibody interactions in body fluids, such as blood serum or CSF from different donor cohorts in comparison. The determination of auto-antigenicity profiles proved to be a valuable source of novel diagnostic biomarker targets. Recent examples for this kind of applications are screenings of patients' blood sera with glaucoma [155], colorectal cancer [156], multiple sclerosis [157, 158], myasthenia gravis [159], neuroblastoma versus Wilms tumor cancers [160], progressive encephalomyelitis [161], anaplastic large cell lymphoma [162], neuro-Behçet disease [163], Wegener's granulomatosis [164]. All mentioned screenings successfully yielded disease-specific auto-antigenicity patterns, useful for discriminating between diseased and healthy donors or

between different cancer types, for example. Of note, however, is that all protein-antibody interactions described so far are exclusively based on antigen-detection with serum antibodies of the IgG class.

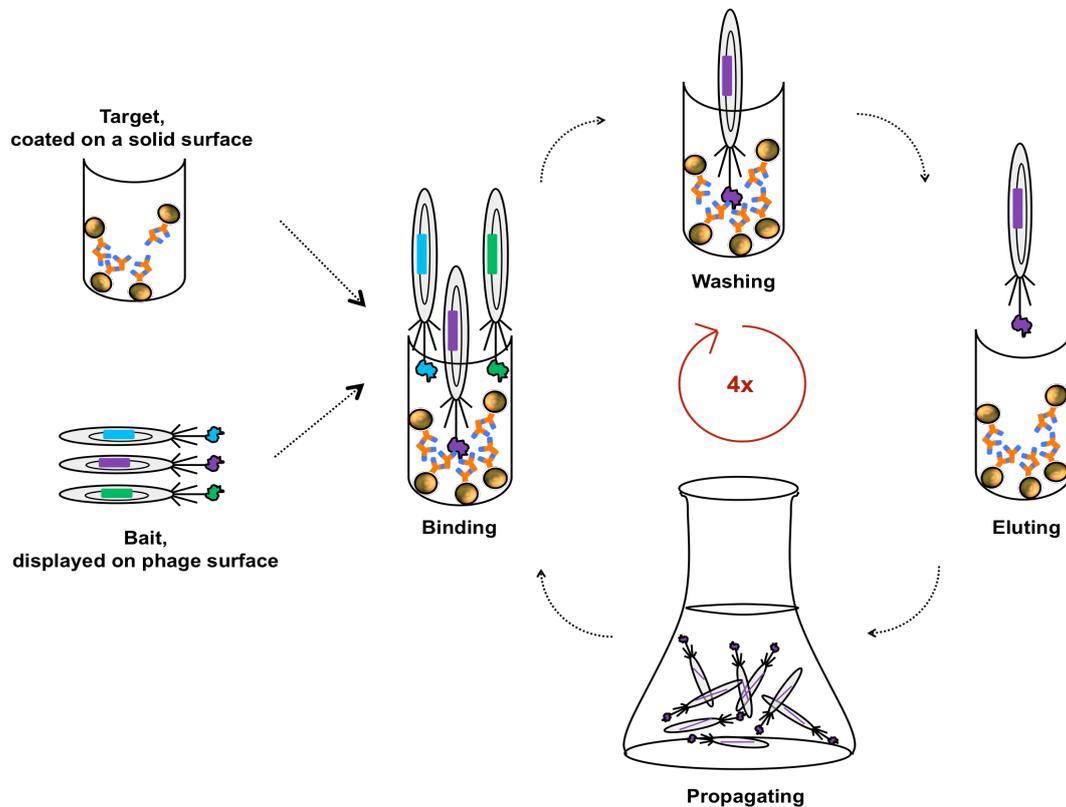
The reason for using anti-IgG detection is obvious, since this major immunoglobulin family is the most abundant in human blood. Furthermore, autoreactive IgGs play a central role in many autoimmune disorders. Yet, also IgA is an interesting study object in the field of autoimmunity, even though it is mainly concentrated in mucous secretions, such as saliva and gastrointestinal tract. In peripheral blood circulation, IgA is the second most abundant isotype, but with a five times higher metabolization rate than IgG [165]. Furthermore, IgA has a direct linkage to autoimmunity, although full mechanism is not entirely understood. Thus, adults, deficient in IgA, were shown to be more susceptible to autoimmune disorders like rheumatoid arthritis, Morbus Crohn or allergies [166, 167]. On this account, a few research groups focus on detection of autoreactive IgA antibodies as well. Potential IgA-based serum biomarkers have been recently described for disease like esophageal cancer [168], autoimmune hemolytic anemia [169], celiac disease [170] and inflammatory bowel disease [171].

### 1.2.2. Filamentous phage display of full-ORF polypeptides

Another powerful platform for high-throughput proteomic screenings is (bacterio)phage surface display. First introduced by Smith *et al.* in 1985, it quickly evolved to one of the most popular presentation scaffolds for polypeptides, due to its highly advantageous physical linkage between genotype and phenotype [172]. Three different kinds of phage display systems have been developed and successfully utilized in molecular biotechnology: T7 phage [173],  $\lambda$ -Phage [174] and filamentous phages of the Ff-class: f1, fd and M13 [175, 176]. Among these, Ff phages, especially M13, are the most widespread. Ff bacteriophages infect gram-negative bacteria, including *E. coli*, in a non-lytic way.

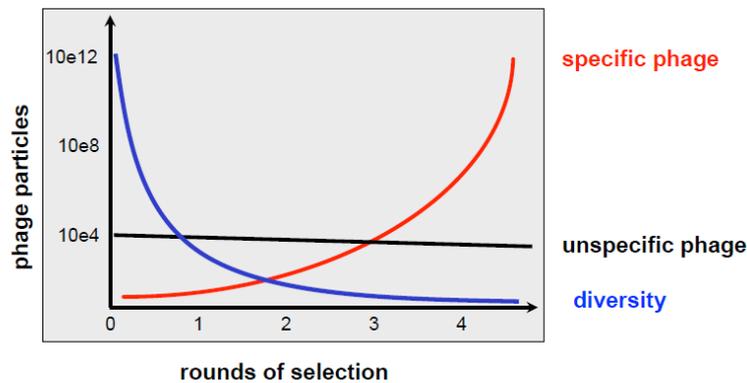
The process of binders selection is called bio-panning in analogy to gold panning, a method used to gather gold particles by agitating earth sediments in water. In principle, phage display mimics the molecular evolutionary selection process *in vitro*. Genes of interest are incorporated into the phage genome in a way to be presented on phage surface, fused to one of the coat proteins. Hence, the polypeptides of interest are being displayed to face their interaction partners in the subsequent selection rounds. In each round, the displayed library is confronted anew with the target of choice, which is usually coated on a solid surface, e.g. nitrocellulose [177], polystyrene bottoms of microtiter plates [178] or magnetic beads [179]. Bait and target undergo following repetitive steps in each selection round: binding, washing,

elution (if required) and finally propagation of selected phages in a bacterial host. Fig. 4 below demonstrates the overall selection procedure schematically.



**Fig. 4 Workflow of the phage display bio-panning process.** In the shown example, targets are antibodies, coated on magnetic beads and baits are proteins, displayed on filamentous phage surface. The genotype-phenotype linkage in phages is indicated by using same color for the GOI and for the POI. Selection with four subsequent rounds is shown. The phage population, displaying a purple protein, binds strong enough to the target and is being selected. On the contrary, phage populations, displaying green and blue proteins do not interact with target and are being washed away. In the end, only the purple population is being propagated and enters the next selection round.

Thereby, an additional washing step is included in each following selection round, in order to increase pressure for the selection of strong binders. Thus, with each additional round the weaker binders become naturally outsourced, while the stronger binders are propagated further. As a consequence, the diversity of the initial phage library declines rapidly during selection, while the specificity of the remaining binders proportionally increases. This main characteristic of the PD selection process is shown in Fig. 5 below.



**Fig. 5 Overview of the phage display selection process.** With the increasing number of selection rounds, also the number of enriched specific phages increases. In contrast, diversity of the phage library and background of non-specific phages decrease. Source: Konthur, Z. [180].

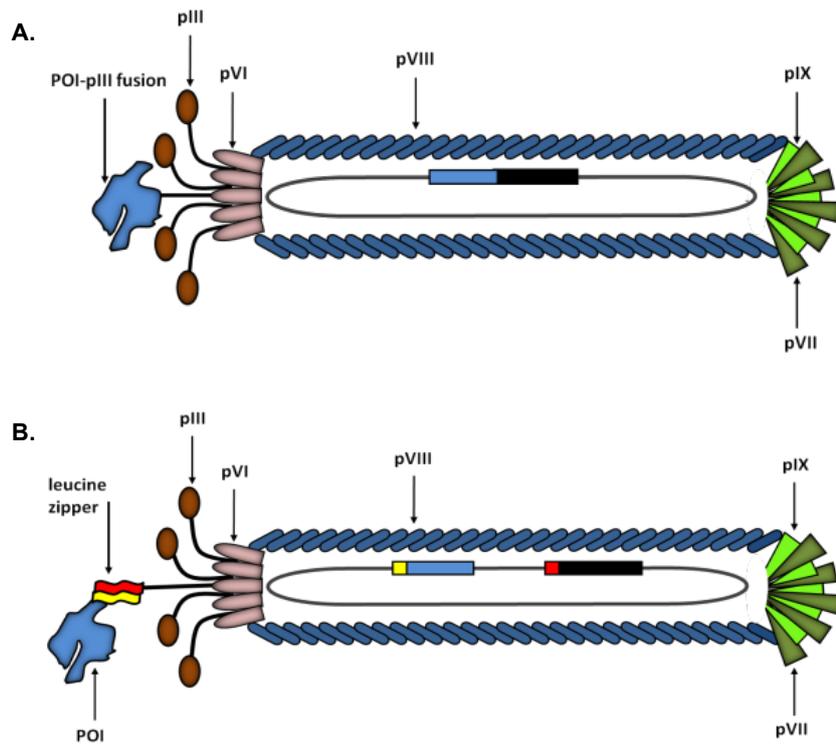
Thus, depending on experimental aim, low number of selection rounds can be chosen to keep library diversity high. Respectively, additional rounds deliver high-affinity binders and a smaller background of non-specific binders.

The M13 virus particle has the shape of a 900 – 1,000 nm long and 6 - 7 nm wide rod. Its single-stranded circular genome (ssDNA) encodes for overall 11 proteins. Five of them are structural coat proteins: one major pVIII, presented with ~2,800 copies per phage, and four minor pIII, pVI, pVII and pIX, each presented with five copies per phage. While pVIII fully covers the longitudinal sides of the virion, the minor coat proteins are situated at its distal ends: pIII and pVI at the one end, pVII and pIX at the opposite (s. Fig. 6) [181]. All five coat proteins have been shown to be appropriate for polypeptide fusion and its presentation on phage surface. However, pIII and pVIII proved to be most effective for these purposes [182]. Yet, pVIII is only suitable for the presentation of shorter peptides, as longer peptides lead to capsid instability. Thus, pIII is often the method of choice in conventional applications and in the present work as well.

Wildtype pIII consists of three domains, D1, D2 and D3, catenated by glycin-rich linkers [183]. It plays a major role in the infection and cell intrusion processes, which occur at the bacterial F-Pilus site [184]. Engineered pIII, applied in phage display, is usually truncated, missing both D1 and D2 or at least a part of D2, which results in diminished infectivity of the recombinant phage particles [185]. In the late 80's so-called phagemid cloning vectors were established and became rapidly popular [186]. They combine features of both bacterial and phage plasmids, such as two origins of replication (bacterial and phage), a phage packaging signal, an antibiotic resistance gene and a multi-cloning site for the polypeptide of interest (POI) to be fused to pIII [187]. Since all other phage genes are absent in the phagemid, a co-infection with a special helper phage is needed to “rescue” the cloned pIII-POI-fusion and to

produce recombinant phage particles. In turn, helper phages are usually genetically engineered to contain deletions in their origin of replication and/or packaging signal sites. Consequently, phagemids and not the helper phage genomes are preferably integrated in the newly produced virions [188]. The choice of an appropriate helper phage allows choosing between different types of POI presentation. For instance, co-infection with the M13K07 helper phage leads to a monovalent display, meaning that on average only one of the five pIII-proteins carries the polypeptide of interest, while the residual four are non-recombinant. In the so-called Hyperphage, on the other hand, the wildtype *pIII*-gene is completely removed from the genome. Hence, only pIII-POI-fusions are integrated into the phage virion leading to a multivalent display, i.e. simultaneous presentation of the recombinant POI on all five pIII-proteins. Consequently, very low infectivity is usually characteristic for multivalent phages, which can potentially hamper the selection process [189]. Thus, each presentation strategy can prove advantageous, depending on project aim. In a bio-panning process with multiple selection rounds, different helper phage applications can be combined, since co-infection is required at every single round of selection.

Filamentous phage display found its broadest application in presentation of combinatorial antibody and peptide libraries [190-192]. Heterogeneous libraries like cDNA-products, however, proved more complicated, due to significant intrinsic limitations. One such limitation is the orientation of the capsid proteins. Most of them are presenting their N-terminus to the outer milieu and their C-terminus is buried within the capsid. Hence, only N-terminal direct fusions of the displayed molecule are feasible. This requires the inserts to be in frame with the downstream *pIII*-gene and to contain no stop codons. Yet, most cDNA-fragments naturally contain translational stop codons in their UTRs. Furthermore, due to the triplet nature of the genetic code, the statistical chance of a gene-coding DNA fragment to contain an ORF to be in frame with *pIII* is 1:3 per definition. Consequently, many strategies have been developed to overcome these bottlenecks and make cDNA phage display more manageable, which we reviewed recently [193]. In 1993 Cramer *et al.* introduced a truly elegant solution to overcome this problem, namely a phagemid vector, called pJuFo that allowed separate expression of pIII and the polypeptide of interest and their subsequent indirect fusion via a leucine-zipper structure [194]. In detail, pIII is expressed with an N-terminal fusion to the leucine zipper domain of the c-Jun, while the POI is expressed with the accordant N-terminal leucine zipper domain of c-Fos. Both, c-Jun and c-Fos are human transcription factors, possessing leucine stretches, able to form very stable leucine zippers [195]. The POI-Fos-fusion and the pIII-Jun-fusion are separately exported into bacterial periplasmic space, where they finally meet and “stick” together. Fig. 6 below shows direct and indirect pIII-POI-fusion in comparison.



**Fig. 6 Schematic representation of bacteriophage M13 and different monovalent display types.**

**A.** Direct fusion of POI to a truncated pIII. **B.** Indirect fusion by means of a leucine zipper structure. Major coat protein pVIII and all minor coat proteins pIII, pVI, pVII and pIX are depicted as well. Source: Georgieva and Konthur [193].

Another issue of heterogeneous polypeptide display is the folding of the presented molecules. In *E. coli* the folding process is dependent on the surrounding milieu. The most common approach, as used for antibody fragments for example, is to equip the POI with a signal peptide, targeting the *E. coli* Sec-translocon [196]. The Sec-export machinery guides polypeptides through the cell membrane in an unfolded state after completed translation. Since pIII needs the periplasmic space for proper folding and association with the rest of the phage coat proteins, Sec is the most appropriate export way for it too [197]. Yet, Sec occurred to be rather unsuitable for other molecules, such as DARPin that could only be successfully displayed when exported through the co-translational SRP-dependent pathway [198]. The third well-known *E. coli* export system, the TAT-pathway, proved appropriate for strictly cytoplasmic proteins, since it translocates mature, already folded polypeptides. Notably, in contrast to Sec and SRP, TAT can only be applied in an indirect POI-pIII-fusion. TAT-mediated phage display was first introduced in 2005 by Paschke and Höhne and is based on the pJuFo-vector [199].

### 1.2.3. Next generation sequencing

The pioneering work of Francis Crick and James D. Watson about decoding nucleic acids structure more than half a century ago steered the research field of molecular biology and biochemistry in a completely new dimension and paved the way for modern life sciences [200]. After Marshal Nirenberg finally deciphered the genetic code in 1962 [201], it was just a matter of time for the first DNA sequencing methods to be developed. Thus, the next milestone in DNA research was the work of Frederick Sanger from 1977 [202]. His DNA sequencing method, based on chain determination and incorporation of labeled ddNTPs, became a worldwide standard and dominated the field for about three decades.

However, recent introduction of the so-called next generation sequencing (NGS) methods revolutionized the DNA sequencing field anew and opened up an entirely novel vista of experimental applications. NGS allows real high-throughput sequencing of millions of DNA fragments in parallel in a very short time without the need of laborious cloning of DNA fragments to generate sequencing templates. NGS platforms and services became commercially available around 2004 – 2005 and still experience rapid development since then. Not only the number of different platforms (i.e. sequencing methods) increases steadily, at the same time the sequencing price per megabase is constantly falling, currently accounting for about 10 US cents per megabase [203]. Consequently, many different applications of NGS were developed, ranging from genome-wide expression surveys (DNA-Seq), through discovery of novel DNA-interacting and –modifying molecules (Methyl-Seq, ChIP-Seq, MAINE-Seq), to transcriptome (RNA-Seq, TRAP) and immune cell analysis (Immune-Seq) and many others (recently reviewed by Shendure and Aiden, 2012 [204]).

The available NGS platforms have different features when it comes down to number of sequence reads, read length and accuracy. In recent years two platforms dominated the international market: Illumina/Solexa Genome Analyzer (Illumina) and 454 FLX (Roche). Roche/454 was the first NGS platform to achieve market access in 2004 [205]. Its methodology is based on so-called pyrosequencing. Pyrophosphates are generated in the process of nucleotide incorporation into the new DNA strand by the DNA polymerase. These are further involved in a side-reaction cascade, at which end a light emission reaction by a luciferase enzyme takes place. This is finally the detection signal for the nucleotide incorporation event. Specific for the 454 sequencing method, is the immobilization of the target DNA molecules on agarose beads in a way that each bead is associated with a single DNA molecule. Additionally, sample amplification is usually performed in emulsion PCR, i.e. in separate oil:water microreactions. Consequently, amplification biases, which usually occur in a conventional open PCR, are strongly minimized [206]. This approach proved advantageous for analyzing *in vivo* conditions, such as natural antibodies repertoires, for

example. Biggest advantage of the 454 system is certainly the relatively large sequence length of up to 500 – 1,000 bases per molecule [207]. Therefore, 454 is usually applied for cDNA-sequencing [208] and in other cases, where insert lengths are not defined and possibly long reads are desirable.

Illumina is currently the primary player on the international NGS market, mainly due to the Solexa sequencing technology, which was commercially acquired in 2006 [209]. The methodology is based on the sequencing-by-synthesis approach. For this, all four nucleotides are added with the DNA polymerase enzyme to the reaction mix. Each nucleotide carries a base-specific fluorescent label and a blocked 3'-OH-group. As a consequence, every incorporation is a unique event, followed by an imaging and detection step. After the image is taken, the blocking agent and the dye at the 3'-end are removed and the DNA strand is ready for yet another incorporation round. Fig. 7 below illustrates the Illumina/Solexa working principle.



Illumina read lengths are naturally defined by the DNA shredding step at the beginning of the sequencing procedure (75 – 150 bp). Biggest advantages of this platform are significantly lower costs per run, compared to 454/Roche, and very high output of sequence data (currently up to 600 Gb of sequence for the HiSeq 2500/2000 sequencer [211]).

In the last few years also entirely novel sequencing approaches, aiming at single molecule sequencing in real-time detection (SMRT), are on the rise [212]. Market leader with this kind of method is currently Pacific Biosciences (USA). Their first sequencer PacBio RS was introduced in late 2010, of which a second improved version is available since April 2013 [213]. The average read length reaches ca. 1,000 bp, but up to 15,000 bp should be theoretically possible, as proposed by the manufacturers [214]. In general, due to a highly competing market and ever growing demand in the scientific communities worldwide, all sequencing platforms experience a steady improvement in regard of read lengths, accuracy, number of sequences obtained per run, sequencing time needed and, last but not least, declining costs. Thus, further exciting developments and applications in this field can be certainly expected in future.

The phage display community has also developed diverse strategies to utilize NGS platforms. The combination of the two high-throughput technologies led to remarkable improvements in performing and evaluating affinity-driven selections. Two major applications proved rather advantageous in this field. Sequencing of the initial pre-selection phage library allows rapid and fairly complete examination of its quality and diversity, on the one hand. On the other, sequencing of selected libraries not only enables identification and annotation of the enriched binders, but also makes enrichment visible, since all molecules in a sample are being sequenced at once. This kind of “deep sequencing” can be informative for prioritization of strong vs. weak binders. First publications, demonstrating successful NGS utilization in combination with PD, appeared around 2009 [215-217]. To date, almost every new publication, dealing with combinatorial libraries and *in vitro* selections, describes implication of an NGS platform as well. Recently, the overall workflow of a “usual” PD-NGS-application was reviewed by Ravn *et al.* [218].

At the beginning of the present work (June, 2009), the conjoined usage of PD and NGS was a novel and exciting application. Meanwhile, however, it became rather frequent. Considering the on-going rapid development of the NGS methods, decreasing sequencing costs and steadily improving possibilities for storage and processing of huge amounts of data bioinformatically, it may be concluded that this methodological symbiosis will become standard in near future.

## 2. Objective

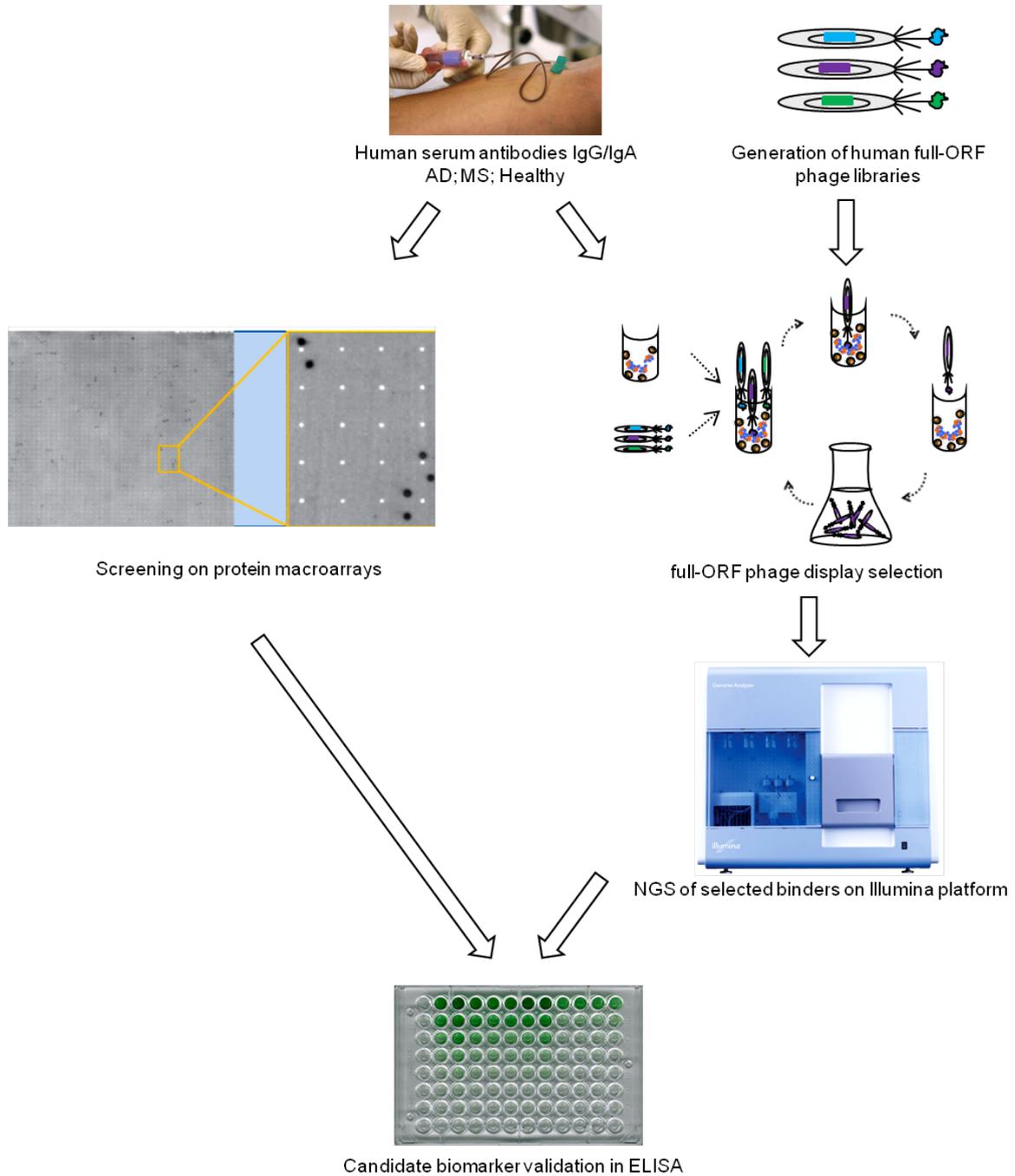
To date, there are no validated uniform biomarkers available for the precise diagnosis of Morbus Alzheimer (AD) or Multiple Sclerosis (MS). Blood serum is an excellent medium for diagnostic purposes, since it is easy to sample and highly informative about organism's health status. The blood brain barrier becomes leaky during neuroinflammation, which occurs in the course of both disorders. Thus, it is expectable that the immune system gets in contact with CNS antigens, resulting from massive neurodegeneration. My thesis is based on the hypothesis that disease-specific autoantibodies are generated and can be detected in blood sera of AD and MS patients. Consequently, major task of my work was to describe intersecting autoimmune profiles between the three different donor cohorts (AD, MS and healthy) and to expose disease-specific autoantigens in AD and MS, which can be used for multiparametric diagnostic tests in future.

Two different proteomics-based high-throughput screening technologies were applied: a human brain cDNA expression library, arrayed on macroarrays, and human full-ORF polypeptides, displayed on the surface of M13 phages. Both presentation scaffolds were screened for interactions with human autoantibodies of the immunoglobulin classes IgG and IgA from the three blood sera cohorts. Phage display selected autoantigens were determined, applying the next generation sequencing platform Illumina Genome Analyzer. Identified positive hits from both screening procedures were further evaluated to prove their potential as applicable diagnostic biomarkers. For this, chosen biomarker candidates with best performances were recombinantly expressed in *E. coli*, affinity-purified and tested in ELISA with larger sera sets.

A secondary, methodology-driven task was the further development of the M13 phage display technique. The functional presentation of heterogeneous full-ORF libraries is non-conventional and was established in the course of my studies. To expand the range of properly displayed proteins, novel phagemid vector series were generated, implementing all three major export pathways in *E. coli*. To prove the feasibility of the newly constructed vectors, test cloning with EGFP and subsequent functionality assays with EGFP-presenting phages were performed.

Finally, results from both screenings were summarized and compared. Advantages and limitations of the two applied high-throughput technologies were discussed in regard to technical applicability, initial source libraries and future application fields.

## Overall experimental workflow



**Fig. 8 Overall experimental workflow.** Pictures sources: [www.picture-alliance/dpa](http://www.picture-alliance/dpa); Protein Macroarrays Manual [150]; [www.illumina.com](http://www.illumina.com); [www.biology.arizona.edu](http://www.biology.arizona.edu)

### 3. Materials

#### 3.1. Consumables

<b>Material</b>	<b>Manufacturer</b>
96-well conical bottom MTP, PP, natural, 0.5 ml	Nunc, Thermo Fisher Scientific, USA
96-well flat bottom MTP, PST, clear, 0.3 ml	Nunc, Thermo Fisher Scientific, USA
96-well flexible PVC flat-bottom MTP	Becton, USA
Amicon® centrifugal filter devices Ultra-15, 10 K	Merck Millipore, IRL
Cryo vials 2 ml	Greiner Bio-One GmbH, Frickenhausen, GER
Culture tubes round-bottom 5 ml	Greiner Bio-One GmbH, Frickenhausen, GER
Dialysis tubing	Carl Roth, Karlsruhe, GER
Disposable cuvettes P948	Carl Roth, Karlsruhe, GER
Disposable pipette tips 10 µl; 200 µl; 1 ml	Gilson, Middleton, USA
Disposable sterile pipettes 5 ml; 10 ml; 25 ml	Corning, New York, USA
Scalpel Cutfix	B Braun, Melsungen, GER
Electroporation cuvettes 1 mm	Bio-Rad Laboratories GmbH, Munich, GER
Erlenmeyer flasks 50 ml; 250 ml; 500 ml; 1 l	Schott, Mainz, GER
Glass bottles 100 ml; 250 ml; 500 ml; 1 l, 3 l	Schott, Mainz, GER
Hybond ECL nitrocellulose membrane	GE Healthcare, Munich, GER
Inoculation loops	Nunc, Roskilde, DK
Kimtech precision wipes, white	Kimberly-Clark Professional, UK
Lazy-L spreader	Sigma-Aldrich GmbH, Munich, GER
Ni-NTA Agarose	Qiagen N.V., Hilden, GER
Nitrile examination gloves	Blossom, USA
NuPAGE 4 – 12% BisTris gel	Invitrogen, Life Technologies, USA
Parafilm M	American National Can™, USA
PCR 8-tube stripes 0,2 ml	Greiner Bio-One GmbH, Frickenhausen, GER
Petri dishes 100 mm	Greiner Bio-One GmbH, Frickenhausen, GER
PP tubes 14 ml; 50 ml	Greiner Bio-One GmbH, Frickenhausen, GER
Pyrex solid glass beads 3 mm	In-house glazier's workshop, MPI-MG
Reaction tubes (0,5 ml; 1,5 ml; 2,0 ml)	Eppendorf AG, Hamburg, GER
Safe-seal tips (10 µl; 20 µl; 200 µl; 1 ml)	BioZym GmbH, Oldendorf, GER
Whatman paper 3 mm	Biometra, Göttingen, GER

### 3.2. Chemicals, buffers and solutions

Unless stated otherwise, buffers and solutions were sterilized by filtration through a 0,2 µm pore size filter, when required.

<b>Chemicals</b>	<b>Manufacturer</b>
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Merck, Darmstadt, GER
29% Acrylamide, 0,8% Bisacrylamide	Carl Roth, Karlsruhe, GER
ABTS tablets	Sigma-Aldrich GmbH, Munich, GER
Acetic acid	Carl Roth, Karlsruhe, GER
Ammonium acetate	Merck, Darmstadt, GER
Ampicillin	Carl Roth, Karlsruhe, GER
APS	BioRad Laboratories GmbH, Munich, GER
Bacto Agar	Becton, Difco Laborastories, Sparks, USA
Bacto Tryptone	Becton, Difco Laborastories, Sparks, USA
Bacto Yeast Extract	Becton, Difco Laborastories, Sparks, USA
Boric acid	Merck, Darmstadt, GER
Bromophenol blue	Sigma-Aldrich GmbH, Munich, GER
BSA	NEB, Ipswich, UK
CaCl <sub>2</sub>	Merck, Darmstadt, GER
Chloramhinicol	Carl Roth, Karlsruhe, GER
Citric acid	Merck, Darmstadt, GER
Coomassie Brilliant Blue G250	Sigma-Aldrich GmbH, Munich, GER
D(+)-Glucose monohydrate	Merck, Darmstadt, GER
DNA loading dye, 6x	Fermentas, Thermo Fisher Scientific, USA
dNTP's	NEB, Ipswich, UK
DTT	Serva, Heidelberg, GER
Ethanol	Merck, Darmstadt, GER
Ethidium bromide	Merck, Darmstadt, GER
Glycerin	Merck, Darmstadt, GER
Glycine	Merck, Darmstadt, GER
Glyco Blue	Applied Biosystems, Carlsbad, USA
H <sub>2</sub> O <sub>2</sub>	Merck, Darmstadt, GER
HCl	Carl Roth, Karlsruhe, GER
Imidazole	Sigma-Aldrich GmbH, Munich, GER
IPTG	Fermentas, Thermo Fisher Scientific, USA
Isobutanol	Merck, Darmstadt, GER
K <sub>2</sub> HPO <sub>4</sub>	Merck, Darmstadt, GER
Kanamycin	Carl Roth, Karlsruhe, GER
KCl	Merck, Darmstadt, GER

K <sub>2</sub> HPO <sub>4</sub>	Merck, Darmstadt, GER
KH <sub>2</sub> PO <sub>4</sub>	Merck, Darmstadt, GER
MES	Sigma-Aldrich GmbH, Munich, GER
MgCl <sub>2</sub>	Merck, Darmstadt, GER
MgSO <sub>4</sub>	Merck, Darmstadt, GER
Na <sub>2</sub> EDTA	Merck, Darmstadt, GER
NaH <sub>2</sub> PO <sub>4</sub>	Merck, Darmstadt, GER
Na <sub>2</sub> HPO <sub>4</sub>	Merck, Darmstadt, GER
Na <sub>3</sub> -citrate	Merck, Darmstadt, GER
NaCl	Sigma-Aldrich GmbH, Munich, GER
NaOH	Merck, Darmstadt, GER
Non-fat dry milk powder	Bio-Rad Laboratories GmbH, Munich, GER
PEG 6000	Merck, Darmstadt, GER
SDS	Bio-Rad Laboratories GmbH, Munich, GER
β-Mercaptoethanol	Merck, Darmstadt, GER
Sucrose	Merck, Darmstadt, GER
TEMED	Invitrogen, Life Technologies, USA
Tris-base	Merck, Darmstadt, GER
Triton X-100	Sigma-Aldrich GmbH, Munich, GER
Tween 20	Sigma-Aldrich GmbH, Munich, GER
UltraPure Agarose	Invitrogen, Life Technologies, USA
Urea	AppliChem, Darmstadt, GER

<b>Buffers</b>	<b>Recipe</b>
AttoPhos buffer	1 mM MgCl <sub>2</sub> ; 100 mM Tris/HCl (pH 9.5)
Blocking buffer for macroarrays	TBS-T + 3% non-fat dry milk powder
Blocking buffer for magnetic beads	PBS + 0.5% BSA + 0.05% Tween 20, pH 7.4
Coating buffer for magnetic beads	0.1 M Boric acid, pH 9.5 with NaOH
Laemmli loading buffer 4x	0.04% Bromophenol blue; 40% (w/v) Glycerin; 8% SDS; 5% β-Mercaptoethanol; 0.2 M Tris-HCl pH 6.8
Laemmli running buffer	0.125 mM Tris-base; 1.25 mM Glycine; 0.5% SDS
MES running buffer 20x, pH 7.3	1 M MES; 1 M Tris-base; 69.3 mM SDS; 20.5 mM EDTA
Native elution buffer	50 mM NaH <sub>2</sub> PO <sub>4</sub> ; 300 mM NaCl; 250 mM Imidazole
Native lysis/binding buffer, pH 8.0	50 mM NaH <sub>2</sub> PO <sub>4</sub> ; 300 mM NaCl; 10 mM Imidazole

Native wash buffer, pH 8.0	50 mM NaH <sub>2</sub> PO <sub>4</sub> ; 300 mM NaCl; 20 mM Imidazole
PBS-T	PBS + 0.2% Tween 20
PBS, pH 7.4	0.0027 M KCl; 0.137 M NaCl; 0.01 M phosphate buffer (Na <sub>2</sub> HPO <sub>4</sub> /NaH <sub>2</sub> PO <sub>4</sub> )
PE buffer for periplasmic extraction	0.1 M Tris-HCl (pH 8.0); 1 mM EDTA; 0.5 M Sucrose
PTM blocking buffer	PBS-T + 2% non-fat dry milk powder
Stripping buffer for protein macroarrays	2% SDS; 65.5 mM Tris/HCl (pH 6.8); 100 mM β-Mercaptoethanol
TAE running buffer 10x, pH 8.2	48.4 g/l Tris-base; 10.9 g/l Glacial acetic acid; 2.92 g/l EDTA
TBE running buffer 50x, pH 8.2	54.0 g/l Tris-base; 27.5 g/l Boric acid; 2.92 g/l EDTA
TBS	10 mM Tris-HCl (pH 7.5); 150 mM NaCl
TBS-T	TBS + 0.1% Tween 20
TE	10 mM Tris-HCl (pH 8.0); 1 mM EDTA
TES buffer for periplasmic extraction	0.2 M Tris-HCl (pH 8.0); 0.5 mM EDTA; 0.5 M Sucrose
Transfer buffer for western blot	4.2 g Tris-base; 19.6 g Glycine in 1.4 l ddH <sub>2</sub> O
Washing and storage buffer for magnetic beads	PBS + 0.1% BSA + 0.05% Tween 20, pH 7.4

Solutions	Recipe
ABTS substrate solution	1x ABTS tablet (10 mg); 10 ml 50 mM Na <sub>3</sub> -citrate; 10 ml 50 mM Citric acid; 10 μl 30% H <sub>2</sub> O <sub>2</sub>
Agarose – gel solution	5 – 10 g/l Agarose in 1x TAE or 0.5x TBE buffer
Ammonium acetate, 7.5 M, pH 7.5	57.8 g Ammonium acetate in 100 ml ddH <sub>2</sub> O
BSA, 4%	4 g BSA in 100 ml PBS
Coomassie blue staining solution	1.25 g Coomassie Brilliant Blue G250; 225 ml Ethanol (technical grade); 50 ml Acetic acid.
Destain solution	20 ml Ethanol (technical grade); 10 ml Acetic acid; 70 ml ddH <sub>2</sub> O
EDTA, 0.5 M, pH 8.0	186.1 g Na <sub>2</sub> EDTA x 2H <sub>2</sub> O in 1 l ddH <sub>2</sub> O
Ethanol, 70% (v/v)	700 ml Ethanol (technical grade) in 1 l ddH <sub>2</sub> O
Glucose, 40% (w/v)	400 g D(+)-Glucose monohydrate in 1 l ddH <sub>2</sub> O
IPTG, 1 M	1.19 g IPTG in 5 ml ddH <sub>2</sub> O
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> stock solution, 3 M, pH 9.5	39.6 g (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> in 100 ml coating buffer

Phage precipitation solution	20% PEG 6000; 2.5 M NaCl in ddH <sub>2</sub> O
SDS separating gel solution	23.8 ml ddH <sub>2</sub> O; 0.4 ml SDS (10%); 10 ml Tris-HCl (0.5 M, pH 6.8); 5.3 ml 29% Acrylamide, 0,8% Bisacrylamide. 400 µl APS; 40 µl TEMED
SDS stacking gel solution	10.4 ml ddH <sub>2</sub> O; 0.8 ml SDS (10%); 20 ml Tris-HCl (1.5 M, pH 8.8); 48 ml 29% Acrylamide, 0,8% Bisacrylamide. 800 µl APS; 25 µl TEMED
Solution 1 for HMFM	3.65 mM MgSO <sub>4</sub> x 7H <sub>2</sub> O; 36 mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mM Na <sub>3</sub> -citrate x 2H <sub>2</sub> O; 44% Glycerin
Solution 2 for HMFM	132 mM K <sub>2</sub> HPO <sub>4</sub> ; 270 mM KH <sub>2</sub> PO <sub>4</sub>
Tris-HCl, 1 M, pH 6.8	12.1 g Tris-base in 100 ml ddH <sub>2</sub> O
Urea, 8 M	48 g Urea in 100 ml ddH <sub>2</sub> O

### 3.3. Growth culture media and additives

Unless stated otherwise, growth culture media were sterilized by autoclaving at 120°C for 20 min. If required, additives were added after cooling down to at least 55°C. Standard concentration of Glucose in all media was 2%, unless stated otherwise.

Growth culture media and additives	Recipe
2xYT Broth	16 g Bacto Tryptone; 10 g Bacto Yeast Extract; 5 g NaCl in 1 l ddH <sub>2</sub> O
2xYT Broth Agar	31 g 2xYT – Broth + 5 g Bacto Agar in 1 l ddH <sub>2</sub> O
Ampicilin, working solution 100 µg/ml	Stock solution: 50 mg/ml in 50% Ethanol
Chloramphenicol, working solution 17 µg/ml	Stock solution: 30 mg/ml in ddH <sub>2</sub> O
HMFM freezing additive, 10x	4 x Solution 1 : 1 x Solution 2
Kanamycin, working solution 60 µg/ml	Stock solution: 34 mg/ml in ddH <sub>2</sub> O
LB	10 g Bacto Tryptone; 5 g Bacto Yeast Extract; 10 g NaCl in 1 l ddH <sub>2</sub> O
SOB	20 g Bacto Tryptone; 5 g Bacto Yeast Extract; 0.5 g NaCl in 1 l ddH <sub>2</sub> O
SOC	100 ml SOB + 1 ml Glucose (40%); 1 ml MgCl <sub>2</sub> (1 M); 1 ml MgSO <sub>4</sub> (1 M)
Top Agar	100 ml heated 2xYT Broth Agar + 100 ml 2xYT Broth

### 3.4. *E. coli* strains

Unless stated otherwise, all used bacterial cells were electro-competent.

Strain name	Genotype	Provider
BL21 Star (DE3)	F- <i>ompT hsdSB</i> (rB-, mB-) <i>gal dcm rne131</i> (DE3)	Invitrogen, Life Technologies, USA
DB3.1	F- <i>gyrA462 endA1 glnV44 Δ(sr1-recA) mcrB mrr hsdS20</i> (r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ) <i>ara14 galK2 lacY1 proA2 rpsL20</i> (Sm <sup>r</sup> ) <i>xyf5 Δleu mtl1</i>	Invitrogen, Life Technologies, USA
DH10B	F- <i>mcrA Δ(mrr-hsd RMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara, leu)7697 araD139 galU galK nupG rpsL λ-</i>	Invitrogen, Life Technologies, USA
HB2151	nal <sup>r</sup> <i>thi-1 ara Δ(lac-proAB) thi</i> [F' <i>pro A+B lacI<sup>q</sup> lacZΔM15</i> ]	Clontech, Mountain View, USA
SCS1/pSE111	<i>hsdR17</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>-</sup> ) <i>recA1 gyrA96 thi-1 relA1 supE44</i>	Stratagene, La Jola, USA
TG1	<i>supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>-</sup> ) [F' <i>traD36 proAB lacI<sup>q</sup> ZΔM15</i> ]	Stratagene, La Jola, USA
XL-1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI<sup>q</sup> ZΔM15 Tn10 (Tet<sup>r</sup>)</i> ]	Stratagene, La Jola, USA

### 3.5. Helper phages

Strain name	Provider
Hyperphage M13K07Δ <i>pIII</i>	Progen Biotechnik, Heidelberg, GER
M13K07 Helper Phage	NEB, Ipswich, UK

### 3.6. Plasmids and OC source libraries

Plasmid name	Producer
pENTR 223/hOHS	Carola Stoschek, AG Konthur, MPI-MG
pENTR/EGFP	Miriam Baradari, AG Konthur, MPI-MG
pENTR/GUS	Invitrogen, Life Technologies, USA
pENTR/OCAA	Imagenes, Source BioScience, Nottingham, UK
pENTR/OCAB	Imagenes, Source BioScience, Nottingham, UK
pIT2-mTKIN	Dr. Volker Sievert, AG Konthur, MPI-MG
pJuFo-B	AG Konthur, MPI-MG

pQE30NST	Dr. Konrad Büssow, AG Konthur, MPI-MG
pRSET-BH6	AG Konthur, MPI-MG
pRSET/EGFP	Dejan Gagoski, AG Konthur, MPI-MG
pUC18	Invitrogen, Life Technologies, USA

All pENTR/full-ORF plasmids, containing the human ORFeome collection (OC), were commercially obtained from Imagenes GmbH (now Source BioScience, UK), pooled in four separate DNA prep-samples:

OC source library	Clone number	DNA concentration [ $\mu\text{g}/\mu\text{l}$ ]
pENTR/OCAA with stop codon	6,240	3.13
pENTR/OCAA w/o stop codon	4,929	0.89
pENTR/OCAB with stop codon	1,200	0.15
pENTR/OCAB w/o stop codon	1,152	0.07

OC clones originated from the following providers: DKFZ, NIH-MGC, HIP and WTSI. Most clones have been generated by a BP reaction between a cDNA fragment, flanked by *attB*-sites, and an appropriate pDONR vector to generate the correspondent pENTR vector. For this, four slightly different pDONR vectors have been used: pDONR201, pDONR221, pDONR223 and pDONR223.1 [179].

### 3.7. Protein macroarrays

hEX select filters with arrayed expression products of cDNA fragments from human fetal brain tissue were commercially obtained from Imagenes GmbH (now Source BioScience, UK). Approx. 24,000 different clones are spotted on each macroarray. Redundancy differs strongly between spotted clones, yet each clone is represented at least twice per filter in a distinct double spotting pattern. The majority of all clones are His<sub>6</sub>-tagged and annotated.

### 3.8. Human blood sera

Human blood sera were commercially obtained from in.vent Diagnostica GmbH, Berlin. 20 sera from each of the three patient cohorts AD, MS and healthy were delivered in 2 x 0.5 ml cryo vials and stored at -80°C. Healthy donors were age-matched to the AD donor group and tested with CERAD test as a criterion for non-diseased. Ethic votes are existing for all 60 donors. Patient identity is anonymous.

**Table 2. Annotations of human AD sera.** All donors were tested negative for genetic predispositions for neurodegenerative diseases.

Lot Nr.	Age	Gender	Weight [kg]	Height [cm]	Year of diagnosis	Sort of test	MRT	CT	Medication	Date of sampling
191491	72	m	85	180	2006	CERAD		2006	Aricept 10 mg	10.09.2009
191492	79	f	65	165	2008	CERAD	2008	2008	Aricept 10 mg	10.09.2009
191493	76	m	64	168	2008	CERAD	2007	2007	Aricept 10 mg	10.09.2009
191494	80	m	80	176	2009	CERAD		2009	Aricept 10 mg	10.09.2009
191680	87	f	64	179	2007	MMSE, CDT		2006	Aricept 5 mg	05.10.2009
191724	72	m	80	165	2003	MMSE, CDT	2004		Aricept 10 mg	08.10.2009
191739	76	m	69	168	2007	MMSE, CDT		2007	Ebixa 20 mg, Melpiron 25 mg	12.10.2009
191817	77	m	78	182	2008	CERAD	2008		Aricept 10 mg	26.10.2009
191828	81	f	64	163	2005	CERAD		2005	Exelon 9 mg	27.10.2009
191829	77	f	60	158	2007	CERAD		2008	Aricept 5 mg	27.10.2009
191830	82	m	73	170	2008	CERAD		2008	Aricept 5 mg	26.10.2009
191836	70	m	65	175	2009	MMSE, CDT	2009		Remenyl 16 mg	27.10.2009
191849	86	m	62	162	2009	CERAD			Remenyl 16 mg	29.10.2009
191862	85	f	60	160	2008	CERAD		2008	Aricept 10 mg	30.10.2009
191863	85	f	61	165	2009	MMSE, CDT		2009	Exelon 9 mg	28.10.2009
191865	81	f	65	168	2009	CERAD	2009		Aricept 5 mg	29.10.2009
191866	77	f	80	152	2007	CERAD		2007	Aricept 10 mg	29.10.2009
191868	79	f	56	159	2007	CERAD	2007		Aricept 10 mg	30.10.2009
191874	84	f	73	164	2009	MMSE, CDT		2008	none	30.10.2009
192010	72	f	70	163	2009	MMSE, CDT		2009	none	06.11.2009

**Table 3. Annotations of human MS sera.** All donors were tested positive for MRT visible lesions.

Lot Nr.	Age	Gender	Weight [kg]	Height [cm]	Year of diagnosis	Subtype	VEP	OCB in CSF	Last relapse	Medication
170171	38	f	50	163	1998	RRMS		positive	1998	Copaxone
170172	39	f	67	164	1992	RRMS	pathologic	positive	2002	Copaxone
170175	47	f	85	170	1986	RRMS			2007	Betaferon
170176	44	f	90	160	2000	RRMS		negative	2005	Copaxone
170178	42	f			2005	RRMS	normal	positive	2007	Rebif 44
170179	44	m	76	178	2002	SPMS			2006	Betaferon
170180	42	f	60	170	1995	RRMS	normal	positive		Rebif 44
170181	45	f	75	178	2004	RRMS	normal		2006	Rebif 22
170184	38	f	59	172	2002	RRMS	pathologic	positive	2004	none
170186	53	m	80	175	1985	RRMS		positive	2004	Copaxone
170187	45	f	63	160	2004	SPMS	normal	positive	2006	Rebif 44
170192	25	f	62	168	2006	RRMS		positive	2006	Betaferon
170195	32	f	52	158	2003	RRMS		positive	2003	Rebif 22
170196	50	m	71	170	1985	RRMS	normal	positive	2007	Copaxone
170197	47	f	80	175	1970	SPMS		positive	90's	Mitoxantron
170198	29	m	95	187	1999	SPMS		positive	2005	Betaferon
170199	50	f	68	164	2007	RRMS		positive	2007	Betaferon
170202	71	f		160	1983	SPMS	pathologic	positive		none
170230	38	f	70	165	2001		pathologic	positive	2007	Rebif 44

170204 30 f 72 171 2004 RRMS pathologic positive 2007 Octagam 10G

**Table 4. Annotations of human Healthy sera.** All donors were tested negative with the CERAD test and had no other diagnosed neurodegenerative diseases at donation point.

Lot Nr.	Age	Gender	Year of test and donation
151002	75	m	2005
151006	77	f	2005
151010	75	f	2005
151014	76	f	2005
151015	73	f	2005
151016	83	m	2005
151017	87	f	2005
151018	77	f	2005
151022	77	m	2005
151034	79	f	2005
155/05	73	m	2005
167/05	75	m	2005
174/05	64	m	2005
181/05	65	f	2005
182/05	70	m	2005
188/05	64	m	2005
193/05	69	f	2005
194/05	68	f	2005
198/05	69	f	2005
204/05	67	m	2005

### 3.9. Antibodies

Antibody	Species origin	Manufacturer
Anti-GAPDH	mouse, monoclonal	RDI, USA
Anti-GFP	rabbit, polyclonal	Invitrogen, Life Technologies, USA
Anti-GFP native 3E6	mouse, monoclonal	MP Biomedicals, USA
Anti-GFP denatured 11E5	mouse, monoclonal	MP Biomedicals, USA
Anti-His <sub>6</sub> -HRP	mouse, monoclonal	Miltenyi Biotec GmbH, GER
Anti-His <sub>6</sub> -RGS	mouse, monoclonal	Qiagen N.V., Hilden, GER
Anti-human-IgA	goat, polyclonal	Pierce, Thermo Scientific, USA
Anti-human-IgA-AP	goat, polyclonal	Sigma Aldrich, Munich, GER
Anti-human-IgA-HRP	rabbit, polyclonal	DakoCytomation, Glostrup, DK
Anti-human-IgG	rabbit, polyclonal	Pierce, Thermo Scientific, USA

Anti-human-IgG-AP	goat, polyclonal	Sigma Aldrich, Munich, GER
Anti-human-IgG-HRP	rabbit, polyclonal	Dako Cytomation, Glostrup, DK
Anti-M13	mouse, monoclonal	Amersham Biotech, UK
Anti-M13-HRP	mouse, monoclonal	GE Healthcare, Munich, GER
Anti-mouse-HRP	rabbit, polyclonal	Sigma Aldrich, Munich, GER
Anti-mouse-IgG	goat, polyclonal	Pierce, Thermo Scientific, USA
Anti-pIII	mouse, monoclonal	MoBiTec, Göttingen, GER
Anti-Tubulin	mouse, monoclonal	BioCarta, San Diego, USA
Anti-rabbit-HRP	goat, polyclonal	AbCam, Cambridge, UK
Anti-Ubiquitin	mouse, monoclonal	Calbiochem, Merck Millipore, GER
Purified human IgA	human, polyclonal	Sigma Aldrich, Munich, GER
Purified human IgG	human, polyclonal	Sigma Aldrich, Munich, GER

### 3.10. Magnetic beads and affinity columns

Product	Manufacturer
Dynabeads® MyOne™ Tosylactivated, 2 ml, 100 mg/ml	Invitrogen, Life Technologies, USA
Ni-NTA Superflow Cartridges ,1 ml	Qiagen N.V., Hilden, GER

### 3.11. DNA and protein markers

Product	Manufacturer
Gene Ruler 1 kb DNA Ladder	Fermentas, Thermo Fisher Scientific, USA
Gene Ruler 100 bp Plus DNA Ladder	Fermentas, Thermo Fisher Scientific, USA
Precision Plus Protein Ladder	Bio-Rad Laboratories GmbH, Munich, GER

### 3.12. Kits

<b>Kit name</b>	<b>Manufacturer</b>
AttoPhos Substrate Set	Roche Diagnostics GmbH, Mannheim, GER
CN/DAB Substrate Kit	Pierce, Thermo Fisher Scientific, USA
Human IgA ELISA Kit	Dunn Labortechnik GmbH, Asbach, GER
Human IgG ELISA Kit	Dunn Labortechnik GmbH, Asbach, GER
Nucleo Spin Extract II	Macherey-Nagel GmbH, Düren, GER
Pierce Silver Stain Kit	Pierce, Thermo Fisher Scientific, USA
Qiagen Plasmid Maxi Kit	Qiagen V.N., Hilden, GER
QIAprep Spin Miniprep Kit	Qiagen V.N., Hilden, GER
QIAquick Gel Extraction Kit	Qiagen V.N., Hilden, GER
QIAquick PCR Purification Kit	Qiagen V.N., Hilden, GER

### 3.13. Enzymes

Unless stated otherwise, enzyme reactions were performed in accordant enzyme buffers, provided by same manufacturer.

<b>Enzyme</b>	<b>Manufacturer</b>
Antarctic Phosphatase FastAP®	NEB, Ipswich, UK
BamHI	NEB, Ipswich, UK
BsrGI	NEB, Ipswich, UK
EcoRI	NEB, Ipswich, UK
Gateway® LR Clonase® II Enzyme Mix	Invitrogen, Life Technologies, USA
HindIII	NEB, Ipswich, UK
Lysozyme	Sigma Aldrich, Munich, GER
NcoI	NEB, Ipswich, UK
NheI	NEB, Ipswich, UK
NotI	NEB, Ipswich, UK
Phusion Hot Start Polymerase	NEB, Ipswich, UK
PMSF	Sigma Aldrich, Munich, GER
Proteases Inhibitor Cocktail	Roche, Basel, SWI
Proteinase K	Invitrogen, Life Technologies, USA

Recombinant Taq Polymerase	In-house production, MPI-MG
SacI	NEB, Ipswich, UK
Sall	NEB, Ipswich, UK
SfiI	NEB, Ipswich, UK
T4 DNA Ligase	NEB, Ipswich, UK
XhoI	NEB, Ipswich, UK
XmaI	NEB, Ipswich, UK

### 3.14. Oligonucleotides

Unless stated otherwise, primers were produced by Eurofins MWG GmbH, Ebersberg, GER. They were delivered in a lyophilized form and diluted in ddH<sub>2</sub>O to an appropriate stock concentration.

Primer name	DNA Sequence 5' → 3'
Fos_seq	GCTCTGCGGTGGTTTGACCG
LMB3	CAGGAAACAGCTATGAC
M13UP (-21)	CGACGTTGTAAAACGACGACGGCCAGT
NcoI-AttR1-for	CAGCCGGCCATGGCCACAAGT
NcoI-CmR-back	ATTTGCCCATGGTAAAAACGGGG
NcoI-CmR-for	CCGTTTTACCATGGGCAAATATT
NheI-pLac-for	CGTTATTATGCTAGCTAGTAACACGAC
NotI-AttR2-back	TCGATTGCGGCCGCTTAACTCTA
pHEN_seq	CTATGCGGCCCCATTCA
pIT2-DsbA-back	ATATGCTATCCCATGGCCGGCTGGGCCTGCTGATGCTGAAAAAGCTAAGACTAA
pIT2-DsbA-for	GCAGTATTAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAAAA ATTTGGTTA
pJuFo-3'	GTAAAACGACGGCCAGT
pJuFo-DsbA-back-1	AGCTAATGCTAACCAAATTTTTTTCATTATGACTGTCTCCTT
pJuFo-DsbA-back-2	CTGATGCTGAAAAAGCTAAGACTAATCCAGCTAATGCTAA
pJuFo-DsbA-back-3	ACTAACATAGAGCTCGGCGATGGCTGCTGATGCTGAAA
pJuFo-XmaI-for	ATTGACACGCCCGGGCGACGGATC
pJuFo-TorA-back-1	TGCCTGAAAGAGATCGTTATTGTTTCATTATGACTGTCTCCTT
pJuFo-TorA-back-2	GCCGCCGAGTTGTGCCAGAAAACGCCGACGTGATGCCTGAAAAGAG
pJuFo-TorA-back-3	TGACGGCCCCAGCATCCCGGCGACGGTTAAGCCGCCGAGTTG

pJuFo-TorA-back-4	TCGCGGCGTTAACAATGACGGCCCCAG
pJuFo-TorA-back-5	CTAACATAGAGCTCGATGGCTTGCGCCGCAGTCGCACGTCGCGGCGTTAA
pJuFo-XbaI-back	AATTGTGTCTAGACCACTTTGTACAAGAAAGCTGA
pQE65	TGAGCGGATAACAATTTACACACAG
pQE276	GGCAACCGAGCGTTCAC
RBS-back	TATGACTGTCTCCTTGGCGACTAGCT

### 3.15. Laboratory equipment

Product	Manufacturer
Agarose gel electrophoresis chambers	In-house production, MPI-MG
ÄKTApurifier	GE Healthcare, Munich, GER
Autoclave FNR 4336E	Tecnomara, Wallisellen, SUI
Autoclave Typ 23	Melag GmbH, Berlin, GER
Bi-Distillation system	In-house production, MPI-MG
BioPhotometer	Eppendorf AG, Hamburg, GER
Branson Sonifier 250	Heinemann GmbH, Schwäbisch Gmünd, GER
CCD camera LAS 1000	Fujifilm, Düsseldorf, GER
Centrifuge 5424	Eppendorf AG, Hamburg, GER
Centrifuge 5810 R with cooling system	Eppendorf AG, Hamburg, GER
Digital pH-/mV-/Thermometer GMH 3510	Greisinger Electronic GmbH, Regenstauf, GER
Electrophoresis power supply: EPS 200, EPS 300, EPS 301, EPS 600	GE Healthcare, Munich, GER
Freezer -80°C Format	Thermo Fisher Scientific, USA
Gel Doc 2000 Gel Documentation System	Bio-Rad Laboratories GmbH, Munich, GER
Incubator Shaker 4430	Innova, Eppendorf AG, Hamburg, GER
King Fisher Flex Magnetic Particle Processor	Thermo Fisher Scientific, USA
Magnet rack Dynal® MPC-S for 6 x 1.5 ml tubes	Invitrogen, Life Technologies, USA
Magnetic stirrer hot plate Stuart SM3	Stuart, Bibby Scientific Lim., Staffordshire, UK
Magnetic stirrer Hotplate MR 3001	Heidolph, Frankfurt am Main, GER
Microcentrifuge SD for PCR tubes	Carl Roth, Karlsruhe, GER
Microflow Laminar Flow Workstation	Kendro Laboratory Products, Langenselbold, GER
MicroPulser™ Electroporator	Bio-Rad Laboratories GmbH, Munich, GER
Minishaker Vortexer MS1	IKA, Staufen, GER

MTP reader Spectramax 200	Molecular Devices, Sunnyvale, USA
Multichannel pipettes 5 – 50 µl, 20 – 200 µl	Abimed, Langenfeld, GER
NanoDrop ND-1000	Thermo Fisher Scientific, USA
NANOpure Diamond Reinstwassersystem	Werner, GER
Peltier Cooled Incubator IPP500	Memmert GmbH, Schwabach, GER
Pipetboy acu	Integra Biosciences AG, Fernwald, GER
Pipettes 200 µl, 1 ml	Eppendorf AG, Hamburg, GER
Pipettes PIPETMAN® 2 µl, 10 µl, 20 µl, 200 µl, 1 ml, 5 ml	Gilson, Middelton, UK
Plate Thermo-Shaker PST-60HL Plus	BioSan, Riga, LV
Plate Thermo-Shaker PST-60HL-4	BioSan, Riga, LV
Polar Star Omega Microplate Reader	BMG Labtech, Ortenberg, GER
Rocky shaker	Fröbel Labortechnik, Wasserburg, GER
Roller TMR-V	IDL, Nidderau, GER
Rotator	Labor-Brand, Gießen, GER
Rotilabo® centrifuge with butterfly-rotor	Carl Roth, Karlsruhe, GER
Self-contained ice flaker 120 kg AF20	Scotsman Ice Systems, USA
Wet blotting device for protein gels TE 70	Hoeffer Inc, Holliston, USA
SUB universal water bath	Grant Instruments, Cambridge, UK
Thermo cycler PTC-200	MJ Research Inc., Quebeck, CAN
Thermomixer Comfort, 1.5 ml and 2 ml	Eppendorf AG, Hamburg, GER
UVL-21 black-ray lamp	Ultra Violet Products, Keswick, AUS
Weighing device Adventure Pro AV812	Ohaus, USA
X Cell SureLock™ Electrophoresis Mini Cell	Invitrogen, Life Technologies, USA
X Cell4 SureLock™ Electrophoresis Midi Cell	Invitrogen, Life Technologies, USA

### 3.16. Software

Product	Developer
Adobe Acrobat	Adobe Systems Inc., San Jose, USA
AIDA Image Analyzer	Raytest, Berlin, GER
BindIt 3.1 Kingfisher Remote Control	Thermo Fisher Scientific, USA
BLAST	NCBI, Bethesda, USA
Endnote X2.0.2	Thomson Reuters Inc., New York City, USA

GraphPad Prism 5	GraphPad Software Inc., La Jolla, USA
Ingenuity Pathway Analyses, IPA®	Ingenuity Systems, Qiagen, USA
Microsoft Office 2007: Word, Excel, Power Point, Access	Microsoft Inc., Redmond, USA
SoftMax Pro 4.8	MindVision Software, Lincoln, USA
TextPad	Helios Software Solutions, Longridge, UK
Vector NTI Advance 11	Invitrogen, Life Technologies, USA

---

---

## 4. Methods

### 4.1. Molecular biology based methods

#### 4.1.1. DNA purification

DNA was extracted and purified by different methods, depending on experimental assay. If required, PCR amplicons were directly purified with the QIAquick PCR Purification Kit.

##### 4.1.1.1. DNA extraction from agarose gel

Prior to ligation, PCR amplicons and cloning vectors were enzymatically digested and fragments were separated by agarose gel electrophoresis, as described in 4.1.4. DNA was visualized under UV light (254 nm) with Gel Doc 2000. DNA bands of interest were cut out from the gel with a clean scalpel using an UV light lamp (365 nm) and the gel piece was weighed. Purification was performed with the Nucleo Spin Extract II Kit according to manufacturer's protocol. Finally, DNA concentration was determined on Nanodrop ND-1000.

##### 4.1.1.2. DNA precipitation with ethanol

Ethanol precipitation was applied after LR recombination or ligation reactions and prior to transformation in *E. coli*, in order to purify DNA material from buffer salts and residual reactants. First, the following reagents were added to the reaction sample:

- 7.5 M ammonium acetate in a 1:10 ratio
- 2 – 3 µl Glyco Blue
- 100% Ethanol (pure grade) in a 2.5:1 ratio

Next, sample was snapped frozen in liquid nitrogen. The blue pellet, containing precipitated DNA, was centrifuged at 12,000 rpm and RT for 10 min. The pellet was washed once again with 20 µl 70% Ethanol (pure grade), dried and finally resuspended in TE buffer in an appropriate volume (usually 5 µl).

##### 4.1.1.3. Plasmid purification

To propagate smaller plasmid amounts, 5 – 6 ml *E. coli* cultures were grown at 180 rpm and 37°C ON. Next day, bacterial cells were pelleted by centrifugation at 4,000 rpm and RT for 30 min. Plasmid extraction was performed with the QIAprep Spin Miniprep Kit according to manufacturer's instructions. Elution volume was usually set at 50 µl. For larger plasmid amounts 100 ml *E. coli* cultures were applied. Purification was performed with the Qiagen Plasmid Maxi Kit according to manufacturer's instructions. Concentration of plasmid DNA was measured on Nanodrop ND-1000.

#### 4.1.2. DNA digestion with restriction enzymes

Restriction of DNA was performed to prepare inserts and vector backbones for ligation, as well as to control cloning results. Standard restriction protocol for a 50 µl reaction:

- 0.5 – 1 µg DNA
- 1 – 5 U restriction enzyme
- 5 µl accordant reaction buffer
- 1:100 BSA (optional)

Reaction sample was thoroughly mixed and incubated at 37°C (or another appropriate temperature, according to the enzyme applied) for at least 2 h, but usually ON. Cleavage efficiency was examined by agarose gel electrophoresis.

#### 4.1.3. Dephosphorylation of digested plasmids

To prevent re-ligation of linearized vectors, dephosphorylation of the 5'-end was performed with the thermosensitive Antarctic Phosphatase, if needed. For this, 1 µl enzyme (5 U) was added to ca. 1 µg digested plasmid, mixed and incubated at 37°C for 10 min. Finally, the enzyme was heat inactivated at 75°C for 5 min.

#### 4.1.4. DNA separation in agarose gel electrophoresis

In order to separate DNA fragments after enzymatic restriction of PCR amplicons and/or cloning vectors prior to ligation, agarose gel electrophoresis was performed. For this, a 1% agarose gel was prepared in TBE buffer (or in TAE buffer to better visualize larger fragments) and Ethidium bromide was added (ca. 1 µg/ml final concentration). DNA samples were mixed with loading dye in a 1:6 ratio and loaded in gel pockets. Electrophoresis was performed in the same buffer at 80 – 120 V for 20 – 40 min. Results were analyzed with Gel Doc 2000 under UV light.

#### 4.1.5. DNA amplification by polymerase chain reaction (PCR)

PCR was applied for analytic as well as preparatory purposes on a thermal cycler. Hot start and heated lid were used for improved amplification efficiency. When error-free amplification was an issue, e.g. insert generation for cloning, the Phusion Hot Start polymerase with proofreading activity was applied. In all other cases, e.g. analytic PCR and sample preparation for NGS, the in-house purified Taq-polymerase was used. Enzymes were added to the reaction samples last, just prior to PCR start. Standard PCR protocol for a 50 µl reaction:

- 1x polymerase reaction buffer
- 4 µl dNTPs (200 µM each)
- 1 µl forward primer (10 pmol)
- 1 µl backward primer (10 pmol)
- 0.5 – 1 µl DNA template (50 – 100 ng)
- 0.5 µl DNA polymerase

Depending on primer characteristics in the accordant PCR, an appropriate annealing temperature was chosen, usually ranging between 50°C – 65°C. Standard PCR program for 30 amplification cycles:

Reaction step	Time	Temperature
1. Initial denaturation (Hot Start)	3 min	95°C
2. Denaturation	1 min	95°C
3. Annealing	1 min	50°C – 65°C
4. Elongation	1 min	72°C
5. Repeat steps 2 – 4	29 cycles	-
6. Final elongation	10 min	72°C
7. Preservation	open end	4°C

#### 4.1.6. DNA ligation

Ligation between a vector backbone and an insert of interest was carried out after enzymatic restriction of both DNA fragments with appropriate endonucleases. The vector amount was set at 50 ng and the insert amount was calculated by using a 1:4 ratio according to the following formula:

$$\text{Insert amount [ng]} = 4 \times \text{Insert size [bp]} \times \frac{\text{Vector amount [ng]}}{\text{Vector size [bp]}}$$

Ligation was usually performed at 16°C ON. Standard ligation protocol for a 10 µl reaction:

- 1 x T4 DNA ligase buffer
- 1 – 2 µl vector DNA
- 1 – 5 µl insert DNA
- 1 µl T4 DNA ligase

All ligation experiments were accompanied by two negative controls: one sample without an insert, to control vector re-ligation and another sample without an enzyme, to control overall reaction background.

#### 4.1.7. Gateway LR recombination

To generate phagemid vectors, the Gateway® Cloning System (Invitrogen, Life Sciences, USA) was applied. In a rapid one-step LR-reaction two inserts, located between special attachment sites in two different plasmids, are recombined, i.e. change places. Thus, no time-consuming restriction/ligation procedures are needed and potential pitfalls, such as undesired restriction sites within the ORF of interest, are not an issue.



**Fig. 9 LR recombination.** Picture source: Gateway Technology Manual, Invitrogen, Life Technologies, May 2010 [219].

Five different destination vectors (pDEST) were generated in this thesis. Two different types of entry clones (pENTR) were applied in different experiments. pENTR/OCAA and pENTR/OCAB, carrying the human full-ORF library, were obtained commercially from Imagenes/Source Biosciences, UK. pENTR/EGFP was previously generated in our research group by Miriam Baradari.

LR reactions were performed according to manufacturer's instructions with each of the five destination vectors and each of the entry clones, or entry clone pools respectively, in equimolar ratios: 50 – 150 fmol DNA. Standard LR reaction protocol for a 10 µl reaction:

- 1 x LR reaction buffer
- 1 – 2 µl pDONR vector
- 1 – 2 µl pENTR clone(s)
- 1 µl LR Clonase mix
- 6 – 4 µl TE buffer

Samples were left at 25°C ON. Next day reaction was stopped with 1 µl Proteinase K at 37°C for 10 min and subsequent enzyme inactivation at 75°C for 5 min. Recombinant DNA was purified via ethanol precipitation. Electro-competent *E. coli* cells of the ccdB-sensitive strain DH10B were finally transformed and plated on 2xYT/Glu agar plates with appropriate antibiotics added. All LR experiments were accompanied by two controls. For a negative control, no LR Clonase was added to the sample. For a positive control of the destination vectors, pENTR/GUS, with an insert for β-Glucoronidase, was applied instead of the other entry clone(s).

#### 4.1.8. DNA Sanger sequencing

Plasmid DNA was sequenced to prove cloning accuracy of newly generated vectors and libraries. For this, 100 – 200 ng DNA and 10 µl of each sequencing primer (10 µM) was given to the in-house sequencing facility (MPI-MG, Berlin) or sent to Eurofins MWG Operon GmbH, Hamburg. Obtained sequencing traces were analyzed with the Contig Express software (VectorNTI, Invitrogen Life Sciences, USA).

#### 4.1.9. Preparation of electro-competent *E. coli* cells

All used *E. coli* strains were transformed by electroporation. To produce electro-competent cells, an ON pre-culture was prepared first: 50 ml of 2xYT medium without additives, inoculated with bacterial cells from a glycerol stock. Next day, 500 ml of pre-heated 2xYT medium was inoculated with the ON pre-culture and grown at 37°C until OD<sub>600</sub> of ~0.5 was reached. Culture was chilled on ice for 30 min and centrifuged at 4,000 rpm and 4°C for 20 min. Bacterial pellets were gently resuspended in 125 ml ice-cold Millipore water and incubated on ice for further 15 min. Another centrifugation step followed, as described above. Now, pellets were resuspended in 50 ml ice-cold Millipore water and kept on ice again for 15 min. After another centrifugation step, pellets were resuspended in 25 ml ice-cold 10% glycerol solution (in Millipore water), kept on ice for 15 min and centrifuged for the last time. Pellets were resuspended in overall 2 ml ice-cold 10% glycerol (in Millipore water). 50 µl aliquots were prepared as shot samples, ready for electroporation, and snapped frozen in liquid nitrogen. Aliquots were stored at -80°C until further use.

Transformation efficiency was determined for each new preparation. For this, one aliquot (50 µl) of electrocompetent cells was transformed by electroporation with 1 ng pUC18 plasmid. 50 µl cell culture were plated on 2xYT/Amp agar plates and incubated at 37°C ON. Next day, cell colonies were counted and transformation efficiency was calculated as [cfu/µg DNA].

#### 4.1.10. Transformation of *E. coli* cells by electroporation

One aliquot of electro-competent *E. coli* cells (50 µl) was ice-thawed and mixed gently with 0.5 - 2 µl plasmid DNA (50 – 100 ng). Incubation on ice for 1 min followed. Next, cells-DNA-mix was transferred to ice-cold electroporation cuvettes (1 mm). Electroporation was immediately performed on a MicroPulser™ Electroporator at 1.5 kV for 5 msec. After that, cell suspension was spilled out with 1 ml SOC medium (without additives) and transferred to a 1.5 ml Eppendorf tube. Culture was incubated at 37°C for 1 h. After regeneration, transformed bacteria were spread on 2xYT/Glu agar plates with the appropriate antibiotics

added, usually in three different dilutions: 50  $\mu$ l, 100  $\mu$ l and residual 900  $\mu$ l (centrifuged down to approx. 100  $\mu$ l volume, prior to plating). Cells were grown at 37°C ON. Negative controls from ligation and/or LR recombination reactions were treated in parallel.

First cloning and transformation efficiencies were verified on next day by colony counting and colony PCR. For detailed analysis, single transformants were picked, grown up to 5 ml ON cultures. Recombinant plasmids were purified and analyzed by PCR, analytical restriction digest and Sanger sequencing.

## **4.2. Protein biochemistry based methods**

### **4.2.1. IPTG-induced protein expression in *E.coli* cells**

Since all used expression vectors contained a *lac* promoter, recombinant expression was induced with IPTG. For this, 100 – 200 ml 2xYT medium with the appropriate antibiotics, but without glucose, was inoculated with the accordant ON culture. Bacteria were grown at 180 rpm and 37°C until  $OD_{600} = 0.6 - 0.8$  and expression was induced with 1 mM IPTG (final concentration). Induced cultures were further agitated at different conditions, depending on experimental procedure: at 16°C or 37°C and 3h or ON.

EGFP was recombinantly expressed as a cytoplasmic protein in BL21 Star *E. coli* cells from a vector without a leader peptide (pRSET/EGFP). Periplasmic EGFP expression was performed in the non-suppressor strain HB2151 from pYG phagemid vectors, containing one of the three leader peptides: pelB, DsbA or TorA.

Human brain cDNA products from the hEX1 library were cytoplasmically expressed from the pQE30NST vector in the SCS1 *E. coli* strain, carrying the helper plasmid pSE111.

Single human full-ORF proteins were available as pENTR223/hOHS clones. For expression, selected clones were cloned via LR recombination into the pRSET-BH6 vector and transformed into BL21 Star *E. coli* cells.

### **4.2.2. Protein extraction under native and denatured conditions**

#### **4.2.2.1. Cytoplasmic protein extraction**

For native protein extraction of cytoplasmically expressed proteins, cell pellet from a 50 ml expression culture was lysed in 3 ml native lysis buffer using 30  $\mu$ l lysozyme (20  $\mu$ g/ml). Proteases inhibitor cocktail in a 1:100 dilution and 1 mM PMSF (inhibitor of serine proteases, final concentration) were added. After 30 min of incubation on ice, suspension was sonicated three times for 20 sec with 20 sec intervals on ice. Lysed cells were centrifuged at 2,200 rpm

and 4°C for 5 min. Cell debris (pellet) was either discarded or resuspended in 8 M Urea and saved for further analysis. Soluble protein fraction (supernatant) was stored at 4°C until use.

For extraction under denatured conditions, cell pellets were resuspended in 500 µl of an 8 M Urea solution. Suspensions were sonicated for 15 min in a water bath. Lysates were centrifuged as described above and soluble fractions were stored for further analysis.

#### **4.2.2.2. Periplasmic protein extraction**

Cells were harvested by centrifugation as described above and pellets were gently resuspended in ice-cold 2 ml TES buffer per 50 ml culture pellet. Next, 3 ml of water-diluted TES buffer (1:5) were added and left on ice for 30 min. After centrifugation at 14,000 rpm and 4°C for 20 min, supernatants with the periplasmic fraction were separated from the spheroblasts in the pellets. Fractions were analyzed on a SDS gel and/or stored at 4°C until further use.

#### **4.2.3. IMAC purification of recombinant His<sub>6</sub>-tagged proteins**

Since all recombinant proteins were His<sub>6</sub>-tagged, immobilized metal ion chromatographic affinity purification with Ni-NTA-Agarose (for EGFP) in batch or with Ni-NTA-columns on an FPLC system (for cDNA products of the hEX1 library and for human full-ORF proteins) was performed.

##### **4.2.3.1. Purification with Ni-NTA-Agarose at gravity flow (batch purification)**

Soluble protein fractions were mixed with 500 µl 50% Ni-NTA-Agarose per 50 ml expression culture. Samples were incubated slowly rotating for one hour at 4°C for binding. Next, suspension was transferred to 5 ml PP columns and flow-through was collected. Protein-bound Ni-NTA-Agarose was washed twice with 4 ml washing buffer. Aliquots of washing flow-through fractions were collected for later analysis. Finally, 1 ml elution buffer was added and mixture was incubated slowly rotating at 4°C for one hour. First 1 ml elution fraction was collected. Elution buffer was added further 2 – 3 times and fractions were collected. Samples were stored at 4°C until use.

##### **4.2.3.1. Purification with Ni-NTA columns on an FPLC system**

All buffers, applied at the ÄKTApurifier FPLC system, were sterile filtered, degassed and kept at 4°C. 1 ml Ni-NTA Superflow cartridge was washed with 15 CV (column volumes) of 0.5 M NaOH solution at a flow rate of 0.5 ml/min and equilibrated with 10 CV of native binding buffer prior to sample loading. Soluble protein fractions were additionally centrifuged to remove larger particles and loaded in a 10 ml loop, connected to the FPLC system and run

over the Ni-NTA column at a flow rate of 0.5 ml/min. After binding, the cartridge was washed with 2 CV washing buffer. Bound protein was eluted in gradient pH (pH 6.5 → pH 4.3) with elution buffer, containing 250 mM Imidazole, at a flow rate of 0.5 ml/min. Elution procedure was monitored by light absorption at 280 nm. All flow-through, washing and elution fractions were collected and stored at 4°C for further analysis.

#### **4.2.4. Concentrating proteins**

To concentrate protein eluates, pooled from the elution peak fractions, two different methods were applied. In case of purified EGFP, pooled eluates were loaded in a dialysis tubing with a cutoff of 3.5 kDa and put on a bed of solid PEG 6000. Liquid volume reduction was monitored and stopped at desired stage. Concentrated protein solution was transferred from the tubing into a PP tube.

Pooled eluates of purified hEX1 cDNA products and full-ORF proteins were concentrated using Amicon® concentration columns with a cutoff of 10 kDa. Protein concentrations were measured on Nanodrop.

#### **4.2.5. Protein separation in SDS PAGE**

Samples of protein fractions were mixed 4:1 with Laemmli loading buffer, boiled at 95°C for 5 min, cooled down and loaded on a self-made 12% SDS gel for general analysis or on a readymade NuPAGE 4% – 12% gradient gel (Invitrogen, Life Sciences), if higher resolution was desired. In the first case Laemmli running buffer was applied. NuPAGE gels were run in 1 x MES buffer. Electrophoresis was performed at 200 V for 45 – 60 min.

#### **4.2.6. Coomassie staining**

For general visualization of overall protein content in the samples, SDS gels were stained with Coomassie brilliant blue R-250 dyeing solution after electrophoresis. For this, gels were incubated in Coomassie for 20 – 30 min under continuous linear shaking. Destaining followed by exchanging Coomassie with a destain solution until protein bands became visible and gel background pale enough. Gels were documented with a CCD camera and stored in water, if required.

#### **4.2.6. Silver staining**

For more sensitive staining of overall protein content in the samples, SDS gels were stained with the Pierce Silver Stain Kit according to manufacturer's instructions and documented with a CCD camera.

#### 4.2.8. Western blot

For immuno-detection, proteins, separated in an SDS-PAGE, were transferred onto a Hybond ECL nitrocellulose membrane in a wet blotting procedure. SDS gel and membrane were put together and arranged between two sheets of Whatmann paper, soaked in transfer buffer, in a blotting clam. Clam was placed in the blotting chamber in a way to direct proteins from gel to membrane and from anode to cathode, accordingly. Transfer was performed at 400 mA for 45 – 60 min. After transfer, membranes were blocked in PTM buffer at 4°C for 2 – 3 h or ON by slowly rocking. Next, membrane was washed twice in PBS-T. Detection was performed with a single HRP-conjugated antibody only or a primary unconjugated and a following secondary HRP-conjugated antibody. All antibodies were diluted in PTM according to supplier's recommendations and incubated with the membrane for one hour by slowly rocking. Membrane was washed twice in PBS after each incubation step. Final detection was performed with the CN/DAB Substrate Kit according to manufacturer's instructions. Color development was stopped with water at desired stage and membrane was scanned for documentation.

#### 4.2.9. Protein ELISA

In all steps 100 µl solution per well was applied. 96-well Maxisorp MTPs (flat polystyrene bottom, clear) were coated with 1 µg antigen per well. Coating was performed at 37°C for 1 h. Washing procedure followed with 1 x PBS-T and 2 x PBS. Next, plate was blocked with 2% PTM and/or with 0.2 mg/ml BSA at 4°C for 2 h or ON. Human sera were diluted 1:100 in PBS. Prior to application, diluted sera were pre-incubated with a supernatant fraction from non-recombinant *E. coli* cell lysates (approx. 450 µg/ml overall protein amount) for 1 h. Since recombinant expression of antigens was carried out in a bacterial host, this pre-incubation step was necessary to block  $\alpha$ -*E. coli* sera antibodies, which were still present in the purified protein samples and could potentially falsify ELISA results. Hybridization was performed for one hour at RT and at slowly linear rocking. Another washing step followed, as described above. Secondary antibodies applied, were  $\alpha$ -hu-IgG-HRP and  $\alpha$ -hu-IgA-HRP, diluted in PTM according to manufacturer's instructions. Incubation was performed for one hour at RT. Plate was washed one last time. ABTS-substrate-solution was prepared according to manufacturer's instructions: one ABTS-tablet was solved in overall 20 ml mix of 50 mM citric acid and 50 mM sodium citrate and 10 µl H<sub>2</sub>O<sub>2</sub> were added. Solution was immediately pipetted to the wells and color development was measured at 405 nm on ELISA reader. Further measurements were taken at 5, 10, 15, 30, 45, 60 and 90 min after adding the substrate. Meanwhile plate was kept in darkness to minimize substrate oxidation by light. Data was analyzed with the SoftMax Pro 4.8 software.

#### 4.2.10. Determination of total immunoglobulin titers in blood sera

Total titers of all immunoglobulin classes were determined in all 60 human blood sera. For this, ELISA kits, based on immunoperoxidase assays, were used (Dunn Labortechnik GmbH, Asbach, GER). Sera samples were diluted as follows: 1:80,000 for IgG ELISA; 1:10,000 for IgA and IgM ELISA; 1:2,000 for IgD ELISA and 1:200 for IgE ELISA. Assays were performed according to manufacturer's instructions.

### 4.3. Protein macroarray technology

#### 4.3.1. Protein macroarrays hybridization with human sera

Dried protein filters with a spotted expression library were first wetted in 70% ethanol (technical grade) for 10 – 15 min, washed in ddH<sub>2</sub>O to remove alcohol and transferred into TBST buffer. Lyophilized bacterial colonies were cautiously wiped off from filter surface with Kimtech wipes under mild pressure. Filters were washed in 25 ml TBST by slowly rocking until no *E. coli* debris were visible and solution remained clear. Next, filters were blocked in 50 ml blocking buffer (BP) for 3 h.

Five human sera from each of the three cohorts were screened on protein macroarrays in duplicate, which is identical serum under same conditions on two different filters.

**Table 5. Human sera, screened on protein macroarrays.** Lot numbers are shown. Sera No. 1 and 2 were screened in a phage display selection as well (s. Table 6).

No.	AD	MS	Healthy
1	191492	170171	151010
2	191494	170175	151014
3	191491	170172	151002
4	191493	170176	151006
5	191724	170178	151015

All sera were diluted 1:100 in 25 ml BP. Each serum dilution was incubated with two filters, back to back, slow rocking at 4°C ON. Next day, serum dilutions were collected in 50 ml falcon tubes and frozen away at -20°C for possible reapplication. Hybridized filters were washed 2 times in 25 ml TBST at RT for 15 min. Incubation with the detection antibody,  $\alpha$ -hu-IgG-AP or  $\alpha$ -hu-IgA-AP (1:1,000 dilution in BP), followed slow rocking at RT for 1 h. Next, filters were washed twice in TBST, once in TBS and transferred into AttoPhos buffer. 2.5 ml AttoPhos substrate (Roche) was diluted in 50 ml AttoPhos buffer and kept in darkness to avoid substrate oxidation by light. Each filter was incubated shortly in 7 ml AttoPhos substrate solution, so that smooth distribution over the whole filter surface was visible. Signal

detection was performed with the CCD camera LAS-1000 (Fujifilm), using a yellow filter. Digital images were saved in a 16-bit grayscale mode.

#### **4.3.2. Data analysis with AIDA Image Analyzer**

The software package AIDA Image Analyzer (Raytest) was used for array analysis. For this, a new grid template was constructed in collaboration with the software provider (Melanie Busse, Imagenes/Source BioScience). Next, clone annotations were imported in an appropriate format and linked to the grid template to match the 5 x 5 spotting pattern around the guiding ink spot. This allowed a rapid identification of positive hits by simply clicking on the white spots, in contrast to the background clones, which remained grey. A hit was considered positive only if both spots in the correct double pattern were visibly white and possessed a comparable color intensity. Background was defined as 10 different dots with the lowest intensities on the current filter and deviated from measured color intensities by the software. Filter evaluations were exported as result tables in an MS Excel format and saved. To simplify further quantitative analysis, intensities, determined by AIDA, were split in three categories: 1 for weak signals (lowest 33%), 2 for medium signals and 3 for strong signals (upper 33%).

### **4.4. Filamentous phage display based methods**

#### **4.4.1. Preparation of M13K07 helper phage**

1:10 dilution series from a M13K07 Helper phage stock (NEB) were prepared first. To 4 ml melt top agar 200  $\mu$ l of TG1 ( in the log growth phase, i.e.  $OD_{600} \sim 0.6$ ) culture and 100  $\mu$ l of phage dilution ( $10^{-7} - 10^{-11}$ ) were added, mixed and distributed evenly on 2xYT agar plates. After incubation at 37°C ON, phage plaques in the bacterial lawn became visible. Phage material from multiple single plaques was picked, inoculated in 4 ml liquid 2YT/Kan medium and grown at 30°C ON. Next day, cells were harvested by centrifugation at 4,000 rpm. For phage precipitation, 20% PEG 6 000/2,5 M NaCl was added to the supernatant in a 1:6 ratio and incubated on ice for 1 h. After a 30 min centrifugation, white phage pellet became visible. Supernatant was discarded and the pellet was left to dry prior resuspension in PBS. Phage suspension was centrifuged 5 – 7 times at 12,000 rpm to remove residual bacteria until supernatant became totally clear. Finally, phage concentration was determined by titration.

#### **4.4.2. Phage titer determination**

In order to determine phage titers, 1:10 dilution series of the newly generated phage suspension were prepared, mixed 1:1 with a fresh TG1 cell culture and incubated at RT for 30 min for infection. 10  $\mu$ l of each dilution step were then dropped on a 2xYT/Kan agar plate,

left to dry and incubated at 37°C ON. Next day, cell colonies were counted. Taking dilution factors into account, phage titers were finally calculated as [pfu/ml]. To analyze phage suspension purity and bacterial background, negative controls were also included. For this, 10 µl uninfected cells, as well as pure PBS, were dropped in parallel and colony growth was monitored on next day.

Since infection with Hyperphage leads to poor infectivity of the recombinant phages, the usual titration assay could not be used in the case of EGFP-presenting phages. Instead, phage titer was determined using a phage-capture-ELISA. 96-well ELISA Maxisorp plates (NUNC) were coated with 10 µg/ml unconjugated mouse α-M13 mAb (Amersham) at RT for 1 h and then blocked with 2% PTM for another hour. Three washing steps with PBST followed. Meanwhile 1:10 dilution series of phage suspension were prepared. As a reference, a non-recombinant M13KO7 helper phage suspension with known titer was used in parallel. Phages were allowed to bind for 1 h, whereupon the plate was washed three more times in PBST. Next, second antibody was applied: HRP-conjugated mouse α-M13 mAb (GE Healthcare) in a 1:5,000 dilution. After an hour of incubation the plate was washed three times and ABTS-substrate (prepared according to the manufacturer instructions, Sigma) was applied. Color development was measured on an ELISA reader at a wavelength of 405 nm: 15, 30, 45 and 60 min after applying the substrate. Data was analyzed with the SoftMax Pro 4.8 software.

#### **4.4.3. Preparation of recombinant M13 phages**

LR reactions were performed with each of the five pYG destination vectors and each of the four full-ORF-containing pENTR-vector pools or pENTR/EGFP respectively in equimolar ratios (50 – 150 fmol of DNA), as described above. Electro-competent cells of the ccdB-sensitive *E. coli* strain DH10B were transformed with 80 – 100 ng recombinant plasmid DNA using the Micropulser Electroporator and incubated in 1 ml SOC medium at 37°C for 1 h.

Meanwhile, large NUNC-plates (24 x 24 cm) were prepared. After cleaning up with 70% ethanol (technical grade), 30 min sterilization under UV light followed. Sterile plates were filled with ~100 ml 2xYT/Glu/Amp agar per plate and left to dry under the Microflow Workstation. 1 – 2 ml cell suspension per one NUNC-plate were spread evenly using small glass beads and incubated at 37°C ON. Next day, plates were examined for absence of undesirable satellite colonies. Further, transformants were counted and fold coverages of the source libraries were estimated. Colonies were then scraped off with a lazy spreader into 100 ml pre-warmed 2xYT/Glu/Amp medium. After a gentle agitation at 37°C for 1 h, cells were harvested by centrifugation and plasmids were rescued from the cell pellets using Qiagen® Plasmid Maxi Kit (Qiagen).

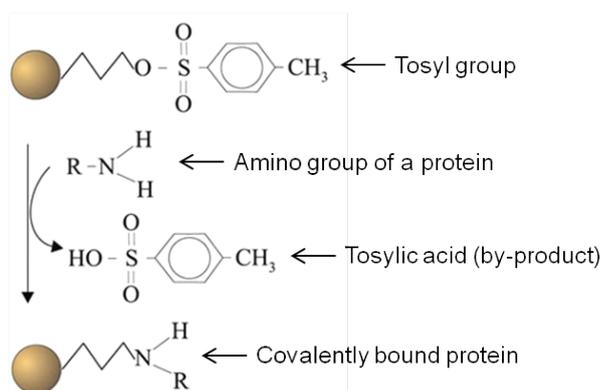
Next, plasmids were re-transformed into the *E. coli* strains XL1Blue or TG1, suitable for phage display, under the same conditions as described above. This time 100 µl cell suspension were plated per one NUNC-plate, hence 10 plates per library were generated. After ON incubation, transformation efficiencies and fold coverages were again estimated: approx.  $10^5$  colonies per plate, if the plate was covered by almost a continuous bacterial lawn, hence  $10^6$  per transformed library. Colonies from one library (10 plates) were scraped off with lazy spreader into 500 ml pre-warmed 2YT/Glu/Amp medium and agitated for 45 min. Cells were centrifuged at 4,000 rpm for 30 min and cell pellets resuspended in a total of 50 ml 2YT/Glu medium with 15% Glycerol added. 500 µl cryo vial aliquots were prepared and glycerol stocks were stored at  $-70^\circ\text{C}$  until further use.

To generate recombinant full-ORF/M13 phages, one glycerol stock aliquot of phagemid containing XL1Blue or TG1 cells was first resuspended in 200 ml pre-warmed 2YT/Glu/Amp medium and agitated at 180 rpm and  $37^\circ\text{C}$  until an  $\text{OD}_{600}$  of 0.4 – 0.6 was reached. Cell suspensions were distributed in four 50 ml falcon tubes. To each tube  $2,5 \times 10^{11}$  pfu M13K07 helper phage was added, mixed by inverting the tube and incubated without shaking at  $37^\circ\text{C}$  for 30 min. Next, cells were harvested at 4,000 rpm for 15 min and resuspended in overall 200 ml pre-warmed 2YT/Amp/Kan medium, containing only 0.1% glucose. Infected cells were grown at 300 rpm and  $37^\circ\text{C}$  ON. Next day, cells were harvested at 4,000 rpm for 30 min. Phages from the supernatant were prepared and titrated as described above and stored at  $4^\circ\text{C}$  until further use.

Recombinant EGFP/M13 phages were generated in the same way with the following difference: 50 ml cell culture was infected with  $10^{10}$  phages of the Hyperphage M13K07 $\Delta$ pIII (Progen Biotechnik).

#### **4.4.4. Loading of tosyl-activated magnetic beads with human autoantibodies**

Tosylactivated Dynabeads<sup>®</sup> MyOne<sup>™</sup> (Invitrogen, Life Technologies) were treated according to manufacturer's instructions. Coupling principle is demonstrated in Fig. 10 below.



**Fig. 10 Coupling principle of tosylactivated magnetic beads.** Hydroxy groups of superparamagnetic polystyrene beads, coated with a polyurethane layer, are activated by reaction with p-toluensulphonyl chloride to produce tosyl groups on bead surface. In the course of bead coupling the sulphonyl ester reacts with the amino group of the target ligand (e.g. antibody). As a result, the tosyl group is replaced with the ligand and tosylic acid is released as a by-product. Hence, target proteins are chemically and physically bound to the magnetic beads surface. Covalent bond formation is accelerated by higher temperatures and alkaline pH conditions. Picture source: modified from Dynabeads<sup>®</sup> MyOne<sup>™</sup> Manual, Invitrogen Dynal, Life Technologies, 2006 [220]

37 mg magnetic beads (375  $\mu$ l) were transferred to a 1.5 ml tube and placed on a magnet rack. After supernatant became clear, it was carefully withdrawn without disturbing the beads pellet. Beads were resuspended in 750  $\mu$ l coating buffer by vortexing. Initial washing step was repeated three more times. Beads were resuspended in a final volume of 150  $\mu$ l.

According to manufacturer's instructions, 25 mg beads can optimally bind 1 mg antibody. Hence, 37 mg beads in 150  $\mu$ l coating buffer were mixed with 1.5 mg pure primary antibody ( $\alpha$ -hu-IgG;  $\alpha$ -hu-IgA or  $\alpha$ -mouse-IgG). For a total coating volume of 936  $\mu$ l following standard protocol was applied:

- 150  $\mu$ l magnetic beads (37 mg)
- 1.5 mg pure primary antibody
- 312  $\mu$ l  $(\text{NH}_4)_2\text{SO}_4$  (1 M final concentration)
- fill up to 936  $\mu$ l with coating buffer and mix by inverting

Coating was performed with slow tilt rotation in a 37°C-room for 24 h. After incubation, tube was placed on a magnet rack and supernatant was withdrawn and stored at 4°C for possible reapplication and WB analysis. Next, coated beads were incubated in blocking buffer, again with slow tilt rotation in a 37°C-room ON. Next day, beads were washed three times in washing buffer and finally resuspended in 750  $\mu$ l ( $C_{\text{beads}} = 50 \mu\text{g}/\mu\text{l}$ ). Suspension was stored at 4°C until further use or loaded immediately with the secondary antibody.

Secondary coating was performed just prior to bio-panning procedure. 1 mg coated beads would optimally capture 1 – 20 µg serum antibody (manufacturer's instructions). Accordingly, 50 µl coated beads (2.5 mg) were placed on the magnet rack and supernatant was withdrawn. Coated beads pellet was directly resuspended in 500 µl human serum, diluted 1:100 in PBS ( $C_{\text{beads}} = 5 \mu\text{g}/\mu\text{l}$ ). Suspension was incubated with slow tilt rotation at RT for 1 h. Next, serum supernatant was withdrawn and stored at 4°C for further use and analysis. Loaded beads were washed three times and resuspended finally in 500 µl sterile PBS ( $C_{\text{beads}} = 5 \mu\text{g}/\mu\text{l}$  or  $5 \times 10^6$  loaded bead particles per µl). Beads were now loaded with human serum auto-antibodies and ready to apply in phage display selection. Human sera were used for at least two coating rounds. Thus, same serum aliquot was applied once with α-hu-IgG-coated beads and once again with α-hu-IgA-coated beads. Cross-reactivities of the commercially obtained α-human antibodies were tested previously.

For positive controls, beads were coated with primary α-mouse-IgG antibody, as described above. Secondary antibodies applied were α-GAPDH, α-Tubulin and α-Ubiquitin. All three antibodies were monoclonal and produced in mouse. 50 µl coated beads (2.5 mg) were resuspended in 500 µl with 1 µl secondary antibody (1:500 dilution).

Coating efficiency was verified in a WB procedure after each coating step. For this, coated and/or loaded beads (2 – 3 µl) were mixed with SDS-loading buffer, boiled for 5 min and pipetted directly into the SDS-gel pockets. For comparison, supernatants from the accordant coating steps were loaded on the gel as well.

#### **4.4.5. Bio-panning procedures**

Prior to selection, full-ORF phage libraries were pre-absorbed with magnetic beads, coated only with the primary antibodies α-hu-IgG and α-hu-IgA. For this, 1.5 ml phage library suspension were mixed with 1.5 ml 2 x PTM, containing 2% BSA. 10 µl accordant beads were added and incubation was performed with slow tilt rotation at RT for 1 h. Finally, beads were withdrawn with a magnet and pre-absorbed phage supernatant was applied in the following bio-panning procedure.

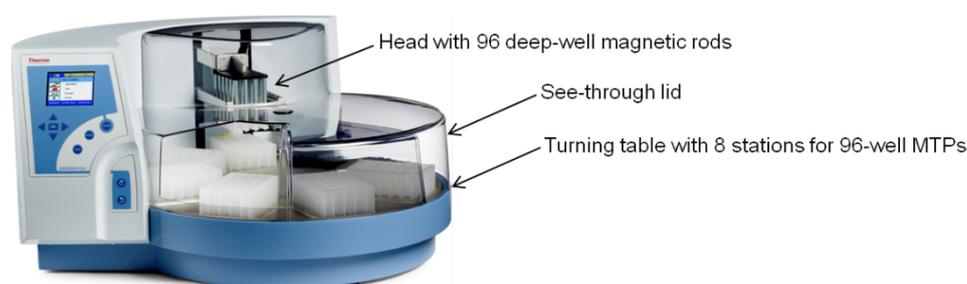
##### **4.4.5.1. Semi-automated selection on a magnetic particle processor**

Four selection rounds were performed with each of the five phage libraries against two autoantibody classes (IgG and IgA), captured from two blood sera from each donor cohort (AD, MS and healthy), respectively. All sera, shown in Table 6 below, were screened on protein macroarrays as well (s. Table 5).

**Table 6. Human sera, screened with phage display.** Serum lot numbers are shown.

No.	AD	MS	Healthy
1	191492	170171	151010
2	191494	170175	151014

Bio-panning was conducted in a semi-automated manner on the Kingfisher Flex magnetic particle processor, depicted in Fig. 11 below.



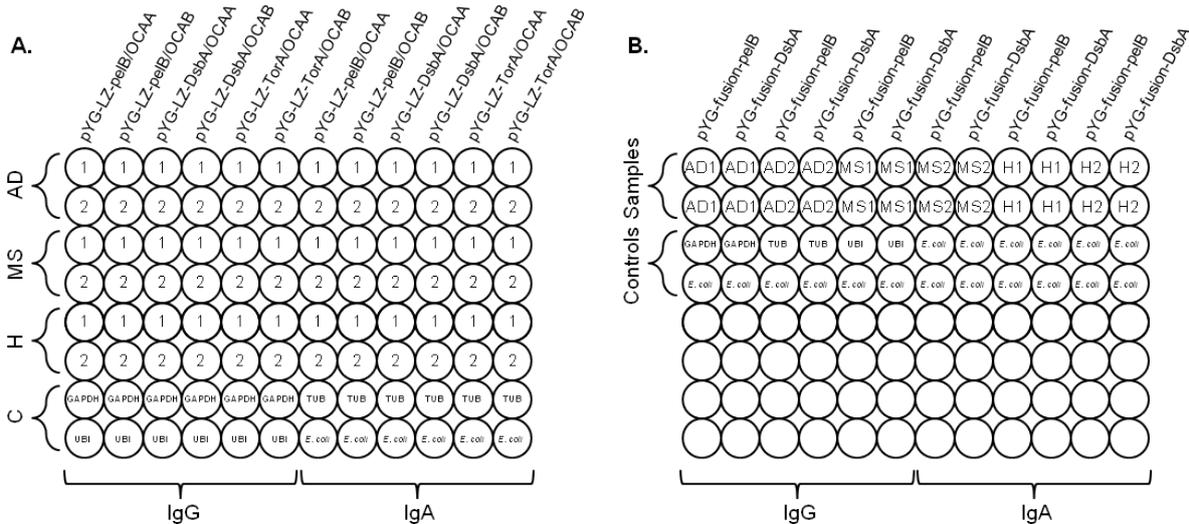
**Fig. 11 KingFisher Flex magnetic particle processor.** Picture source: modified from Thermo Scientific KingFisher Flex User Manual, Thermo Scientific, 2010 [221].

First, 96-well PP NUNC-plates with the interacting partners were prepared according to the pipetting schema in Fig. 12. MTP well volume was set at 200  $\mu$ l. Plates were placed in the accordant station positions on the turning table of the KingFisher processor and a program for the 1<sup>st</sup> selection round was started.

**Table 7. Program protocol for the 1<sup>st</sup> selection round on the KingFisher magnetic particle processor.**

Station	Plate	Charging	Protocol step	Time [min]
1	Bead plate	Beads, loaded with human sera autoantibodies in 1 x PTM	Blocking beads	60
2	Phage plate	Pre-absorbed full-ORF/pYG phage libraries in 1 x PTM	Beads-phages interaction	60
3	Wash plate	PBST	Washing beads with bound phages	10
4	Release plate	PBS	Waiting position	max. 120
5	<i>E. coli</i> plate	XL1 Blue cell culture with OD <sub>600</sub> = 0.5 – 0.7	Infection with selected phages	60

With each subsequent selection round, one additional washing step in PBST was included. *E. coli* plate had to be prepared during antecedent steps in a way to match temporally the required growth stage of the bacteria with the release step of the beads-bound phages. Also positive and negative controls were included to monitor selection procedure. Fig. 12 below shows the resulting interaction schema, upon which preparation of plates was based.



**Fig. 12 MTP loading schema during selection on the KingFisher particle processor.** A circle represents one MTP well. AD1 is AD serum No. 1, which is 191491, and so forth (s. Table 5). C = control samples. Positive controls were beads loaded with: mouse  $\alpha$ -GAPDH (Glyceraldehyde 3-phosphate dehydrogenase), mouse  $\alpha$ -TUB (Tubulin) or mouse  $\alpha$ -UBI (Ubiquitin) antibodies. Negative controls were pure *E. coli* cultures without any beads or phages. **A. Bio-panning with pYG-LZ phage libraries.** Each full-ORF/pYG-LZ phage library (upper identifier) contained both clones with and without stop codons. Example: in well A1 (upper left corner) pYG-LZ-pelB phages, presenting the full-ORF library OCAA (clones with and w/o stops), encountered magnetic beads, loaded with IgG autoantibodies from the AD1 (hence 191491) human serum. **B. Bio-panning with pYG-fusion phage libraries.** In these vector constructs only clones without C-terminal stop codons were applied. Hence, each full-ORF/pYG-fusion phage library contained clones w/o stops from both OCAA and OCAB. Example: in well A1 (upper left corner) pYG-fusion-pelB phages, presenting the full-ORF libraries OCAA and OCAB (clones w/o stops), encountered magnetic beads, loaded with IgG autoantibodies from the AD1 (i.e. 191491) human serum.

After 60 min incubation of beads-bound phages in bacterial culture for infection, the *E. coli* plate was removed from the KingFisher robot. Next steps were performed manually. 20  $\mu$ l of 10 x Amp/Glu solution per well were added and plates were incubated in a plate thermoshaker at 37°C for 2.5 h. Next, 200  $\mu$ l of pre-warmed 2xYT/Amp/Glu medium per well was added and mixed. Half of the volume, hence 200  $\mu$ l, were transferred to a new MTP and incubated in a plate thermoshaker (back up culture MTP) at 30°C ON. Original MTP with the residual 200  $\mu$ l per well was infected with M13K07 helper phage to further propagate selected recombinant phages. For this,  $10^{10}$  cfu M13K07 per well were added and incubated without shaking at RT for 30 min. Next, plate was centrifuged at 2,000 rpm for 3 min. Supernatants were withdrawn and the cell/bead pellets were resuspended in 220  $\mu$ l fresh 2xYT/Amp/Kan medium, containing only 0.1% glucose, per well. Finally, culture plate was incubated in a plate thermoshaker at 30°C ON.

Next day, all 200  $\mu\text{l}$  from the back up MTP were transferred to PCR stripes. 20  $\mu\text{l}$  10 x HMFM solution were added per tube and stripes were stored at  $-70^{\circ}\text{C}$  until needed. On the other hand, ON culture plate was centrifuged at 2,000 rpm for 3 min. Supernatants, now containing propagated recombinant phages from the 1<sup>st</sup> round, were processed further. 95  $\mu\text{l}$  were transferred to another MTP for later ELISA analysis and stored at  $4^{\circ}\text{C}$  until use. 10  $\mu\text{l}$  were transferred to PCR-tubes for later titration analysis and also stored at  $4^{\circ}\text{C}$ . Remaining 90 – 100  $\mu\text{l}$  were transferred to a new NUNC plate and mixed with 100  $\mu\text{l}$  2 x PTM, now ready for the 2<sup>nd</sup> selection round (Phage plate). Prior to bio-panning start, loaded beads were rotated slowly for one hour at RT for resuspension and equilibration. 2<sup>nd</sup> selection round was performed in analogy to the 1<sup>st</sup> round, yet with one additional washing step in PBST. 3<sup>rd</sup> and 4<sup>th</sup> selection rounds were performed subsequently in the following days according to the described protocol gradually increasing rounds of washing.

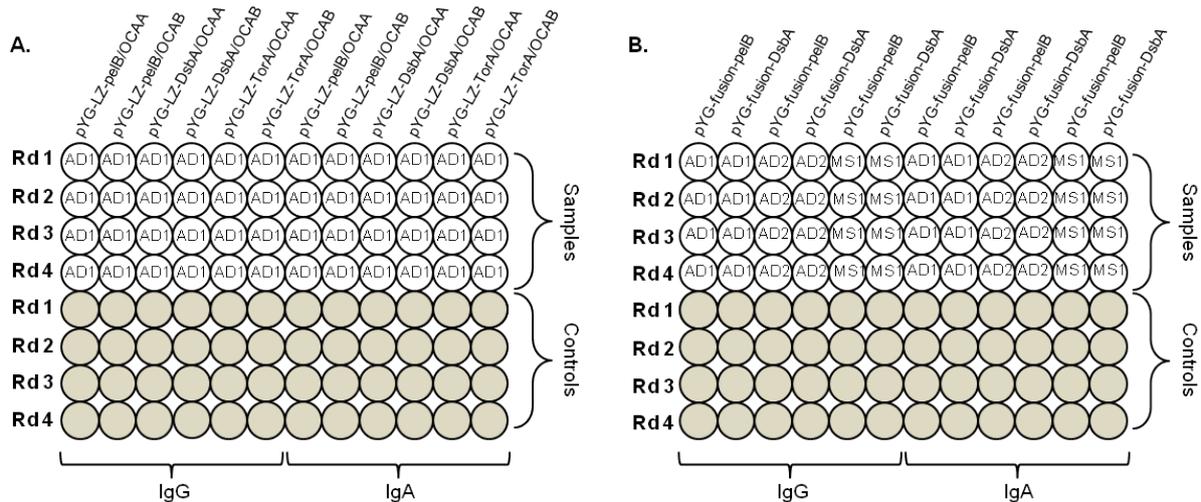
#### 4.4.5.2. Polyclonal phage ELISA

Polyclonal phage ELISAs were performed to monitor enrichment. To prevent bacterial contamination and/or proteolytic degradation, the assay was usually carried out on next day after the accordant selection round. Wells were filled according to the schema, depicted in Fig. 13. Again, MTP well volume was set at 200  $\mu\text{l}$ . Plates were placed in the accordant station positions on the turning table of the KingFisher processor and an ELISA program was started.

**Table 8. Program protocol for ELISA on the KingFisher magnetic particle processor.**

Station	Plate	Charging	Protocol step	Time [min]
1	Bead plate	Beads, loaded with human sera autoantibodies (rows A – D) or only with $\alpha$ -human antibodies (rows E – H) in 1 x PTM	Blocking beads	60
2	Phage plate	Enriched phages from the accordant selection round	Beads-phages interaction	60
3	Wash plate 1	PBST	Washing beads with bound phages	10
4	Wash plate 2	PBST	Washing beads with bound phages	10
5	Antibody plate	$\alpha$ -M13-HRP antibody in PTM (1:5 000)	Binding phage detection antibody	60
6	ABTS plate	ABTS substrate solution	Releasing beads in substrate	30

A MaxiSorp ELISA plate (ABTS plate) was prepared while steps 1 – 4 were run. First it was blocked with PTM for 2 – 3 h and then washed twice in PBST and once in PBS. ABTS substrate was prepared according to manufacturer's instructions and plate was filled with 200  $\mu\text{l}$  per well. ABTS plate was finally placed on its position in the KingFisher robot just shortly before Step 6 was started. Color development lasted for 30 min. Finally, beads were withdrawn from solution and measurements were performed on an ELISA reader at 405 nm.



**Fig. 13 MTP loading schema for ELISA on the KingFisher particle processor.** A circle represents one MTP well. AD1 is AD serum No. 1, which is 191491, and so forth (s. Table 5). Rd 1 is 1<sup>st</sup> round of selection. Samples (rows A – D) are identical to the accordant ones during bio-panning. Negative controls (rows E – H, shaded) are same phage libraries, as in samples, yet with beads coated only with  $\alpha$ -human antibodies (no serum autoantibodies). Equal beads-concentration in samples and controls wells. **A. ELISA with enriched pYG-LZ phage libraries.** One plate sufficed for one serum sample only. Thus, an example with AD1 is shown. Overall 6 ELISAs were finally performed to cover all samples. **B. ELISA with enriched pYG-fusion phage libraries.** One plate sufficed for three sera samples. An example with AD1, AD2 and MS1 is shown. Consequently, overall two ELISAs were performed to cover all samples.

To evaluate enrichment in the three positive controls ( $\alpha$ -GAPDH,  $\alpha$ -Tubulin,  $\alpha$ -Ubiquitin coated beads), ELISA plates were directly coated with  $\alpha$ -GAPDH,  $\alpha$ -Tubulin or  $\alpha$ -Ubiquitin antibodies respectively (1  $\mu$ g per well). Accordant phage samples from each selection round were added. Detection was performed with  $\alpha$ -M13-HRP antibody. To determine real ELISA data, control-values (wash buffer instead of GAPDH-phage samples) were subtracted from sample-values, respectively. Measurements were usually carried out in duplicate.

#### 4.5. Next generation sequencing on Illumina Genome Analyzer

Recombinant phage DNA from all initial pYG libraries, as well as enriched phages from the 2<sup>nd</sup> and 4<sup>th</sup> selection rounds, were deep sequenced (NGS) on Illumina Genome Analyzer.

##### 4.5.1. Preparation of phage-derived full-ORF gene inserts for sequencing

Sequencing samples were prepared using a standard open PCR protocol. For this,  $10^7$  recombinant phages from each library were used as a template in a 100  $\mu$ l reaction. Amplification was performed in 30 cycles and at an annealing temperature of 55°C in a

thermal cycler with a Taq DNA polymerase. Used oligonucleotide pairs were: [LMB3 x pHEN\_seq] for the pYG-fusion vectors and [Fos\_seq x pJuFo-3'] for the pYG-LZ vectors.

Amplicons were agarose gel purified using Qiaquick<sup>®</sup> Gel Extraction Kit. To eliminate dispensable backbone sequences, purified amplicons were digested with *BsrGI* and gel purified anew. *BsrGI* restriction sites are flanking closely the full-ORF inserts in all pYG vector constructs. Finally, DNA concentrations were measured on Nanodrop and samples were sent for further processing and sequencing to Source BioScience (Nottingham, UK).

#### **4.5.2. Processing of sequencing samples and applied NGS protocols**

Samples were processed, analyzed and finally sequenced at Source BioScience (Nottingham, UK) on the Illumina Genome Analyzer. First, DNA quality and quantity have been supervised on a MCE<sup>®</sup>-202 MultNA Microchip Electrophoresis System (Schimadzu). Next, inserts have been shredded into 75 bp long fragments, adapters were ligated and finally, samples were loaded on an Illumina chip. Sequencing has been performed according to Nextera or Illumina Truseq protocols.

#### **4.5.2. Bioinformatical raw data processing**

Sequencing raw data was processed in collaboration with Dr. Steffen Hennig (Imagenes GmbH, Berlin). For hits assignment, results have been mapped to the sequences of the source OC libraries. Normalization was performed on target sequences length level, hence coverages per hit gene have been determined as well.

#### **4.6. Statistical analysis of final ELISA results**

Results from the final ELISA experiments with autoantigens were analyzed statistically with the non-parametric Wilcoxon-Mann-Whitney test applying GraphPad Prism software. This test is a modification of the usual t-test and more specific for non-normal, thus non-Gaussian distributions. Signal distribution groups (20 sera per group respectively) were analyzed pairwise and separately for each of the two antibody classes. The interval of confidence was set at 95% and two-sided P-values were estimated. Hence, P-values under 0.05 were considered significant. P-values of less than 0.005 were considered as especially significant.

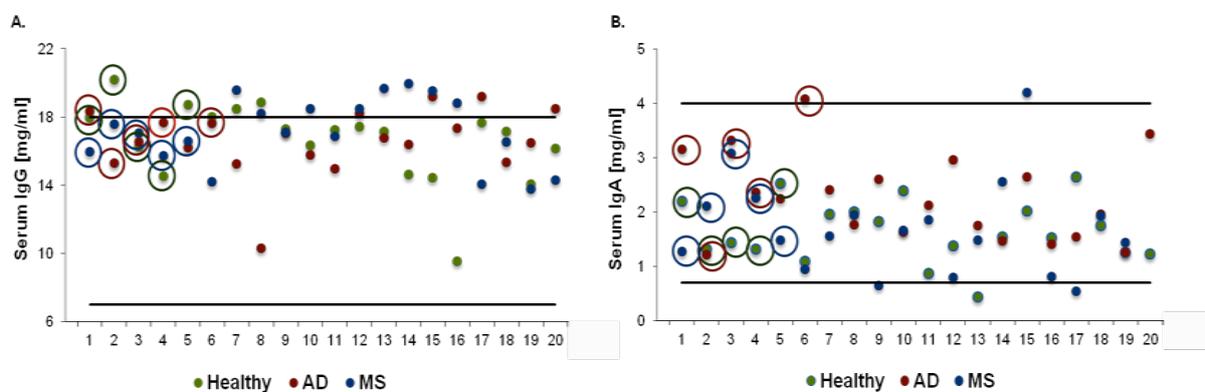
Another non-parametric statistical test applied, was the Kruskal-Wallis test, extended with the so called Dunn correction. Also this test is specific for data distributions of a non-Gaussian type. Thus, it is a modification of its parametric equivalent ANOVA one-way analysis of variance, which is usually combined with the Bonferroni correction, and an extension to Wilcoxon-Mann-Whitney, since it compares multiple data groups simultaneously. Both

corrections are applied to control the so called familywise error rate, which often occurs in statistical tests with multiple hypotheses. The interval of confidence was set again at 95%. The Kruskal-Wallis test was used to reconfirm results, previously obtained with the Wilcoxon-Mann-Whitney test.

## 5. Results

### 5.1. Determination of total immunoglobulin titers in blood sera

Prior to all screenings, total immunoglobulin concentrations were measured in all 60 blood sera. This was made to test, if there are any notable differences between the different cohorts. Another reason for this analysis was to ascertain that IgG and IgA titer in the 15 sera, screened in this study, were in the normal statistical range for the accordant antibody class. Measurements were performed with commercially attained ELISA Ig kits, as described in section 4.2.10. Scatter plots in Fig. 14 below show achieved results for IgG and IgA.

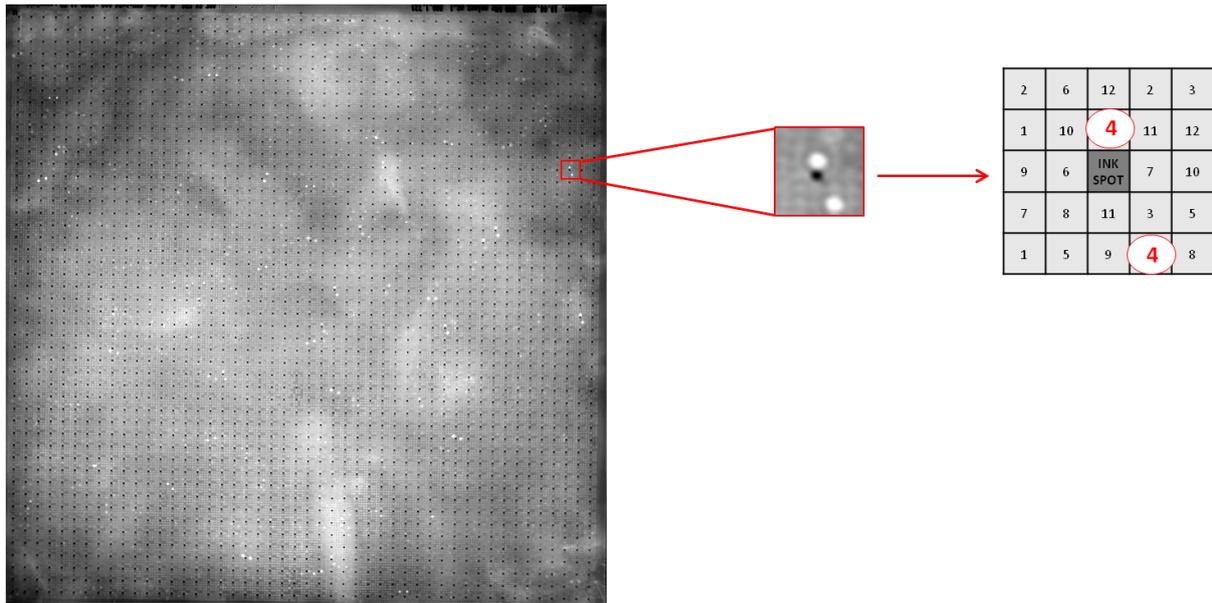


**Fig. 14 Total IgG and IgA titer in donors' blood sera.** X-axes represent cumulative number of blood sera. Y-axes show determined immunoglobulin concentrations in [mg/ml]. Encircled dots represent the five sera from each of the three cohorts, which were screened in this study. Lines in black show the common range of values for the accordant Ig titer in human blood sera. **A. Serum IgG titer.** Common range for IgG concentration is between 7 – 18 mg/ml. **B. Serum IgA titer.** Common range for IgA concentration is between 0.7 – 4 mg/ml

Serum titers of the IgM, IgD and IgE classes were determined as well (data not shown). Analysis showed no notable variations between cohorts. Thus, titers of all five classes and in all samples were in the common range for the accordant antibody or aberrated only slightly.

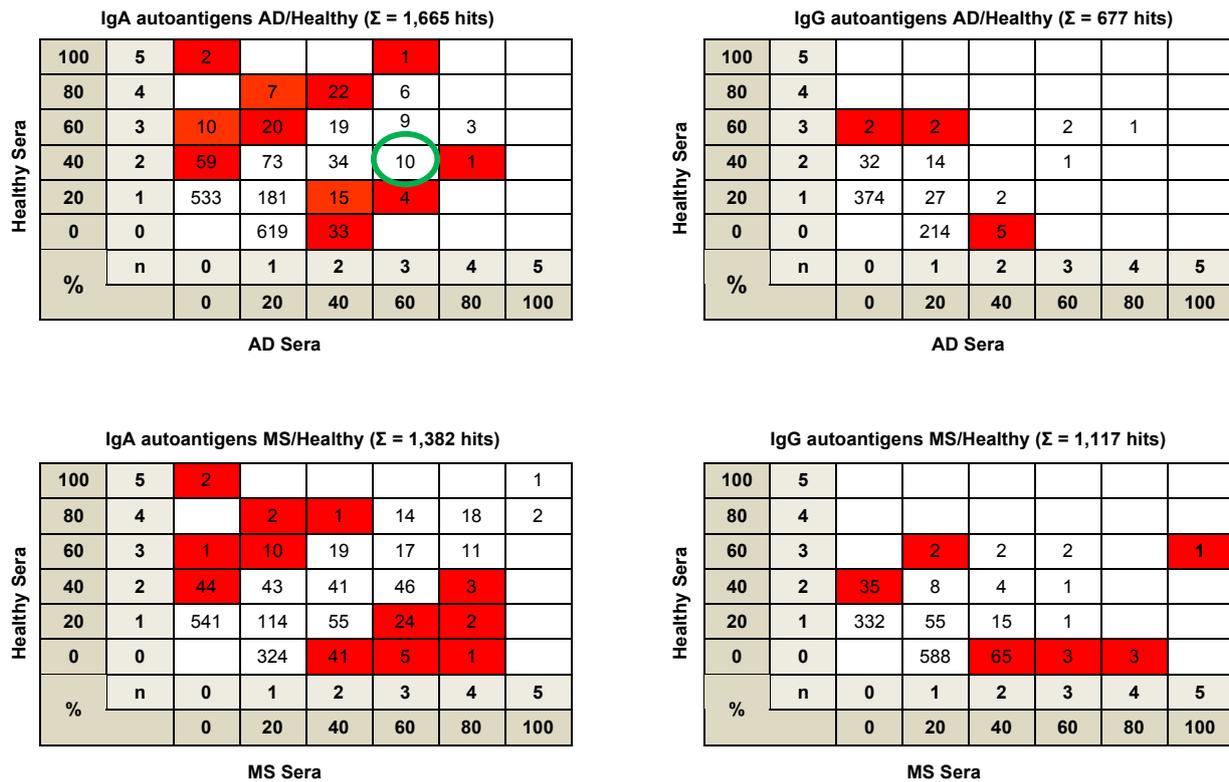
### 5.2. High-throughput autoantibody screening on protein macroarrays

Five sera from each of the three donor cohorts were screened separately on cDNA macroarrays for IgG- and IgA-specific autoantigenicity profiles. Final aim was to identify disease-specific autoantigens, which could be potentially useful for AD- and/or MS-diagnostic purposes. Here, hits are defined as arrayed cDNA fragments, showing interaction with the accordant serum autoantibody class. Fig. 15 below shows an example of a hybridized filter and the accordant spotting pattern.



**Fig. 15 Example of a hybridized high-density protein macroarray.** A hEXselect filter is shown, which was hybridized with the healthy human serum No. 154014 and detected with  $\alpha$ -human-IgA-AP antibody. A “hit”, or a “positive clone”, is here a white signal in duplicate, arranged in a specific pattern around a dark guiding ink spot. The magnified detail shows a hit, which is pattern No. 4 in the 5 x 5 arraying pattern.

During data analysis the focus was laid on differential hits that highlighted in trials with diseased sera, but not with healthy ones. Thus, initial examination of overall positive clones was performed by plotting hits from diseased sera against those from healthy sera: AD vs. Healthy and MS vs. Healthy, respectively. It should be taken into consideration that no sera were pooled during screening. Also detections of the two different antibody classes, IgA and IgG, were performed separately. Consequently, four different plots were the outcome of this analysis, as shown in Fig. 16 below.



**Fig. 16. Summary plots of macroarray screening results.** In each table positive clones from diseased sera (x-axes) are plotted against healthy sera (y-axes). Both upper tables show **AD/Healthy** and both lower tables show **MS/Healthy** plotting. Both left tables show hits from **IgA** screening and both right tables show hits from **IgG** screening. **n** is the cumulative number of screened sera. Accordant percentage is shown at the side. Numbers in the grid are non-redundant counts of positive clones. For example, the green circled grid, showing 10, means that ten different clones were found to be positive, hence immunoreactive, with 3 AD sera (60%) and with 2 healthy sera (40%) in a screening with IgA-detecting antibodies. Most interesting hits are highlighted in red, since they show largest differentiation between diseased and healthy samples.

Next, the most interesting hits, highlighted in red in the tables above, were analyzed in detail. To filter potentially interesting genes, following criteria were considered:

- **Frame prediction:** Is the positive clone annotated “in frame”? Hence, can the hit cDNA fragment be linked to a real gene with biological relevance?
- **Clone redundancy:** How many different clones, covering the same gene, are spotted on the filter? And how many of these were actually hit during screening?
- **Clones in duplicate:** Does the positive clone have a real duplicate, spotted on the filter? And was it hit as well?

According to these criteria positive clones were prioritized and finally divided in three groups: not specific, potentially specific and highly specific for the accordant donor cohort. For example, a **CCDC50** gene (Coiled coil domain containing 50) is represented by three

different clones, each spotted once on the macroarray. Two of the clones are annotated in-frame, one is off-frame. All three clones were hit in two different AD sera, but in no healthy sera, in an IgA-detecting screening. Hence, on the upper left table they belong to the clones in grid 33, highlighted in red. Such genes were determined highly specific, in this case for AD. Table 9 below summarizes results from this analysis.

**Table 9. Positive clones analysis from macroarray screening.** Numbers represent count of positive clones, assigned specific.

	AD		H vs. AD		MS		H vs. MS	
	IgA	IgG	IgA	IgG	IgA	IgG	IgA	IgG
<b>Highly specific</b>	11	1	2	0	13	3	1	0
<b>Potentially specific</b>	9	0	20	9	9	13	12	13

A second analysis of the screening data was performed in parallel as well. This time, all in-frame hits were selected and assigned to the accordant genes at the beginning (total hit genes). Next, a differential score for each antibody class was designed in order to prioritize positive hit genes, based on the following formula:

$$[(Hits\ disease - Hits\ healthy) > 0] \times Sera\ disease\ with\ hits$$

Thus, the difference of all hit clones per gene between a diseased and a healthy serum was calculated. Thereby, it was not significant, which clones were hit, as long as they belonged to one gene and were in-frame. If the difference was bigger than zero, meaning that a gene is positive with more diseased sera than with healthy ones, it was further multiplied with the actual number of diseased sera, with which the accordant gene was hit. Consequently, genes, hit with more diseased sera, achieved a higher differential score, hence a higher priority. For example, if a gene was hit twice with 1 AD serum, once with 3 other AD sera and once with 2 Healthy sera, the accordant gene would achieve a differential score of 12:

$$(5 - 2) \times 4 = 3 \times 4 = 12$$

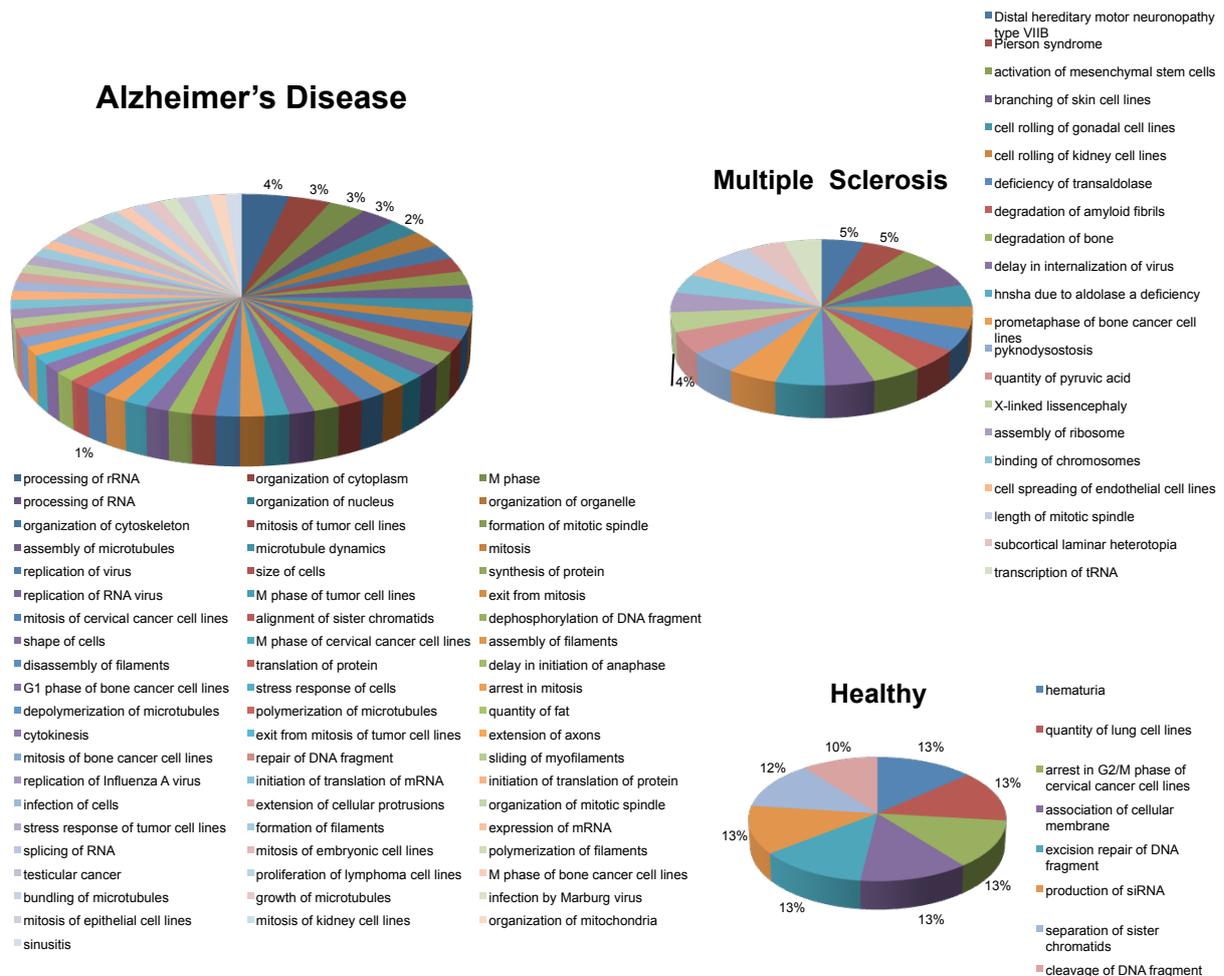
All genes with a score over a threshold value of two were considered potentially interesting. From these, the ones were further selected, which occurred in only one cohort. Hence, these were named “unique” for the accordant donor group. Table 10 below summarizes the results from this analysis.

**Table 10. Hit genes analysis from macroarray screening.**

	Positive clones	Positive clones	Hit genes	Hit genes		Hit genes	
	total	in frame	total	diff. score > 2	unique		
<b>AD</b>	2,036	971	378	142	88		
<b>MS</b>	2,737	1,125	370	170	72		
<b>Healthy</b>	2,469	1,113	368	vs. AD 177	vs. MS 151	vs. AD 12	vs. MS 6

Finally, the results from both analyses were compared to each other. Naturally, most of the highly specific hits from the first analysis would be present in the list of unique genes from the second analysis as well. Some of these most promising candidates were then selected for later recombinant expression and ELISA validation, as shown in chapter 5.4.

The third kind of analysis carried out, was an examination of the interaction pathways, represented by the hit genes in the different cohorts. This was based on the assumption that in complex multifactorial diseases, such as AD and MS, it is fairly unlikely that one or few very prominent biomarker exist. A more plausible expectation would be to identify proteins, belonging to certain physiological pathways and/or interaction networks that are affected by particular disease conditions. To prove this hypothesis, pathway analyses were performed with the Ingenuity software (IPA, Qiagen). For this, unique hit genes from the three accordant groups were imported in the IPA software and a general pathway analysis was run with the three data sets. In the healthy cohort all 18 genes were imported as one data set. P-values for each identified pathway were calculated by the software. Pivotal criteria were the number of hit genes, belonging to one discrete pathway, as well as the size of the pathway itself. Pathways were not sorted and are therefore a mixture of biochemical, physiological and disease relevant interaction networks, as described in the IPA data bank. Pathways with a p-value over 0.01 were considered significant and are depicted as pie diagrams in Fig. 17 below.



**Fig. 17 IPA analyses of the results from macroarray screening.** Percentage pie diagrams represent distribution of discrete pathways (biochemical, physiological, disease relevant, etc.) of unique hit genes (differential score >2) in the accordant donor cohort. Presented data are negative logarithms of the p-values, calculated by IPA. Only pathways with p-values  $>10^{-2}$  were included. Arrows mean same percentage in the shown direction up to the next value. Data sets: AD set with 88 genes, MS set with 72 genes and Healthy set with 18 genes.

With this, analyses of the macroarrays screening results were completed. From the list of the unique hits some most promising AD and MS biomarker candidates were selected for bacterial recombinant expression and further validation via antigen ELISA (chapter 5.4.). All interesting candidates are listed in the supplementary material.

### 5.3. Semi-automated selection of human autoantigens, presented on M13 phages

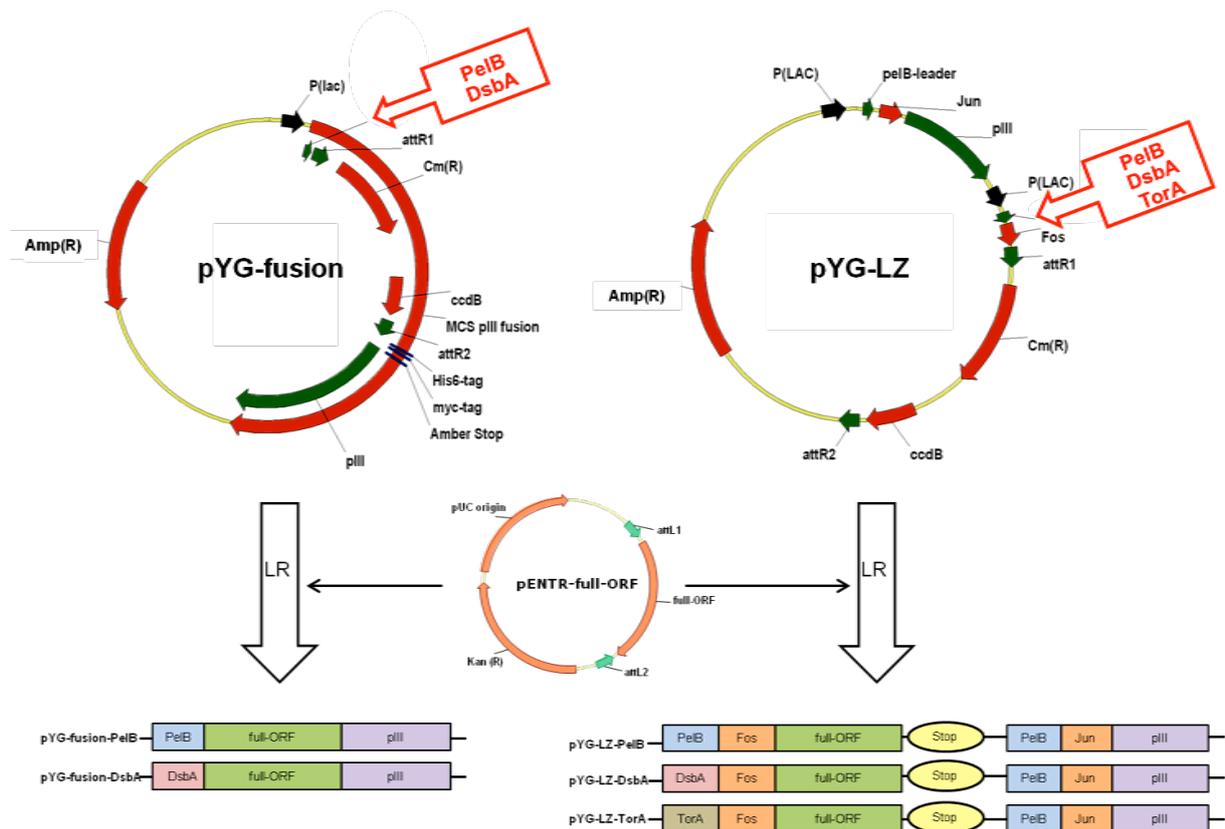
#### 5.3.1. Generation of human full-ORF phagemid libraries

Prior to bio-panning, novel phagemid libraries had to be generated. These were phage display compatible bacterial cell stocks, transformed with phagemid vectors, containing human full-ORF inserts from the source OC libraries. Since the full-ORF inserts resided in

Gateway pENTR plasmids, ready for LR recombination, new destination vectors (pDEST) were created first and called pYG vectors. Two series of the pYG vectors were generated in a way to be applicable for both LR reaction and later for POI display on the minor pIII phage coat protein. The pYG-fusion series led to a direct catenation of the POI with pIII via a C-terminal fusion. The pYG-LZ series, on the other hand, linked the POI indirectly to pIII via a leucine zipper structure (Fos-Jun). In this case both interaction partners were expressed and exported separately to the periplasmic space, where hydrophobic linkage spontaneously took place and bound both polypeptides without a direct peptide linkage.

Furthermore, in both constructs two variants with two different leader sequences for periplasmic expression were designed: with the pelB sequence for the Sec export pathway and with DsbA for the SRP-dependent pathway. In the pYG-LZ series also a variant with the TorA leader sequence was additionally implemented, which utilizes the bacterial TAT export machinery. This was not possible in the pYG-fusion construct, since pIII needs the periplasmic space for proper folding and assembly with the other coat proteins and has to be exported in an unfolded state through the Sec pore. In conclusion, five different pYG destination vectors were generated, as shown in Fig. 18 below.

Another major feature of the pYG destination vectors was a *ccdB* cassette, comprised of a *ccdB* gene and a chloramphenicol resistance gene (Cm(R)). The *ccdB* product, toxic for *ccdB*-sensitive *E. coli* strains like DH10B, is a standard selection feature in the Gateway system. Its functionality was important for minimizing the background of non-recombinant parental clones in the newly generated phage libraries after LR reaction.

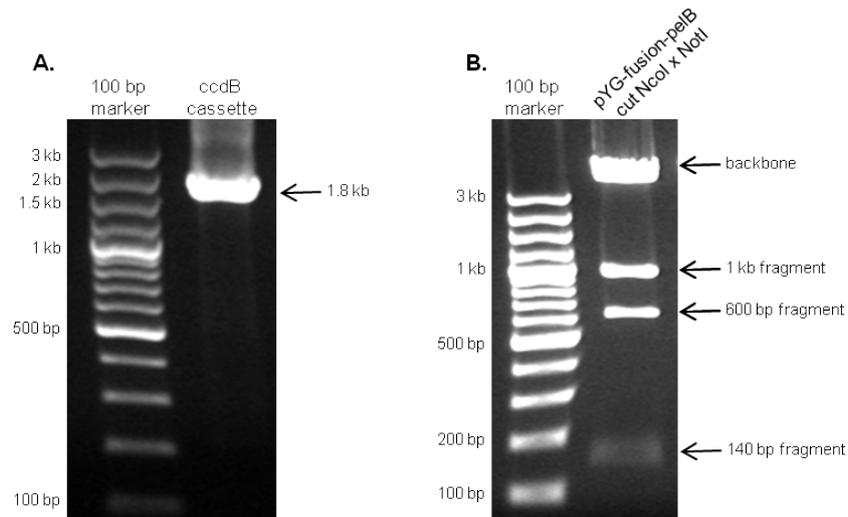


**Fig. 18 Cloning strategy for the pYG-full-ORF phage vectors.** pYG-fusion and pYG-LZ destination vectors were constructed at first. Cloning procedures included incorporation of the leader sequences (red arrow boxes) in both vector types, a *ccdB* cassette in the pYG-fusion construct and a stop codon, downstream from the full-ORF POI, in the pYG-LZ construct. After validation of the pYG destination vectors, LR reactions with each of the pENTR-full-ORF libraries were performed. The final five pYG-full-ORF library constructs, ready for phage display, are shown at figure bottom.

Following chapters describe the construction of the pYG destination vectors and the subsequent generation of the pYG-full-ORF libraries in detail, including results from validating experiments.

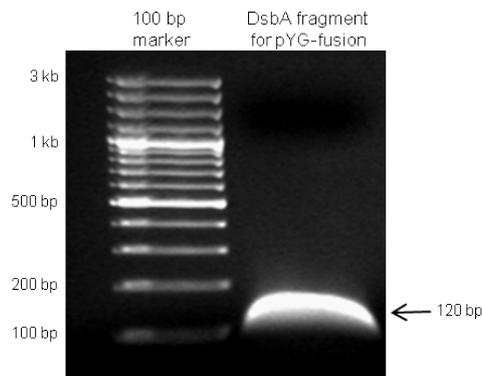
### 5.3.1.1. Construction of destination pYG-vector series

pYG-fusion vectors originated from a pIT2 backbone with an inserted mTKIN (*Thermomyces* kinesin-3) gene between a *pelB* leader sequence and a truncated *pIII* gene. To generate pYG-fusion-*pelB*, the mTKIN insert was replaced with the *ccdB* cassette, flanked by LR attachment sites *attR1* and *attR2*, using *NcoI* x *NotI* restriction. Since both enzymes had intrinsic recognition sites in the *Cam(R)* gene, cloning had to be performed in two subsequent steps. Applied primer pairs were: *NcoI*-*AttR1*-for x *NcoI*-*CmR*-back for the first step and *NcoI*-*CmR*-for x *NotI*-*AttR2*-back for the second.



**Fig. 19 PCR for cloning the pYG-fusion-pelB vector. A. Analytical PCR after cloning.** A PCR product of the complete *ccdB* cassette is shown amplified from the ready-made pYG-fusion-pelB plasmid. Primer: *NcoI*-AttR1-for x *NotI*-AttR2-back. **B. Analytical digest after cloning.** Restriction with *NcoI* x *NotI* of the pYG-fusion-pelB vector leads to four fragments, since there are two recognition sites for each enzyme.

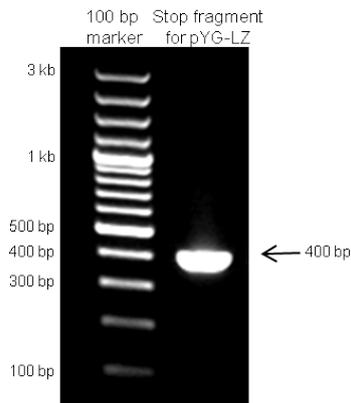
Next, the pYG-fusion-DsbA destination vector was generated. For this, the *pelB* leader sequence in pYG-fusion-pelB was replaced with the amplified fragment of the DsbA leader polypeptide, using the primer pairs pIT2-DsbA-for x pIT2-DsbA-back and subsequent restriction with *HindIII* x *SfiI*.



**Fig. 20 PCR for cloning the pYG-fusion-DsbA vector.** PCR product of the DsbA fragment, prior to digestion and ligation into the pYG-fusion vector backbone is shown.

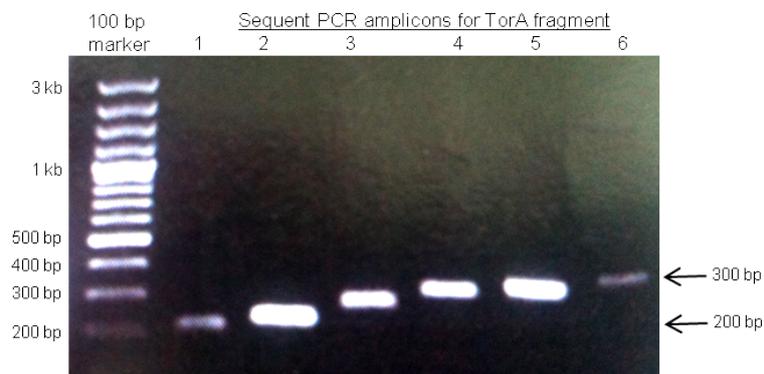
pYG-LZ destination vectors originated from the pJuFo-B vector (reading frame B), which already contained all necessary features: *ccdB* cassette, LR attachment sites, truncated pIII with an N-terminal *pelB* and Jun sequences, as well as a *pelB* leader and a Fos sequence N-terminal to the *ccdB* cassette. To allow expression of full-ORF inserts without a C-terminal stop codon, a Stop-Serine-Stop (TAG-AGT-TAA) sequence was introduced C-terminal to the

ccdB cassette on the pJuFo backbone. This was made applying primer pairs pJuFo-XmaI-for x pJuFo-XbaI-back and subsequent restriction with XmaI x XbaI.



**Fig. 21 PCR for cloning the pYG-LZ-peIB vector.** A PCR product of the fragment, containing the Stop-Serine-Stop sequence prior to digestion and ligation is shown.

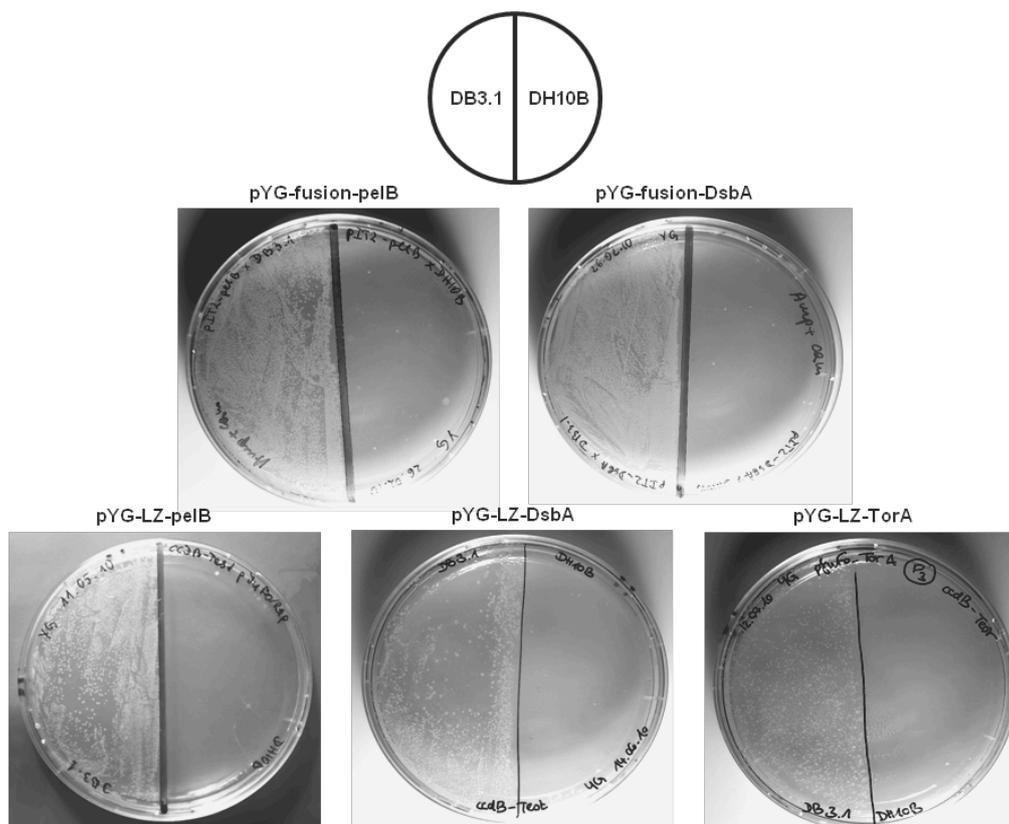
To generate pYG-LZ-DsbA and pYG-LZ-TorA, the peIB leader sequence N-terminal to the Fos-ccdB-cassette-fragment, was replaced with inserts, containing the DsbA and the TorA sequences, respectively. Due to the lack of appropriate restriction sites in this vector region, relatively large inserts had to be amplified, containing also the promoter sequence pLac and the following RBS (ribosome binding site) site. Hence, the two inserts were generated in consecutive PCR reactions, where the prior amplicon was used as a template in the following reaction and so on. Forward primer in all PCR reactions was NheI-pLac-for. Sequent backwards primers were RBS-back, DsbA-back 1 → 4 and TorA-back 1 → 5, respectively. Fig. 21 below shows all TorA-fragments as an example. The complete TorA insert was the 300 bp large amplicon No.6 on the agarose gel picture below (Fig. 22). Both DsbA and TorA inserts were restricted with NheI x SacI.



**Fig. 22 Sequent PCR for cloning the pYG-LZ-DsbA/TorA vectors.** Only TorA generation is shown as an example. First amplicon (200 bp) served as a template for the second and so on. Final TorA insert (300 bp) was digested and cloned into the vector backbone.

All pYG destination vectors were propagated in the *ccdB*-resistant *E. coli* strain DB3.1. After transformation, cells were spread on 2xYT agar plates, containing both Ampicillin and Chloramphenicol for selection. Single transformants were analyzed applying colony PCR and analytical digestion. Newly generated plasmids were rescued and sequenced (Sanger). Special attention was paid to correct reading frames of future expression products.

In order to generate heterogenic libraries of high quality, the background of parental non-recombinant clones has to be as low as possible. During cloning procedures, this is ensured by the negative selection features, chosen in the cloning strategy. The *ccdB* gene codes for a suicide expression product, which inhibits the bacterial gyrase enzyme in sensitive strains. Usually, this selection system is highly efficient, providing up to 99% positive selection, and is therefore a standard feature in numerous systems, including Gateway®. To verify *ccdB* functionality in the newly generated pYG destination vectors, so-called *ccdB*-tests were performed with each of the five vectors. For this, same amount of pYG-plasmid-DNA was transformed into DH10B (*ccdB* sensitive) and into DB3.1 (*ccdB* resistant) cells of similar efficiencies in parallel. After incubation, same amount from both cell suspensions was spread on split 2xYT/Amp/Cam plates. Survival rates were monitored on next day.



**Fig. 23 Results of *ccdB* tests with pYG destination vectors.** Photographic images show satisfactory results of *ccdB* functionality tests with each of the five pYG destination vectors.

### 5.3.1.2. LR reactions and phagemid library validation

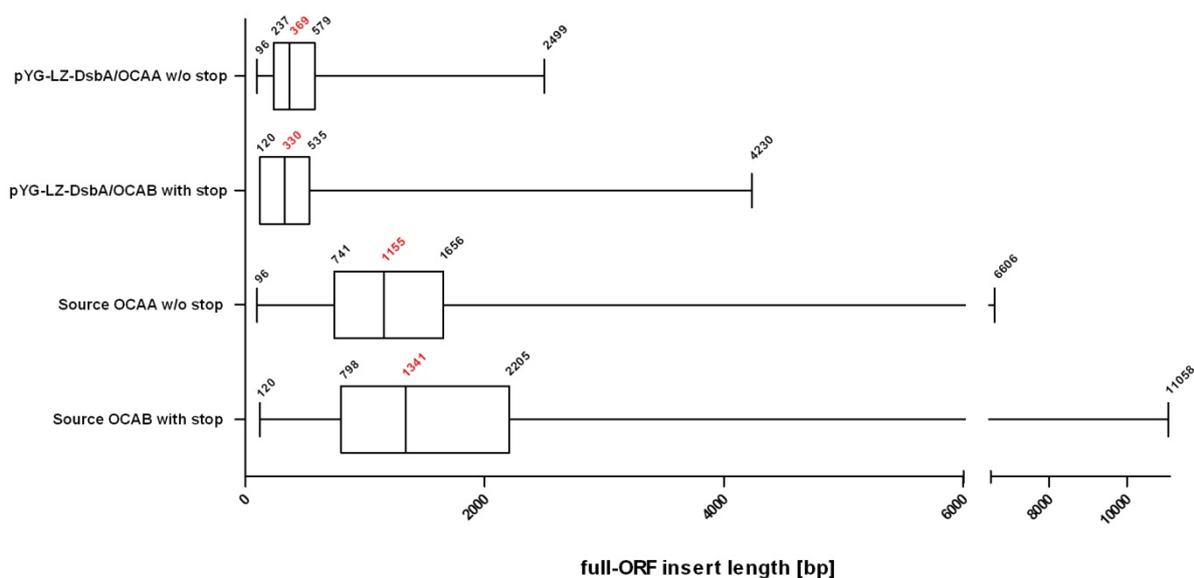
LR reactions were carried out with each of the five pYG destination vectors and each of the four pooled pENTR/OC libraries, respectively. Recombinant plasmids were then transformed in DH10B cells for positive selection and plated out completely on large 2xYT/Amp agar plates. On next day, transformants were counted in order to estimate achieved fold coverages. Table 11 below shows results from this first amplification step.

**Table 11. Amplification of pYG/OC libraries in DH10B cells.** Colony counts and accordant fold coverages are shown for each recombinant pYG/OC library after LR recombination and subsequent transformation in DH10B cells. Numbers in brackets show original size of the source OC libraries. Only libraries without stops were recombined with the two pYG-fusion destination vectors.

	pENTR/OCAA-w/o-stop (4,929 clones)		pENTR/OCAA-with-stop (6,240 clones)		pENTR/OCAB-w/o-stop (1,152 clones)		pENTR/OCAB-with-stop (1,200 clones)	
	Colony count	Fold coverage	Colony count	Fold coverage	Colony count	Fold coverage	Colony count	Fold coverage
pYG-fusion-pelB	$3.1 \times 10^4$	6.3	–	–	$4.7 \times 10^4$	41	–	–
pYG-fusion-DsbA	$3.6 \times 10^4$	7.3	–	–	$2.7 \times 10^4$	23	–	–
pYG-LZ-pelB	$4.6 \times 10^4$	9.3	$1.3 \times 10^4$	2.1	$1.3 \times 10^4$	10.8	$1.9 \times 10^4$	15.8
pYG-LZ-DsbA	$5.7 \times 10^4$	11.6	$4.6 \times 10^4$	7.3	$2.4 \times 10^4$	20.6	$7.1 \times 10^4$	59
pYG-LZ-TorA	$1.4 \times 10^5$	28.4	$5 \times 10^4$	8	$4.8 \times 10^4$	41.6	$8.5 \times 10^4$	70.8

To gain a deeper insight into libraries' quality, background and diversity, overall 130 single clones from two different plates were randomly picked and sequenced (Sanger): 62 clones from the pYG-LZ-DsbA/OCAA plate and 68 clones from the pYG-LZ-DsbA/OCAB plate. Sequencing results revealed that 71% of the OCAA clones were unique. 39% of the OCAB clones were also picked just once. Furthermore, only one clone appeared to carry an empty backbone without any inserts and none was found to be a parental clone, still comprising a ccdB cassette. All identified inserts were in the correct reading frame and without any mutations or other errors in the DNA sequence.

It is known for LR reactions that although even very long fragments of up to 10 kb can be successfully recombined, reaction speed and efficacy decrease steadily with increasing insert size. Therefore, I hypothesized that one major bottleneck would manifest exactly at this cloning stage. To prove this, insert size distribution of the sequenced 130 clones were calculated and compared to the distribution in the original OC libraries. Results from this analysis are depicted as box-and-whiskers plots in Fig. 24 below.



**Fig. 24 Full-ORF inserts size distributions.** Two upper plots show distributions in the two recombined pYG-LZ-DsbA/OC libraries among sequenced 62 and 68 clones, respectively. Two lower plots show distributions in the two accordant OC source libraries among all clones available. Numbers correspond to the standard five distribution parameters: minimum value, 25% percentile, median value (highlighted in red), 75% percentile and maximum value.

In conclusion, inserts size distribution analysis confirmed prior expectations. As Fig. 24 shows, while the median values of the two source libraries inserts are about 1,248 bp, in the sequenced clones these amount only to approx. 340 bp.

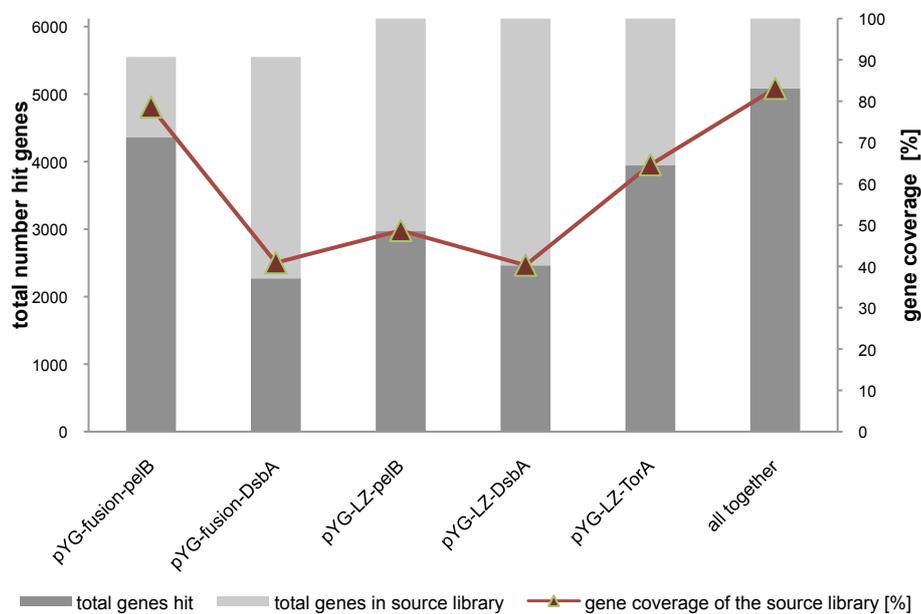
In the next step, pYG/full-ORF plasmids were rescued and re-transformed in XL1 Blue or TG1 *E. coli* cells, convenient for phage display. As expected, further significant multiplication of the fold coverages was achieved, as shown in Table 12 below.

**Table 12. pYG/OC libraries amplification in XL1 Blue/TG1 cells.** Colony counts and accordant fold coverages are shown for each recombinant pYG/OC library after re-transformation in XL1 Blue or TG1 cells. Numbers in brackets show original size of the source OC libraries. Only libraries without stop codons were comprised in the two pYG-fusion destination vectors.

	pENTR/OCAA-w/o-stop (4,929 clones)		pENTR/OCAA-with-stop (6,240 clones)		pENTR/OCAB-w/o-stop (1,152 clones)		pENTR/OCAB-with-stop (1,200 clones)	
	Colony count	Fold coverage	Colony count	Fold coverage	Colony count	Fold coverage	Colony count	Fold coverage
pYG-fusion-pelB	$2 \times 10^6$	405	–	–	$8.6 \times 10^5$	747	–	–
pYG-fusion-DsbA	$4 \times 10^6$	811	–	–	$2.3 \times 10^6$	1,997	–	–
pYG-LZ-pelB	$9 \times 10^5$	183	$9 \times 10^5$	144	$10^6$	868	$10^6$	833
pYG-LZ-DsbA	$9.7 \times 10^5$	197	$10^6$	160	$9.5 \times 10^5$	825	$10^6$	833
pYG-LZ-TorA	$10^6$	203	$10^6$	160	$10^6$	868	$10^6$	833

To produce full-ORF containing phages, recombinant XL1 Blue or TG1 cells were infected with the M13K07 helper phage. These initial libraries were PCR amplified, using phage DNA as template, and deep sequenced on an Illumina platform. Thereby, the pYG-fusion-pelB library was produced once in TG1 and once in XL1 Blue cells. Both phage preparations were sequenced. For all other libraries only the XL1 Blue preparations were sequenced for analysis. Approximate read numbers, achieved by NGS usually ranged between 1.5 Mio and 7.5 Mio reads per library and selection stage. From both pYG-fusion libraries the initial (pre-selection) stage, as well as the 2<sup>nd</sup> and the 4<sup>th</sup> selection rounds were sequenced. From the three LZ libraries, beside the initial stage, only the 4<sup>th</sup> selection round was sequenced.

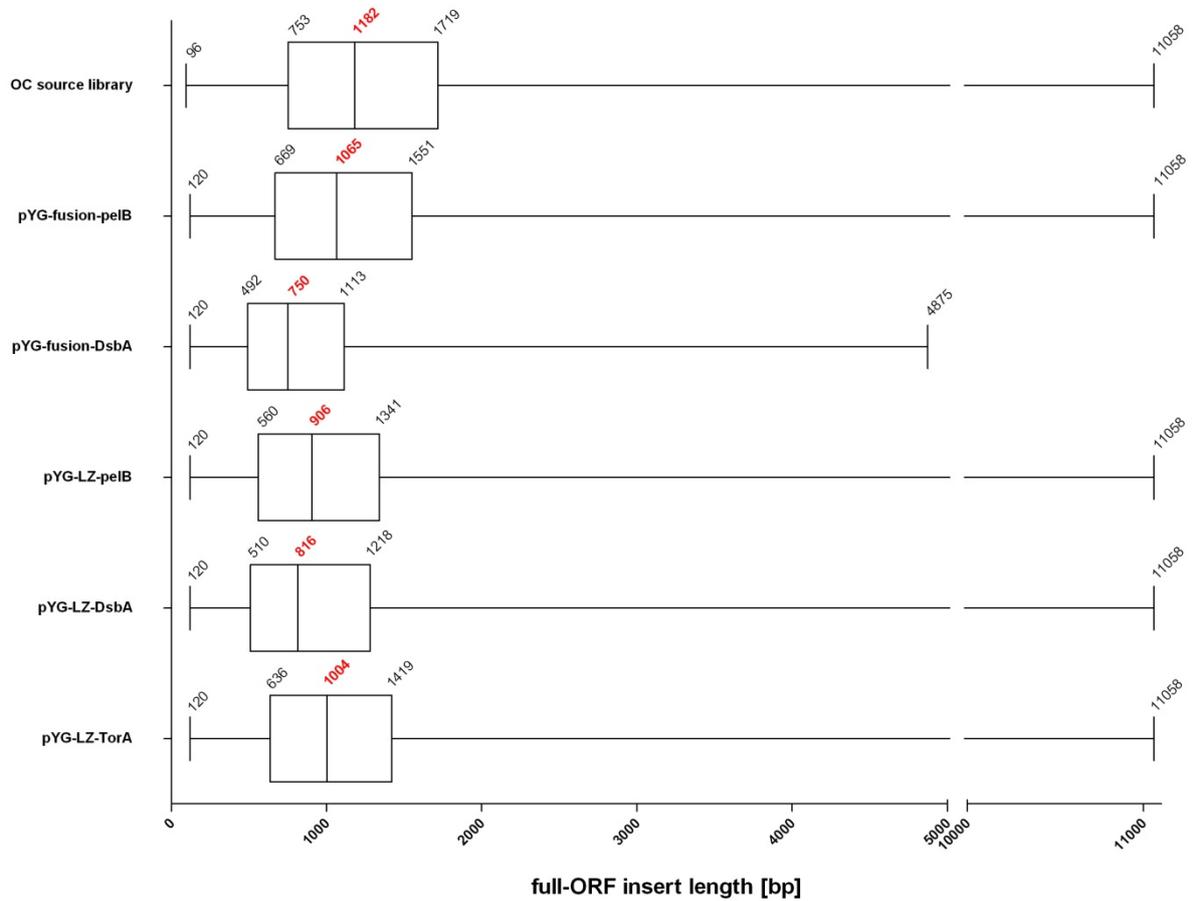
Analysis of the gene coverage in the initial libraries, i.e. how many genes from the original OC library could be re-identified in phage DNA by NGS, revealed following results:



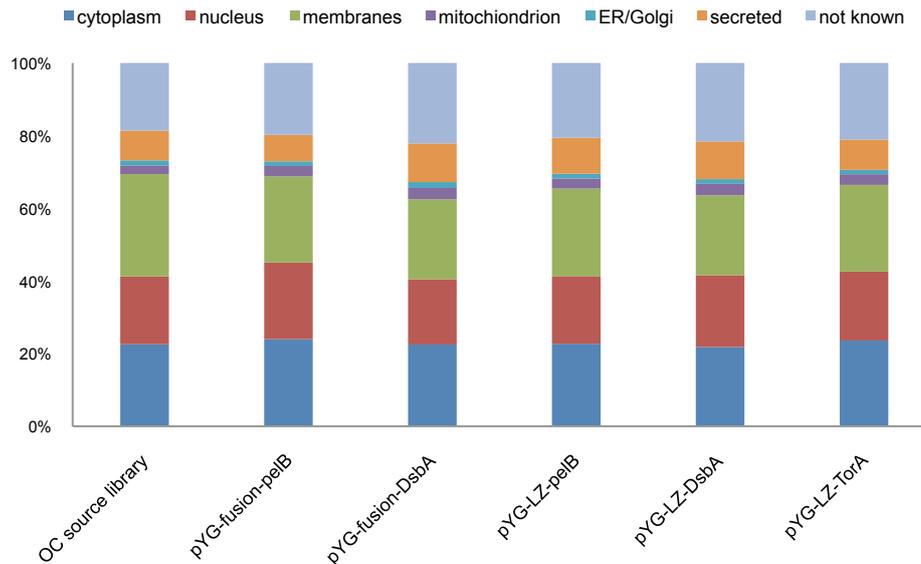
**Fig. 25 Gene coverages in initial pYG/OC libraries.** Stacked bar chart shows how many genes in total were hit by NGS (left y-axis) in the single libraries and in overall in general (last bar). Further, which percentile ratio of the original source OC genes was covered (right y-axis and secondary line in red). For example, 5,087 genes were identified by NGS in all pYG/OC libraries in total (dark grey area of the “all together” bar). Since the source OC library consists of overall 6,119 genes, 1,032 genes were not found (light grey area). Thus, overall gene coverage in the initial phage libraries accounted for 83% (red triangle) of the source library.

The initial libraries were analyzed in two further aspects as well: overall distribution of the full-ORF insert size (Fig. 26) and distribution of inserts’ native cellular compartments in a eukaryotic cell (Fig. 27). Intended purpose for this kind of analysis was to test, if cloning procedures had any undesired impact on libraries composition and diversity. On the other hand, it was interesting to see, if there are any noticeable differences between the five vector

systems applied. For example, if a certain leader peptide or one of the pIII-fusion construct types would lead to any biases. All five libraries were compared to each other and to the source OC library.



**Fig. 26 Full-ORF inserts size (bp) distributions in all pYG/OC initial libraries.** Box-and-whiskers plots show all five pYG/OC libraries and the source OC library (upper plot) in comparison. Numbers correspond to the standard five distribution parameters: minimum value, 25% percentile, median value (highlighted in red), 75% percentile and maximum value.



**Fig. 27 Distribution of predicted native subcellular localizations of the full-ORF proteins.** Percentaged stacked bar chart shows distributions in all five pYG/OC libraries and the source OC library (first bar on the left) in comparison. Annotations source was the UniProt Knowledgebase UniProtKB, version from September 2011 [222].

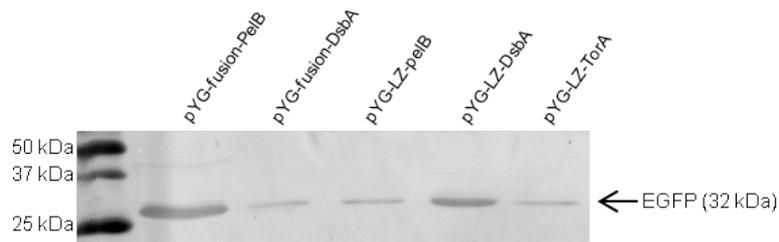
With this, analysis of the initial pYG/full-ORF libraries on genetical scale was accomplished. Next task, prior to selection, was to test expression efficiencies of the pYG-vectors and the functionality of the polypeptides, presented on phage surface. For this a well-studied model protein was chosen: EGFP (enhanced green fluorescent protein).

### 5.3.1.3. Phagemid libraries evaluation with EGFP

GFP and its variants can fold correctly and attain fluorescence only if allowed to mature in cell cytoplasm [223]. Thus, periplasmic expression functions well when using the TAT secretion pathway of *E. coli*. However neither Sec-, nor SRP-pathways are suitable in this case, since they would export the unfolded GFP-polypeptide to the periplasm, which would be therefore non-functional and hence, non-fluorescent. To demonstrate the validity of these findings for the pYG vector system as well, EGFP was expressed from all five vectors in the non-suppressor *E. coli* strain HB2151 (*supE*). Thus, the Amber codon (TAG) between the insert and the pIII-protein in the two pYG-fusion-vectors was read as a translational stop signal, allowing the expression of a single EGFP polypeptide.

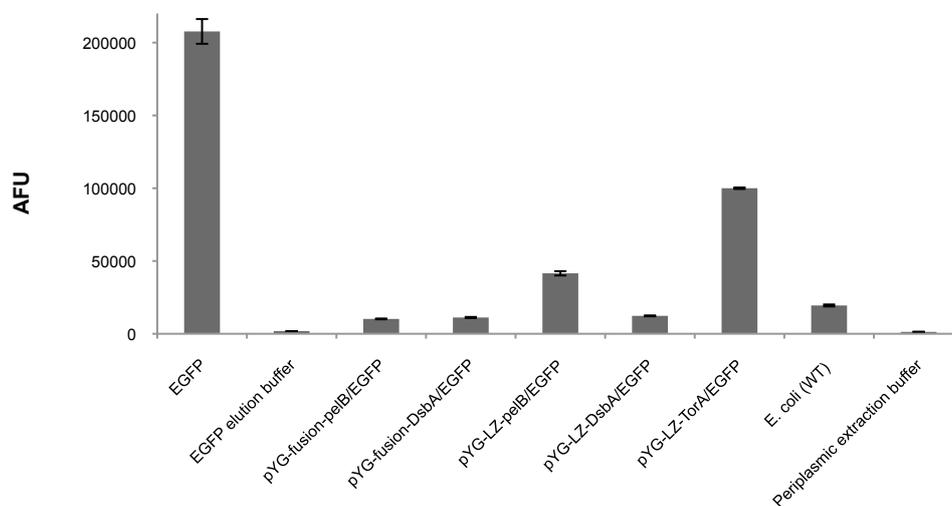
As expected, EGFP was strongly expressed from all five vectors: expression products were found in cell lysates, periplasm fractions and a good deal in the spheroblasts. After centrifugation of the induced ON cultures first differences between the five samples could be visually observed: only the pYG-LZ-TorA cell pellet was bright green. While the periplasm fraction was only slightly greenish, the spheroblasts of this sample remained intensive in

color. The other four samples did not show any visible color changes at any stage of the expression or periplasm extraction processes.



**Fig. 28 Western blot analysis of EGFP expression in *E. coli* periplasm.** EGFP was expressed from each pYG-vector and extracted from periplasmic space. Same sample volume was loaded per lane. Detection was performed with  $\alpha$ -GFP 11E5 antibody (mouse) and subsequent HRP-conjugated  $\alpha$ -mouse antibody.

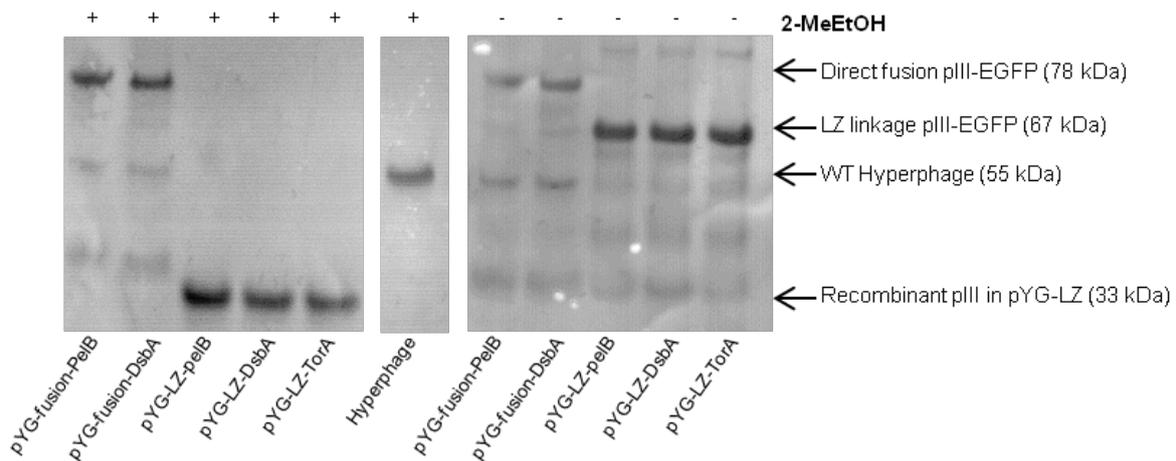
Fluorescence measurements of the periplasm fractions finally confirmed previous observations: significant fluorescence could be detected in the TorA-sample only, as pictured in Fig. 29 below. All fluorescence measurements were performed in triplicate on an Omega Star microplate reader. Standard errors were calculated from standard deviations divided by the square root of the sample size.



**Fig. 29 Fluorescence measurements of EGFP periplasm fractions.** Arbitrary fluorescence units (AFU) are assigned on the y-axis. Recombinant EGFP (4.56  $\mu$ g/ml), expressed in *E. coli* and IMAC purified (Ni-NTA-agarose) served as a positive reference (first bar on the left). Next, measurements in EGFP periplasm extractions, expressed from all five pYG vectors are shown (bars 3 – 7). Periplasm fraction of non-recombinant *E. coli* XL1 Blue culture (bar 8) and accordant buffers (bars 2 and 9) served as background or negative controls, respectively.

Given the results from periplasmic EGFP expression, I assumed that this scenario will recur also when presenting EGFP on phage surface. Thus, all five pYG/EGFP vectors were

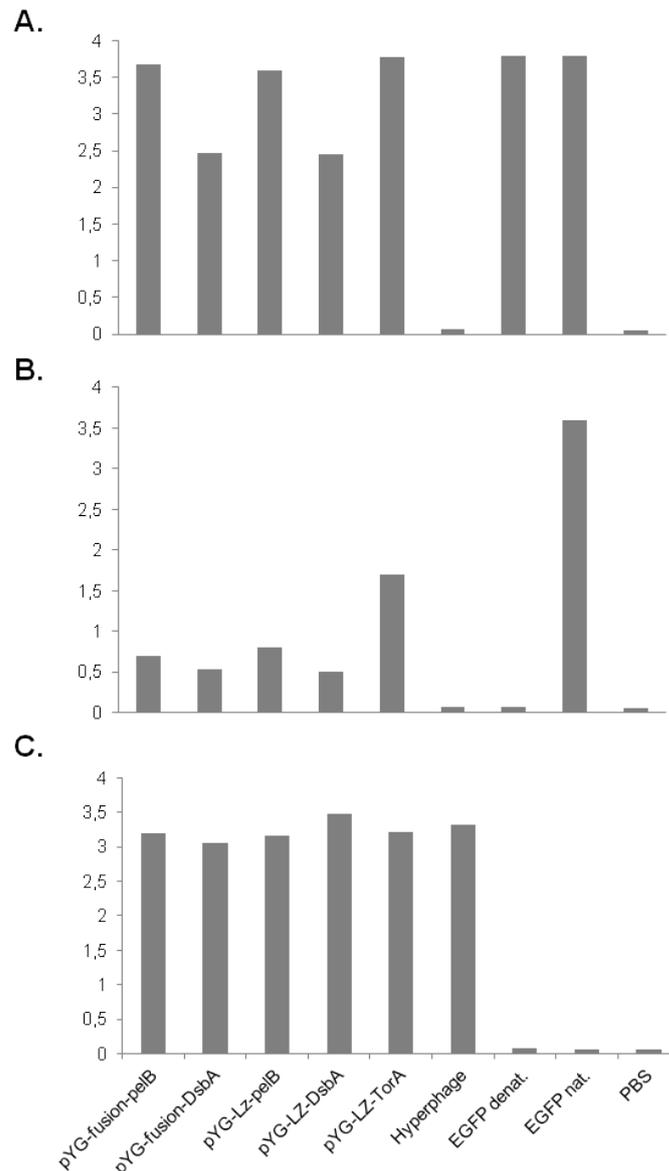
transformed in the suppressor *E. coli* strain XL1Blue (*supE*), suitable for phage display. To achieve multivalent presentation of EGFP, infection was performed with Hyperphage, which led to up to five EGFP molecules per phage particle and virtually no wild type pIII on the recombinant phages [189]. Resulting EGFP-pIII-constructs were first analyzed on western blots under reducing and non-reducing conditions, applying SDS loading buffer with and without the reducing agent 2-Mercaptoethanol (Fig. 29).



**Fig. 30 Western blot analysis of EGFP presenting pYG phages.** Samples on the left were treated with 2-Mercaptoethanol (+) prior to boiling and samples on the right were not (-). Lane 6 in the middle was charged with non-recombinant Hyperphage for comparison. Approx.  $10^8$  phage particles were loaded per lane. Detection was performed with a primary mouse  $\alpha$ -pIII antibody and a secondary HRP-conjugated  $\alpha$ -mouse antibody.

Thus, EGFP could be successfully presented on phage surface, expressed from all five pYG constructs. As expected, 2-MeEtOH treatment had no influence on band patterns in the two pYG-fusion samples: the pIII-EGFP polypeptide migrated steadily at ca. 78 kDa. On the contrary, the three pYG-LZ samples showed a clear shift from ca. 33 kDa (recombinant pIII alone) when using 2-MeEtOH, to a ca. 67 kDa band (pIII-EGFP LZ linkage) in a non-reducing loading buffer, where the leucine zipper structure remained intact.

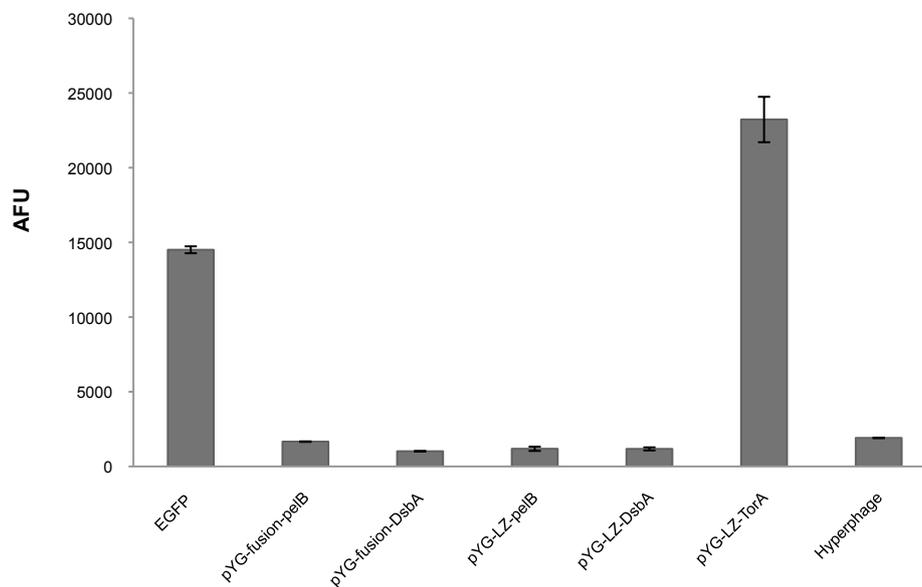
After having proved that EGFP was expressed and presented on all five phage type surfaces, the question still remained, if according to my hypothesis EGFP was properly folded and thus functional on the pYG-LZ-TorA phages, but not in the other four constructs. For this purpose EGFP-phages were further analyzed by ELISA- and fluorescence-assays.



**Fig. 31 ELISA analysis of EGFP presenting pYG phages.** Y-axis shows absorbance at 405 nm. In all three measurements first five bars correspond to the five pYG constructs. Approx.  $5 \times 10^8$  phages were loaded per well. Non-recombinant Hyperphage (bar 6) and PBS (bar 9) were used as negative controls. Purified EGFP (4.56  $\mu\text{g/ml}$ ) served as a positive control and was loaded in both denatured (EGFP denat., bar 7) and native (EGFP nat., bar 8) forms. **A. EGFP detection with a polyclonal  $\alpha$ -EGFP antibody**, which recognizes both denatured and native EGFP. **B. EGFP detection with a monoclonal  $\alpha$ -EGFP antibody**, which recognizes only the conformational epitope 3E6 and thus only the folded EGFP form. **C. Phage detection with monoclonal  $\alpha$ -pIII antibody.**

As expected, the polyclonal EGFP antibody (Fig. 31/A) was able to recognize both EGFP forms, denatured (EGFP denat.) and native (EGFP nat.). Also phage-bound EGFP gave strong signals with this antibody, demonstrating once again the availability of the molecule on the surfaces of all five phage types. Yet, pYG-fusion-DsbA and pYG-LZ-DsbA produced distinctly weaker signals, compared to the other three phage samples, with both  $\alpha$ -EGFP

antibodies. The monoclonal  $\alpha$ -EGFP antibody (Fig. 31/B), which recognizes the conformational epitope 3E6, bound just as good to the native EGFP control sample, as the polyclonal one and virtually not at all to its denatured counterpart. This result, also demonstrated recently by Velappan et al. [224], validated the usage of this advantageous antibody for the EGFP-phage-samples. All five phage samples revealed well detectable signals, yet pYG-LZ-TorA had the strongest one, about three times higher as the other four. Hence, the epitope was obviously present in its native state in all five samples and sufficiently exposed to be recognized by the antibody. However, the TorA-sample seemed to comprise much more molecules with the correctly built epitope. The presence of the correct 3E6-epitope would not necessarily mean that the overall polypeptide is properly folded or even functional and fluorescent. To further prove my hypothesis, fluorescence rates of the five EGFP-phage-samples were measured next.



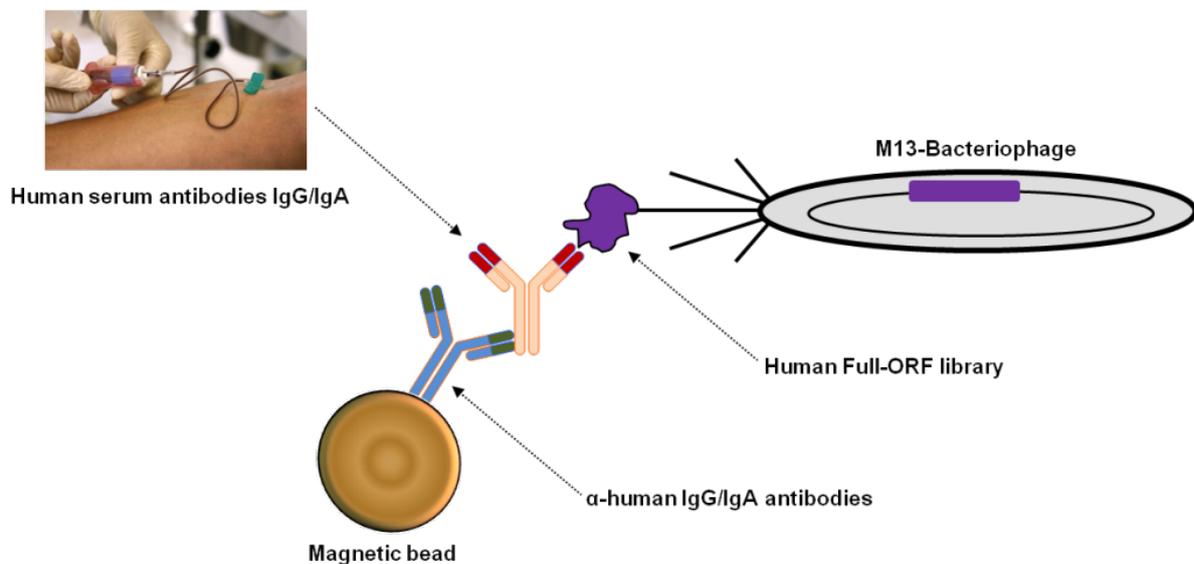
**Fig. 32 Fluorescence measurements of pYG/EGFP phage suspensions.** Arbitrary fluorescence units (AFU) are assigned on the y-axis. Recombinant EGFP (0.456  $\mu$ g/ml), expressed in *E. coli* and IMAC purified (Ni-NTA-agarose) served as a positive reference (first bar). Next, measurements in EGFP presenting pYG-phage suspensions shown (bars 2 – 6). Non-recombinant Hyperphage (bar 7) was used as a background control. Measurements in triplicate.

As shown in Fig. 32 above, only the pYG-LZ-TorA/EGFP phages showed significant fluorescence compared with the residual four samples, which only had a background fluorescence signal, comparable to the WT-Hyperphage sample.

Taking all four EGFP-assays together (Fig. 29 – 32), it could be concluded that the pYG-LZ-TorA vector was the only one in the pYG series, able to produce strongly fluorescent bacteria and fluorescent phages.

### 5.3.2. Autoantigens selection procedures

Potential disease-relevant autoantigens from the human full-ORF OC library were selected according to a semi-automated bio-panning protocol on a KingFisher Flex magnetic particle processor. For this, prior to selection, tosyl-activated magnetic Dynabeads were loaded with human IgG and IgA autoantibodies in a two-step coating procedure. A four round selection was finally performed with the full-ORF presenting pYG phages. For better demonstration, Fig. 33 below shows schematically the overall structure of the developed bait-target (human full-ORF library – human autoantibodies) construct.



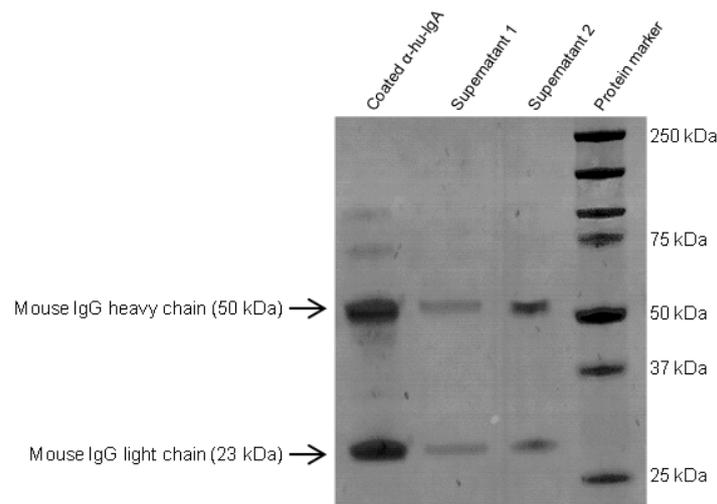
**Fig. 33 Schematic representation of the bait-target construct during bio-panning.** Commercially obtained  $\alpha$ -human IgG and IgA antibodies are coated on magnetic beads. In a subsequent secondary coating process human IgG and IgA antibodies are captured from patient sera (AD, MS and healthy controls). During the phage display selection human autoantibodies (target) encounter human full-ORF proteins (bait), presented via pIII-fusions on the surface of M13 bacteriophages. Source of picture in the upper left corner: [www.picture-alliance/dpa](http://www.picture-alliance/dpa)

#### 5.3.2.1. Loading magnetic beads with human autoantibodies

Tosyl-activated MyOne® Dynabeads were used as a solid phase to immobilize serum autoantibodies from the three donor cohorts AD, MS and Healthy. Autoantibodies of the IgG and IgA classes were applied separately. To capture immunoglobulins from the blood sera samples, magnetic beads were first coated with primary anti-human antibodies:  $\alpha$ -hu-IgG (rabbit) and  $\alpha$ -hu-IgA (goat). Another fraction was coated with  $\alpha$ -mouse-IgG (goat) for positive control of enrichment of GAPDH, Tubulin and Ubiquitin during selection. All three polypeptides were present in the initial OC libraries. In a second loading step, which was performed just prior to selection start, sample beads fraction was further coated with the

accordant human IgG and IgA autoantibodies from donors' blood sera. The beads fraction for the positive controls was coated with the mouse antibodies  $\alpha$ -GAPDH,  $\alpha$ -Tubulin and  $\alpha$ -Ubiquitin.

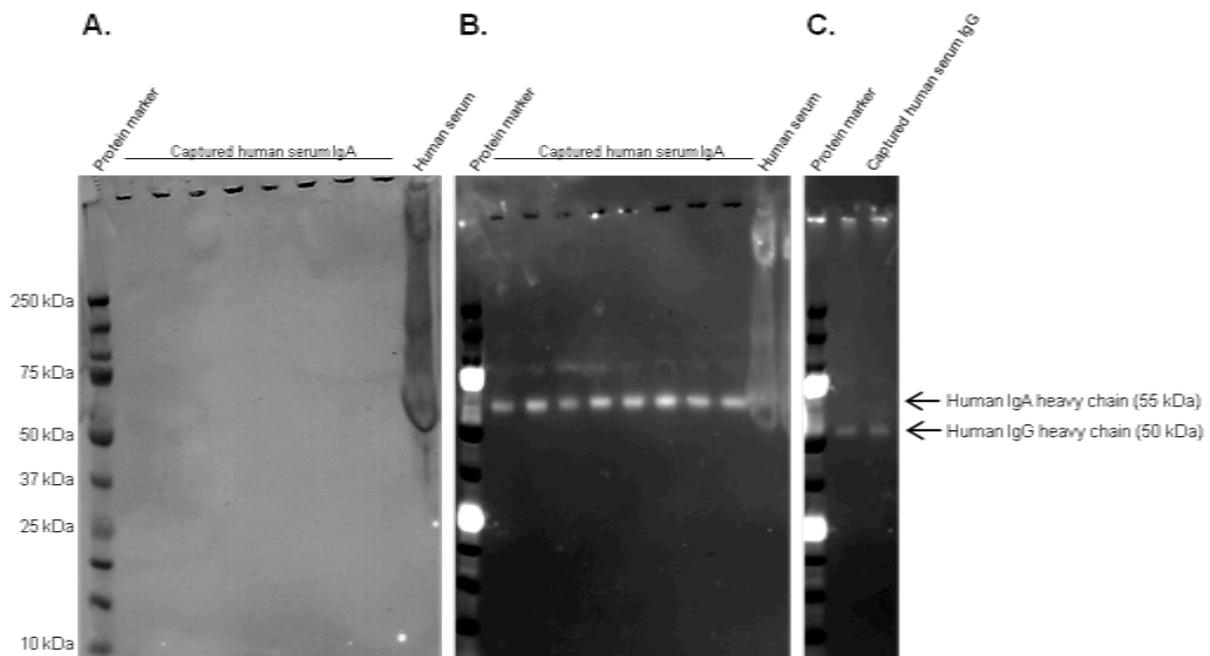
Coupling efficiency of the primary antibodies was monitored in a western blot procedure, as shown exemplary in Fig. 34 below. For this, coated beads were treated with reducing SDS loading buffer, boiled and pipetted directly into the gel pockets. During electrophoresis, uncoupled antibodies migrated into the gel, while blank bead particles remained in the gel pockets.



**Fig. 34 Western blot analysis of beads coating efficiency with the primary antibody.** An example is shown of beads, coated with mouse  $\alpha$ -hu-IgA antibodies. Beads were loaded in lane 1 (50  $\mu$ g). Lane 2 was loaded with 5  $\mu$ l of the supernatant, left after beads withdrawal from coating solution. Lane 3 was loaded with 5  $\mu$ l of the supernatant from subsequent washing step. Detection was performed with HRP-conjugated  $\alpha$ -mouse antibody, recognizing both heavy and light chains of mouse IgG. Values in brackets indicate expected approximate molecular size.

This kind of qualitative analysis was performed after each new coating procedure to assure sufficient beads loading with capture antibodies. As expected, a small fraction of uncoated antibodies remained in the supernatant and a further small amount was lost during subsequent washing steps. Nevertheless, the bulk of conjugates obviously remained bound to the beads and in this case coating was considered satisfactory.

Coated beads were also analyzed after the second loading step, i.e. incubation of beads, coated with primary  $\alpha$ -human antibodies, in human serum. Samples for WB were treated in the same way as described above. This time, however, serum autoantibodies were detected with appropriate counterparts. Also cross-reactivity of the primary antibodies was tested.

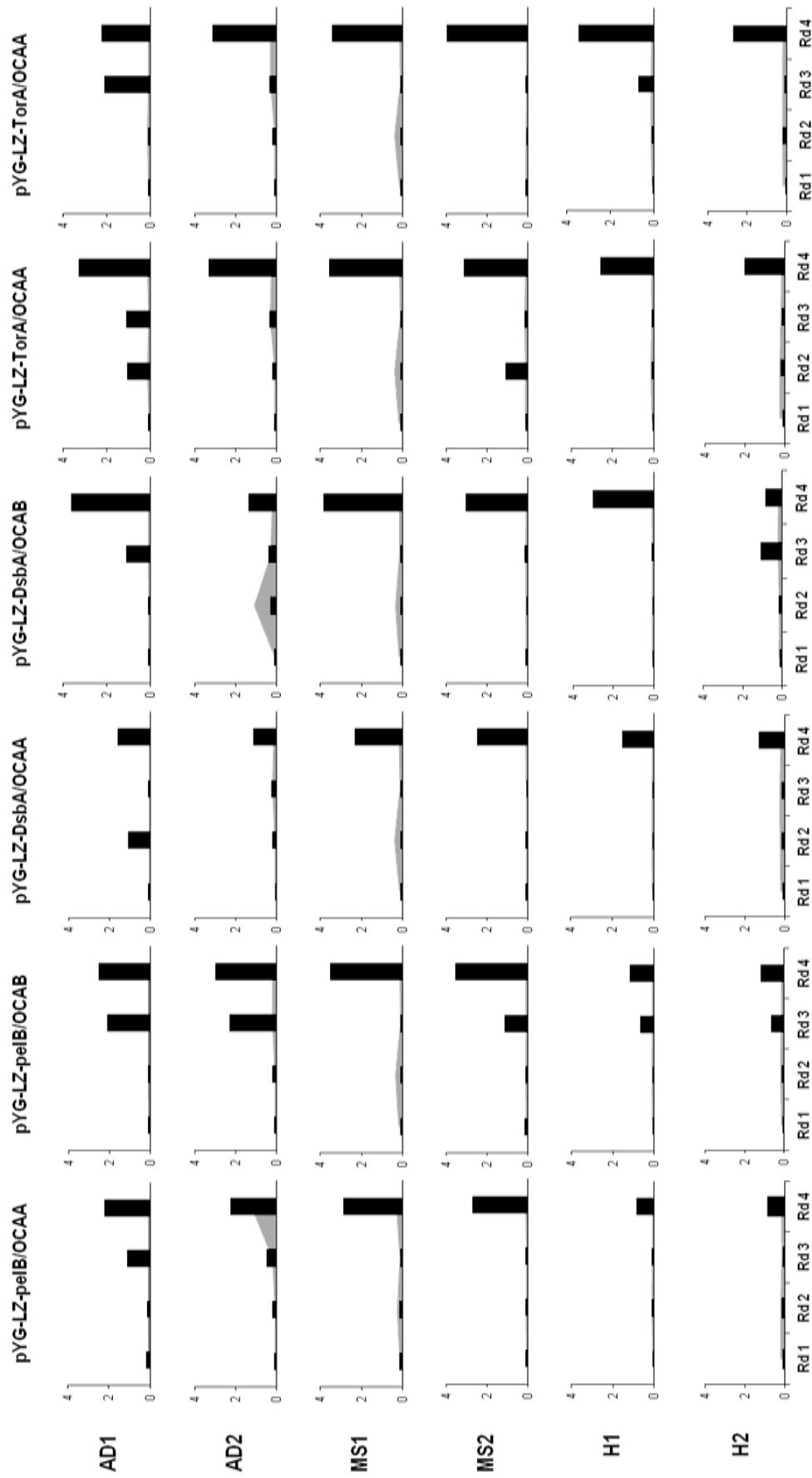


**Fig. 35 Western blot analysis of coating beads with human serum autoantibodies and cross-reactivity test of the primary  $\alpha$ -human antibodies.** Beads, coated with  $\alpha$ -hu-IgA (A. and B.) or  $\alpha$ -hu-IgG (C.) antibodies were incubated in a human serum to capture accordant autoantibodies. **A. Detection with HRP-conjugated  $\alpha$ -hu-IgG antibody.** No IgA autoantibodies could be detected in the bead samples, but naturally in the serum control (last lane). **B. Detection with AP-conjugated  $\alpha$ -hu-IgA antibody.** Same blot was used as in A. to demonstrate successful loading with IgA autoantibodies. **C. Detection with AP-conjugated  $\alpha$ -hu-IgG antibody.** Also loading with IgG autoantibodies proved successful.

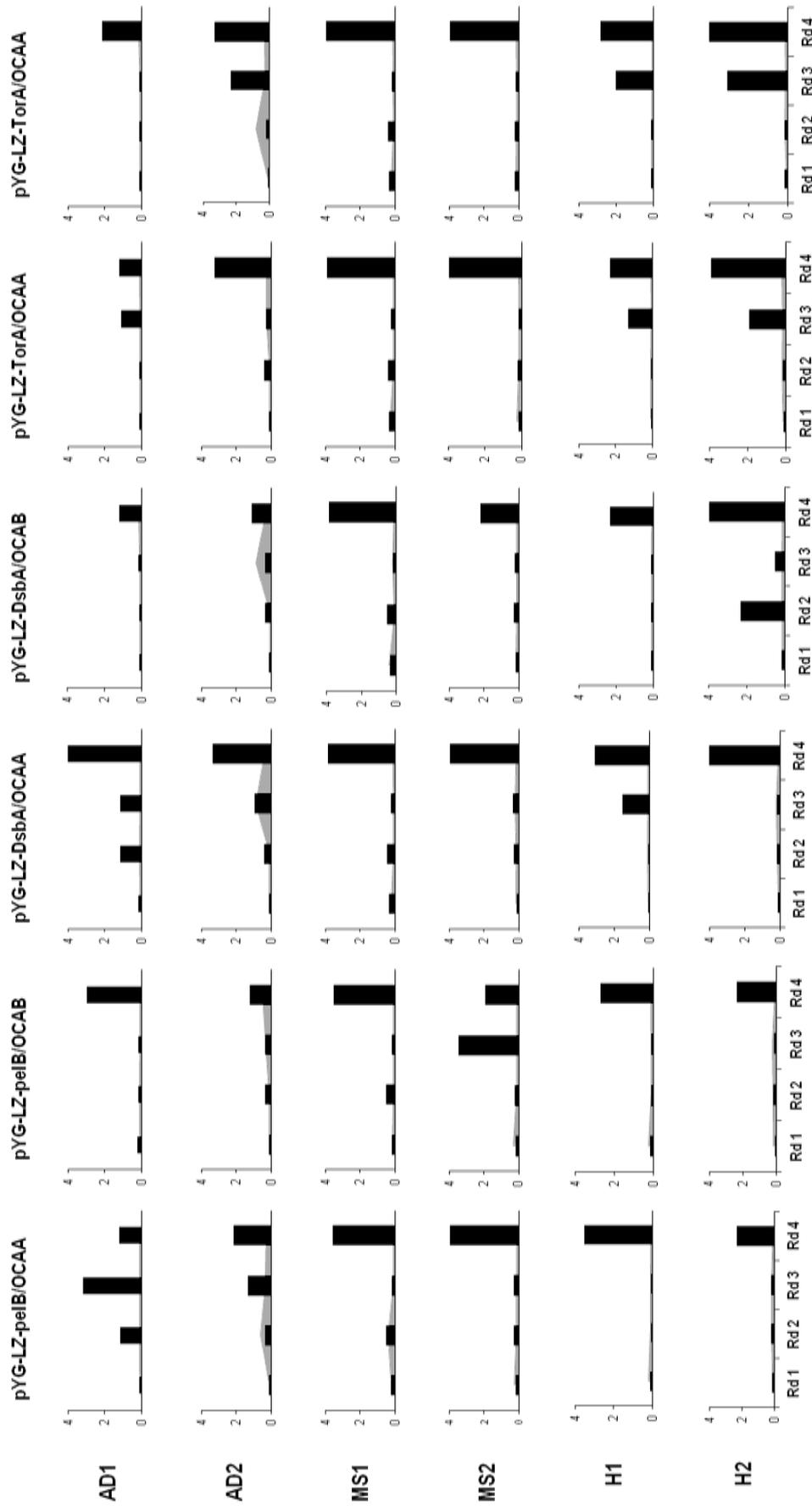
In conclusion, magnetic beads coating was considered satisfactory and ready to be applied as bait in phage display. Since the applied antibodies showed no cross-reactivity towards the other immunoglobulin class, reliable discrimination between IgG- and IgA-derived autoantigens during selection was expected.

### 5.3.2.2. Polyclonal phage ELISA

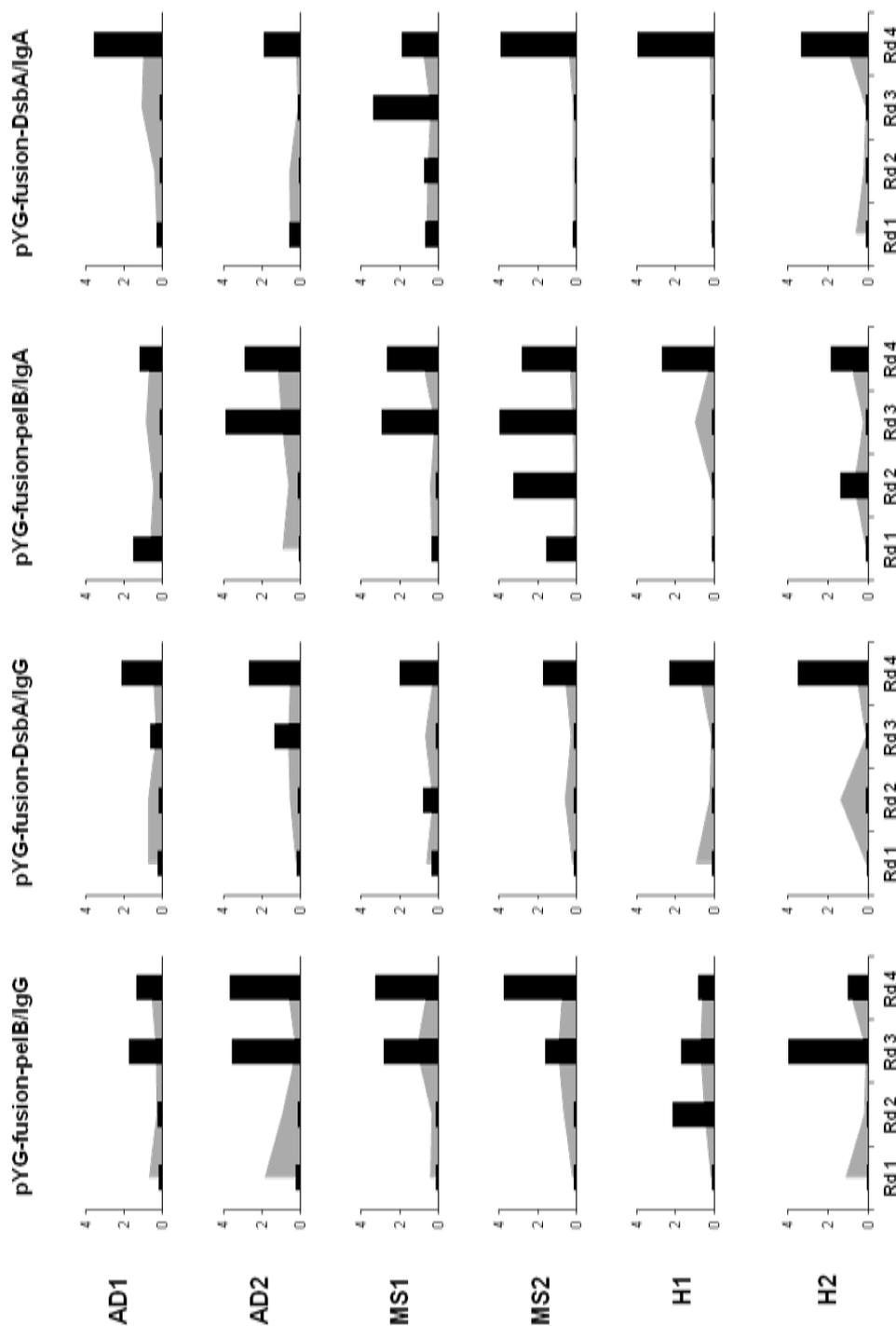
After each of the four bio-panning rounds, selection process was monitored in each sample, applying polyclonal phage ELISA, as described in section 4.4.5.2. Fig. 36 – 38 below illustrate achieved results.



**Fig. 36 Polyclonal phage ELISA results of bio-panning rounds with pYG-LZ vectors and serum IgG autoantibodies.** The six screened sera are shown on the left. Upper identifier shows accordant pYG/full-ORF phage library. Rd1 – Rd4 represent bio-panning selection rounds. Bar charts show arbitrary absorbance units, measured at 405 nm (y-axis). Samples are represented as black bars and backgrounds are depicted as grey areas.



**Fig. 37 Polyclonal phage ELISA results of bio-panning rounds with pYG-LZ vectors and serum IgA autoantibodies.** The six screened sera are shown on the left. Upper identifier shows accordant pYG/full-ORF phage library. Rd1 – Rd4 represent bio-panning selection rounds. Bar charts show arbitrary absorbance units, measured at 405 nm (y-axis). Samples are represented as black bars and backgrounds are depicted as grey areas.

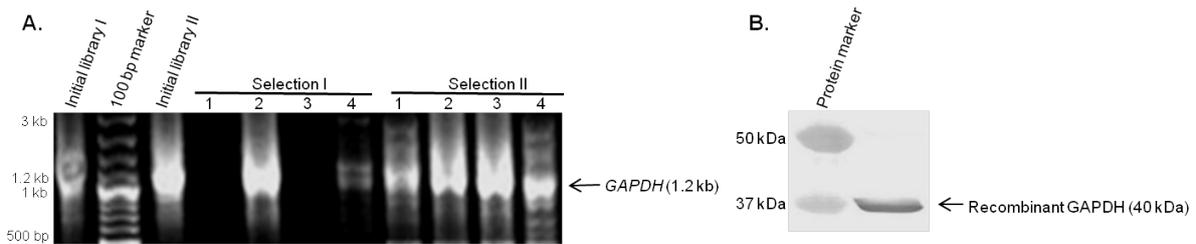


**Fig. 38 Polyclonal phage ELISA results of bio-panning rounds with pYG-fusion vectors and serum IgG/IgA autoantibodies.** The six screened sera are shown on the left. Upper identifier shows accordant pYG/full-ORF phage library. Rd1 – Rd4 represent bio-panning selection rounds. Bar charts show arbitrary absorbance units, measured at 405 nm (y-axis). Samples are represented as black bars and backgrounds are depicted as grey areas

Irrespective of vector system, antibody or blood serum, enrichment became visible usually in the 4<sup>th</sup> selection round. In ca. 25% of cases, it could already be observed in the 3<sup>rd</sup> round. Finally, 8% of the samples showed stronger signals in the 3<sup>rd</sup> than in the 4<sup>th</sup> round.

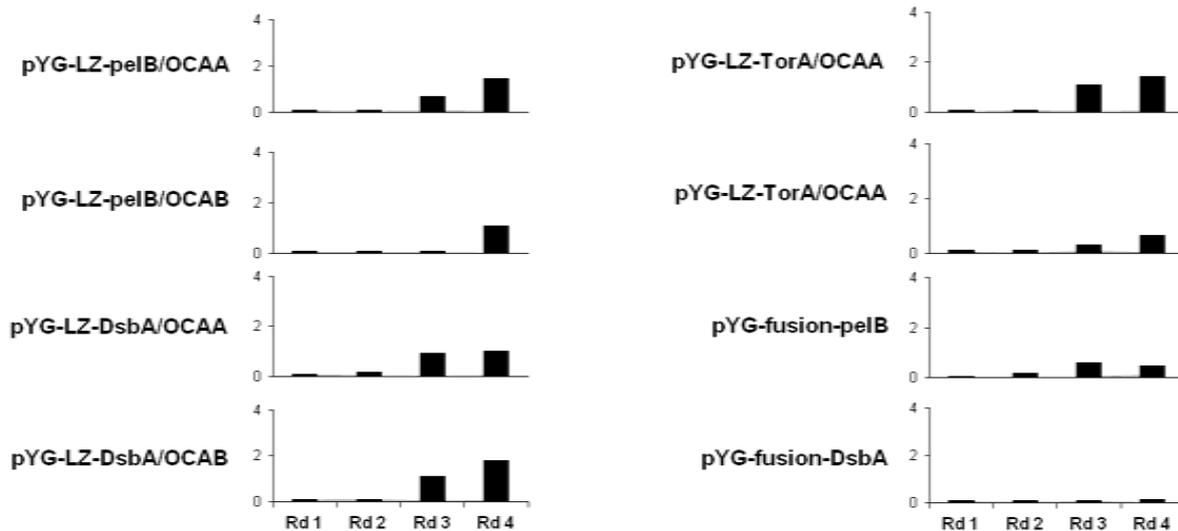
The three positive selection controls with  $\alpha$ -GAPDH,  $\alpha$ -Tubulin and  $\alpha$ -Ubiquitin mouse antibodies could not be validated in the same phage ELISA set up. Since the only commercially available  $\alpha$ -M13-HRP antibody originated from mouse as well, cross-reactivity with the primary  $\alpha$ -mouse antibody was to be expected. Furthermore, enrichment was

monitored exemplary only for the GAPDH control sample. First, open PCR reactions with a vector-specific forwards primer (Fos\_seq) and an insert-specific backwards primer (GAPDH\_back) were performed. Templates were two randomly chosen initial libraries and their correspondent *E. coli* colonies from each of the four selection rounds. Furthermore, recombinant human GAPDH was detected on a western blot in order to test the  $\alpha$ -GAPDH antibody, prior to its application in phage ELISA. Results from these two preliminary experiments are shown in Fig. 39 below.



**Fig. 39 Preliminary tests for GAPDH enrichment validation. A. Colony PCR with Fos\_seq x GAPDH\_back primers.** Initial library I: pYG-LZ-peIB/OCAA. Initial library II: pYG-LZ-TorA/OCAB. Selection I and selection II correspond to initial libraries respectively. Numbers represent selection rounds. **B. Western blot of recombinant human GAPDH**, expressed in HEK293 cells. Detection antibodies: mouse monoclonal  $\alpha$ -GAPDH and  $\alpha$ -mouse-HRP. This western blot experiment was performed by Sunniva Förster during her diploma thesis, which she completed in our research group under my supervision [225].

Thus, the presence of the *GAPDH* gene in phages of the initial libraries, as well as in cells from the bio-panning process was confirmed and general functionality of the  $\alpha$ -GAPDH antibody was approved. Next, phage ELISA was performed to visualize probable GAPDH enrichment in the pYG libraries and to validate overall selection procedure. For this, *E. coli* cultures from selection back up plates were grown up and infected with Hyperphage to produce GAPDH-presenting phages. In this case, ELISA plates were coated with the mouse  $\alpha$ -GAPDH antibody (0.5  $\mu$ g/well), incubated with GAPDH-presenting Hyperphages (50  $\mu$ l/well) and finally detected with the HRP-conjugated mouse  $\alpha$ -M13 antibody. Fig. 40 below shows achieved results.

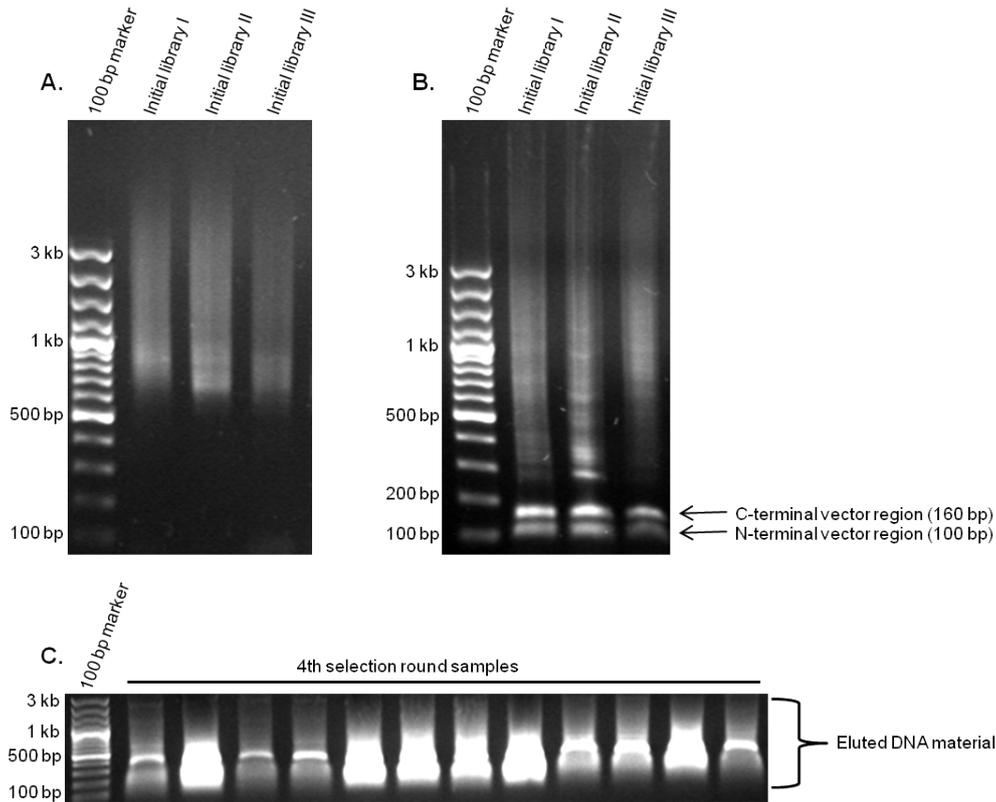


**Fig. 40 Validation of the GAPDH enrichment process.** Results from polyclonal phage ELISA with GAPDH-presenting phages are shown. Identifiers on the left represent accordant phage library. Rd1 – Rd4 represent bio-panning selection rounds. Bar charts show arbitrary absorbance units, measured at 405 nm (y-axis). Sample are shown as black bars and represent average values from double measurements. All measured background values were nearly 0 and therefore not visible.

GAPDH was also found per NGS with a relatively small amount of reads in all initial libraries: 13 reads in pYG-fusion-pelB, 2 reads in pYG-fusion-DsbA, 145 reads in pYG-LZ-pelB, 7 reads in pYG-LZ-DsbA and 24 reads in pYG-LZ-TorA. As Fig. 40 shows, in all selections, with the exception of pYG-fusion-DsbA, GAPDH-presenting phages should have been produced during the bio-panning process. Enrichment could be observed in all pYG-LZ libraries, whereas both pYG-fusion libraries demonstrated the lowest ELISA signals and no visible enrichment. Furthermore, it should be taken into consideration that overall signals were 1.5 – 2 times lower than in bead ELISAs with serum antibodies (s. Fig. 36 – 38). Still, in conclusion, the bio-panning process could be assumed as successful.

### 5.3.3. Preparation of DNA samples for sequencing on Illumina Genome Analyzer

Full-ORF DNA inserts were amplified in an open PCR reaction, as described in section 4.5.1. Applied templates were phages of the initial libraries and *E. coli* cells from the 4<sup>th</sup> selection round. Both PCR primers bound on vector backbone, thus generating ORF-inserts, flanked by vector sequences of 100 bp and 160 bp, respectively. In order to get rid of these unnecessary DNA pieces, which could potentially hamper sequencing process, restriction with the *BsrGI* endonuclease was performed after each amplification round. Finally, samples were loaded anew on an agarose gel and run very shortly to minimize the size of gel pieces to be cut out. The whole DNA material with a size between 200 bp and 4 kb was finally eluted. Fig. 41 below demonstrates an example of an NGS sample preparation process.



**Fig. 41 DNA sample preparation for Illumina sequencing. A. PCR amplification results of three initial libraries:** pYG-LZ-peIB (Initial library I), pYG-LZ-DsbA (Initial library II) and pYG-LZ-TorA (Initial library III). **B. Same three initial libraries as in A. after digestion with *BsrGI*.** Both ORF-flanking regions are clearly visible. **C. PCR amplification results of 12 enriched libraries from the 4<sup>th</sup> selection round.** Shown examples are all 6 pYG-LZ-libraries, enriched against IgG and IgA autoantibodies from the MS-Serum 170171 (s. Fig. 12.A). DNA was eluted between ca. 200 bp and ca. 4 kb.

Final DNA concentration in the eluted samples ranged between 20 ng/ $\mu$ l and 90 ng/ $\mu$ l. Median value was estimated at 43 ng/ $\mu$ l. 30  $\mu$ l of each eluate was sent to Source BioScience/Imagenes for further sample processing and subsequent sequencing on an Illumina platform.

### 5.3.4. Enrichment analysis of NGS results

Achieved read numbers per library ranged between one thousand and 17.9 million reads with a median value of 2.3 million reads. To cull genes, which should have been enriched during selection, a cut off of minimum 100 reads per gene was fixed in the results from the 4<sup>th</sup> selection round. Next, read numbers per gene between the initial library and the 4<sup>th</sup> round were compared. Genes with a higher number of reads in the 4<sup>th</sup> round than in the initial library were considered potentially enriched. Finally, genes with a difference of at least 1,000 reads between the initial library and the accordant 4<sup>th</sup> round were designated truly enriched

and thus, highly interesting for further analysis. Table 13 below summarizes NGS analysis results up to this point for each donor cohort.

**Table 13. Analysis of enriched genes in phage display.** Shown numbers are cumulative gene numbers in the accordant cohort, non-redundant in respect of pYG-library and screening antibody. “Uniques” represent genes, identified only in one cohort with an enrichment of over 1,000 reads. All uniques with an enrichment of over 1 000 reads are listed in the supplementary material.

	4 <sup>th</sup> round $\geq$ 100 reads	Reads 4 <sup>th</sup> round > Reads initial library	[Reads 4 <sup>th</sup> round – Reads initial library] $\geq$ 1,000	Uniques with enrichment $\geq$ 1,000 reads
AD	910	753	190	74
MS	782	602	155	12
H	571	363	104	9

Thus, a comparable number of enriched genes could be identified with NGS in each of the three cohorts. Next analysis step was to take a closer look at the enriched genes and finally to select appropriate biomarker candidates for expression and re-validation in ELISA assay.

## 5.4. Validation of identified biomarker candidates

### 5.4.1. Selection of biomarker candidates for recombinant bacterial expression

Criteria, applied for selecting potentially interesting candidates for single expression and re-validation from the phage display screening were the following:

- Unique genes: Only genes, enriched in one of the three cohorts and in none of the residual two respectively, were considered.
- High degree of enrichment: Only the top 20 unique genes with the highest difference of read numbers between initial library and accordant 4<sup>th</sup> round library were further analyzed.
- Biological coherence in the accordant disease context: Each of the unique Top 20 genes was examined in regard of any known publications, suggesting its functional correlation to the correspondent disorder (Sources: Pub Med, UniProt). Beside direct disease-linkage also indications such as neurodegeneration, neurological and mental disorders, age, neuroinflammation, immunological context in general and autoimmunity in particular were taken into consideration.
- Positive sera number: Was the hit gene positive, hence enriched, in one or in both screened sera?

- Correlation to the macroarray screening: Is the hit gene also expressed on the macroarray filter? If yes, is it annotated in frame and how did it perform in the macroarray screening?

Also hit antigens from the macroarray screenings were selected for further validation. Priorities were set as already described in chapter 5.2. Finally, candidates for expression and re-validation were selected from the “Highly specific and unique” group, listed in tables 19 – 21 in the supplementary material.

Following Table 14 summarizes potential biomarker candidates, which were finally chosen from both screenings according to the priority criteria, described above. These were 9 antigens from Alzheimer’s screening (AD) and 3 antigens from Multiple Sclerosis screening (MS).

**Table 14. Potential biomarker candidates, chosen for re-validation.**

Biomarker candidate	Screening platform	Disease Cohort	Screening autoantibody	No. positive sera
CCDC50	Macroarray	AD	IgA	2
PRDX1	Macroarray	AD	IgA	2
TANK	Macroarray	AD	IgA	2
TRAF4	Macroarray	AD	IgG	2
ANXA2	Phage Display	AD	IgA	1
DTNBP1	Phage Display	AD	IgA	1
GDI1	Phage Display	AD	IgA	1
NDRG4	Phage Display	AD	IgA	1
PAX6	Phage Display	AD	IgA	1
ANKHD1	Macroarray	MS	IgA	2
DEAF1	Macroarray	MS	IgA	2
IMPACT	Macroarray	MS	IgA	2

#### 5.4.2. Recombinant bacterial expression, IMAC purification and antigen ELISA of selected biomarker candidates

Frame and construct accuracy were confirmed for all 12 constructs via Sanger sequencing prior to expression. This was then performed in the appropriate *E. coli* strain, as described in chapter 4.2.1. Since expression products could be presumably very diverse in regard of polypeptide length, folding characteristics, expression and degradation rate and/or possible

toxic effects on the host, three different expression conditions after IPTG induction were tested for each of the candidates, in order to ascertain the most optimal one. These were:

- I. 3 h at 37°C and 220 rpm
- II. ON at 37°C and 200 rpm
- III. ON at 16°C and 200 rpm

CCDC50 was the only gene, which could not be expressed under any of the three conditions. GDI1 and PAX6 showed only slightly better expression yields under condition I (3h/37°C) compared to condition II (ON/37°C). The residual 10 genes were expressed best under condition II (ON/37°C). PRDX1, TRAF4, ANXA2 and DTNB1 could be successfully expressed under condition III (ON/16°C) as well, but with a much worse performance than under I or II. Extensive degradation products were observed only in the case of IMPACT under all three conditions. Thus, it was concluded that under condition II (ON/37°C) the best or at least very good results could be reached for all 12 expressible antigens. Consequently, this condition was used in all following large-scale expression assays.

All antigens were expressed in the cytoplasm and cell lysis was performed under denaturing conditions applying 8 M Urea buffer. Next, expression and purification results for 10 autoantigens will be presented in detail: 8 from AD screenings and 2 from MS screenings. CCDC50 (AD candidate) was excluded, since no expression could be achieved at all. IMPACT (MS candidate) was excluded, because of massive degradation products, already visible in cell lysates prior to purification (data not shown).

Antigen ELISAs were performed with all available 60 human sera (20 sera per donor group). 1 µg antigen was coated per well and 100 µl filling volume, respectively (s. Chapter 4.2.9. for details). Measurements were made at 405 nm every five minutes in a time scale between 5 min and 60 min after adding the ABTS substrate. However, preliminary antibody tests showed best performances at the 60 min point (s. Chapter 5.4.3.). All measurements were made in duplicate on two separate ELISA plates. Wells, coated with purified human IgG or IgA served as positive controls. Wells, coated with an antigen and directly incubated with the detecting HRP-antibody, i.e. without human serum, served to determine general ELISA background. Both controls were performed anew on each ELISA plate. All following graphs show adjusted signal values, resulting from subtracting the ELISA background and subsequent estimation of the average value from the double measurement.

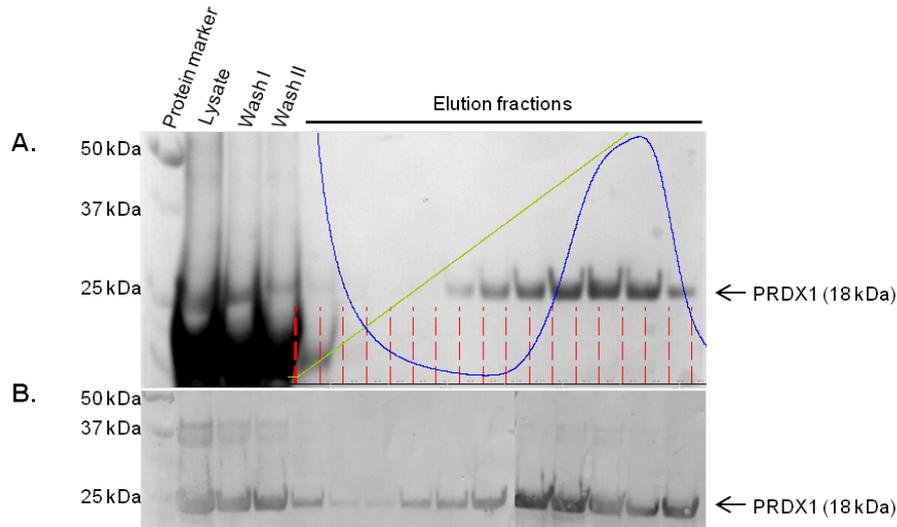
#### **5.4.2.1. PRDX1**

The PRDX1 (peroxiredoxin 1) gene is represented by 20 different cDNA fragments on the macroarray filters. 15 clones are annotated “in frame” and 5 are annotated “off frame”, respectively. 6 “in frame” clones were hit positive with two different AD-sera in an IgA-

screening. 2 “off frame” clones were hit positive with the same AD-sera, but also with one MS-serum, again in an IgA-screening. Thus, PRDX1 was assigned “highly potential” in the macroarray screening. The PRDX1-full-ORF could be identified in the initial phage display libraries, but no enrichment was observed in any of the cohorts.

Peroxiredoxins are an essential part of the oxidative stress pathway and broadly expressed in brain tissues. Elevated expression levels have been shown in several neurodegenerative diseases, including Alzheimer’s [226]. Furthermore, special rat neurons, resistant to A $\beta$ -toxicity, demonstrated elevated PRDX1 expression levels as well. But also in post mortem cortical tissues of AD-patients increased levels of PRDX1 have been estimated [227]. These findings suggest a potential role of peroxiredoxins and PRDX1 in particular in possible neuroprotective mechanisms in the course of AD.

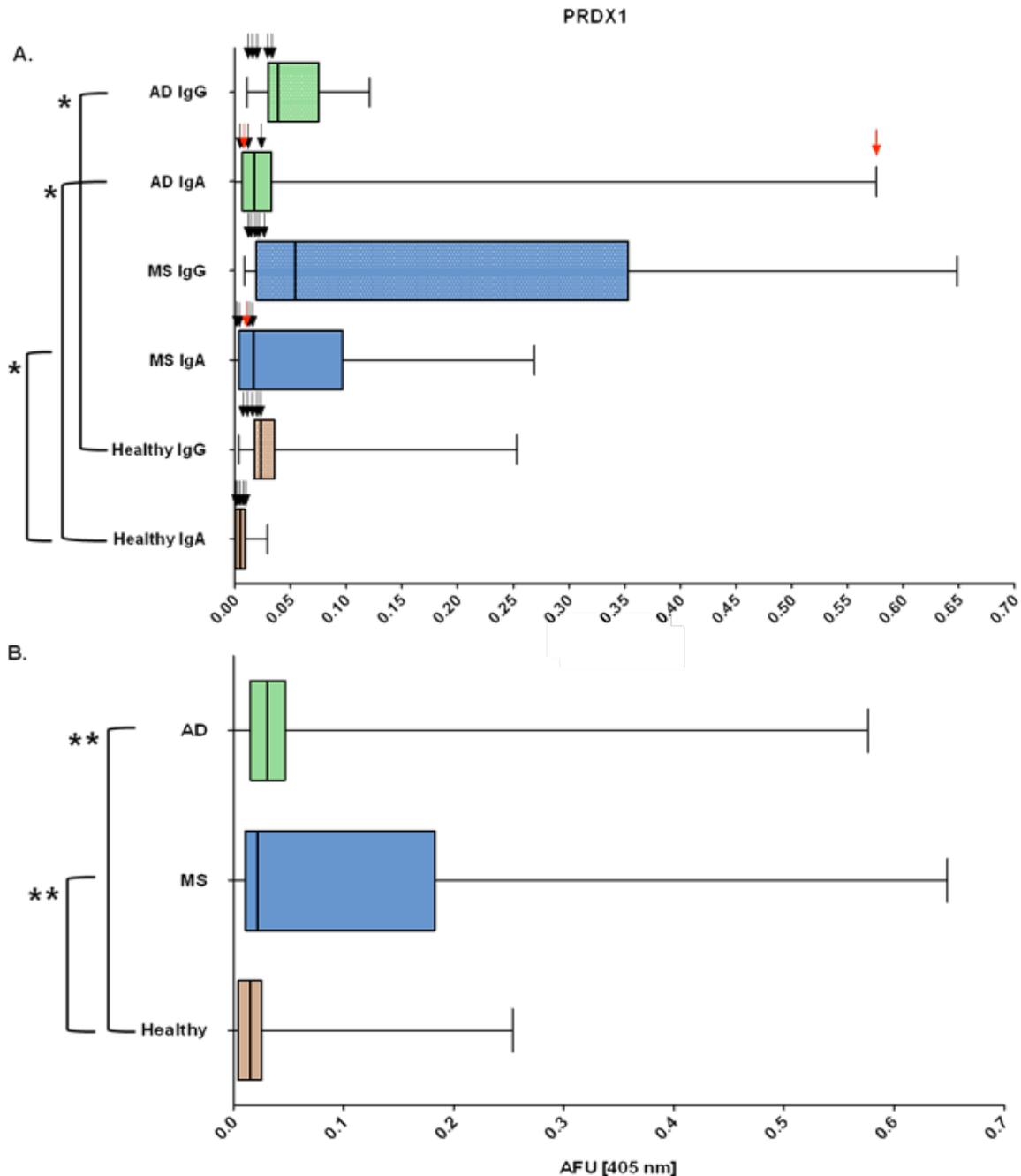
Accordingly, PRDX1 was chosen in this study for recombinant expression and further re-validation as a potential biomarker candidate for Alzheimer’s disease. The selected cDNA clone consisted of the whole mRNA sequence, including also the 5’-UTR region. Thus, the 330 bp long 5’-UTR region, the 117 bp long CDS region and the 36 bp long His-tag sequence amounted to a 483 bp long insert, resulting in a ca. 18 kDa big expression product. Following Fig. 42 below summarizes the results from PRDX1 expression in *E. coli* and subsequent FPLC purification via a His-tag. The FPLC chromatogram is shown only for this antigen as a demonstration.



**Fig. 42 PRDX1 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, two subsequent washing steps and elution fractions are shown. Accordant FPLC chromatogram overlays gel picture, corresponding roughly to it. Elution fractions are highlighted in red, relative protein absorption curve at 280 nm is shown in blue and gradually increasing elution buffer curve (i.e. decreasing pH gradient) is shown in green. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, PRDX1 could be sufficiently overexpressed and successfully purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Fractions under the elution peak were pooled together and concentrated to an overall volume of approx. 1 ml. Final protein concentration accounted for 940  $\mu$ g/ml.

For re-validation, ELISA plates were coated with the PRDX1 antigen. Incubation with all available human blood sera and detection with HRP-conjugated  $\alpha$ -hu-IgG and  $\alpha$ -hu-IgA, respectively, followed. ELISA signal distributions in the different six groups, differentiated by donor cohort and antibody class, were statistically analyzed pairwise with the non-parametric Wilcoxon-Mann-Whitney test for each of the two antibodies. Interval of confidence was set at 95%. Two-sided P-values under 0.05 were considered significant and those under 0.01 as especially significant. Confirmation of statistical results was performed with the Kruskal-Wallis test, extended with a Dunn correction. Next figure shows summarized ELISA results as box-and-whiskers-plots.



**Fig. 43 Signal distribution and statistical analysis of antigen ELISA with PRDX1.** Both x-axes show arbitrary fluorescence units (AFU), measured at 405 nm. Plot borders correspond to the standard five distribution parameters: minimum value, 25% percentile, median value, 75% percentile and maximum value. On the left side, only group pairs are shown, which yielded significant P-values in the Wilcoxon-Mann-Whitney test ( $>0.05$ ). Pairs with a P-value between 0.05 and 0.01 (significant) are highlighted with one star symbol (\*). Those with a P-value, smaller than 0.01 (especially significant), are highlighted with two star symbols (\*\*). **A. ELISA signal distributions in each cohort with each antibody class.** Total number of values in each plot accounts for 20 signals from 20 human sera from the accordant cohort, tested with the accordant antibody. Arrows show signal values from the same five sera in the macroarray screening. Arrows, highlighted in red, are sera with hits in the macroarray

screening. **B. ELISA signal distributions in each cohort in both antibody classes together.** Total number of values in each plot accounts for 40 signals from 20 human sera from the accordant cohort.

Above Fig. 43 shows distinctly that although the PRDX1 antigen was selected in an IgA-screening, in ELISA format it obviously reacted also with IgG serum antibodies. Furthermore, also sera from the MS and Healthy cohorts, which showed no reactions during macroarray screening, were reactive in this experimental set up. Statistical analysis revealed that the only three pairs of data groups with a significant P-value were AD ↔ Healthy with both antibody classes and MS ↔ Healthy with IgA (Fig. 43/B.).

In order to compare results from the three cohorts on a more comprehensive level, signal values from measurements with the particular antibody class were pooled together. This led to three signal distribution groups, each including 40 values (Fig. 43/B.). Interestingly, achieved P-values in the two pairs AD ↔ Healthy and MS ↔ Healthy were now one order of magnitude smaller. Hence, even more significant distinction between the pairs was observed with a striking differentiation between Diseased ↔ Non-diseased. On the other hand, no significant P-values were estimated in the pairs AD ↔ MS.

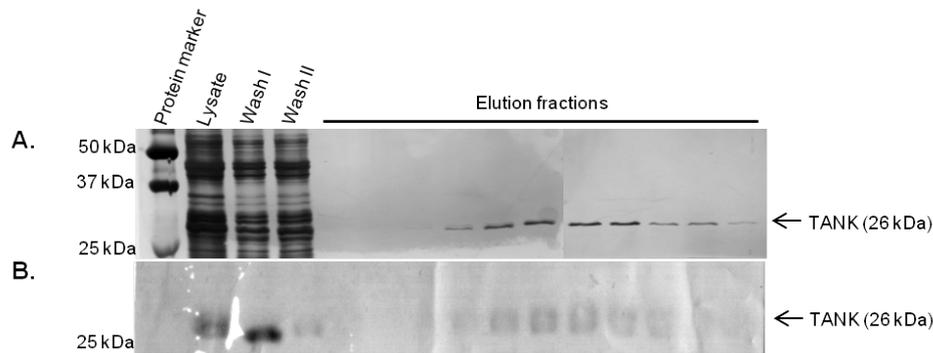
#### 5.4.2.2. TANK

The TANK (TRAF family member-associated NFκB activator) gene is represented by only one “in frame” cDNA clone on the macroarray filter, which is spotted twice as a real duplicate. Both identical clones were hit positive with two different AD-sera in an IgA-screening. Thus, also TANK was assigned “unique and highly potential” in the macroarray screening. Interestingly, this clone has been also hit positive in an earlier IgA-screening in our research group with pooled sera from patients with Primary Glomerulonephritis, which is an autoimmune disorder. The TANK-full-ORF is not present in the source library.

So far, TANK has not been found to be correlated directly to AD or any other neurodegenerative disease. However, as regulator of the TRAF-proteins it is part of the NFκB-pathway, thus playing a co-role in numerous immune response processes, such as inflammation and apoptosis [228]. Furthermore, as a modulator of synaptic plasticity and function, NFκB has been implemented in memory and learning [229, 230]. Finally, TANK was very recently shown to be ubiquitously expressed in mouse DRG sensory neurons and to have elevated transcription and translation levels after peripheral nerve injury [231]. Thus, there are evidences available, confirming the potential role of TANK in neurological and possibly psychiatric processes as well.

The TANK cDNA clone included a 669 bp long piece from the C-terminal CDS region, representing approx. 52% of it. Thus, the expected molecular mass of the expression product

was approx. 26 kDa. Following Fig. 44 summarizes the results from TANK expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 44 TANK expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, TANK could be overexpressed and successfully purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Fractions under the elution peak were pooled together and concentrated to an overall volume of approx. 500  $\mu$ l. Final protein concentration accounted for only 21  $\mu$ g/ml. Subsequent attempts to gain more expression product and to increase purification efficiency did not bring desired results. Therefore, TANK was excluded from ELISA re-validation.

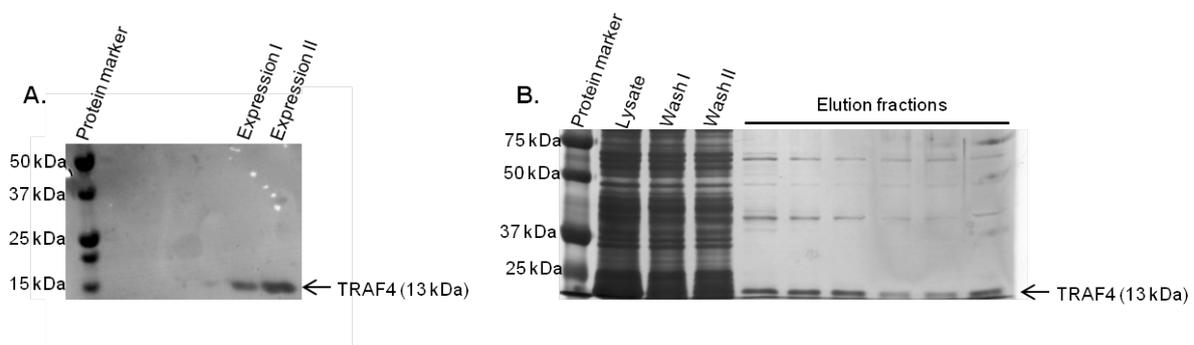
#### 5.4.2.3. TRAF4

The TRAF4 (TNF-receptor associated factor 4) gene is represented by nine different cDNA clones on the macroarray filters. 6 clones are annotated “in frame” and 3 are annotated “off frame”. 2 “in frame” clones were hit positive with two different AD-sera in an IgG-screening and once with same serum in an Ig-screening. Thus, TRAF4 was assigned “highly potential” for IgG in the macroarray screening. The TRAF4-full-ORF could also be identified in the initial phage display libraries in a very low number of reads: between 5 and 25 reads per library. In one of the sequencing rounds a slight enrichment in both AD sera could be estimated in the 2<sup>nd</sup> selection round: 228 and 243 reads per serum. In the 4<sup>th</sup> round only one single read was identified for the TRAF4 gene. In selections with MS and Healthy sera no reads were found at any stage.

TRAF4 belongs to the TRAF-protein family and with this to the NF $\kappa$ B signaling pathway, analog to TANK, as described above. In CNS tissues TRAF proteins have been shown to interact with neurotrophin receptors and to play important role in neuronal growth and

apoptosis [232]. Furthermore, TRAF proteins are major modulators in the JNK signaling pathway as well. Like the NF $\kappa$ B pathway, it is a central regulatory mechanism in numerous processes, related to cell cycle and inflammation. These are in particular: apoptosis, T cell differentiation, cellular stress response, cell growth and proliferation, cytokine production and several others [233, 234]. Finally, the JNK pathway has been recently directly related to tau pathology in AD [235].

The TRAF4 cDNA clone included a 1,326 bp long piece from the C-terminal CDS region, representing approx. 94% of it. However, Sanger sequencing validation revealed that only the first 300 bp of the insert were in the correct reading frame. Then a frame shift followed, leading to 27 random base pairs and finally a stop codon. Thus, the expected size of the expression product was approx. 13 kDa, taking the N-terminal His-tag into consideration. Following Fig. 45 summarizes the results from TRAF4 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 45 TRAF4 expression and affinity purification.** **A.** Western blot with lysates from two clones, overexpressing the TRAF4 construct, prior to purification. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP. **B.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown.

Thus, TRAF4 gene fragment could be successfully overexpressed in *E. coli*. Yet purification was considered insufficient. As demonstrated in Fig. 45/B, elution fractions still contained plenty of other proteins and/or di-/trimers of the TRAF4 polypeptide. Consequently, this antigen was excluded from ELISA validation.

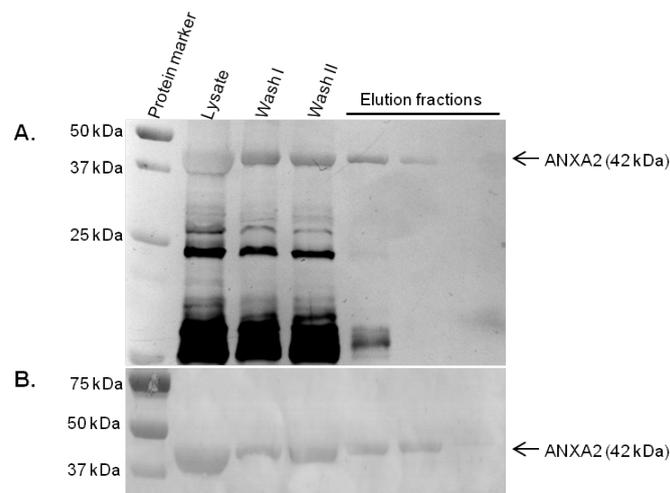
#### 5.4.2.4. ANXA2

The ANXA2 (annexin A2) full-ORF was found strongly enriched in an IgA-selection with one of the AD sera. Only 46 reads were identified in the initial library, 6,074 reads were found after the 2<sup>nd</sup> selection round and finally 37,873 reads were found after the 4<sup>th</sup> round. No other serum showed any enrichment effects. The gene is represented by 3 clones on the

macroarrays, all of them being annotated “in frame”. None of them was hit positive in any of the macroarray screenings.

Proteins of the annexin family are  $\text{Ca}^{2+}$ -dependent phospholipid-binding polypeptides with a general role in cellular growth, shape and motility. Further intracellular functions are organization and transport of vesicles, exo-/endocytosis and ion channel formation [236]. In the extracellular space, annexins are involved in signal transduction during apoptosis and inflammation [237]. Very recently, annexin A5 was shown to have elevated protein levels in blood plasma of Alzheimer’s and Dementia with Lewy Bodies patients, suggesting its potential as a biomarker [238]. Also ANXA2 has already been directly correlated to neurodegeneration [239].

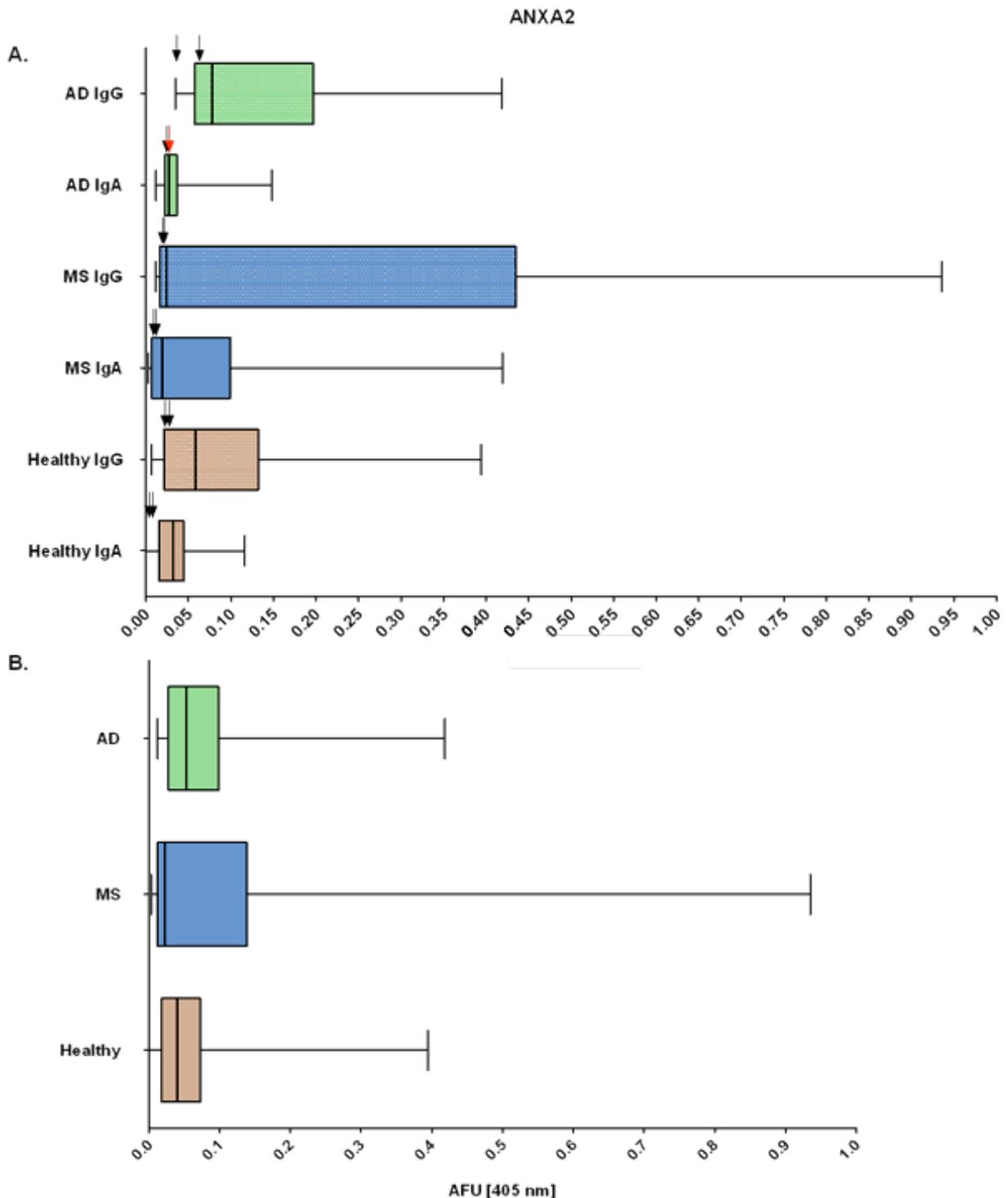
The full-ORF ANXA2-gene could be identified as a transcript variant 2 (RefSeq analysis). The cloned gene was 1,020 bp long. Together with the N-terminal Avi-His<sub>6</sub>-tag (66 bp) and the following LR-linker (57 bp) it made up to a 1,143 bp long insert or a 42 kDa big expression product. Following Fig. 46 summarizes the results from ANXA2 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 46 ANXA2 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, ANXA2 could be overexpressed and successfully purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Fractions under the elution peak were pooled together and concentrated to an overall volume of approx. 1 ml. Final protein concentration accounted for 1,626  $\mu\text{g/ml}$ .

ELISA re-validation with ANXA2 as coating antigen was performed as described above. Achieved signals were again statistically analyzed under same conditions. Next figure shows summarized ELISA results as box-and-whiskers-plots.



**Fig. 47 Signal distribution and statistical analysis of antigen ELISA with ANXA2.** Both x-axes show arbitrary fluorescence units (AFU), measured at 405 nm. Plot borders correspond to the standard five distribution parameters: minimum value, 25% percentile, median value, 75% percentile and maximum value. **A. ELISA signal distributions in each cohort with each antibody class.** Total number of values in each plot accounts for 20 signals from 20 human sera from the accordant cohort,

tested with the accordant antibody. Arrows show signal values from the same two sera in the phage display screening. Arrows, highlighted in red, are sera with enriched antigens in the phage display screening. **B. ELISA signal distributions in each cohort in both antibody classes together.** Total number of values in each plot accounts for 40 signals from 20 human sera from the accordant cohort.

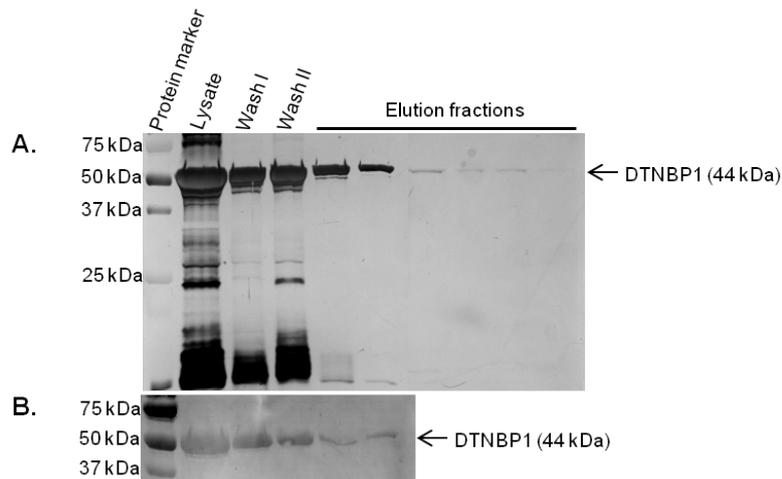
Above Fig. 47 shows distinctly that although the ANXA2 antigen was selected in an IgA-screening, in ELISA format it not only reacted with IgG serum antibodies, but signals were even stronger as with IgA and overall more sera were IgG- than IgA-reactive. Furthermore, also sera from the MS and Healthy cohorts, which showed no enrichment in phage display, were reactive in this experimental set up. Statistical analysis revealed that none of the compared pairs could achieve a significant P-value, smaller than 0.05. Also comparison of consolidated data (Fig. 47/B.) did not reveal any statistically significant differences between any of the data pairs.

#### 5.4.2.5. DTNBP1

The DTNBP1 (dystrobrevin binding protein 1, also known as dysbindin) full-ORF was found strongly enriched in an IgA-selection with one of the Alzheimer sera: 465 reads were identified in the initial library, 3,668 reads after the 2<sup>nd</sup> selection round and finally 97,316 reads after the 4<sup>th</sup> round. Also after the 4<sup>th</sup> round of an IgA-selection with the second AD serum 1,624 reads of the DTNBP1 gene were found. Both healthy sera showed no enrichment. The gene is represented by only 1 “in frame” clone on the macroarrays and was not hit positive in any of the screenings.

DTNBP1 is strongly expressed in brain and neuronal tissues, being responsible for morphological extension of neurites (dendrites and axons), trafficking of synaptic vesicles, neurotransmitter release and other neuron functions, related to information forwarding [240, 241]. The gene gained much attention with the discovery that genetic defects lead to altered glutamatergic transmission in brain and was associated with susceptibility to schizophrenia [242]. No direct linkage of DTNBP1 to AD has been established so far. However, recently it was associated with variations in hippocampal and prefrontal grey matter volume in humans [243], altered cognitive abilities [244] and psychosis [245].

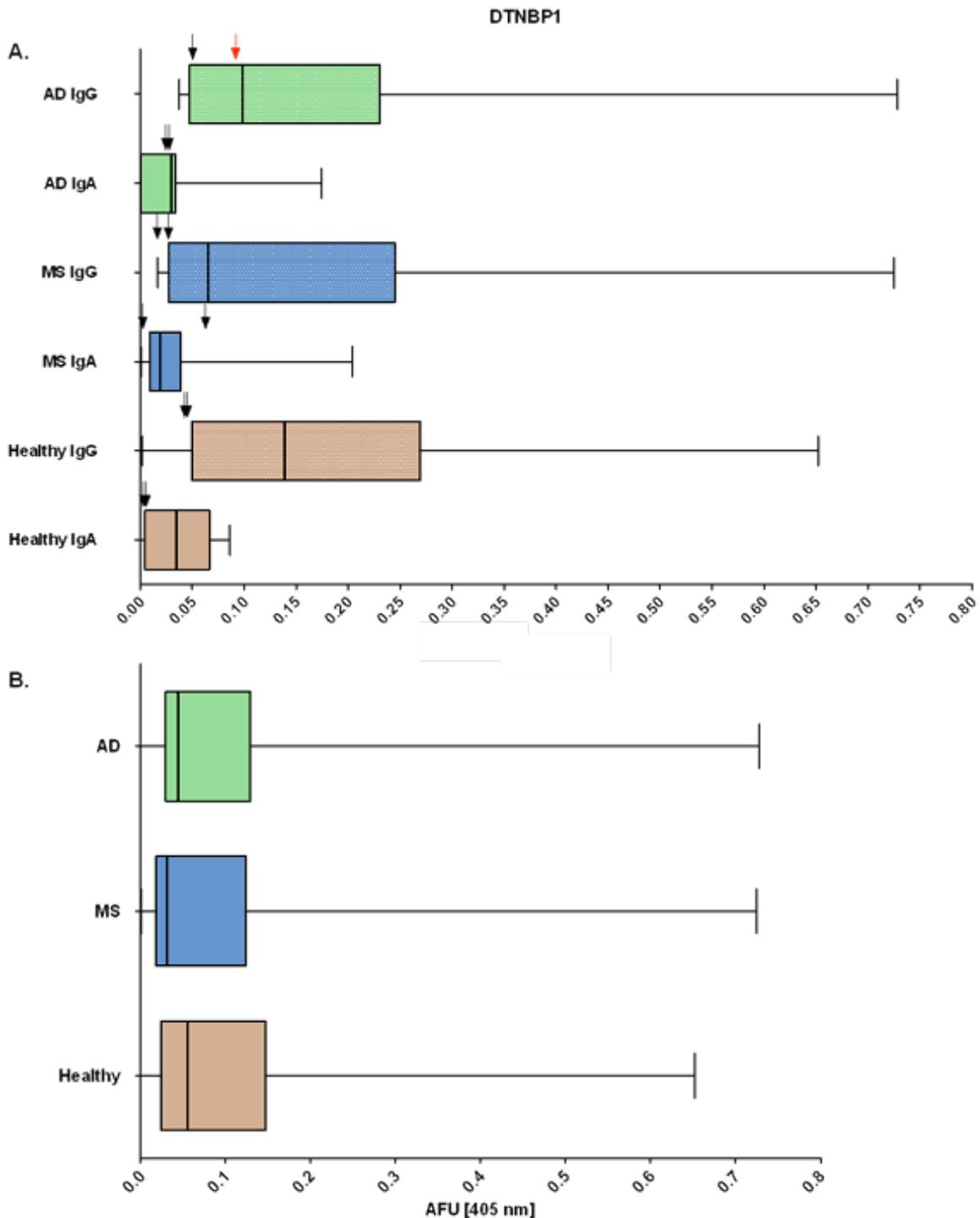
The full-ORF DNTBP1-gene could be identified as a transcript variant 1 (RefSeq analysis). The cloned gene was 1,053 bp long, which results in a protein of 39.5 kDa. Together with the N-terminal tag-sequence it made up to a ca. 44 kDa big expression product. Following Fig. 46 summarizes the results from DTNB1 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 48 DTNBP1 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, DTNBP1 could be overexpressed and successfully purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Clean fractions under the elution peak were pooled together and concentrated to an overall volume of approx. 1 ml. Final protein concentration accounted for 557  $\mu$ g/ml.

ELISA re-validation with DTNBP1 as coating antigen was performed as described above. Achieved signals were again statistically analyzed under same conditions. Next figure shows summarized ELISA results as box-and-whiskers-plots.



**Fig. 49 Signal distribution and statistical analysis of antigen ELISA with DTNBP1.** Both x-axes show arbitrary fluorescence units (AFU), measured at 405 nm. Plot borders correspond to the standard five distribution parameters: minimum value, 25% percentile, median value, 75% percentile and maximum value. **A. ELISA signal distributions in each cohort with each antibody class.** Total number of values in each plot accounts for 20 signals from 20 human sera from the accordant cohort, tested with the accordant antibody. Arrows show signal values from the same two sera in the phage display screening. Arrows, highlighted in red, are sera with enriched antigens in the phage display screening. **B. ELISA signal distributions in each cohort in both antibody classes together.** Total number of values in each plot accounts for 40 signals from 20 human sera from the accordant cohort.

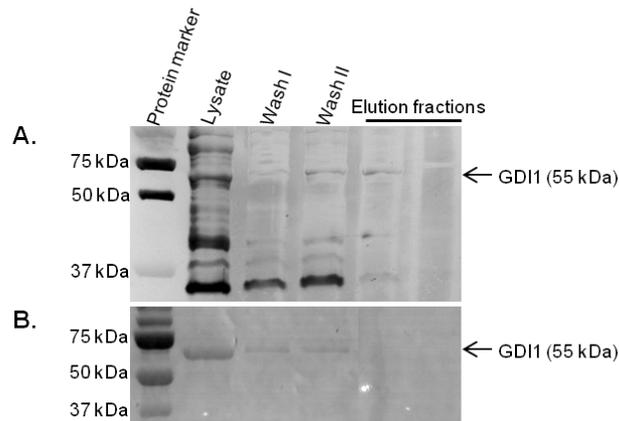
Above Fig. 49 shows distinctly that although the DTNPB1 antigen was selected in an IgA-screening, in ELISA format it not only reacted with IgG serum antibodies, but signals were even stronger as with IgA and overall more sera were IgG- than IgA-reactive. Furthermore, also sera from the MS and Healthy cohorts, which showed no enrichment in phage display, were reactive in this experimental set up. Surprisingly strong were in this case especially signals in IgG-ELISA with healthy sera. Statistical analysis revealed that none of the compared pairs could achieve a significant P-value, smaller than 0.05. Also comparison of consolidated data (Fig. 49/B.) did not reveal any statistically significant differences between any of the data pairs.

#### 5.4.2.6. GDI1

The GDI1 (GDP dissociation inhibitor 1) full-ORF was found strongly enriched in an IgA-selection with one of the Alzheimer sera: 576 reads were identified in the initial library, 3,772 reads after the 2<sup>nd</sup> selection round and finally 4,045 reads after the 4<sup>th</sup> round. Other sera showed no enrichment. The gene is represented by 10 clones on the macroarrays, all annotated “in frame”. In an IgA-screening one clone was hit positive with the same AD serum, with which the GDI1-full-ORF was enriched in phage selection. Another GDI1-clone was hit in an IgG-screening with a different AD serum. MS and healthy sera did not show any positive reaction in any screening.

General function of the GDI proteins is negative regulation of the GDP-GTP exchange at proteins of the Rab family. In turn, Rab are small G-proteins, which are involved in intracellular vesicular trafficking. GDI1 is a brain specific protein, predominantly expressed in neural and sensory tissues [246]. Mutations in GDI1 lead to a X-linked mental retardation [247]. Recently its absence was also linked to a decline in synaptic plasticity and to memory deficit [248]. On the other hand, Rab proteins have been shown to be upregulated in postmortem brains of AD patients [249]. Moreover, it could be recently demonstrated that Rab11 vesicles are transporter of A $\beta$  and thus central part of the intracellular A $\beta$  clearance in lysosomes [250].

The full-ORF GDI1-gene could be identified as a transcript variant 1 (RefSeq analysis). The cloned gene was 1,341 bp long (51 kDa). Together with the N-terminal tag-sequence it made up to a ca. 55 kDa big expression product. Following Fig. 48 summarizes the results from GDI1 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 50 GDI1 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, GDI1 gene fragment could be successfully expressed in *E. coli*. Yet purification was considered unsuccessful, since aim protein could only be detected in the flowthrough and wash fractions, but not in the elution fractions. Consequently, this antigen was excluded from ELISA validation.

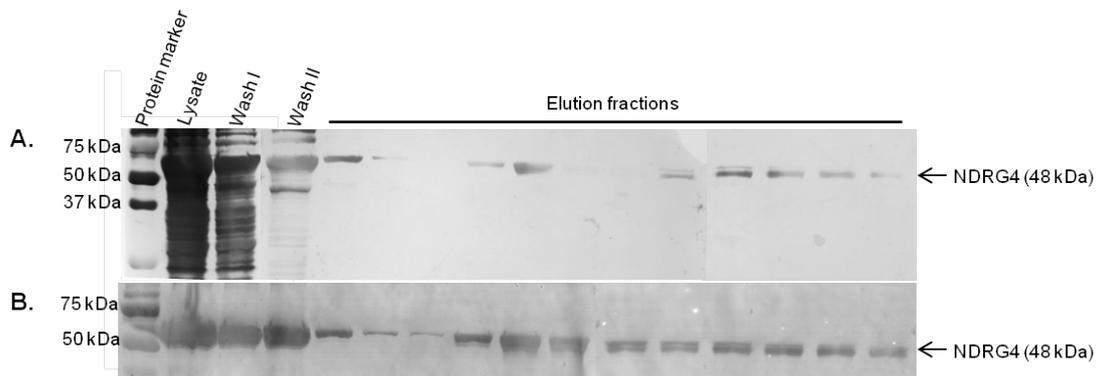
#### 5.4.2.7. NDRG4

The NDRG4 (N-myc downstream regulated gene 4 protein) full-ORF was found strongly enriched in an IgA-selection with one of the Alzheimer sera: 510 reads were identified in the initial library, 2,747 reads after the 2<sup>nd</sup> selection round and finally 104,691 reads after the 4<sup>th</sup> round. With the second AD serum 2,290 reads were found after the 4<sup>th</sup> selection round. Reads in the 4<sup>th</sup> selection round were also identified with healthy sera (871 and 1,258 reads per serum), but no enrichment was visible. The gene is represented by 6 clones on the macroarrays, all annotated “off frame”. None of them was hit positive during screening.

Proteins of the NDRG family belong to the  $\alpha/\beta$ -hydrolase superfamily. Their exact functions has not been fully elucidated, yet all family members seem to play central role in cell proliferation, cell cycle, cell differentiation and stress response [251]. Unlike other family members, which are ubiquitous, NDRG4 is almost exclusively expressed in brain and heart tissues and predominantly in astrocytes [252]. There they are part of the cell cycle progression and survival mechanisms and have been correlated to glioblastoma development [253]. Further interesting function of the NDRG proteins is their involvement in inflammatory processes, in particular in allergy and anaphylaxis [254]. Finally, NDRG4 was shown to be elevated in AD affected brains [255]. It was also linked to neurodegeneration

and axon survival in response to glucocorticoids [256]. Also NDRG4 was associated with neuronal differentiation and neurite formation [257, 258].

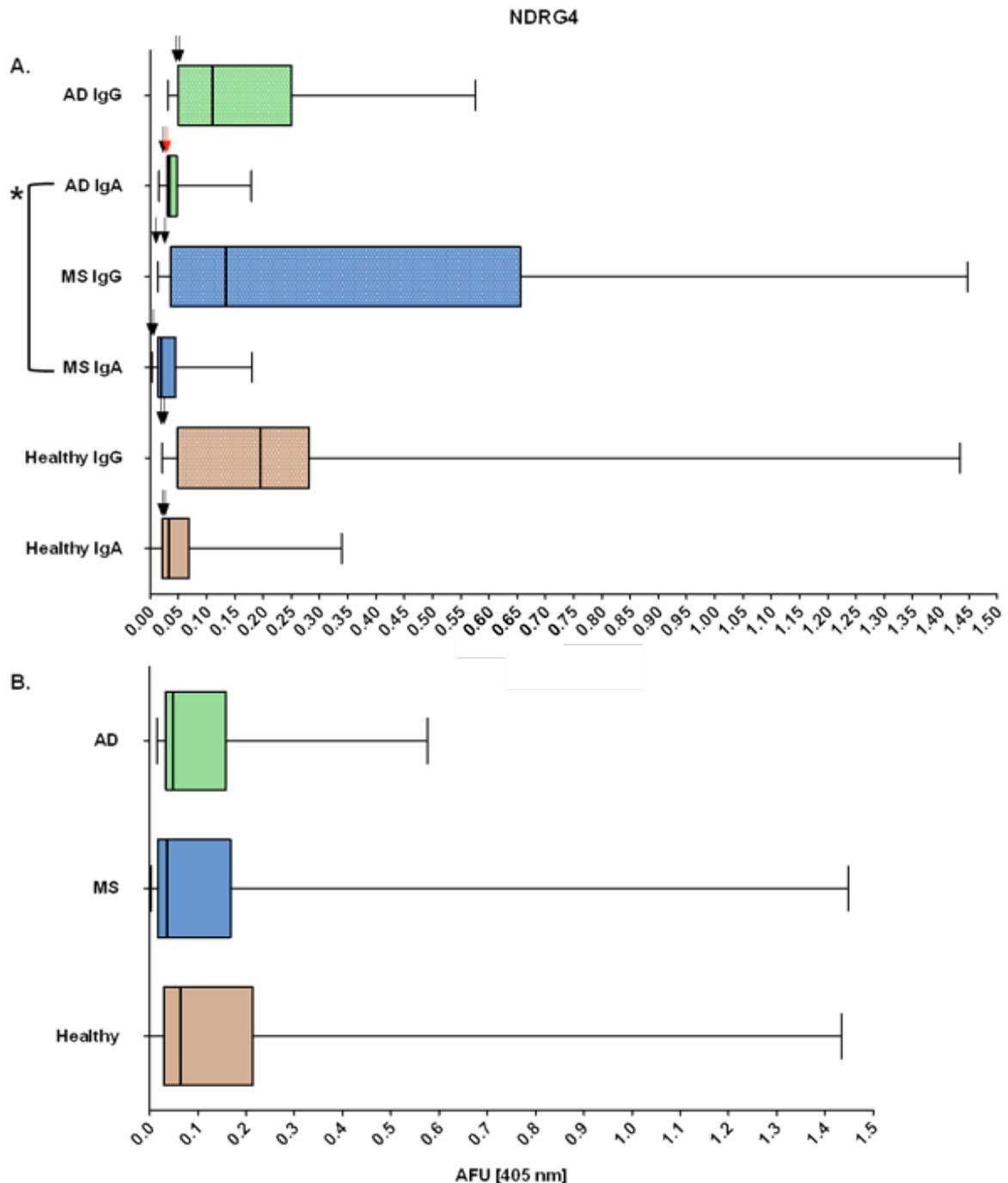
The full-ORF NDRG4-gene could be identified as a transcript variant 7 (RefSeq analysis). The cloned gene was 1,173 bp long (43 kDa). Together with the N-terminal tag-sequence, a ca. 47 kDa big expression product was to be expected. Following Fig. 51 summarizes the results from NDRG4 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 51 NDRG4 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent two column washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, NDRG4 could be overexpressed and successfully purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Clean fractions under the elution peak were pooled together and concentrated to an overall volume of approx. 2 ml. Final protein concentration accounted for 2,730  $\mu$ g/ml.

ELISA re-validation with NDRG4 as coating antigen was performed as described above. Achieved signals were again statistically analyzed under same conditions. Next figure shows summarized ELISA results as box-and-whiskers-plots.



**Fig. 52 Signal distribution and statistical analysis of antigen ELISA with NDRG4.** Both x-axes show arbitrary fluorescence units (AFU), measured at 405 nm. Plot borders correspond to the standard five distribution parameters: minimum value, 25% percentile, median value, 75% percentile and maximum value. On the left side, only group pairs are shown, which yielded significant P-values in the Wilcoxon-Mann-Whitney test ( $>0.05$ ). Pairs with a P-value between 0.05 and 0.01 (significant) are highlighted with one star symbol (\*). **A. ELISA signal distributions in each cohort with each antibody class.** Total number of values in each plot accounts for 20 signals from 20 human sera from the accordant cohort, tested with the accordant antibody. Arrows show signal values from the same two sera in the phage display screening. Arrows, highlighted in red, are sera with enriched antigens in

the phage display screening. **B. ELISA signal distributions in each cohort in both antibody classes together.** Total number of values in each plot accounts for 40 signals from 20 human sera from the accordant cohort.

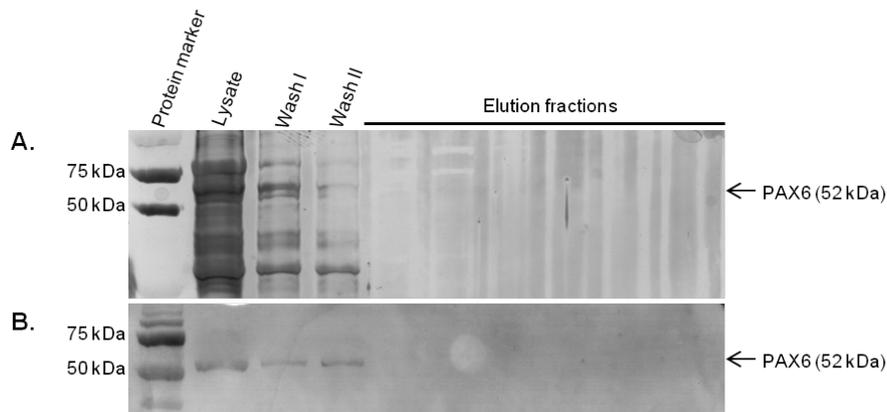
Above Fig. 52 shows distinctly that although the NDRG4 antigen was selected in an IgA-screening, in ELISA format it not only reacted with IgG serum antibodies, but signals were even stronger as with IgA and overall more sera were IgG- than IgA-reactive. Furthermore, also sera from the MS and Healthy cohorts, which showed no enrichment in phage display, were reactive in this experimental set up. Again, fairly strong were especially signals in IgG-ELISA with healthy sera. Statistical analysis revealed that only one single pair could achieve a significant P-value, smaller than 0.05: AD ↔ MS. However, comparison of consolidated data (Fig. 52/B.) did not reveal any statistically significant differences between any of the data pairs.

#### 5.4.2.8. PAX6

The PAX6 (paired box 6) full-ORF was found strongly enriched in an IgA-selection with one of the Alzheimer sera: 327 reads were identified in the initial library, 3,358 reads after the 2<sup>nd</sup> selection round and finally 32,784 reads after the 4<sup>th</sup> round. Other sera showed no enrichment. The gene is represented by 3 clones on the macroarrays, all annotated “off frame”.

PAX6 is a transcription factor, involved in the development of tissues of an ectodermal origin, such as eyes, other sensor organs, CNS, pancreas. Hence, deficiencies have been linked to numerous defects of the eye, as well as to mental retardation and autism [259]. It is strongly expressed in the fetal development stage. In adults, PAX6 expression is limited to eye and brain tissues. Interestingly, mutations in PAX6 have been shown to correlate with an early-onset form of diabetes mellitus, suggesting its role in the insulin pathway as a mediator between the insulin receptor and its target genes [260]. One such gene is NQO1, a quinone oxidoreductase, responsible for detoxification of quinones. Quinones are highly redox-active molecules, leading to oxidative stress, and directly related to aging processes and Alzheimer’s Disease. Thus, NQO1 expression was also co-localized with AD pathology [261]. To sum up, PAX6 was shown to play a major role in the hypothetical neuroprotective pathway Insulin → NQO1 in aging and AD [262].

The full-ORF PAX6-gene could be identified as a transcript variant 1 (RefSeq analysis). The cloned gene was 1,266 bp long (47 kDa). Together with the N-terminal tag-sequence, a ca. 52 kDa big expression product was to be expected. Following Fig. 53 summarizes the results from PAX6 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 53 PAX6 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, PAX6 gene fragment could be successfully, yet insufficiently, expressed in *E. coli*. Accordingly, also purification was considered insufficient, since target protein could only be detected in the flowthrough and wash fractions and not in the elution fractions. Consequently, this antigen was excluded from ELISA validation.

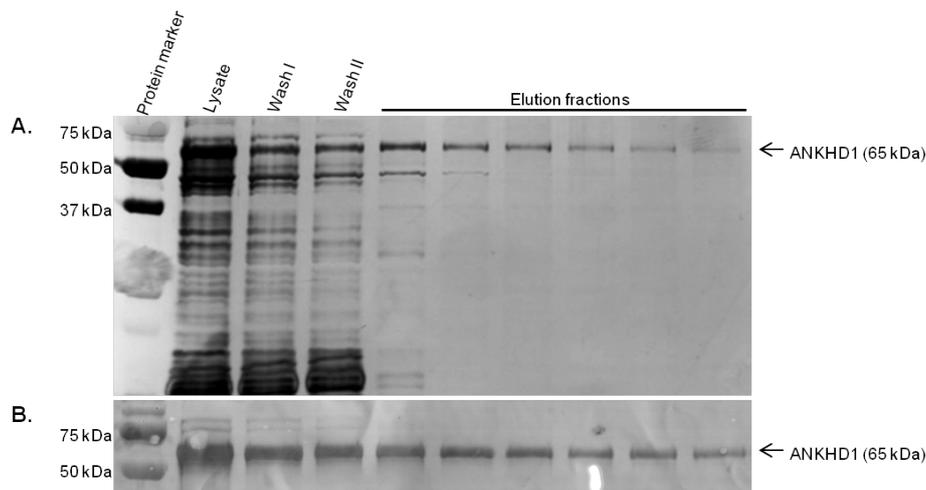
#### 5.4.2.9. ANKHD1

The ANKHD1 (ankyrin repeat and KH containing 1) gene is the first of three antigens, identified as highly potential MS biomarkers in macroarray screenings. It is represented by four different cDNA fragments on the macroarray filters. 3 clones are annotated “in frame” and 1 is annotated “off frame”, respectively. 1 “in frame” clone was hit positive with two different MS-sera in an IgA-screening. The ANKHD1-full-ORF is not available in the source library.

ANKHD1 contains multiple ankyrin repeats in its polypeptide structure and may therefore play a role as a scaffold protein. It is ubiquitously expressed with highest expression levels in brain, cervix and spleen. It has been linked to several cancer types, being highly upregulated in acute leukemia [263]. Evidence exists that ANKHD1 may have an antiapoptotic function in cell survival processes [264]. No correlation to MS, neurodegeneration or inflammation in general could be estimated so far.

The ANKHD1 cDNA clone included theoretically a 2,262 bp long piece from the CDS region, which together with the N-terminal His-tag would lead to a 84 kDa big expression product. However, Sanger sequencing was not able to sequence through the whole insert. Thus, it could not be confirmed, if eventually a downstream stop codon would generate a smaller polypeptide. Expression analysis finally revealed a product of approx. 65 – 70 kDa. Following

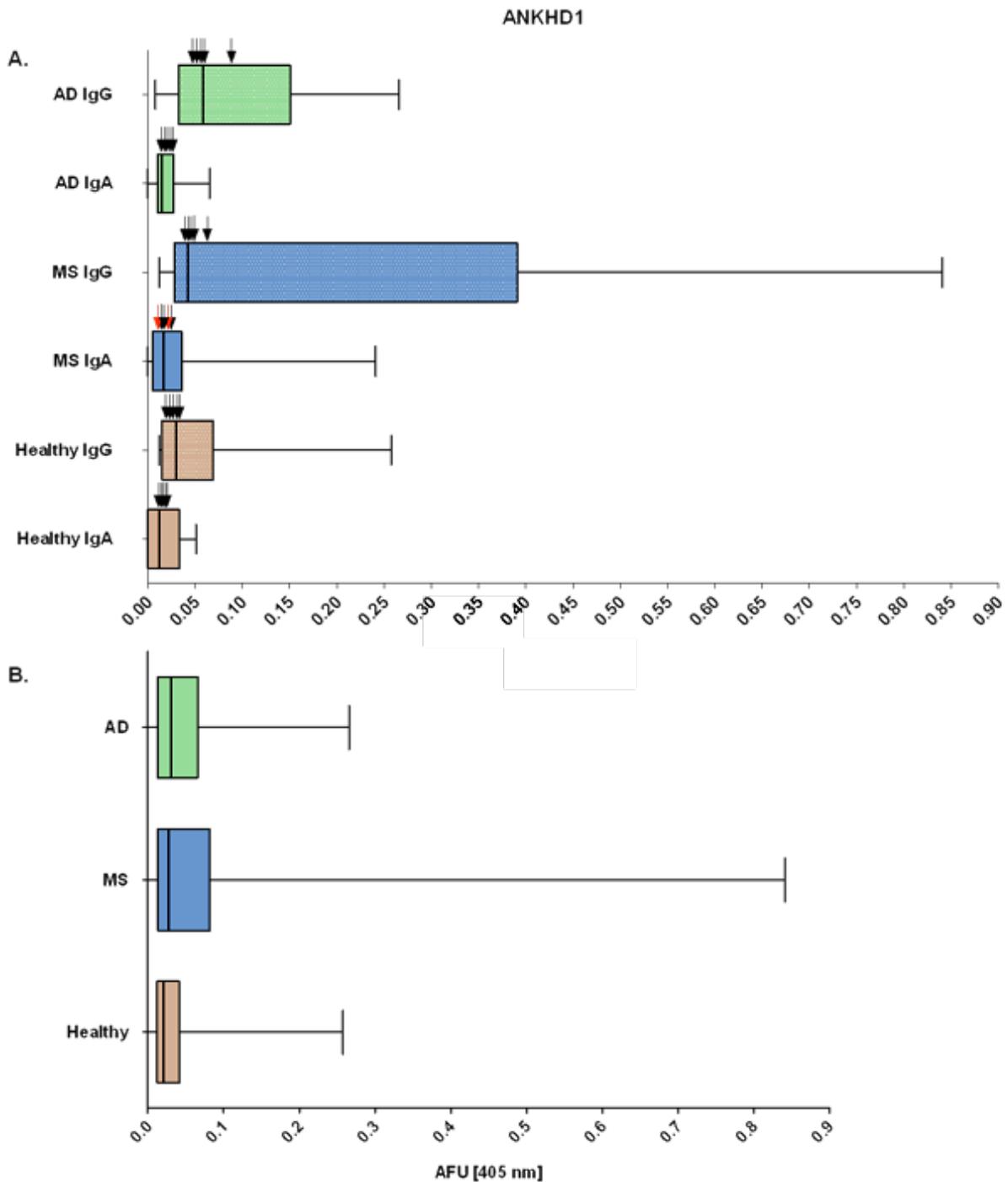
Fig. 54 summarizes the results from ANKHD1 expression in *E. coli* and subsequent FPLC purification via a His<sub>6</sub>-tag.



**Fig. 54 ANKHD1 expression and affinity purification.** **A.** Silver stained SDS gel with fractions from the FPLC affinity purification via a His<sub>6</sub>-tag. Lysate flowthrough, subsequent washing steps and elution fractions are shown. **B.** Western blot with FPLC fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, ANKHD1 could be theoretically expressed and purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Clear elution fractions were pooled together and concentrated to an overall volume of approx. 1 ml. Final protein concentration accounted for 163  $\mu$ g/ml.

ELISA re-validation with NDRG4 as coating antigen was performed as described above. Achieved signals were again statistically analyzed under same conditions. Next figure shows summarized ELISA results as box-and-whiskers-plots.



**Fig. 55 Signal distribution and statistical analysis of antigen ELISA with ANKHD1.** Both x-axes show arbitrary fluorescence units (AFU), measured at 405 nm. Plot borders correspond to the standard five distribution parameters: minimum value, 25% percentile, median value, 75% percentile and maximum value. **A. ELISA signal distributions in each cohort with each antibody class.** Total number of values in each plot accounts for 20 signals from 20 human sera from the accordant cohort, tested with the accordant antibody. Arrows show signal values from the same two sera in the phage display screening. Arrows, highlighted in red, are sera with enriched antigens in the phage display screening. **B. ELISA signal distributions in each cohort in both antibody classes together.** Total number of values in each plot accounts for 40 signals from 20 human sera from the accordant cohort.

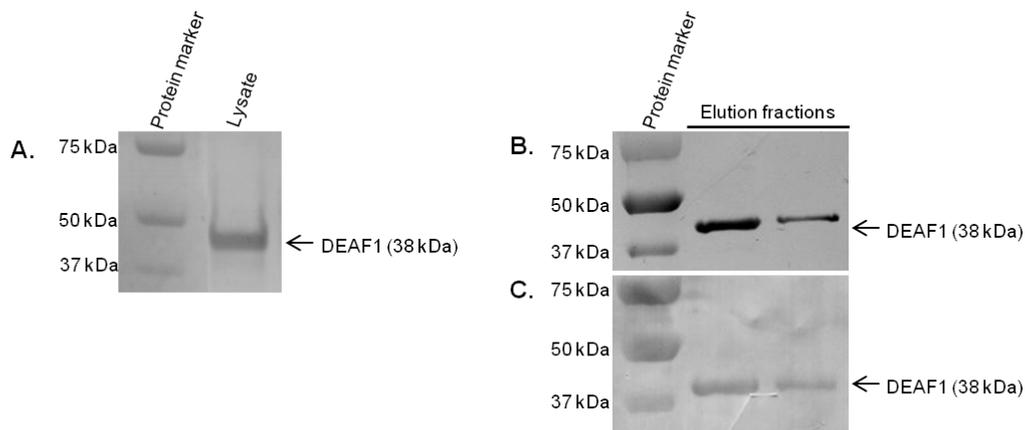
Above Fig. 55 shows distinctly that although the ANKHD1 antigen was selected in an IgA-screening, in ELISA format it not only reacted with IgG serum antibodies, but signals were even stronger as with IgA and overall more sera were IgG- than IgA-reactive. Furthermore, also sera from the MS and Healthy cohorts, which showed no enrichment in phage display, were reactive in this experimental set up. Again, fairly strong were especially signals in IgG-ELISA with healthy sera. Statistical analysis revealed that none of the compared pairs could achieve a significant P-value, smaller than 0.05. Also comparison of consolidated data (Fig. 55/B.) did not reveal any statistically significant differences between any of the data pairs.

#### 5.4.2.10. DEAF1

The DEAF1 (deformed epidermal autoregulatory factor 1) gene, the second MS antigen, is represented by 8 different cDNA fragments on the macroarray filters. 7 clones are annotated “off frame” and only one is annotated “in frame”, respectively. The “in frame” clone was hit positive with two different MS-sera in an IgA-screening. The DEAF1-full-ORF is available in the source library, but no reads were found in any of the initial or the selected libraries.

DEAF1 is a transcriptional factor, regulating many different genes, expressed in various tissues. Two isoforms are localized intracellular in nucleus and cytoplasm and two others are secreted polypeptides. It has been shown to play an important role in the regulating pathway of serotonin production and was subsequently related to depression and suicide [265]. Beside cancer [266, 267], it has also been linked to other disease conditions, such as diabetes type I [268] and defects of the neuronal tube [269]. Recently, DEAF1 was shown to have a positive impact on IFN $\gamma$  production and secretion as a response to viral infections [270]. No correlation to MS, neurodegeneration or inflammation in general could be estimated so far.

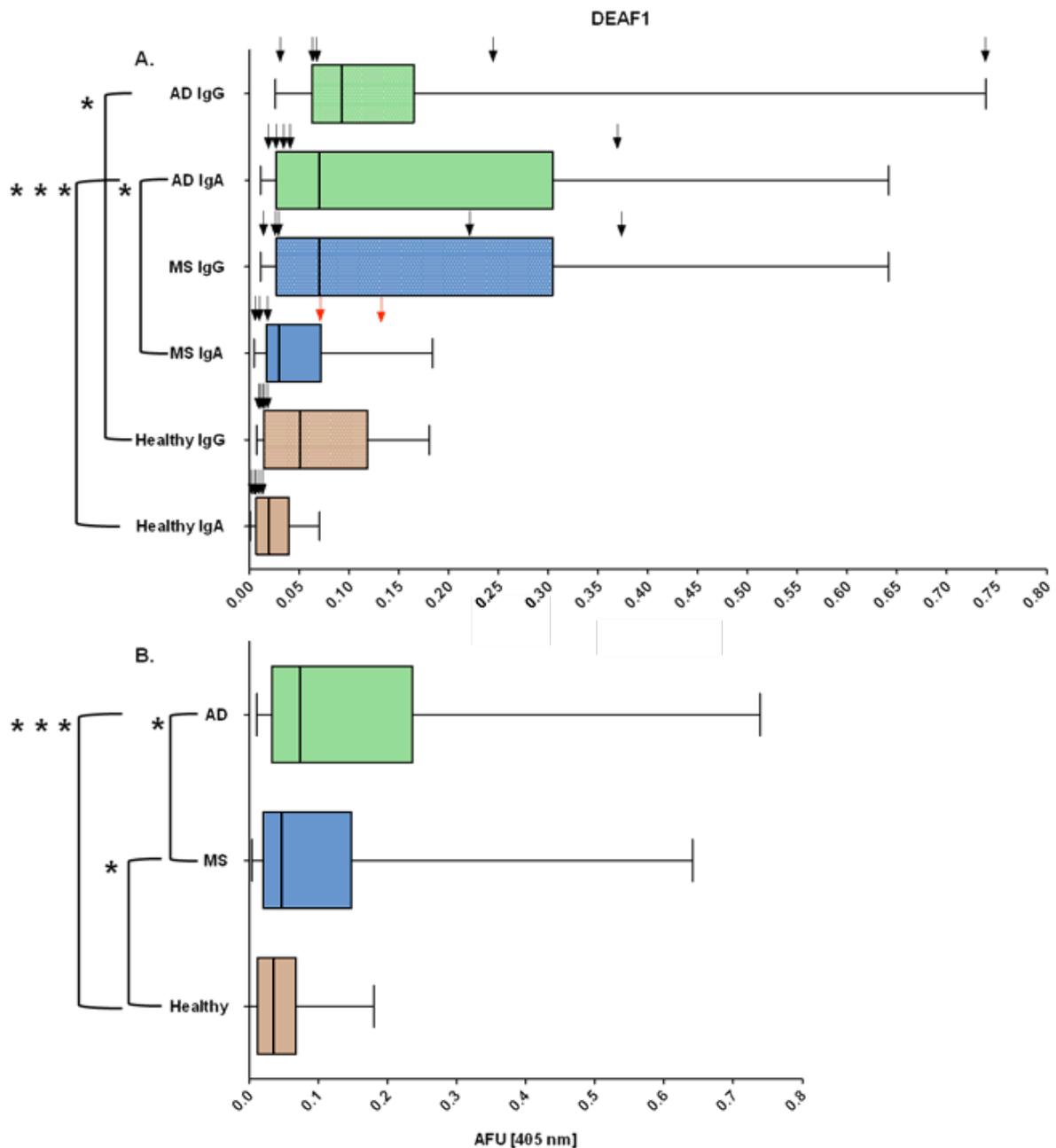
The DEAF1 cDNA clone included theoretically a 960 bp long piece from the C-terminal CDS region, representing approx. 57% of it. Together with the N-terminal His-tag (36 bp), an approx. 37 kDa big expression product was to be expected. Following Fig. 56 summarizes the results from DEAF1 expression in *E. coli* and subsequent batch-purification with Ni-Agarose via a His<sub>6</sub>-tag.



**Fig. 56 DEAF1 expression and affinity purification.** **A.** Western blot with cell lysate of overexpressed DEAF1. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP. **B.** Silver stained SDS gel with fractions from the affinity batch-purification with Ni-Agarose via a His<sub>6</sub>-tag. Two elution fractions are shown. **C.** Western blot with the elution fractions, ordered in the same sequence as on the silver stained gel. Detecting antibodies: mouse  $\alpha$ -His and  $\alpha$ -mouse-HRP.

Thus, DEAF1 could be overexpressed and successfully purified from denatured *E. coli* lysates. Antigen purity level was estimated from sensitive silver staining and considered high enough to be applied as a coating agent in ELISA. Both elution fractions were pooled together and concentrated to an overall volume of approx. 1 ml. Final protein concentration accounted for 324  $\mu$ g/ml.

ELISA re-validation with DEAF1 as coating antigen was performed as described above. Achieved signals were again statistically analyzed under same conditions. Next figure shows summarized ELISA results as box-and-whiskers-plots.



**Fig. 57 Signal distribution and statistical analysis of antigen ELISA with DEAF1.** Both x-axes show arbitrary fluorescence units (AFU), measured at 405 nm. Plot borders correspond to the standard five distribution parameters: minimum value, 25% percentile, median value, 75% percentile and maximum value. On the left side, only group pairs are shown, which yielded significant P-values in the Wilcoxon-Mann-Whitney test ( $>0.05$ ). Pairs with a P-value between 0.05 and 0.01 (significant) are highlighted with one star symbol (\*). Those with a P-value, smaller than 0.001 (extremely significant), are highlighted with three star symbols (\*\*\*). **A. ELISA signal distributions in each cohort with each antibody class.** Total number of values in each plot accounts for 20 signals from 20 human sera from the accordant cohort, tested with the accordant antibody. Arrows show signal values from the same two sera in the phage display screening. Arrows, highlighted in red, are sera with enriched antigens in the phage display screening. **B. ELISA signal distributions in each cohort in both**

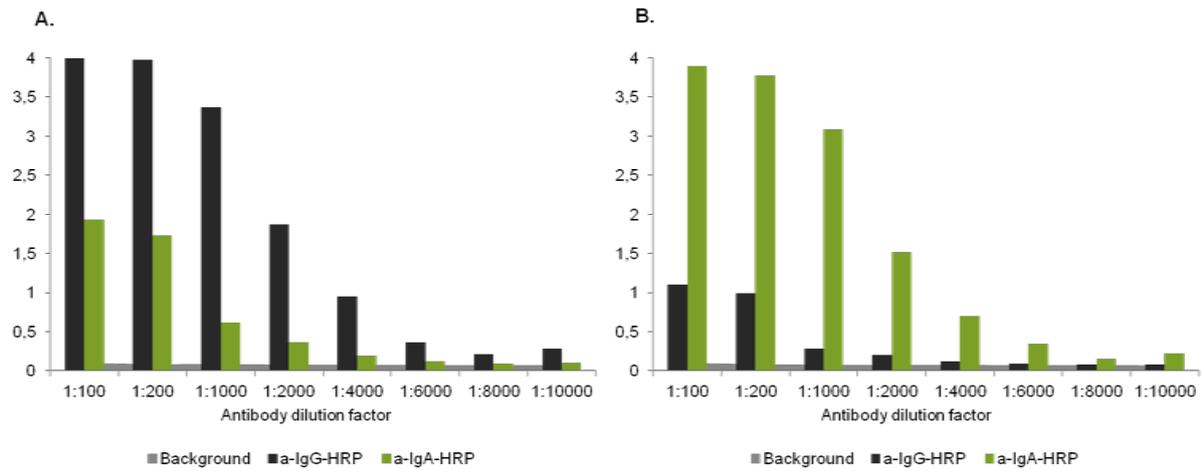
**antibody classes together.** Total number of values in each plot accounts for 40 signals from 20 human sera from the accordant cohort.

Above Fig. 57 shows distinctly that although the DEAF1 antigen was selected in an IgA-screening, in ELISA format it obviously reacted also with IgG serum antibodies. Furthermore, also sera from the AD and Healthy cohorts, which showed no reactions during macroarray screening, were reactive in this experimental set up. Surprisingly, the AD cohort revealed fairly strong signals with both antibody classes. Thus, statistical analysis revealed an extremely small, hence significant P-value of 0.0005 in the comparison AD ↔ Healthy with IgA. This pair showed also a good P-value on IgG level. Also for the AD ↔ MS pair with IgA a significant P-value was estimated. As expected, this trend sustained also when comparing the consolidated cohort groups (Fig. 57/B.). Again, a very small P-value of 0.0001 was observed for the AD ↔ Healthy pair. The AD ↔ MS pair, but this time also MS ↔ Healthy, showed distinct significance.

#### **5.4.3. Cross-reactivity tests of HRP-conjugated anti-human antibodies**

Prior to antigen ELISA, secondary HRP-conjugated  $\alpha$ -hu-IgG and  $\alpha$ -hu-IgA antibodies were tested with commercially available pure human IgG and IgA proteins. The aim of these preliminary experiments was to estimate the most appropriate dilution factor, at which signals would be in an acceptable measurable range, as well as the appropriate time scale for substrate conversion. Furthermore, undesirable cross-reactivities of the detecting antibodies had to be excluded. Hence, two different ELISA experiments were arranged.

For the first test, two 96-well ELISA plates were coated with 1  $\mu$ g per well purified human IgG or IgA. Each plate was incubated with each of the HRP-conjugated  $\alpha$ -human antibodies in a dilution series between 1:100 and 1:10,000. Finally, signals were measured at 60 min.



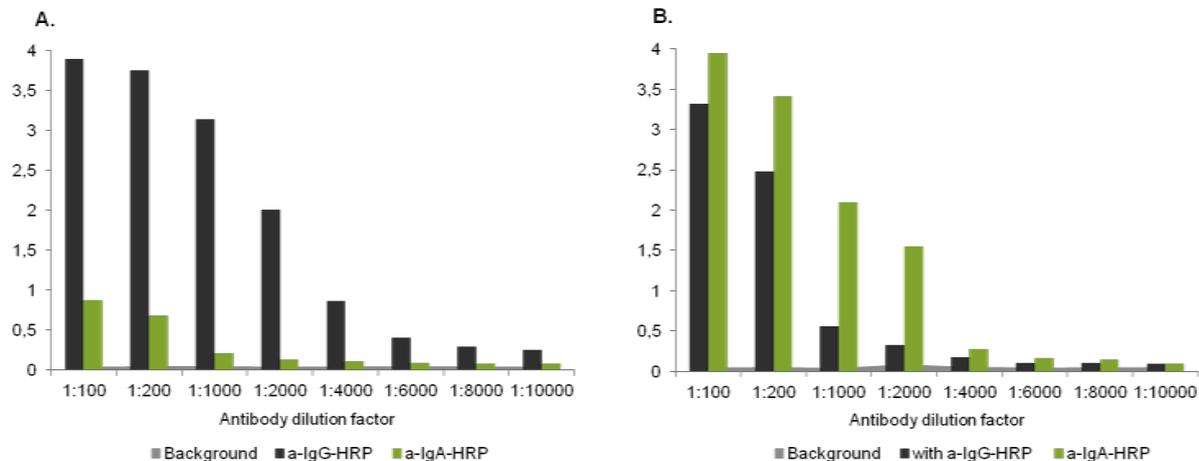
**Fig. 58 Test ELISA with purified Ig-proteins and HRP-conjugated  $\alpha$ -human antibodies.** All measurements were made at 60 min after adding the ABTS substrate. Y-axes show arbitrary absorbance units, measured at 405 nm. X-axes show antibody dilution series. Detecting antibodies:  $\alpha$ -hu-IgG-HRP in dark grey and  $\alpha$ -hu-IgA in green. ELISA background signal is shown as a light grey area. **A.** Plate coated with purified human IgG. **B.** Plate coated with purified human IgA.

Thus, both HRP-conjugated antibodies showed cross-reactivities, which were strongest at high concentrations of the detecting antibodies (dilution < 1:1,000). However, at this dilution stage the accordant antibody also reached its saturation state. Thus, it could be concluded that a dilution factor of 1:2,000, which was also manufacturer's recommendation, a good balance between the specific signal and a weak cross-reactivity background was achieved. After subtracting ELISA background (light grey areas on the charts) from both values at 1:2,000 dilution, following cross-reactivity percentage ratios remained: 15% for the  $\alpha$ -IgG-HRP antibody and 9% for the  $\alpha$ -IgA-HRP antibody.

Having proved the overall functionality of the detecting antibodies with purified Ig protein, the question still remained, if these results will recur also with human sera. Here, it had to be taken into account that sera are highly heterogeneous media, where normally all antibody classes are present in very different proportions. Since IgG and IgA are the two most abundant classes in blood serum, they were most likely to provoke undesired cross-reactivities. Also the third prevalent antibody IgM (approx. 1.5 mg/ml blood serum) could be a potential source of error. However,  $\alpha$ -IgM cross-reactivities were not tested. Finally, reactivities against IgD and IgE were considered marginal, since their concentrations in serum are fairly low (approx. 30  $\mu$ g/ml and 0.05  $\mu$ g/ml, respectively).

For the second test, two 96-well ELISA plates were coated with 1  $\mu$ g  $\alpha$ -hu-IgG and  $\alpha$ -hu-IgA, respectively. Each plate was incubated first with a human serum from the healthy-cohort (diluted 1:100) and then with each of the HRP-conjugated  $\alpha$ -human antibodies, again in a dilution series between 1:100 and 1:10,000. Finally, signals were measured at different time

points after adding the substrate: between 5 and 60 min. Next Fig. 59 shows obtained results at 60 min.



**Fig. 59 Test ELISA with serum Ig-proteins and HRP-conjugated  $\alpha$ -human antibodies.** All measurements were made at 60 min after adding the ABTS substrate. Y-axes show arbitrary absorbance units, measured at 405 nm. X-axes show antibody dilution series. Detecting antibodies:  $\alpha$ -hu-IgG-HRP in dark grey and  $\alpha$ -hu-IgA in green. ELISA background signal is shown as a light grey area. **A.** Plate coated with  $\alpha$ -hu-IgG and subsequently incubated with human serum. **B.** Plate coated with  $\alpha$ -hu-IgA and subsequently incubated with human serum.

The overall picture from the first ELISA test replicated also with incorporated sera incubation step. Interestingly, this time the  $\alpha$ -IgA-HRP antibody showed less cross-reactivity against serum IgG (Fig. 59/A) than against purified IgG (Fig. 58/A). On the contrary,  $\alpha$ -IgG-HRP was much more reactive against serum IgA (Fig. 59/B) than against purified IgA (Fig. 58/B). Especially at higher concentrations of the detecting antibody, e. g. dilution factor less than 1:1,000, very high cross-reactivity signals were observed. These differences between the human serum test and the test using purified human IgG and human IgA may be explained by the fact that for testing human sera, additional anti-human IgG and IgA antibodies for coating were used, which may have different cross-reactivities than the secondary HRP labeled antibodies. However, for the purpose of this study, measurements at such high antibody concentrations were not relevant and could be therefore ignored. In the chosen range of 1:2,000 antibody dilution following cross-reactivity percentage ratios were calculated after subtracting the general ELISA background: 5% for the  $\alpha$ -IgG-HRP antibody and 16% for the  $\alpha$ -IgA-HRP antibody.

In conclusion, both HRP-conjugated  $\alpha$ -human antibodies were proved functional and with an acceptable low cross-reactivity. Thus, in case of  $\alpha$ -hu-IgG-HRP a possible failure ratio in the range of 5% – 15% could be expected. In case of  $\alpha$ -hu-IgA-HRP this was in the range of 9% – 16%.

## 6. Discussion

In my thesis I identified and characterized novel potential autoantigens in human blood sera from Morbus Alzheimer and Multiple Sclerosis patients, which could be used as biomarkers for diagnostic purposes in future. For this, I applied two different screening platforms: protein macroarrays, which are well established and commercially available, and selections with human full-ORF-polypeptides, presented by M13 phages. Thereby, the full-ORF phage display technology was first developed in the course of this study. In detail, I generated and evaluated novel expression vectors and initial phagemid libraries, creating the basis for all further experimental procedures. Next, I performed the bio-panning in a semi-automated way with human autoantibodies, immobilized on magnetic beads. Enriched phage clones were identified with an Illumina NGS device. Finally, I selected the most interesting candidates from both screenings. After their recombinant expression and affinity purification, I re-evaluated these potential biomarkers in ELISA assays with larger sera groups and analyzed statistically their performance.

In this chapter, experimental results will be discussed and suggestions for further development of the applied methods will be made. Findings will be put into perspective, concerning current literature state. Finally, the two high-throughput screening platforms will be compared in regard to their technical capacities, advantages and limitations.

### 6.1 Results from protein macroarray screening

Five sera were chosen randomly from each of the three donor cohorts and screened on protein macroarrays in order to identify human antigens, immunoreactive against IgG and IgA autoantibodies. The AD cohort consisted of 20 sera from donors between 72 and 87 years old. From these, the five screened sera originated from four males and one female in the age of 72, 79, 76, 80 and 72. Since the healthy cohort was chosen as a control group to AD, it was accordingly age-matched. Thus, overall healthy donors' age ranged between 64 and 87 and the five tested sera originated from four females and one male in the age of 75, 77, 75, 76 and 73. Psychiatric tests have been applied to diagnose AD patients, as well as to exclude disorder in the healthy cohort. However, it is impossible to predict, if the healthy probands would show indications of Alzheimer's or any other kind of dementia later in their lives. Thus, it is at least theoretically conceivable that some of the people, assigned non-diseased, already suffered from an early AD stadium at the point of blood withdrawal. Furthermore, the denotation "healthy" should not be misleading, since it refers only to AD and means in this context merely "tested negative for Alzheimer's".

The MS cohort differed stronger from the other two, due to characteristic disease specification. Here overall donors' age ranged between 29 and 71. Yet 50% of all 20 probands were between 38 and 47 years old and the median value of the group amounted for 43 years. Another specific feature of this cohort was the relatively high number of female donors, namely 16 out of 20. In comparison, the gender ratio in AD and Healthy was more balanced, accounting for 11 females and 9 males in both groups. Finally, the five screened MS sera originated from five females in the age of 38, 39, 47, 42 and 44. All of them have been diagnosed with the disease subtype RRMS.

To my knowledge, none of the 60 donors was diagnosed with any other disease at the point of blood withdrawal. Yet, non-symptomatic etiopathologies cannot be definitely excluded. Especially cancer, allergies, infectious, autoimmune, and other disorders, directly correlating to immune response, can have a distinct impact on autoantibody profiles and with this on screening results. Moreover, also past diseases can leave long-lasting immunological marks in blood serum. Especially in elderly people this can be a significant point of concern. Naturally, it is not possible to test probands for all kinds of medical aberrations. Yet, a more thorough characterization of donors' overall health status and medical history prior to a screening trial, would undoubtedly lead to a more precise data interpretation.

Macroarray screening results proved well reproducible and thus reliable. Screening replication, meaning the same serum in same dilution and with the same detecting antibody, but on another macroarray filter, usually showed signal overlapping of over 90%. However, the weakest positive hits were not trailed further, anyway. The strongest and thus, most interesting signals, showed steady reproducibility of nearly 100%.

All sera were screened in a 1:100 dilution and were not adjusted to any uniform IgG or IgA concentration, like it has been done by Cepok, S. *et al.* [158]. Independently of the detecting antibody and donor cohort, each screened serum manifested a unique overall signal pattern. Thus, also sera from healthy donors were as immunoreactive, as any of the diseased sera. This finding was not surprising and demonstrates once more the individual and highly diverse nature of each immune system. Furthermore, overall serum IgG/IgA concentrations did not correlate with signal numbers, detected in the accordant serum during screening. Thus, for example, AD serum 191492 with the lowest general IgA concentration (s. Fig. 14/B, AD serum No. 2) showed very strong IgA-immunoreactivity with almost 800 hit clones. On the other hand, healthy serum 151006 with the highest general IgG titer (s. Fig. 14/A, Healthy serum No. 2) showed relatively moderate IgG-immunoreactivity with overall 198 positive hit clones. Hence, it could be concluded that no obvious correlation exists between total immunoglobulin titer and serum immunoreactivity against the expressed cDNA library,

spotted on the macroarray filters. Accordingly, normalization of serum Ig concentrations prior to screening would not necessarily lead to normalization of signal counts and/or intensities and was therefore not performed in this study.

Another issue, often discussed in relation to high-throughput screenings, is pooling of samples from one comparison group. Beyond doubt, pooling sera together can be advantageous, since more material can be screened in one run. On the other hand, screening single samples is more costly in terms of labor, time and, last but not least, finances. Moreover, amplification of signal intensity can be theoretically expected with pooled samples, if presumably more sera contain antibodies against the same clone. On that account, published data from macroarray screenings often originate from experiments with pooled liquid samples, like in Erdag *et al.* [157], for example. However, mixing of samples harbors certain risks as well. As already discussed above, serum antibody patterns are highly individual and diverse. Thus, the risk is relatively high to include by chance a very immunoreactive sample to the pool, which will outshine the less reactive ones and will finally deliver strong signals, specific for this one serum, but not for the cohort as a whole. To eliminate this undesirable effect, all sera in my study were screened separately.

Comparable numbers of positively hit clones were found in all three cohorts during the macroarray screening: 2,036 for AD, 2,737 for MS and finally, 2,469 for Healthy (s. Table 10). Approx. 45% of these positive clones in each of the three cohorts were annotated “in frame”, allowing their direct linkage to real genes. This ratio was expected, since also 40% of all spotted clones on the macroarray filter are annotated “in frame” as well. Also numbers of positively hit genes were comparable: 378 for AD, 370 for MS and finally, 368 for Healthy. After prioritizing of hit genes with the formula, depicted on page 81, numbers of hit genes with a differential score over the cut-off of 2, still did not differ significantly between the three cohorts. However, the unique genes in these last analysis groups, meaning genes with a diff. score  $> 2$ , which were hit only in one cohort and not at all in the other two, revealed remarkable differences between diseased and non-diseased: 88 for AD, 72 for MS and only 18 for Healthy. These highly interesting results finally led to the following conclusions:

- No dissimilarities are visible in the general immunoreactivities of AD, MS and Healthy sera, since neither total numbers of positive clones, nor of hit genes revealed significant differences between the three groups.
  - After ruling out the most interesting genes and subsequently assigning unique ones for each cohort, in the two diseased cohorts still remained approx. four times more antigens, as in the Healthy group.
  - Hence, although hit autoantigen numbers were almost the same in all three groups, in the Healthy cohort over 90% of these was obviously non-

specific. On the contrary, screened AD and MS sera were distinctly immunoreactive against cohort-specific autoantigens. For AD, 62% of the most interesting hit genes were cohort-unique and for MS, these were 42%.

Thus, this first screening result delivered a general confirmation of my hypothesis that also primarily non-autoimmune disorders like Alzheimer's, harbor immunological phenomena, which are comparable to those in a real autoimmune disease, such as Multiple Sclerosis. Furthermore, since healthy sera show comparable immunoreactivities, as diseased ones, it should be theoretically possible to identify autoantigens, present in the non-diseased and absent in the diseased cohort. Such negative biomarkers can eventually prove as valuable as real biomarkers in diagnostics. In my study the number of screened sera was too low to definitely assign such an "antimarker". Kijanka *et al.* [156], for example, screened much larger cohorts of sera from donors with colorectal cancer (43 sera) and without (40 sera) on the same macroarrays. They were able to identify not only 18 specific biomarkers in the diseased sera, but also 4 antimarkers in the healthy ones with a significant P-value ( $< 0.05$ ).

Another kind of analysis, performed with the data from macroarray screenings, was pathway analysis with the IPA software tool (s. Fig. 16). The idea behind, was to examine, in which biochemical and/or (patho)physiological pathways the hit genes are involved. It was interesting to see, if any distinct pathway would emerge in one particular cohort. In the healthy group only 8 random pathways were determined by IPA. This was not surprising, since the analyzed data set consisted of only 18 cohort-unique genes and no strong correlation between them was expected. Hence, this group was tested as a negative comparative group. On the other hand, the data sets of the two diseased cohorts consisted of a comparable number of cohort-unique genes: 88 for AD and 72 for MS. Nevertheless, analysis results differed strongly between them. First, the overall number of determined pathways in the AD group was almost exactly three times higher, as in the MS group: 61 pathways in AD and 21 pathways in MS. Second, no overlapping of pathways between AD and MS was found. But also no "prominent" pathways could be estimated in any of the groups. In case of AD, however, a remarkably high number of pathways were determined, belonging to intracellular organization processes, such as organization of cytoplasm, organization of nucleus or organization of cytoskeleton. Another noteworthy trend in this group was presence of pathways, involved in cell cycle and cell development. Remarkably, no metabolic pathways were represented in the AD distribution. These data correlate well with results published recently by Manavalan *et al.* [271]. They determined differentially expressed proteins in brain-regions of AD patients, which are usually strongly affected by the disorder (hippocampus and parietal cortex), and assigned these to functional pathways via

IPA-analysis. Also their data show a high ratio of proteins, involved in cell organization and cell development, but a weak representation of canonic metabolic pathways. Interestingly, in the proteome of the cerebellum, which is assumed to remain relatively unaffected during disease course [272], the ratio of metabolic pathways was calculated to be 10% higher. In conclusion, a distinct pathways profile could be assigned to the identified AD-unique genes, which differs strongly from the other two cohorts. Yet, more sera have to be screened and analyzed in order to ascertain these findings and to render more precisely the hypothetical pathways involved in AD.

An interesting aspect, implicating macroarray screening results, are hit clones assigned “off frame”. After all, they account for approx. 40% – 45% of all clones, spotted on the filter, as well as of all hit clones in the screenings. First, it should be taken into consideration that these frame assignments are not completely accurate. Hence, it is necessary to re-clone and thoroughly re-sequence hit clones of interest. In this study, this was done for all clones, which were chosen for further experimental work after screening. Yet, also confirmed “out of frame” clones harbor considerable potential and can deliver interesting findings. The generated polypeptides are indeed artificial and do not make any biological sense. But if they provoke significant immunoreactivity, they obviously present one or more epitopes, recognized by the tested antibody. This fact can be already a sufficient reason for constructing a peptides-based diagnostic assay, for example. After all, biological background of applied bait agents is not relevant for diagnostic purposes. Yet, deeper analyses of identified immunoreactive peptides can be undertaken as well, in an attempt to reveal possible reasons for the interaction. Thus, it often turns out that the artificial sequence forms an immunoreactive epitope of a real protein by chance, demonstrating a case of molecular mimicry. If the imitated protein is not present on the filter, such out of frame clones can finally expand the repertoire of the screened cDNA library. Such a case was described by Cepok, *et al.* [158]. Screening of CSF samples from MS patients on macroarrays resulted in identification of 21 clones, with strong immunoreactivity in at least two MS samples. Yet, half of them were annotated off-frame. Following epitope mapping and thorough analysis of the peptide sequences finally revealed two epitopes, which could be assigned to two different antigens of the Epstein-Barr virus (EBV). Although still not entirely understood, the association between MS etiology and the presence of antibodies against EBV-antigens in serum and CSF is well documented [273, 274]. Thus, analysis of hit off-frame clones led to coherent results in this case.

## 6.2 Results from phage display screening

The idea of cloning human full-ORF libraries into five different phagemid vectors originated from the consideration that functional presentations of highly heterogenic libraries on phage surface can prove problematic. Expression and presentation hurdles are usually due to diversities in proteins' features, e.g. folding characteristics, solubility, host toxicity and others [193]. Since presentation of the protein of interest (POI) as a fusion to the pIII phage coat protein requires its export into periplasmic space, an opportunity is given, to utilize the three different *E. coli* secretion pathways and thus expand the range of properly folded POIs. Therefore, in my newly generated pYG vector series all three pathways were implicated: Sec, SRP-dependent and Tat. Another variation in the series was the art of catenation of POI and pIII. Two vectors were cloned for a direct polypeptide-fusion (pYG-fusion) and three other for an indirect fusion via a leucine zipper (pYG-LZ). Thus, my vector series suggests a solution to the particular problematic of functional polypeptide folding in periplasmic space.

The newly generated destination vectors were tested per Sanger sequencing after each cloning step. Furthermore, special attention was paid to the functionality of the ccdB-protein to ensure low background of parental clones and with this, high quality of the pYG-full-ORF libraries (Fig. 22). Finally, functionality of the pYG-vectors was tested with EGFP, which served in this study as an example of a full-ORF-polypeptide. These results demonstrated clearly the full applicability of all five vectors, both for expression, as well as for functional POI-presentation on M13 phage surface. EGFP was a perfect example for this purpose, since its fluorescence depends on the correct conformation of the chromophore, which in turn can be correctly formed only in the cytoplasm. Hence, I assumed to find strong fluorescent EGFP, only if expressed from the pYG-LZ-TorA vector, since only the Tat secretion pathway exports cytoplasmically folded polypeptides. Thus, I was finally able to confirm my hypothesis that different vector systems are required in order to encompass also difficult to present proteins with special folding requirements in a heterogenic ORF library.

All five vectors exhibited strong EGFP expression levels, as expected (Fig. 27). Yet, only cell pellets of the pYG-LZ-TorA sample were visibly green. Also fluorescence measurements in the periplasm fractions confirmed this observation (Fig. 28). These results were not surprising, since polypeptides with a pelB leader sequence are kept unfolded in the cytoplasm by chaperones like secB until the actual translocation across the cell membrane takes place (post-translational translocation) [275]. On the other hand, translation of polypeptides containing a DsbA leader sequence is aborted shortly after the synthesis of the DsbA-peptide. The whole ribosomal complex together with the attached SRP particles migrates to the Sec-translocon, where protein synthesis continues in parallel to secretion (co-

translational translocation) [276, 277]. Consequently, the pelB- and the DsbA-EGFP proteins have actually no chance to form properly in the cytoplasm in contrast to TorA-EGFP. Thus, my data confirmed previous results described earlier. The still brightly green spheroblasts after the periplasm extraction of TorA-EGFP were probably due to the fact that the Tat export machinery acts rather slowly. At least half of the overexpressed EGFP remains still in the cytoplasm or even stuck in the cell membrane. As shown by Thomas *et al.*, the TAT-export efficiency of GFP seems to account for only about 50% [278].

Interesting visualization of the two different catenation types between EGFP and pIII delivered western blot experiments with and without the reducing agent 2-Me-EtOH (Fig. 29). While the two different reducing conditions had no impact on the band pattern of pYG-fusion-EGFP constructs, a clear shift was seen in case of pYG-LZ-EGFP. These findings were also expected, as in the pYG-LZ construct both polypeptides are expressed separately and are held together via a leucine zipper and not a real covalent bond like in pYG-fusion. The hydrophobic bonds in LZ are furthermore enforced by two disulfide bridges residing at both ends of the zipper, which get naturally destroyed by 2-MeEtOH. Thus, the EGFP-pIII linkage in pYG-LZ can only be visualized by using non-reducing conditions and hence keeping the leucine zipper structure intact.

The usage of two different  $\alpha$ -EGFP antibodies, able to distinguish between the folded and the denatured EGFP-forms, proved very advantageous for the ELISA experiments, shown in Fig. 30. Beside expected results, already described on page 95, also some other intriguing observations were made. Thus, in a detection with the polyclonal  $\alpha$ -EGFP antibody, recognizing both EGFP states, the two pYG-DsbA constructs showed weaker signals than all other samples (Fig. 30/A.). Interestingly, Thie *et al.* [279] witnessed the same phenomenon when comparing ScFv-presentation on phages with PelB- and DsbA-leader sequences. On the other hand, Steiner *et al.* [198] could show higher display levels of DARPin when going through the SRP- instead of the Sec-pathway. As Thie *et al.* rightfully pointed out, DARPins contain predominantly  $\alpha$ -helices while the prevailing secondary structures in ScFvs are  $\beta$ -sheets. Also inner membrane proteins (IMPs), made mainly of  $\alpha$ -helices, are typically targeted through the SRP-dependent translocon, for example [277, 280]. In conclusion, it can be speculated that the SRP-system is probably more suitable for polypeptides not containing  $\beta$ -sheets in their structure. As the tertiary EGFP structure is a  $\beta$ -barrel, made of 11  $\beta$ -sheets and one  $\alpha$ -helix, our data seem to confirm the hypothesis discussed by Thie *et al.*

Still, the presence of the correct 3E6-epitope, demonstrated by the ELISA experiments, would not necessarily mean that the overall polypeptide is properly folded as well or even functional and fluorescent. It is not well known, why GFP is not able to fold accurately in the

periplasm, although the formation of the chromophore is a self-catalytic process and does not require any cofactors or chaperones. Since there are two cysteine residues present within its  $\beta$ -barrel structure (C49 and C71), it is possible that under the oxidizing conditions in the periplasm wrong intermolecular disulfide bonds may form, thus leading to a misfolded polypeptide, as suggested by Aronson *et al.* [281]. Yet, a properly formed  $\beta$ -barrel structure is crucial for fluorescence functionality, as it seems to shield the chromophore from the aqueous environment. As proposed by Reid *et al.* [282], the loss of fluorescence is probably due to quenching effects in the misfolded GFP, since the mature chromophore is rather stable and remains chemically intact even under denaturing conditions. Potentially, such molecule could possess a properly folded 3E6 epitope, recognized by the applied antibody, and still be non-functional. However, fluorescence measurements of the EGFP-phage suspensions showed that the only sample with measurable values is once again pYG-LZ-TorA (Fig. 31). In all other samples only background fluorescence could be detected. Thus, in contrast to Velappan *et al.* [224], I did not observe any significant fluorescence in phage samples, based on Sec-vectors. This could be due to the fact that I used unmodified EGFP, since my intention was to demonstrate an example of expression and presentation of a usual protein from a full-ORF library. On the contrary, the superfolder GFP (sfGFP), applied by Velappan *et al.*, is a much more stable and robust GFP-variant. This is certainly of importance in experiments aiming to show proper or rather improper folding under different conditions.

Cloning of the human full-ORF collection into the five pYG-vectors and subsequent propagation of the phagemid libraries in *ccdB*-sensitive DH10B *E. coli* cells brought two desired consequences. First, lethality of the *ccdB*-expression product led to positive selection for real transformants and elimination of parental clones' background. Second, libraries were amplified, reaching fold coverages of at least 6 times per library (Table 11). This effect was increased further, after re-transformation of the rescued libraries into phage display compatible, *ccdB*-resistant XL1Blue and TG1 strains (Table 12). After this amplification step, achieved fold coverages were very high, ranging between 160 and 1 997 times per library. Furthermore, since same amount of plasmid-DNA and cells with similar transformation efficiencies were used, it was not surprising that the smaller OCAB libraries showed bigger fold coverages as their OCAA counterparts. Finally, the multiple cloning of the library into five different vectors in parallel, further contributed to a significantly increased statistical probability of overrepresenting each gene in the final initial library, which is a desirable effect.

Analysis of the polypeptide size distributions of the full-ORFs after the first cloning step into DH10B cells demonstrated that smaller inserts were slightly favored (Fig. 23 and 25). Apparently, this bottleneck typically appears early in the cloning process, namely in the LR

reaction [283]. The vector type itself did not seem to play any role. As shown in Fig. 25, DNA inserts in the range of 750 bp – 1,100 bp were most successful, while the median insert size in the source library was at 1,182 bp. This undesirable bias can be eventually overcome by improving LR reaction conditions, like applying smaller DNA amounts, increasing reaction times or changing reaction temperature.

Native subcellular localization has a major impact on proteins' nature. Therefore, it was interesting to analyze, how these distributions look like in the different initial libraries. As Fig. 26 reveals, the distributions in the phage libraries remained very similar to the one in the OC source library. No unique patterns were observed in the different vector types. Thus, it could be concluded that cloning procedures executed no significant impact or bias on libraries' compositions as a whole and the complexity of the source collection was successfully reproduced. Here, it should be taken into consideration that different full-ORF-clones originated from different providers and were non-systematically generated. Thus, the source collection reflects in no way any *in vivo* situation. In general, the source OC library represents approx. 25% of the human protein coding genome.

To my knowledge, presentation of a preassigned human full-ORF library on M13 phages, as well as its selection against serum autoantibodies, immobilized on magnetic beads, was shown for the first time in this study. This kind of antigen phage display has grave advantages in comparison to cDNA presenting methods, which we recently reviewed [193]. Particularly the laborious and time consuming filtering for real ORFs, which naturally account for only one third of a cDNA library, can be avoided. However, the most important feature of a defined full-ORF library is certainly the presence of the whole CDS region, as well as the absence of undesired UTR regions, which can lead to artificial epitopes generation and selection of false positive binders. Furthermore, for identification of selected targets per sequencing, very short fragments of the inserted gene are required. This finally allowed the application of the Illumina NGS platform, since its methodology is based on fragmentation of the DNA material in ca. 75 bp long fragments, prior to sequencing. In case of cDNA, preferably the whole insert has to be sequenced, in order to assign thoroughly the expressed polypeptide. This was usually done with the 454 pyrosequencing platform from Roche [284]. Yet, sequencing with Illumina is still considerably cheaper and sample preparation is more trivial. However, it should be taken into consideration that a full-ORF library is by nature an artificial composition and does not reflect any *in vivo* situation. A cDNA library, on the other hand, represents a real transcriptional state of a cell or a tissue at the time point of library generation. This characteristic is certainly one of the strongest arguments for working with cDNA and with this, a limitation of the source library, applied here.

The usage of magnetic beads as solid phase for immobilizing sera autoantibodies allowed a semi-automated way of bio-panning on the King Fischer magnetic particle processor. This proved very advantageous in regard of simplified and rapid bio-panning, as well as polyclonal phage ELISA procedures and allowed simultaneous handling of multiple samples. Another benefit was definitely the increased results reliability and with this, the decreased risk of manual aberrations. Thus, it can be expected that the ongoing trend towards high-throughput scale analyses will also increase automation steps in screening protocols, especially in such well-automatable techniques, like phage display [285, 286].

In general, enrichment was achieved and visualized by polyclonal phage ELISA in all samples, hence, with all pYG/full-ORF libraries (Fig. 34 – 36). No differences were observed between selections against IgG or IgA autoantibodies. With the pYG-LZ libraries nearly all samples showed highest signals in the 4<sup>th</sup> selection round. Only two samples showed in the 3<sup>rd</sup> round stronger signals than in the 4<sup>th</sup> selection round (Fig. 35/pYG-LZ-peIB/OCAA/AD1 and Fig. 35/pYG-LZ-peIB/OCAB/MS2). Interestingly, ELISA results of pYG-fusion samples obviously differed from pYG-LZ (Fig. 36). Here, half of the samples showed strong signals already in the 3<sup>rd</sup> round. Six of them even had weaker signals in the following 4<sup>th</sup> round. Usually, four to six selection rounds are accepted as a standard in the phage display process [287]. However, depending on experimental intention and library characteristics also a lower number of rounds can be sufficient or even recommendable. Thus, in combinatorial antibody libraries, where presented agents are uniform in their overall structure, a more stringent selection procedure with many rounds can be advantageous. This is especially estimable, if the intention is to find a high-affinity binder with best binding characteristics. But in other cases excessive rounds can lead to an undesired phenomenon of “overselecting” and possible loss of potentially interesting candidates or simply have no additional benefit. Although it is difficult, to make a general recommendation, t’Hoen *et al.* suggest applying multiple selection rounds when working with smaller libraries and sequencing limited numbers of clones [288]. This observation correlates well with my results, given the high number of samples, showing enrichment in the 4<sup>th</sup> round, but not in the previous ones.

Still, the question remains, why the pYG-fusion constructs showed earlier enrichment levels, as their pYG-LZ counterparts. To my knowledge, this is the first time that both pIII-catenation types are used in parallel and in comparison, so no published references can be given at that point. However, it can be speculated, if the much stronger binding of the POI to pIII in pYG-fusion samples leads to a better retention of selected binders and thus to a more rapid enrichment. On the contrary, it is thinkable that the indirect fusion via a leucine zipper gets at times mechanically destroyed, leading to a loss of selected phage particles. Thus, more selection rounds may be necessary, to make enrichment with pYG-LZ-vectors visible. To

confirm these implications, it would be interesting to sequence phage populations from all four rounds, in order to determine the best point of enrichment for both constructs. In conclusion, the usual limitations of the phage display system, often leading to enrichment of false positive target candidates, should be taken into consideration as well. These were recently reviewed anew by Vodnik *et al.* [289].

Illumina sequencing results of the initial libraries, revealed that over 80% of all full-ORF genes in the original source library could be re-found in the pYG-libraries as well (Fig. 24). This is already a good rate, but still further improvable, since final goal is to be able to find all genes in the expression libraries after all cloning procedures. Since sequencing was performed by a company, it is difficult to say, if given results are partially artifacts or really reflect libraries' state. My analysis of the final sequencing results revealed that also genes, not present in the initial libraries could be identified later in the selected ones. This confirms my speculations that a too stringent sequence and/or mapping cut-off(s) have been possibly set and led to artificially restricted libraries' coverage. Unfortunately, it was not possible to estimate a distinct error source.

Very interesting results were gained from sequencing phage DNA from enriched libraries. First, as expected and already shown in the macroarray screenings, healthy sera proved as immunoreactive as diseased ones. Thus, overall numbers of enriched genes against IgG and IgA autoantibodies were comparable in all three cohorts (Table 13). Noticeable differences became visible after filtering out enriched cohort-unique genes: 74 in the AD group, 12 in the MS group and 9 in the Healthy group. Thus, as already seen in the macroarray screenings, Alzheimer's sera showed a relatively high ratio of unique autoantigens, with a potential for good biomarker candidates. Also Healthy sera data were coherent with those from macroarray screenings, which is a relatively low number of enriched cohort-unique antigens. Surprisingly, MS data set was this time more comparable with the Healthy, as with the AD results. In conclusion, also from screening with the phage display sufficient number of cohort-unique autoantigens could be identified. To prove their potential as diagnostic biomarkers, some promising candidates were separately expressed, affinity purified and tested in ELISA with all available 60 human sera.

### **6.3 Results from biomarkers validation in ELISA**

In summary, overall 15 sera (5 per donor group) were screened on protein macroarrays and 6 of these (2 per donor group) were screened in phage display as well. This is certainly not a statistically representative number of screened samples and with this, probably the biggest deficit of this study. However, estimated results demonstrate clearly the applicability and potential of both systems as diagnostic screening platforms. Thus, autoantigens, analyzed

here, have an exemplary role, showing the broad range of possibilities, which are still open for investigation. Furthermore, since both platforms are now well established and full-ORF-libraries ready to use in phage display, many more sera can be screened in future.

For recombinant expression and affinity purification overall 12 biomarker candidates were chosen (Table 14). From these only one originated from an IgG-screening. All others were identified in an IgA screening. This was not an intended effect, but resulted rather from the priority criteria for hit and enriched clones in general. Thus, because of the relatively low number of selected antigens, this finding should not be overvalued. In my opinion, not a sufficient number of sera was screened to be able to conclude, which of the two antibody classes shows more specific autoimmunoreactivity and if there is any disease correlation to a certain Ig class. Screening for autoantigens against both IgG and IgA antibodies has been shown in only a few studies so far. Thus, recently Sakaguchi *et al.* detected strong IgG and IgA crossreactivities against multiple antigens in cases of bullous dermatosis and proposed epitope-spreading mechanisms [290]. Also Pedersen *et al.* [291] and Blixt *et al.* [292] were able to show that both IgG and IgA antibodies were highly specific to one combined epitope of an O-glycan and protein backbone in cases of colorectal cancer. In conclusion, also my data suggest strong immunoreactivity of serum IgA autoantibodies in AD and MS patients, on the one hand. On the other hand, my ELISA results clearly show that no class-specific antigens could be identified and thus, correlate well with findings, described by others.

From the 12 chosen candidates, 10 could be successfully expressed and from these, 6 were also sufficiently purified and subsequently applied as coating agents in ELISA assays: PRDX1, ANXA2, DTNBP1, NDRG4, ANKHD1 and DEAF1. In case of CCDC50, in the time scale of my thesis, it was not possible to find out, why no expression products could be achieved. Yet, it was one of the most interesting macroarray hits and thus a very potential biomarker candidate for AD. Since Sanger sequencing of the chosen clone showed a correct reading frame and presence of the whole CDS-region in the cDNA fragment, it is worth considering re-cloning of the ORF-polypeptide in a new expression vector and repeating the experiments. All three available CCDC50 clones are expressed successfully in *E. coli* on the macroarray filter. Hence, no strong toxic effects for the bacterial host should be expected. For the four candidates, which were successfully expressed, but insufficiently purified, i.e. TANK, TRAF4, GDI1 and PAX6, certainly better results can be achieved. Here usual improvements of the affinity purification procedure are recommendable, such as slowing down the flow through rate, especially during sample loading and elution. Also altered concentrations of the eluting agent imidazole could bring desirable effects in some cases. Finally, changing the Ni-column to a Cd-loaded one with a lower affinity, could improve

purification quality. This could be advantageous especially in samples with relatively high unspecific background, such as in TRAF4 and GDI1.

Coming back to the six positive results, intriguing and partly surprising results were achieved in the ELISA experiments. These findings demonstrate once again, the importance of larger sample sets in screenings. Here analyses with appropriate statistical tests could be applied for the first time, since each of the three cohorts contained all 20 sera available and the groups were large enough for reliable calculations of distributions and coherent P-values. Therefore, when enlarging sample groups also in macroarray and phage display screenings, same tests should be applied as well. I assume that for a serious assignment of a real disease-related biomarker a minimum of ten screened sera has to be set. However, the more sera in the comparison groups are screened and analyzed, the bigger the significance and statistical reliability of the identified biomarker would be. One of the biggest limitations is certainly the availability of well annotated serum material. But also time, laboratory staff and costs, for example for next generation sequencing, have to be taken into consideration.

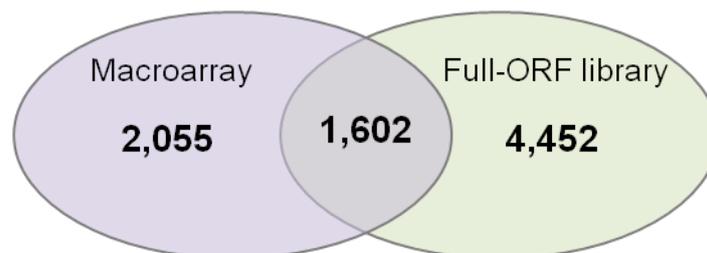
Of all tested antigens, three showed significant P-values in the pairwise group comparisons: PRDX1 (Fig. 43), NDRG4 (Fig. 52) and DEAF1 (Fig. 130). PRDX1 was one of the most promising candidates, as it was very successfully expressed and purified. Although identified in a screening with IgA antibodies, in ELISA it strongly reacted with the IgG antibody as well. This was observed by virtually all tested antigens and confirmed my conclusion that no explicit distinction can be made between IgG- and IgA-specific autoantigens. But, as already discussed above, an artificial impact, due to the very low number of screened samples, cannot be excluded. In order to omit this effect, distributions with all signals, summed up from one cohort, were calculated, i.e. without splitting by the detecting antibody (shown under B. in all ELISA analysis figures). As expected, estimated significances recurred in the summarized representations as well. Thus, significantly low P-values were estimated in the comparison AD ↔ Healthy, but also in MS ↔ Healthy. Since PRDX1 was identified as a potential AD biomarker, the second result was rather unexpected. It can be speculated, if presence of α-PRDX1 antibodies in patients' sera could be an indication of neurodegeneration as a whole. As peroxiredoxins are an important part of the oxidative stress pathway, this is certainly thinkable. Yet, many more tests with larger sample groups and in comparison with other neurodegenerative disorders will be necessary to confirm this.

Highly interesting results showed analysis of the ELISA tests with DEAF1. Here, the lowest and with this, the most significant P-values were estimated. However, these were found not in the expected comparative group MS ↔ Healthy, but rather in the AD ↔ Healthy group. No correlation of DEAF1 to neurodegeneration and/or inflammation is known so far. Hence, at

this stage no serious biological interpretation can be made about the presence of autoantibodies in patients' sera.

In conclusion, no autoantigenicity profiles could be compiled with the low number of analyzed antigens. However, the analyzed 12 antigens make up only a relatively small part of all hits, generated in both screenings (s. Supplementary material). Thus, many more potential biomarker candidates can be further recombinantly expressed and tested in ELISA, preferably with even larger sera sets. Another important issue is the biological correlation of the identified antigens. Beyond doubt, this is very interesting from a biological point of view. But of note, a good diagnostic biomarker does not have to be directly linked to the accordant disease. Technically, it only has to deliver high specificities and sensitivities, when compared to healthy samples. And these can only be reliably estimated in bigger cohorts. However, the first step of screening samples on a high-throughput scale is a very effective way to filter potential candidates. And both, phage display and protein arrays, are powerful and productive technologies for this, as successfully demonstrated in my thesis.

A final subject of discussion is the overlap between the autoantigens found in the macroarray screening and the enriched ones in the phage selection. No correlation could be estimated by any of the identified hits in any of the three cohorts. To a great extent, the identified antigens from the macroarray screening were simply not present in the full-ORF library and vice versa. When looking at the primal overlap between the two source libraries, depicted in Fig. 60 below, 44% of all macroarray genes are also present in the full-ORF library and 26% of the full-ORF genes are also spotted on the macroarrays.



**Fig. 60 Intersection between shared genes spotted on the protein macroarrays and genes represented in the human full-ORF library.** From all spotted clones on the macroarrays only clones in frame, hence, real genes were considered.

In general, such a relatively minor overlap between the two source libraries is definitely desirable. The combination of the two screening platforms, as applied in this study, led to a remarkable expansion of the initial gene set. Thus, I was finally able to screen as much as 8,109 different human genes altogether. However, it is questionable, if it can be expected to find shared genes as significant hits in both screening types. Then the art and nature of the

presented polypeptides differ considerably between the two displaying scaffolds. Polypeptides, linked on the macroarray filters are cDNA-products, partly also containing UTR regions. Furthermore, they are in a denatured form. On the other hand, full-ORF polypeptides are full-length proteins, which are theoretically presented in a native form by at least one of the five vector constructs. Thus, it can be expected that different epitopes are formed during the two screening types and the chance is relatively small to identify exactly the same hits at the end. This consideration was finally confirmed by the comparisons I made between the results from both screenings. In conclusion, it should be taken into consideration that since both platforms are based on bacterial expression systems, no posttranslational modifications take place. This is certainly one disadvantage of the two techniques, when working with eukaryotic, e.g. human proteins. To address this problem, several combined systems for phage display have been already introduced by others. Thus, solutions for phosphorylation and phosphopantetheinylation [293] as well as for glycosylation [294, 295] were recently proposed.

#### **6.4 Could I achieve the aims of my thesis?**

With the development of the novel phage display protocols for functional presentation of human full-ORF proteins, for semi-automated antigen selection against human autoantibodies and finally for identification of the selected autoantigens via NGS, I could successfully establish a powerful high-throughput technique for rapid screening of biomarker candidates in human blood sera. Generally, I could demonstrate for the first time combined application of preassigned human full-ORF libraries and filamentous phage display. The newly generated full-ORF-phagemid libraries are ready for phage generation and are a valuable resource that can be used with all possible kinds of sera or other body fluids, applicable in screening processes.

Furthermore, I successfully applied two different screening platforms in parallel and managed to identify interesting and potential biomarker candidates in blood sera of AD and MS patients. First, I was able to confirm fully my starting hypothesis that primarily non-autoimmune neurodegenerative disorders, such as AD, expose specific autoantigenicity profiles, which can be used for diagnostic purposes. Results from both screening procedures entirely correlate at this point. Yet, due to the restricted number of screened sera as well as analyzed antigens in total, no disease-specific autoantigen patterns could be estimated. In future, number of screened samples per cohort will have to be enlarged. In addition, more antigens, already identified as potential candidates in this study (listed in the supplementary material), can be recombinantly expressed, purified and finally tested in ELISA assays. Thus, together with the autoantigens, already analyzed in this study, in future it should be possible

to compile a set of highly potential AD biomarkers for a possible application in multiparametric diagnostic assays.

Finally, I was able to identify two very promising AD biomarker candidates, PRDX1 and DEAF1, which performed very well in the ELISA experiments and showed significant P-values, especially in the statistical comparison with the healthy cohort. To my knowledge, this is the first time that the two proteins were shown to be probable disease-specific autoantigens in sera of AD patients.

## 7. Summary

Neurodegenerative diseases such as Morbus Alzheimer (AD) and Multiple Sclerosis (MS) affect millions of people each year. Since AD is directly related to higher age and human population gets constantly older, a significant increase of AD-cases is to be expected in the next decades. With this, also health policy costs for treatment and patients care will inevitably raise. Both disorders are incurable and characterized by a dramatic decline of life quality, causing heavy burdens to patients, family members and caretakers. Accordingly, both, pharmaceutical industry and academic science undertake great efforts in exploring treatment and diagnostic possibilities. Yet, in both cases, despite many significant achievements, major aspects of disease cause, risk factors or medication alternatives are still unclear.

Particularly, differential diagnosis of AD and MS proved to be a notably difficult area. Main disease manifestations like dementia by AD or muscle dysfunctions by MS, are shared by other neurodegenerative disorders and cannot be easily assigned. In addition, biopsying brain or spinal nerve tissue in a living person is not executable. Hence, clinical diagnosis of AD still mainly relies on psychological tests and medical history by proxy. Diagnostic procedures for MS include also MRI and spinal liquor analysis. However, the heterogeneity of the manifold MS-subforms and symptoms is an additional challenge. Thus, novel diagnostic tools, based on biochemical information, are desperately needed. These should allow precise, preferably early diagnosis and with this, better prognostic and therapeutic chances.

Blood serum is an ideal medium for diagnostic purposes. Withdrawal can be easily performed and is not harming for patients. Because of its role as a main transport system, it is a precious information carrier about organism's health status. Furthermore, all major players of the immune system, e.g. specialized cells, antibodies and molecular messengers circulate there. Thus, in the field of autoimmune disorders, many diagnostic platforms are based on detection and analysis of serum autoantibodies. But also diseases, which do not have a distinct autoimmune background, can potentially lead to pathologic changes in the natural autoimmunity.

According to my thesis hypothesis, AD is one such example. Although no distinct immunological reasons are currently known, it seems to provoke autoimmune reactions. As massive brain tissue degradation and progressive leakage in the brain blood barrier take place, it can be hypothesized that inflammatory processes lead to autoantibody production in order to eliminate necrotic organic material. Thus, AD patients' blood sera could feature specific autoantibody patterns, distinguishable from sera of healthy individuals. On the other

hand, MS is a classical example of an autoimmune disorder and served here as a comparative reference to AD.

Main aim of my work was to analyze AD-, MS- and Healthy blood sera for disease specific autoantigens that could be applied as novel biomarkers for diagnostic purposes. This was done using two different high-throughput screening technologies: protein macroarrays with spotted cDNA expression clones on the one hand and phage display of human full-ORF libraries in combination with a next generation sequencing platform on the other.

Five sera from each of the three cohorts were screened on protein macroarrays for reaction against IgG and IgA autoantibodies. For the second screening I generated new phage vector series first, compatible with the Gateway cloning system. Special focus was laid on functional presentation of the full-ORF-polypeptides on phage surface. For this, I implemented all three major *E. coli* secretion pathways. This step led to a considerable expansion of the range of properly folded and presented proteins. On the example of EGFP, I could finally demonstrate usability and functionality of the new vector series. Next, full-ORF phagemid libraries were generated and bio-panning with four-round selections with two sera from each of the three cohorts were performed in a semi-automatic way. Again, autoantigen detection was performed with IgG and IgA sera autoantibodies. DNA from both initial and selected phages was deep sequenced on an Illumina Genome Analyzer platform. In both screenings numerous disease-specific AD- and MS-autoantigens could be identified. Potential candidates with best performances were recombinantly expressed in *E. coli* and affinity purified. These were finally re-validated with an enlarged set of 20 sera from each cohort, in ELISA assays.

In conclusion, my results fully confirmed my initial hypothesis that non-autoimmune disorders like AD evoke autoimmune interactions that are detectable in proteomics-based high-throughput screenings. The novel disease-related autoantigens are potentially promising biomarker candidates for multiparametric diagnostic assays. Furthermore, I could demonstrate that IgA autoantibodies exhibit pronounced immunoreactivity in blood serum and are therefore well-suited for this kind of screenings besides IgG. Finally, I succeeded in adapting the phage display technology for functional presentation of full-ORF polypeptides. The newly generated pYG vector series and human full-ORF-phagemid libraries are valuable resources that can be applied in diverse selections in future.

## 8. Zusammenfassung

Millionen von Menschen erkranken jährlich an neurodegenerativen Erkrankungen wie Morbus Alzheimer (AD) und Multiple Sklerose (MS). Aufgrund der direkten Relation zwischen AD und höherem Alter sowie der fortschreitenden Veralterung der Bevölkerung, kann in den nächsten Jahrzehnten eine erhebliche Zunahme an AD-Fällen erwartet werden. Folglich ist auch ein Anstieg der Gesundheitskosten für die Behandlung und Pflege unvermeidlich. Beide Erkrankungen sind nicht heilbar, charakterisieren sich durch eine dramatische Abnahme der Lebensqualität und verursachen schwere Belastungen für Patienten, Familienmitglieder und Pfleger. Entsprechend groß sind die Bemühungen in der pharmazeutischen Industrie und in der akademischen Forschung, neue therapeutische und diagnostische Ansätze zu finden. Trotz zahlreicher bedeutender Entdeckungen sind in beiden Fällen Krankheitsursachen, Risikofaktoren und Medikationsalternativen immer noch unklar.

Vor allem die Differenzialdiagnose von AD und MS erwies sich als besonders schwierig. Zentrale klinische Manifestationen wie Demenz bei AD oder Muskelfunktionsstörungen bei MS kommen auch bei anderen neurodegenerativen Erkrankungen vor und sind daher nicht leicht zuzuordnen. Außerdem sind Biopsien aus dem Hirn- oder Rückenmarksgewebe lebender Menschen nicht durchführbar. Daher beruht die klinische AD Diagnose immer noch hauptsächlich auf psychologischen Tests und Angehörigenanamnese. Diagnostische Maßnahmen für MS beinhalten auch MRI und die Analyse der Wirbelsäulenflüssigkeit. Allerdings ist die Heterogenität der unterschiedlichen MS-Subformen und –Symptome eine zusätzliche Schwierigkeit. Somit werden neue, auf biochemische Informationen basierende diagnostische Werkzeuge dringend gebraucht. Diese sollen eine präzise, möglichst frühe Diagnose und damit auch bessere prognostische und therapeutische Perspektiven bieten können.

Blutserum ist ein ideales Medium für diagnostische Zwecke. Die Entnahme ist einfach durchzuführen und ist nicht invasiv für Patienten. Als Haupttransportsystem ist Blut ein wertvoller Informationsträger für den Gesundheitsstatus des Organismus. Des Weiteren zirkulieren im Blut alle Hauptkomponente des Immunsystems, wie z.B. spezialisierte Zellen, Antikörper und molekulare Messenger. Auf dem Gebiet der Autoimmunerkrankungen basieren daher viele diagnostische Plattformen auf dem Nachweis und der Analyse von Serum-Autoantikörpern. Aber auch Krankheiten ohne erkennbaren autoimmunen Hintergrund können unter Umständen zu pathologischen Veränderungen der angeborenen Autoimmunität führen.

Der Hypothese meiner Dissertation zufolge ist AD ein Beispiel dafür. Obwohl bis dato keine deutlichen immunologischen Zusammenhänge bekannt sind, scheint AD Autoimmunreaktionen auszulösen. Aufgrund des massiven Abbaus des Hirngewebes und der zunehmenden Durchlässigkeit der Blut-Hirn-Schranke kann vermutet werden, dass Entzündungsprozesse zur Produktion von Autoantikörpern führen, die nekrotisches organisches Material beseitigen. Demzufolge könnten Blutseren von AD Patienten spezifische Autoantikörperprofile aufweisen, die sich von Seren gesunder Probanden unterscheiden lassen. MS ist andererseits ein klassisches Beispiel einer Autoimmunerkrankung und diente hier als Vergleichsreferenz zu AD.

Die Hauptaufgabe meiner Arbeit war es, Blut Seren von AD-, MS- und gesunden Probanden auf krankheitsspezifische Autoantigene zu analysieren, die als neue Biomarker für diagnostische Zwecke dienen könnten. Dafür wurden zwei unterschiedliche Hochdurchsatz-Screening Technologien eingesetzt: Protein Macroarrays mit gespotteten cDNA Expressionsklonen und Phagen-Display von humanen full-ORF Bibliotheken in Kombination mit einer Sequenzierungsplattform der nächsten Generation.

Fünf Seren aus jeder der drei Kohorten wurden auf Protein Macroarrays auf Reaktivitäten gegen IgG und IgA Autoantikörper gescreened. Für das zweite Screening generierte ich zunächst neue, mit dem Gateway-Klonierungssystem kompatible Phagen-Vektoren. Dabei wurde ein besonderer Fokus auf die funktionale Präsentation der full-ORF-Polypeptide auf der Phagenoberfläche gelegt. Dafür implementierte ich alle drei *E. coli* Hauptsekretionswege, was zu einer deutlichen Erweiterung der Menge an korrekt gefalteten und präsentierten Proteine führte. Am Beispiel mit EGFP konnte ich schließlich die Verwendbarkeit und Funktionalität der neuen Vektorreihe demonstrieren. Als nächstes wurden full-ORF-Phagen-Bibliotheken generiert und Bio-Panning mit vier Selektionsrunden und zwei Seren aus jeder der drei Kohorten im semi-automatischen Verfahren durchgeführt. Erneut wurden die Autoantigene mit IgG und IgA Autoantikörpern detektiert. Die DNS von Ausgangs-, wie auch von selektierten Phagen wurde mittels der Illumina Genome Analyzer Plattform durchsequenziert. Bei beiden Screenings konnten krankheitsspezifische AD- und MS-Autoantigene identifiziert werden. Die potentiellen Kandidaten mit den besten Ergebnissen wurden in *E. coli* exprimiert und mittels Affinitätschromatographie aufgereinigt. Diese wurden schließlich mit größeren Sätzen von 20 Seren aus jeder Kohorte im ELISA-Experiment revalidiert.

Zusammenfassend bestätigen meine Ergebnisse meine Hypothese, dass auch nicht-autoimmune Erkrankungen wie AD autoimmune Interaktionen hervorrufen können, die in Proteomics-basierten Hochdurchsatz-Screenings nachweisbar sind. Die neuen 10 AD und 3

MS Autoantigene sind vielversprechende Biomarker Kandidaten für multiparametrische diagnostische Tests. Des Weiteren konnte ich zeigen, dass IgA Autoantikörper eine ausgeprägte Immunreaktivität im Blutserum aufweisen und somit neben IgG sehr gut für Screenings dieser Art geeignet sind. Schließlich gelang es mir, die Phagen Display Technologie für die funktionale Präsentation von full-ORF-Polypeptiden zu adaptieren. Die neu generierte pYG Vektorreihe und die humanen full-ORF-Phagemid-Bibliotheken sind wertvolle Ressourcen, die zukünftig in verschiedensten Selektionen eingesetzt werden können.

## 9. References

1. Eurostat.  
[http://epp.eurostat.ec.europa.eu/statistics\\_explained/index.php/Mortality\\_and\\_life\\_expectancy\\_statistics](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Mortality_and_life_expectancy_statistics). 2010.
2. Robustillo, S.A., et al., *EU Employment and social situation. A quarterly review*. 2013, Eurostat: Luxemburg.
3. Lanzieri, G., *The greying of the baby boomers. A century-long view of ageing in European populations in Statistics in focus*. 2011, Eurostat: Luxemburg.
4. WHO, *Dementia: a public health priority*. 2012, WHO: Geneva.
5. Batsch, N. and M. Mittelman, *World Alzheimer Report 2012. Overcoming the stigma of dementia*. 2012, Alzheimer's Disease International: London.
6. Prince, M., et al., *The global prevalence of dementia: a systematic review and metaanalysis*. *Alzheimers Dement*, 2013. **9**(1): p. 63-75 e2.
7. Ballard, C.G., et al., *The natural history of psychosis and depression in dementia with Lewy bodies and Alzheimer's disease: persistence and new cases over 1 year of follow-up*. *J Clin Psychiatry*, 2001. **62**(1): p. 46-9.
8. Talassi, E., et al., *Personality changes in Alzheimer's disease*. *Aging Ment Health*, 2007. **11**(5): p. 526-31.
9. Robins Wahlin, T.B. and G.J. Byrne, *Personality changes in Alzheimer's disease: a systematic review*. *Int J Geriatr Psychiatry*, 2011. **26**(10): p. 1019-29.
10. Keating, N. and N. Gaudet, *Quality of life of persons with dementia*. *J Nutr Health Aging*, 2012. **16**(5): p. 454-6.
11. Lopez, O.L. and S.T. Dekosky, *Clinical symptoms in Alzheimer's disease*. *Handb Clin Neurol*, 2008. **89**: p. 207-16.
12. Alonso, A. and M.A. Hernan, *Temporal trends in the incidence of multiple sclerosis: a systematic review*. *Neurology*, 2008. **71**(2): p. 129-35.
13. Chiaravalloti, N.D. and J. DeLuca, *Cognitive impairment in multiple sclerosis*. *Lancet Neurol*, 2008. **7**(12): p. 1139-51.
14. Hoffmann, S., M. Tittgemeyer, and D.Y. von Cramon, *Cognitive impairment in multiple sclerosis*. *Curr Opin Neurol*, 2007. **20**(3): p. 275-80.
15. Compston, A. and A. Coles, *Multiple sclerosis*. *Lancet*, 2008. **372**(9648): p. 1502-17.
16. Rosati, G., *The prevalence of multiple sclerosis in the world: an update*. *Neurol Sci*, 2001. **22**(2): p. 117-39.
17. Trisolini, M., et al., *Global economic impact of multiple sclerosis*. 2010, RTI International.
18. Getsios, D., et al., *An economic evaluation of early assessment for Alzheimer's disease in the United Kingdom*. *Alzheimers Dement*, 2012. **8**(1): p. 22-30.
19. Cummings, J.L., et al., *Alzheimer's disease: etiologies, pathophysiology, cognitive reserve, and treatment opportunities*. *Neurology*, 1998. **51**(1 Suppl 1): p. S2-17; discussion S65-7.
20. Wenk, G.L., *Neuropathologic changes in Alzheimer's disease*. *J Clin Psychiatry*, 2003. **64 Suppl 9**: p. 7-10.
21. Seab, J.P., et al., *Quantitative NMR measurements of hippocampal atrophy in Alzheimer's disease*. *Magn Reson Med*, 1988. **8**(2): p. 200-8.
22. Double, K.L., et al., *Topography of brain atrophy during normal aging and Alzheimer's disease*. *Neurobiol Aging*, 1996. **17**(4): p. 513-21.
23. Knopman, D.S., et al., *Neuropathology of cognitively normal elderly*. *J Neuropathol Exp Neurol*, 2003. **62**(11): p. 1087-95.
24. Jendroska, K., et al., *Amyloid beta-peptide and its relationship with dementia in Lewy body disease*. *J Neural Transm Suppl*, 1997. **51**: p. 137-44.
25. Jendroska, K., et al., *Amyloid beta-peptide and the dementia of Parkinson's disease*. *Mov Disord*, 1996. **11**(6): p. 647-53.

26. Claeysen, S., et al., *Alzheimer culprits: cellular crossroads and interplay*. Cell Signal, 2012. **24**(9): p. 1831-40.
27. Nalivaeva, N.N. and A.J. Turner, *The amyloid precursor protein: A biochemical enigma in brain development, function and disease*. FEBS Lett, 2013.
28. Glenner, G.G. and C.W. Wong, *Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein*. Biochem Biophys Res Commun, 1984. **120**(3): p. 885-90.
29. Hiltunen, M., T. van Groen, and J. Jolkkonen, *Functional roles of amyloid-beta protein precursor and amyloid-beta peptides: evidence from experimental studies*. J Alzheimers Dis, 2009. **18**(2): p. 401-12.
30. De Strooper, B. and W. Annaert, *Proteolytic processing and cell biological functions of the amyloid precursor protein*. J Cell Sci, 2000. **113** ( Pt 11): p. 1857-70.
31. Tyan, S.H., et al., *Amyloid precursor protein (APP) regulates synaptic structure and function*. Mol Cell Neurosci, 2012. **51**(1-2): p. 43-52.
32. Castellani, R.J. and M.A. Smith, *Compounding artefacts with uncertainty, and an amyloid cascade hypothesis that is 'too big to fail'*. J Pathol, 2011. **224**(2): p. 147-52.
33. Hardy, J., *The amyloid hypothesis for Alzheimer's disease: a critical reappraisal*. J Neurochem, 2009. **110**(4): p. 1129-34.
34. Mullane, K. and M. Williams, *Alzheimer's therapeutics: continued clinical failures question the validity of the amyloid hypothesis-but what lies beyond?* Biochem Pharmacol, 2013 **85**(3): p. 289-305.
35. Berlau, D.J., et al., *Dissociation of neuropathologic findings and cognition: case report of an apolipoprotein E epsilon2/epsilon2 genotype*. Arch Neurol, 2007. **64**(8): p. 1193-6.
36. Lee, G. and C.J. Leugers, *Tau and tauopathies*. Prog Mol Biol Transl Sci, 2012. **107**: p. 263-93.
37. Weingarten, M.D., et al., *A protein factor essential for microtubule assembly*. Proc Natl Acad Sci U S A, 1975. **72**(5): p. 1858-62.
38. Caceres, A., S. Potrebic, and K.S. Kosik, *The effect of tau antisense oligonucleotides on neurite formation of cultured cerebellar macroneurons*. J Neurosci, 1991. **11**(6): p. 1515-23.
39. Ballatore, C., V.M. Lee, and J.Q. Trojanowski, *Tau-mediated neurodegeneration in Alzheimer's disease and related disorders*. Nat Rev Neurosci, 2007. **8**(9): p. 663-72.
40. Rapoport, M., et al., *Tau is essential to beta -amyloid-induced neurotoxicity*. Proc Natl Acad Sci U S A, 2002. **99**(9): p. 6364-9.
41. Vossel, K.A., et al., *Tau reduction prevents Abeta-induced defects in axonal transport*. Science, 2010. **330**(6001): p. 198.
42. Alonso, A.C., et al., *Mechanism of tau-induced neurodegeneration in Alzheimer disease and related tauopathies*. Curr Alzheimer Res, 2008. **5**(4): p. 375-84.
43. Small, S.A. and K. Duff, *Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis*. Neuron, 2008. **60**(4): p. 534-42.
44. Pritchard, S.M., et al., *The toxicity of tau in Alzheimer disease: turnover, targets and potential therapeutics*. J Cell Mol Med, 2011. **15**(8): p. 1621-35.
45. Arriagada, P.V., K. Marzloff, and B.T. Hyman, *Distribution of Alzheimer-type pathologic changes in nondemented elderly individuals matches the pattern in Alzheimer's disease*. Neurology, 1992. **42**(9): p. 1681-8.
46. Price, J.L., et al., *Neuropathology of nondemented aging: presumptive evidence for preclinical Alzheimer disease*. Neurobiol Aging, 2009. **30**(7): p. 1026-36.
47. Sunderland, T., et al., *Decreased beta-amyloid1-42 and increased tau levels in cerebrospinal fluid of patients with Alzheimer disease*. JAMA, 2003. **289**(16): p. 2094-103.
48. Spies, P.E., et al., *Reviewing reasons for the decreased CSF Abeta42 concentration in Alzheimer disease*. Front Biosci, 2012. **17**: p. 2024-34.

49. Mulder, C., et al., *Amyloid-beta(1-42), total tau, and phosphorylated tau as cerebrospinal fluid biomarkers for the diagnosis of Alzheimer disease*. Clin Chem, 2010. **56**(2): p. 248-53.
50. Petzold, A., et al., *Neurofilament ELISA validation*. J Immunol Methods, 2010. **352**(1-2): p. 23-31.
51. Verwey, N.A., et al., *A worldwide multicentre comparison of assays for cerebrospinal fluid biomarkers in Alzheimer's disease*. Ann Clin Biochem, 2009. **46**(Pt 3): p. 235-40.
52. del Campo, M., et al., *Recommendations to standardize preanalytical confounding factors in Alzheimer's and Parkinson's disease cerebrospinal fluid biomarkers: an update*. Biomark Med, 2012. **6**(4): p. 419-30.
53. Vanderstichele, H., et al., *Standardization of preanalytical aspects of cerebrospinal fluid biomarker testing for Alzheimer's disease diagnosis: a consensus paper from the Alzheimer's Biomarkers Standardization Initiative*. Alzheimers Dement, 2012. **8**(1): p. 65-73.
54. Morris, J.C. and J.L. Price, *Pathologic correlates of nondemented aging, mild cognitive impairment, and early-stage Alzheimer's disease*. J Mol Neurosci, 2001. **17**(2): p. 101-18.
55. Shaw, L.M., et al., *Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects*. Ann Neurol, 2009. **65**(4): p. 403-13.
56. De Meyer, G., et al., *Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people*. Arch Neurol, 2010. **67**(8): p. 949-56.
57. Kang, J.H., et al., *Simultaneous analysis of cerebrospinal fluid biomarkers using microsphere-based xMAP multiplex technology for early detection of Alzheimer's disease*. Methods, 2012. **56**(4): p. 484-93.
58. Hansson, O., et al., *Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study*. Lancet Neurol, 2006. **5**(3): p. 228-34.
59. Benadiba, M., et al., *New molecular targets for PET and SPECT imaging in neurodegenerative diseases*. Rev Bras Psiquiatr, 2012. **34 Suppl 2**: p. S125-36.
60. Klunk, W.E., *Amyloid imaging as a biomarker for cerebral beta-amyloidosis and risk prediction for Alzheimer dementia*. Neurobiol Aging, 2011. **32 Suppl 1**: p. S20-36.
61. Aizenstein, H.J., et al., *Frequent amyloid deposition without significant cognitive impairment among the elderly*. Arch Neurol, 2008. **65**(11): p. 1509-17.
62. Jagust, W., et al., *What does fluorodeoxyglucose PET imaging add to a clinical diagnosis of dementia?* Neurology, 2007. **69**(9): p. 871-7.
63. Jack, C.R., Jr., et al., *11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnesic mild cognitive impairment*. Brain, 2008. **131**(Pt 3): p. 665-80.
64. Schoonenboom, N.S., et al., *CSF and MRI markers independently contribute to the diagnosis of Alzheimer's disease*. Neurobiol Aging, 2008. **29**(5): p. 669-75.
65. Jack, C.R., Jr., *Alzheimer disease: new concepts on its neurobiology and the clinical role imaging will play*. Radiology, 2012. **263**(2): p. 344-61.
66. Jack, C.R., Jr., et al., *Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade*. Lancet Neurol, 2010. **9**(1): p. 119-28.
67. McKhann, G., et al., *Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease*. Neurology, 1984. **34**(7): p. 939-44.
68. Alzheimer's Association. *New Diagnostic Criteria and Guidelines for Alzheimer's Disease*. 2013; Available from: [http://www.alz.org/research/diagnostic\\_criteria/](http://www.alz.org/research/diagnostic_criteria/).
69. Aprahamian, I., et al., *The accuracy of the Clock Drawing Test compared to that of standard screening tests for Alzheimer's disease: results from a study of Brazilian elderly with heterogeneous educational backgrounds*. Int Psychogeriatr, 2010. **22**(1): p. 64-71.
70. Palmqvist, S., et al., *Practical suggestions on how to differentiate dementia with Lewy bodies from Alzheimer's disease with common cognitive tests*. Int J Geriatr Psychiatry, 2009. **24**(12): p. 1405-12.

71. Ala, T.A., et al., *The Mini-Mental State exam may help in the differentiation of dementia with Lewy bodies and Alzheimer's disease*. Int J Geriatr Psychiatry, 2002. **17**(6): p. 503-9.
72. Borroni, B., et al., *Combined biomarkers for early Alzheimer disease diagnosis*. Curr Med Chem, 2007. **14**(11): p. 1171-8.
73. Lublin, F.D. and S.C. Reingold, *Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis*. Neurology, 1996. **46**(4): p. 907-11.
74. Olek, M.J., *Multiple sclerosis--Part 2. Treatment strategies*. J Am Osteopath Assoc, 1999. **99**(12): p. 611-9.
75. Olek, M.J., *Multiple sclerosis--Part 1. Overview, pathophysiology, diagnostic evaluation, and clinical parameters*. J Am Osteopath Assoc, 1999. **99**(11): p. 574-88.
76. Koch, M., et al., *The natural history of primary progressive multiple sclerosis*. Neurology, 2009. **73**(23): p. 1996-2002.
77. Miller, D.H., D.T. Chard, and O. Ciccarelli, *Clinically isolated syndromes*. Lancet Neurol, 2012. **11**(2): p. 157-69.
78. McDonald, W.I., et al., *Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis*. Ann Neurol, 2001. **50**(1): p. 121-7.
79. Polman, C.H., et al., *Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria*. Ann Neurol, 2011. **69**(2): p. 292-302.
80. Miller, D.H., et al., *Differential diagnosis of suspected multiple sclerosis: a consensus approach*. Mult Scler, 2008. **14**(9): p. 1157-74.
81. Eckstein, C., S. Saidha, and M. Levy, *A differential diagnosis of central nervous system demyelination: beyond multiple sclerosis*. J Neurol, 2012. **259**(5): p. 801-16.
82. Halliday, A.M., W.I. McDonald, and J. Mushin, *Visual evoked response in diagnosis of multiple sclerosis*. Br Med J, 1973. **4**(5893): p. 661-4.
83. Graves, J. and L.J. Balcer, *Eye disorders in patients with multiple sclerosis: natural history and management*. Clin Ophthalmol, 2010. **4**: p. 1409-22.
84. Asselman, P., D.W. Chadwick, and D.C. Marsden, *Visual evoked responses in the diagnosis and management of patients suspected of multiple sclerosis*. Brain, 1975. **98**(2): p. 261-82.
85. Scolding, N., et al., *Immunological and inflammatory disorders of the central nervous system*. First edition ed. 1999, Oxford: Butterworth-Heinemann
86. Fjeldstad, A.S., N.G. Carlson, and J.W. Rose, *Optical coherence tomography as a biomarker in multiple sclerosis*. Expert Opin Med Diagn, 2012. **6**(6): p. 593-604.
87. Sakai, R.E., et al., *Vision in multiple sclerosis: the story, structure-function correlations, and models for neuroprotection*. J Neuroophthalmol, 2011. **31**(4): p. 362-73.
88. Tintore, M., et al., *Brainstem lesions in clinically isolated syndromes*. Neurology, 2010. **75**(21): p. 1933-8.
89. Tintore, M., et al., *Baseline MRI predicts future attacks and disability in clinically isolated syndromes*. Neurology, 2006. **67**(6): p. 968-72.
90. Swanton, J.K., et al., *MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study*. Lancet Neurol, 2007. **6**(8): p. 677-86.
91. Kim, S.S., et al., *Limited utility of current MRI criteria for distinguishing multiple sclerosis from common mimickers: primary and secondary CNS vasculitis, lupus and Sjogren's syndrome*. Mult Scler, 2013.
92. Dobson, R., et al., *Cerebrospinal fluid oligoclonal bands in multiple sclerosis and clinically isolated syndromes: a meta-analysis of prevalence, prognosis and effect of latitude*. J Neurol Neurosurg Psychiatry, 2013.
93. Filippini, G., et al., *Sensitivities and predictive values of paraclinical tests for diagnosing multiple sclerosis*. J Neurol, 1994. **241**(3): p. 132-7.

94. Hutchinson, M., *CSF oligoclonal bands are important in the diagnosis of multiple sclerosis, unreasonably downplayed by the McDonald Criteria 2010: Commentary*. *Mult Scler*, 2013. **19**(6): p. 719-20.
95. Alvarez-Cermeno, J.C. and L.M. Villar, *Multiple sclerosis: Oligoclonal bands-a useful tool to avoid MS misdiagnosis*. *Nat Rev Neurol*, 2013. **9**(6): p. 303-4.
96. Link, H. and Y.M. Huang, *Oligoclonal bands in multiple sclerosis cerebrospinal fluid: an update on methodology and clinical usefulness*. *J Neuroimmunol*, 2006. **180**(1-2): p. 17-28.
97. Kuenz, B., et al., *Cerebrospinal fluid B cells correlate with early brain inflammation in multiple sclerosis*. *PLoS One*, 2008. **3**(7): p. e2559.
98. Rinker, J.R., 2nd, K. Trinkaus, and A.H. Cross, *Elevated CSF free kappa light chains correlate with disability prognosis in multiple sclerosis*. *Neurology*, 2006. **67**(7): p. 1288-90.
99. Kanter, J.L., et al., *Lipid microarrays identify key mediators of autoimmune brain inflammation*. *Nat Med*, 2006. **12**(1): p. 138-43.
100. Fassbender, K., et al., *Increased release of interleukin-12p40 in MS: association with intracerebral inflammation*. *Neurology*, 1998. **51**(3): p. 753-8.
101. Malmestrom, C., et al., *IL-6 and CCL2 levels in CSF are associated with the clinical course of MS: implications for their possible immunopathogenic roles*. *J Neuroimmunol*, 2006. **175**(1-2): p. 176-82.
102. Matuszevicius, D., et al., *Interleukin-17 mRNA expression in blood and CSF mononuclear cells is augmented in multiple sclerosis*. *Mult Scler*, 1999. **5**(2): p. 101-4.
103. Sharief, M.K. and R. Hentges, *Association between tumor necrosis factor-alpha and disease progression in patients with multiple sclerosis*. *N Engl J Med*, 1991. **325**(7): p. 467-72.
104. Sellebjerg, F., et al., *Increased cerebrospinal fluid concentrations of the chemokine CXCL13 in active MS*. *Neurology*, 2009. **73**(23): p. 2003-10.
105. Szczuczinski, A. and J. Losy, *CCL5, CXCL10 and CXCL11 chemokines in patients with active and stable relapsing-remitting multiple sclerosis*. *Neuroimmunomodulation*, 2011. **18**(1): p. 67-72.
106. Rejdak, K., et al., *CSF nitric oxide metabolites are associated with activity and progression of multiple sclerosis*. *Neurology*, 2004. **63**(8): p. 1439-45.
107. Comabella, M., et al., *Cerebrospinal fluid chitinase 3-like 1 levels are associated with conversion to multiple sclerosis*. *Brain*, 2010. **133**(Pt 4): p. 1082-93.
108. Brettschneider, J., et al., *Tau protein level in cerebrospinal fluid is increased in patients with early multiple sclerosis*. *Mult Scler*, 2005. **11**(3): p. 261-5.
109. Salzer, J., A. Svenningsson, and P. Sundstrom, *Neurofilament light as a prognostic marker in multiple sclerosis*. *Mult Scler*, 2010. **16**(3): p. 287-92.
110. Petzold, A., et al., *Axonal damage accumulates in the progressive phase of multiple sclerosis: three year follow up study*. *J Neurol Neurosurg Psychiatry*, 2005. **76**(2): p. 206-11.
111. Berger, T., et al., *Antimyelin antibodies as a predictor of clinically definite multiple sclerosis after a first demyelinating event*. *N Engl J Med*, 2003. **349**(2): p. 139-45.
112. Tejada-Simon, M.V., et al., *Cross-reactivity with myelin basic protein and human herpesvirus-6 in multiple sclerosis*. *Ann Neurol*, 2003. **53**(2): p. 189-97.
113. Lamers, K.J., H.P. de Reus, and P.J. Jongen, *Myelin basic protein in CSF as indicator of disease activity in multiple sclerosis*. *Mult Scler*, 1998. **4**(3): p. 124-6.
114. Graber, J.J. and S. Dhib-Jalbut, *Biomarkers of disease activity in multiple sclerosis*. *J Neurol Sci*, 2011. **305**(1-2): p. 1-10.
115. Muldoon, L.L., et al., *Immunologic privilege in the central nervous system and the blood-brain barrier*. *J Cereb Blood Flow Metab*, 2013. **33**(1): p. 13-21.
116. Alvarez, J.I., T. Katayama, and A. Prat, *Glial influence on the blood brain barrier*. *Glia*, 2013. **61**(12): p. 1939-58.
117. Juhler, M. and E.A. Neuwelt, *The blood-brain barrier and the immune system, in Implications of the blood-brain barrier and its manipulation*, E.A. Neuwelt, Editor. 1989, Plenum Press: New York.

118. Matsui, M., K.J. Mori, and T. Saida, *Cellular immunoregulatory mechanisms in the central nervous system: characterization of noninflammatory and inflammatory cerebrospinal fluid lymphocytes*. *Ann Neurol*, 1990. **27**(6): p. 647-51.
119. Naparstek, Y., et al., *Activated T lymphocytes produce a matrix-degrading heparin sulphate endoglycosidase*. *Nature*, 1984. **310**(5974): p. 241-4.
120. Wekerle, H., *Breaking ignorance: the case of the brain*, in *Current concepts in autoimmunity and chronic inflammation*, A. Radbruch and P.E. Lipsky, Editors. 2006, Springer Verlag: Heidelberg.
121. Hayes, G.M., M.N. Woodroffe, and M.L. Cuzner, *Microglia are the major cell type expressing MHC class II in human white matter*. *J Neurol Sci*, 1987. **80**(1): p. 25-37.
122. Aloisi, F., F. Ria, and L. Adorini, *Regulation of T-cell responses by CNS antigen-presenting cells: different roles for microglia and astrocytes*. *Immunol Today*, 2000. **21**(3): p. 141-7.
123. Rodriguez, M., *Advances in multiple sclerosis and demyelinating disease*, ed. M. Rodriguez. 2008, Heidelberg: Springer Verlag.
124. Tuppo, E.E. and H.R. Arias, *The role of inflammation in Alzheimer's disease*. *Int J Biochem Cell Biol*, 2005. **37**(2): p. 289-305.
125. Grammas, P., *Neurovascular dysfunction, inflammation and endothelial activation: implications for the pathogenesis of Alzheimer's disease*. *J Neuroinflammation*, 2011. **8**: p. 26.
126. Radbruch, A. and P.E. Lipsky, *Current concepts in autoimmunity and chronic inflammation*, ed. A. Radbruch and P.E. Lipsky. 2006, Heidelberg: Springer Verlag.
127. Takeda, S., et al., *Increased blood-brain barrier vulnerability to systemic inflammation in an Alzheimer disease mouse model*. *Neurobiol Aging*, 2013. **34**(8): p. 2064-70.
128. Zlokovic, B.V., *Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders*. *Nat Rev Neurosci*, 2011. **12**(12): p. 723-38.
129. Minagar, A. and J.S. Alexander, *Blood-brain barrier disruption in multiple sclerosis*. *Mult Scler*, 2003. **9**(6): p. 540-9.
130. Assini, A., et al., *Plasma levels of amyloid beta-protein 42 are increased in women with mild cognitive impairment*. *Neurology*, 2004. **63**(5): p. 828-31.
131. Graff-Radford, N.R., et al., *Association of low plasma Abeta42/Abeta40 ratios with increased imminent risk for mild cognitive impairment and Alzheimer disease*. *Arch Neurol*, 2007. **64**(3): p. 354-62.
132. Abdullah, L., et al., *The influence of diagnosis, intra- and inter-person variability on serum and plasma Abeta levels*. *Neurosci Lett*, 2007. **428**(2-3): p. 53-8.
133. Motta, M., et al., *Altered plasma cytokine levels in Alzheimer's disease: correlation with the disease progression*. *Immunol Lett*, 2007. **114**(1): p. 46-51.
134. Engelhart, M.J., et al., *Inflammatory proteins in plasma and the risk of dementia: the rotterdam study*. *Arch Neurol*, 2004. **61**(5): p. 668-72.
135. Liao, P.C., et al., *Proteomics analysis of plasma for potential biomarkers in the diagnosis of Alzheimer's disease*. *Proteomics Clin Appl*, 2007. **1**(5): p. 506-12.
136. Ray, S., et al., *Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins*. *Nat Med*, 2007. **13**(11): p. 1359-62.
137. Britschgi, M. and T. Wyss-Coray, *Blood protein signature for the early diagnosis of Alzheimer disease*. *Arch Neurol*, 2009. **66**(2): p. 161-5.
138. Kienzl, E., et al., *A broader horizon of Alzheimer pathogenesis: ALZAS--an early serum biomarker?* *J Neural Transm Suppl*, 2002(62): p. 87-95.
139. Jellinger, K.A., et al., *Biomarkers for early diagnosis of Alzheimer disease: 'ALZheimer ASSociated gene'--a new blood biomarker?* *J Cell Mol Med*, 2008. **12**(4): p. 1094-117.
140. Nagele, E., et al., *Diagnosis of Alzheimer's disease based on disease-specific autoantibody profiles in human sera*. *PLoS One*, 2011. **6**(8): p. e23112.
141. Colasanti, T., et al., *Autoantibodies in patients with Alzheimer's disease: pathogenetic role and potential use as biomarkers of disease progression*. *Autoimmun Rev*, 2010. **9**(12): p. 807-11.

142. Hagman, S., et al., *Disease-associated inflammatory biomarker profiles in blood in different subtypes of multiple sclerosis: prospective clinical and MRI follow-up study*. J Neuroimmunol, 2011. **234**(1-2): p. 141-7.
143. Festoff, B.W., et al., *Soluble thrombomodulin levels in plasma of multiple sclerosis patients and their implication*. J Neurol Sci, 2012. **323**(1-2): p. 61-5.
144. Harberts, E., et al., *Translocator protein 18 kDa (TSPO) expression in multiple sclerosis patients*. J Neuroimmune Pharmacol, 2013. **8**(1): p. 51-7.
145. Quintana, F.J., et al., *Antigen microarrays identify unique serum autoantibody signatures in clinical and pathologic subtypes of multiple sclerosis*. Proc Natl Acad Sci U S A, 2008. **105**(48): p. 18889-94.
146. Nagele, E.P., et al., *Natural IgG autoantibodies are abundant and ubiquitous in human sera, and their number is influenced by age, gender, and disease*. PLoS One, 2013. **8**(4): p. e60726.
147. Glokler, J., T. Schutze, and Z. Konthur, *Automation in the high-throughput selection of random combinatorial libraries--different approaches for select applications*. Molecules, 2010. **15**(4): p. 2478-90.
148. Bussow, K., et al., *A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library*. Nucleic Acids Res, 1998. **26**(21): p. 5007-8.
149. Bussow, K., et al., *A human cDNA library for high-throughput protein expression screening*. Genomics, 2000. **65**(1): p. 1-8.
150. *Protein Macroarrays Manual*. 2010; Available from: [http://www.lifesciences.sourcebioscience.com/media/290406/sbs\\_ig\\_manual\\_protein\\_array\\_v1.pdf](http://www.lifesciences.sourcebioscience.com/media/290406/sbs_ig_manual_protein_array_v1.pdf).
151. Kowenz-Leutz, E., et al., *Crosstalk between C/EBPbeta phosphorylation, arginine methylation, and SWI/SNF/Mediator implies an indexing transcription factor code*. EMBO J, 2010. **29**(6): p. 1105-15.
152. O'Connell, D.J., et al., *Integrated protein array screening and high throughput validation of 70 novel neural calmodulin-binding proteins*. Mol Cell Proteomics, 2010. **9**(6): p. 1118-32.
153. Wobst, H., et al., *UCHL1 regulates ubiquitination and recycling of the neural cell adhesion molecule NCAM*. FEBS J, 2012. **279**(23): p. 4398-409.
154. Mahlknecht, U., O.G. Ottmann, and D. Hoelzer, *Far-Western based protein-protein interaction screening of high-density protein filter arrays*. J Biotechnol, 2001. **88**(2): p. 89-94.
155. Dervan, E.W., et al., *Protein macroarray profiling of serum autoantibodies in pseudoexfoliation glaucoma*. Invest Ophthalmol Vis Sci, 2010. **51**(6): p. 2968-75.
156. Kijanka, G., et al., *Human IgG antibody profiles differentiate between symptomatic patients with and without colorectal cancer*. Gut, 2010. **59**(1): p. 69-78.
157. Erdag, E., et al., *Switch-associated protein 70 antibodies in multiple sclerosis: relationship between increased serum levels and clinical relapse*. Inflamm Res, 2012. **61**(9): p. 927-30.
158. Cepok, S., et al., *Identification of Epstein-Barr virus proteins as putative targets of the immune response in multiple sclerosis*. J Clin Invest, 2005. **115**(5): p. 1352-60.
159. Becker, A., et al., *Myasthenia gravis: analysis of serum autoantibody reactivities to 1827 potential human autoantigens by protein macroarrays*. PLoS One, 2013. **8**(3): p. e58095.
160. Schmitt, J., et al., *Autoantibody signature differentiates Wilms tumor patients from neuroblastoma patients*. PLoS One, 2011. **6**(12): p. e28951.
161. Shugaiv, E., et al., *Progressive encephalomyelitis with rigidity and myoclonus: a syndrome with diverse clinical features and antibody responses*. Eur Neurol, 2013. **69**(5): p. 257-62.
162. Patel, S., et al., *RNASET2--an autoantigen in anaplastic large cell lymphoma identified by protein array analysis*. J Proteomics, 2012. **75**(17): p. 5279-92.
163. Vural, B., et al., *Anti-neuronal and stress-induced-phosphoprotein 1 antibodies in neuro-Behcet's disease*. J Neuroimmunol, 2011. **239**(1-2): p. 91-7.

164. Thurner, L., et al., *Wegener's granuloma harbors B lymphocytes with specificities against a proinflammatory transmembrane protein and a tetraspanin*. J Autoimmun, 2010. **36**(1): p. 87-90.
165. Woof, J.M. and M.A. Kerr, *The function of immunoglobulin A in immunity*. J Pathol, 2006. **208**(2): p. 270-82.
166. Jacob, C.M., et al., *Autoimmunity in IgA deficiency: revisiting the role of IgA as a silent housekeeper*. J Clin Immunol, 2008. **28 Suppl 1**: p. S56-61.
167. Monteiro, R.C., *The role of IgA and IgA Fc receptors as anti-inflammatory agents*. J Clin Immunol, 2010. **30 Suppl 1**: p. S61-4.
168. Cheng, Y., et al., *Circulating autoantibody to ABCC3 may be a potential biomarker for esophageal squamous cell carcinoma*. Clin Transl Oncol, 2013. **15**(5): p. 398-402.
169. McGann, P.T., et al., *IgA-mediated autoimmune hemolytic anemia in an infant*. Pediatr Blood Cancer, 2010. **56**(5): p. 837-9.
170. Byrne, G., et al., *Mutagenesis of the catalytic triad of tissue transglutaminase abrogates coeliac disease serum IgA autoantibody binding*. Gut, 2007. **56**(3): p. 336-41.
171. Kazemi-Shirazi, L., et al., *IgA autoreactivity: a feature common to inflammatory bowel and connective tissue diseases*. Clin Exp Immunol, 2002. **128**(1): p. 102-9.
172. Smith, G.P., *Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface*. Science, 1985. **228**(4705): p. 1315-7.
173. Danner, S. and J.G. Belasco, *T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries*. Proc Natl Acad Sci U S A, 2001. **98**(23): p. 12954-9.
174. Huse, W.D., et al., *Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda*. Science, 1989. **246**(4935): p. 1275-81.
175. McCafferty, J., et al., *Phage antibodies: filamentous phage displaying antibody variable domains*. Nature, 1990. **348**(6301): p. 552-4.
176. Rakonjac, J., et al., *Filamentous bacteriophage: biology, phage display and nanotechnology applications*. Curr Issues Mol Biol, 2011. **13**(2): p. 51-76.
177. Hawlisch, H., et al., *Site-specific anti-C3a receptor single-chain antibodies selected by differential panning on cellulose sheets*. Anal Biochem, 2001. **293**(1): p. 142-5.
178. D'Mello, F. and C.R. Howard, *An improved selection procedure for the screening of phage display peptide libraries*. J Immunol Methods, 2001. **247**(1-2): p. 191-203.
179. Walter, G., Z. Konthur, and H. Lehrach, *High-throughput screening of surface displayed gene products*. Comb Chem High Throughput Screen, 2001. **4**(2): p. 193-205.
180. Konthur, Z. and R. Cramer, *High-throughput applications of phage display in proteomic analyses*. Drug discovery today: TARGETS, 2003. **2**: p. 261-270.
181. Rodi, D., S. Mandava, and L. Makowski, *Filamentous bacteriophage structure and biology*, in *Phage display in biotechnology and drug discovery*, S. Sidhu, Editor. 2005, CRC Press: Boca Raton, Florida.
182. Felici, F., et al., *Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector*. J Mol Biol, 1991. **222**(2): p. 301-10.
183. Stengele, I., et al., *Dissection of functional domains in phage fd adsorption protein. Discrimination between attachment and penetration sites*. J Mol Biol, 1990. **212**(1): p. 143-9.
184. Deng, L.W., P. Malik, and R.N. Perham, *Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pilus of Escherichia coli*. Virology, 1999. **253**(2): p. 271-7.
185. Armstrong, J., R.N. Perham, and J.E. Walker, *Domain structure of bacteriophage fd adsorption protein*. FEBS Lett, 1981. **135**(1): p. 167-72.
186. Mead, D.A. and B. Kemper, *Chimeric single-stranded DNA phage-plasmid cloning vectors*. Biotechnology, 1988. **10**: p. 85-102.
187. Breitling, F., et al., *A surface expression vector for antibody screening*. Gene, 1991. **104**(2): p. 147-53.

188. Vieira, J. and J. Messing, *Production of single-stranded plasmid DNA*. Methods Enzymol, 1987. **153**: p. 3-11.
189. Rondot, S., et al., *A helper phage to improve single-chain antibody presentation in phage display*. Nat Biotechnol, 2001. **19**(1): p. 75-8.
190. Geyer, C.R., et al., *Recombinant antibodies and in vitro selection technologies*. Methods Mol Biol, 2012. **901**: p. 11-32.
191. Rentero, I. and C. Heinis, *Screening of large molecule diversities by phage display*. Chimia (Aarau), 2011. **65**(11): p. 843-5.
192. Konthur, Z., M. Hust, and S. Dubel, *Perspectives for systematic in vitro antibody generation*. Gene, 2005. **364**: p. 19-29.
193. Georgieva, Y. and Z. Konthur, *Design and screening of M13 phage display cDNA libraries*. Molecules, 2011. **16**(2): p. 1667-81.
194. Cramer, R. and M. Suter, *Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production*. Gene, 1993. **137**(1): p. 69-75.
195. Halazonetis, T.D., et al., *c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities*. Cell, 1988. **55**(5): p. 917-24.
196. Hust, M. and S. Dubel, *Phage display vectors for the in vitro generation of human antibody fragments*. Methods Mol Biol, 2005. **295**: p. 71-96.
197. Rapoza, M.P. and R.E. Webster, *The filamentous bacteriophage assembly proteins require the bacterial SecA protein for correct localization to the membrane*. J Bacteriol, 1993. **175**(6): p. 1856-9.
198. Steiner, D., et al., *Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display*. Nat Biotechnol, 2006. **24**(7): p. 823-31.
199. Paschke, M. and W. Hohne, *A twin-arginine translocation (Tat)-mediated phage display system*. Gene, 2005. **350**(1): p. 79-88.
200. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid*. Nature, 1953. **171**(4356): p. 737-8.
201. Nirenberg, M.W. and J.H. Matthaei, *The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides*. Proc Natl Acad Sci U S A, 1961. **47**: p. 1588-602.
202. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
203. Wetterstrand, K.A. *DNA sequencing costs: data from the NHGRI large-scale genome sequencing program*. 2012 1. Oktober 2012]; Available from: <http://www.genome.gov/sequencingcosts>.
204. Shendure, J. and E. Lieberman Aiden, *The expanding scope of DNA sequencing*. Nat Biotechnol, 2012. **30**(11): p. 1084-94.
205. Mardis, E.R., *Next-generation DNA sequencing methods*. Annu Rev Genomics Hum Genet, 2008. **9**: p. 387-402.
206. Schutze, T., et al., *A streamlined protocol for emulsion polymerase chain reaction and subsequent purification*. Anal Biochem, 2011. **410**(1): p. 155-7.
207. Gilles, A., et al., *Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing*. BMC Genomics, 2011. **12**: p. 245.
208. Di Niro, R., et al., *Rapid interactome profiling by massive sequencing*. Nucleic Acids Res, 2010. **38**(9): p. e110.
209. Pollack, A., *A DNA Chip Maker Acquires Gene-Sequencing Company*. New York Times, 2006.
210. Tucker, T., M. Marra, and J.M. Friedman, *Massively parallel sequencing: the next big thing in genetic medicine*. Am J Hum Genet, 2009. **85**(2): p. 142-54.
211. Illumina, I. 2013; Available from: <http://www.illumina.com/systems/sequencing/ilmn>.
212. Eid, J., et al., *Real-time DNA sequencing from single polymerase molecules*. Science, 2009. **323**(5910): p. 133-8.

213. Pacific Biosciences of California, I. *Photo Release -- Pacific Biosciences Launches the PacBio(R) RS II Sequencing System*. 2013 11. April 2013]; Available from: <http://investor.pacificbiosciences.com/releasedetail.cfm?ReleaseID=755828>.
214. Quail, M.A., et al., *A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers*. BMC Genomics, 2012. **13**: p. 341.
215. Glanville, J., et al., *Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire*. Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20216-21.
216. Dias-Neto, E., et al., *Next-generation phage display: integrating and comparing available molecular tools to enable cost-effective high-throughput analysis*. PLoS One, 2009. **4**(12): p. e8338.
217. Ravn, U., et al., *By-passing in vitro screening--next generation sequencing technologies applied to antibody display and in silico candidate selection*. Nucleic Acids Res, 2010. **38**(21): p. e193.
218. Ravn, U., et al., *Deep sequencing of phage display libraries to support antibody discovery*. Methods, 2013. **60**(1): p. 99-110.
219. Invitrogen, L.T., *Gateway Technology. A universal technology to clone DNA sequences for functional analysis and expression in multiple systems*. . 2010.
220. Invitrogen Dynal, L.T., *Dynabeads® MyOne™ Tosylactivated Manual*. 2006: Oslo.
221. Scientific, T.F., *Thermo Scientific KingFisher Flex User Manual*. 2010: Vantaa.
222. UniProt, C., *Ongoing and future developments at the Universal Protein Resource*. Nucleic Acids Res, 2011. **39**(Database issue): p. D214-9.
223. Feilmeier, B.J., et al., *Green fluorescent protein functions as a reporter for protein localization in Escherichia coli*. J Bacteriol, 2000. **182**(14): p. 4068-76.
224. Velappan, N., et al., *A comprehensive analysis of filamentous phage display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins*. Nucleic Acids Res, 2010. **38**(4): p. e22.
225. Förster, S., *Discovery of novel serum biomarkers by cDNA phage display*. Diploma thesis in biology, Humboldt Universität zu Berlin, 2011.
226. Krapfenbauer, K., et al., *Aberrant expression of peroxiredoxin subtypes in neurodegenerative disorders*. Brain Res, 2003. **967**(1-2): p. 152-60.
227. Cumming, R.C., et al., *Increase in expression levels and resistance to sulfhydryl oxidation of peroxiredoxin isoforms in amyloid beta-resistant nerve cells*. J Biol Chem, 2007. **282**(42): p. 30523-34.
228. Cheng, G. and D. Baltimore, *TANK, a co-inducer with TRAF2 of TNF- and CD 40L-mediated NF-kappaB activation*. Genes Dev, 1996. **10**(8): p. 963-73.
229. Albenis, B.C. and M.P. Mattson, *Evidence for the involvement of TNF and NF-kappaB in hippocampal synaptic plasticity*. Synapse, 2000. **35**(2): p. 151-9.
230. Meffert, M.K., et al., *NF-kappa B functions in synaptic signaling and behavior*. Nat Neurosci, 2003. **6**(10): p. 1072-8.
231. Salerno, K.M., et al., *TRAF family member-associated NF-kappa B activator (TANK) expression increases in injured sensory neurons and is transcriptionally regulated by Sox11*. Neuroscience, 2013. **231**: p. 28-37.
232. Ye, X., et al., *TRAF family proteins interact with the common neurotrophin receptor and modulate apoptosis induction*. J Biol Chem, 1999. **274**(42): p. 30202-8.
233. Waetzig, V. and T. Herdegen, *Context-specific inhibition of JNKs: overcoming the dilemma of protection and damage*. Trends Pharmacol Sci, 2005. **26**(9): p. 455-61.
234. Vlahopoulos, S. and V.C. Zoumpourlis, *JNK: a key modulator of intracellular signaling*. Biochemistry (Mosc), 2004. **69**(8): p. 844-54.
235. Wang, D., et al., *beta2 adrenergic receptor, protein kinase A (PKA) and c-Jun N-terminal kinase (JNK) signaling pathways mediate tau pathology in Alzheimer disease models*. J Biol Chem, 2013. **288**(15): p. 10298-307.
236. Gerke, V., C.E. Creutz, and S.E. Moss, *Annexins: linking Ca<sup>2+</sup> signalling to membrane dynamics*. Nat Rev Mol Cell Biol, 2005. **6**(6): p. 449-61.

237. van Genderen, H.O., et al., *Extracellular annexin A5: functions of phosphatidylserine-binding and two-dimensional crystallization*. Biochim Biophys Acta, 2008. **1783**(6): p. 953-63.
238. Sohma, H., et al., *Evaluation of annexin A5 as a biomarker for Alzheimer's disease and dementia with lewy bodies*. Front Aging Neurosci, 2013. **5**: p. 15.
239. Mishra, M., et al., *Gene expression analysis of frontotemporal lobar degeneration of the motor neuron disease type with ubiquitinated inclusions*. Acta Neuropathol, 2007. **114**(1): p. 81-94.
240. Numakawa, T., et al., *Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia*. Hum Mol Genet, 2004. **13**(21): p. 2699-708.
241. Di Pietro, S.M., et al., *BLOC-1 interacts with BLOC-2 and the AP-3 complex to facilitate protein trafficking on endosomes*. Mol Biol Cell, 2006. **17**(9): p. 4027-38.
242. Talbot, K., et al., *Dysbindin-1 is reduced in intrinsic, glutamatergic terminals of the hippocampal formation in schizophrenia*. J Clin Invest, 2004. **113**(9): p. 1353-63.
243. Trost, S., et al., *The DTNBP1 (dysbindin-1) gene variant rs2619522 is associated with variation of hippocampal and prefrontal grey matter volumes in humans*. Eur Arch Psychiatry Clin Neurosci, 2013. **263**(1): p. 53-63.
244. Carr, G.V., et al., *Loss of dysbindin-1 in mice impairs reward-based operant learning by increasing impulsive and compulsive behavior*. Behav Brain Res, 2013. **241**: p. 173-84.
245. Fatjo-Vilas, M., et al., *Dysbindin-1 gene contributes differentially to early- and adult-onset forms of functional psychosis*. Am J Med Genet B Neuropsychiatr Genet, 2011. **156B**(3): p. 322-33.
246. Bachner, D., et al., *Expression patterns of two human genes coding for different rab GDP-dissociation inhibitors (GDIs), extremely conserved proteins involved in cellular transport*. Hum Mol Genet, 1995. **4**(4): p. 701-8.
247. D'Adamo, P., et al., *Mutations in GDI1 are responsible for X-linked non-specific mental retardation*. Nat Genet, 1998. **19**(2): p. 134-9.
248. Bianchi, V., et al., *Forebrain deletion of alphaGDI in adult mice worsens the pre-synaptic deficit at cortico-lateral amygdala synaptic connections*. PLoS One, 2012. **7**(1): p. e29763.
249. Ginsberg, S.D., et al., *Upregulation of select rab GTPases in cholinergic basal forebrain neurons in mild cognitive impairment and Alzheimer's disease*. J Chem Neuroanat, 2011. **42**(2): p. 102-10.
250. Li, J., et al., *Differential regulation of amyloid-beta endocytic trafficking and lysosomal degradation by apolipoprotein E isoforms*. J Biol Chem, 2012. **287**(53): p. 44593-601.
251. Melotte, V., et al., *The N-myc downstream regulated gene (NDRG) family: diverse functions, multiple applications*. FASEB J, 2010. **24**(11): p. 4153-66.
252. Zhou, R.H., et al., *Characterization of the human NDRG gene family: a newly identified member, NDRG4, is specifically expressed in brain and heart*. Genomics, 2001. **73**(1): p. 86-97.
253. Schilling, S.H., et al., *NDRG4 is required for cell cycle progression and survival in glioblastoma cells*. J Biol Chem, 2009. **284**(37): p. 25160-9.
254. Taketomi, Y., et al., *Impaired mast cell maturation and degranulation and attenuated allergic responses in Ndr1-deficient mice*. J Immunol, 2007. **178**(11): p. 7042-53.
255. Mitchelmore, C., et al., *NDRG2: a novel Alzheimer's disease associated protein*. Neurobiol Dis, 2004. **16**(1): p. 48-58.
256. Nichols, N.R., et al., *Glucocorticoid regulation of glial responses during hippocampal neurodegeneration and regeneration*. Brain Res Brain Res Rev, 2005. **48**(2): p. 287-301.
257. Takahashi, K., et al., *Ndr2 promotes neurite outgrowth of NGF-differentiated PC12 cells*. Neurosci Lett, 2005. **388**(3): p. 157-62.
258. Ohki, T., et al., *Inhibition of neurite outgrowth by reduced level of NDRG4 protein in antisense transfected PC12 cells*. Brain Res Dev Brain Res, 2002. **135**(1-2): p. 55-63.
259. Davis, L.K., et al., *Pax6 3' deletion results in aniridia, autism and mental retardation*. Hum Genet, 2008. **123**(4): p. 371-8.

260. Nishi, M., et al., *A case of novel de novo paired box gene 6 (PAX6) mutation with early-onset diabetes mellitus and aniridia*. Diabet Med, 2005. **22**(5): p. 641-4.
261. SantaCruz, K.S., et al., *Regional NAD(P)H:quinone oxidoreductase activity in Alzheimer's disease*. Neurobiol Aging, 2004. **25**(1): p. 63-9.
262. Lu, Y., X. He, and S. Zhong, *Cross-species microarray analysis with the OSCAR system suggests an INSR->Pax6->NQO1 neuro-protective pathway in aging and Alzheimer's disease*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W105-14.
263. Traina, F., et al., *ANKHD1, ankyrin repeat and KH domain containing 1, is overexpressed in acute leukemias and is associated with SHP2 in K562 cells*. Biochim Biophys Acta, 2006. **1762**(9): p. 828-34.
264. Miles, M.C., et al., *Molecular and functional characterization of a novel splice variant of ANKHD1 that lacks the KH domain and its role in cell survival and apoptosis*. FEBS J, 2005. **272**(16): p. 4091-102.
265. Goswami, D.B., et al., *Transcriptional expression of serotonergic regulators in laser-captured microdissected dorsal raphe neurons of subjects with major depressive disorder: sex-specific differences*. J Neurochem, 2010. **112**(2): p. 397-409.
266. Barker, H.E., et al., *Deaf-1 regulates epithelial cell proliferation and side-branching in the mammary gland*. BMC Dev Biol, 2008. **8**: p. 94.
267. Manne, U., et al., *Altered subcellular localization of suppressin, a novel inhibitor of cell-cycle entry, is an independent prognostic factor in colorectal adenocarcinomas*. Clin Cancer Res, 2001. **7**(11): p. 3495-503.
268. Yip, L., et al., *Deaf1 isoforms control the expression of genes encoding peripheral tissue antigens in the pancreatic lymph nodes during type 1 diabetes*. Nat Immunol, 2009. **10**(9): p. 1026-33.
269. Hahm, K., et al., *Defective neural tube closure and anteroposterior patterning in mice lacking the LIM protein LMO4 or its interacting partner Deaf-1*. Mol Cell Biol, 2004. **24**(5): p. 2074-82.
270. Ordureau, A., et al., *DEAF1 Is a Pellino1-interacting Protein Required for Interferon Production by Sendai Virus and Double-stranded RNA*. J Biol Chem, 2013. **288**(34): p. 24569-80.
271. Manavalan, A., et al., *Brain site-specific proteome changes in aging-related dementia*. Exp Mol Med, 2013. **45**: p. e39.
272. Larner, A.J., *The cerebellum in Alzheimer's disease*. Dement Geriatr Cogn Disord, 1997. **8**(4): p. 203-9.
273. Munger, K.L., et al., *Anti-Epstein-Barr virus antibodies as serological markers of multiple sclerosis: a prospective study among United States military personnel*. Mult Scler, 2011. **17**(10): p. 1185-93.
274. Lossius, A., et al., *Epstein-Barr virus in systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis-association and causation*. Viruses, 2012. **4**(12): p. 3701-30.
275. Driessen, A.J., *SecB, a molecular chaperone with two faces*. Trends Microbiol, 2001. **9**(5): p. 193-6.
276. Houben, E.N., et al., *Early encounters of a nascent membrane protein: specificity and timing of contacts inside and outside the ribosome*. J Cell Biol, 2005. **170**(1): p. 27-35.
277. Luirink, J., et al., *Biogenesis of inner membrane proteins in Escherichia coli*. Annu Rev Microbiol, 2005. **59**: p. 329-55.
278. Thomas, J.D., et al., *Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli*. Mol Microbiol, 2001. **39**(1): p. 47-53.
279. Thie, H., et al., *SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli*. N Biotechnol, 2008. **25**(1): p. 49-54.
280. Bibi, E., *Is there a twist in the Escherichia coli signal recognition particle pathway?* Trends Biochem Sci, 2012. **37**(1): p. 1-6.
281. Aronson, D.E., L.M. Costantini, and E.L. Snapp, *Superfolder GFP is fluorescent in oxidizing environments when targeted via the Sec translocon*. Traffic, 2011. **12**(5): p. 543-8.

282. Reid, B.G. and G.C. Flynn, *Chromophore formation in green fluorescent protein*. *Biochemistry*, 1997. **36**(22): p. 6786-91.
283. Hartley, J.L., G.F. Temple, and M.A. Brasch, *DNA cloning using in vitro site-specific recombination*. *Genome Res*, 2000. **10**(11): p. 1788-95.
284. Rubelt, F., et al., *Onset of immune senescence defined by unbiased pyrosequencing of human immunoglobulin mRNA repertoires*. *PLoS One*, 2012. **7**(11): p. e49774.
285. Kumaran, J., C.R. Mackenzie, and M. Arbabi-Ghahroudi, *Semiautomated panning of naive camelidae libraries and selection of single-domain antibodies against peptide antigens*. *Methods Mol Biol*, 2012. **911**: p. 105-24.
286. Turunen, L., et al., *Automated panning and screening procedure on microplates for antibody generation from phage display libraries*. *J Biomol Screen*, 2009. **14**(3): p. 282-93.
287. Dennis, M., *Selection and screening strategies*, in *Phage display in biotechnology and drug discovery*, S.S. Sidhu, Editor. 2005, CRC Press. p. 143-164.
288. t Hoen, P.A., et al., *Phage display screening without repetitious selection rounds*. *Anal Biochem*, 2012. **421**(2): p. 622-31.
289. Vodnik, M., et al., *Phage display: selecting straws instead of a needle from a haystack*. *Molecules*, 2011. **16**(1): p. 790-817.
290. Sakaguchi, M., et al., *Three Cases of Linear IgA/IgG Bullous Dermatitis Showing IgA and IgG Reactivity With Multiple Antigens, Particularly Laminin-332*. *JAMA Dermatol*, 2013.
291. Pedersen, J.W., et al., *Seromic profiling of colorectal cancer patients with novel glycopeptide microarray*. *Int J Cancer*, 2011. **128**(8): p. 1860-71.
292. Blixt, O., et al., *A high-throughput O-glycopeptide discovery platform for seromic profiling*. *J Proteome Res*, 2010. **9**(10): p. 5250-61.
293. Yen, M. and J. Yin, *High-throughput profiling of posttranslational modification enzymes by phage display*. *Biotechniques*, 2007. **43**(1): p. 31, 33, 35 passim.
294. Durr, C., et al., *The Escherichia coli glycophage display system*. *Glycobiology*, 2010. **20**(11): p. 1366-72.
295. Celik, E., et al., *A filamentous phage display system for N-linked glycoproteins*. *Protein Sci*, 2010. **19**(10): p. 2006-13.

## 10. Supplementary

Gene IDs in all following tables correspond to gene symbols, approved by HGNC (Human Genome Nomenclature Committee).

### 10.1. Lists of autoantigens from protein macroarray screenings

**Table 15. Positive hits from Alzheimer's Disease screening.**

Priority	Antibody	Gene ID	Gene Description
Highly specific and unique	IgG	TRAF4	TNF receptor-associated factor 4
Highly specific and unique	IgA	CCDC50	coiled-coil domain containing 50
Highly specific and unique	IgA	GCN1L1	GCN1 general control of amino-acid synthesis 1-like 1
Highly specific and unique	IgA	LIG3	ligase III, DNA, ATP-dependent
Highly specific and unique	IgA	MAGED2	melanoma antigen family D, 2
Highly specific and unique	IgA	MYL6B	myosin, light chain 6B, alkali, smooth muscle and non-muscle
Highly specific and unique	IgA	PRDX1	peroxiredoxin 1
Highly specific and unique	IgA	PRDX4	peroxiredoxin 4
Highly specific and unique	IgA	RPL26	ribosomal protein L26
Highly specific and unique	IgA	TANK	TRAF family member-associated NFKB activator
Highly specific and unique	IgA	TPX2	microtubule-associated, homolog ( <i>Xenopus laevis</i> )
Highly specific and unique	IgA	ZNF33B	zinc finger protein 33B
Unique	IgG	AGXT2L2	alanine-glyoxylate aminotransferase 2-like
Unique	IgG	BBS4	Bardet-Biedl syndrome 4
Unique	IgG	CRTC2	CREB regulated transcription coactivator 2
Unique	IgG	DIRAS3	DIRAS family, GTP-binding RAS-like 3
Unique	IgG	ELAVL1	ELAV (embryonic lethal, abnormal vision, <i>Drosophila</i> )-like 1
Unique	IgG	PKM2	pyruvate kinase, muscle
Unique	IgG	TBC1D3	TBC1 domain family, member 3
Unique	IgG	THAP3	THAP domain containing, apoptosis associated protein 3
Unique	IgG	WBP11	WW domain binding protein 11
Unique	IgG	ZSCAN18	zinc finger and SCAN domain containing 18
Unique	IgA	ALDH5A1	aldehyde dehydrogenase 5 family, member A1
Unique	IgA	APC	adenomatous polyposis coli
Unique	IgA	BUB1	budding uninhibited by benzimidazoles 1 homolog (yeast)
Unique	IgA	C11orf31	Selenoprotein H (SelH)
Unique	IgA	CACNA1H	calcium channel, voltage-dependent, T type, alpha 1H subunit
Unique	IgA	CALM2	calmodulin 2 (phosphorylase kinase, delta)
Unique	IgA	CAPZB	capping protein (actin filament) muscle Z-line, beta

Unique	IgA	CBX4	chromobox homolog 4
Unique	IgA	CDC25B	cell division cycle 25 homolog B (S. pombe)
Unique	IgA	CEP57	centrosomal protein 57kDa
Unique	IgA	CHD3	chromodomain helicase DNA binding protein 3
Unique	IgA	CLIP3	CAP-GLY domain containing linker protein 3
Unique	IgA	COQ4	coenzyme Q4 homolog (S. cerevisiae)
Unique	IgA	DEK	DEK oncogene
Unique	IgA	DMTF1	cyclin D binding myb-like transcription factor 1
Unique	IgA	ECHS1	enoyl CoA hydratase, short chain, 1, mitochondrial
Unique	IgA	EIF4H	eukaryotic translation initiation factor 4H
Unique	IgA	EXOC1	exocyst complex component 1
Unique	IgA	F11R	F11 receptor
Unique	IgA	FAU	Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV) ubiquitous
Unique	IgA	FSCN1	fascin homolog 1, actin-bundling protein
Unique	IgA	FZR1	fizzy/cell division cycle 20 related 1 (Drosophila)
Unique	IgA	GPSM1	G-protein signaling modulator 1
Unique	IgA	GPX4	glutathione peroxidase 4 (phospholipid hydroperoxidase)
Unique	IgA	HSPA4	heat shock 70kDa protein 4
Unique	IgA	KIAA1614	KIAA1614
Unique	IgA	KIF19	kinesin family member 19
Unique	IgA	KIF3C	kinesin family member 3C
Unique	IgA	KIF4A	kinesin family member 4A
Unique	IgA	KPNB1	karyopherin (importin) beta 1
Unique	IgA	LITAF	lipopolysaccharide-induced TNF factor
Unique	IgA	MAP4K5	mitogen-activated protein kinase kinase kinase kinase 5
Unique	IgA	MAPT	microtubule-associated protein tau
Unique	IgA	MED8	mediator complex subunit 8
Unique	IgA	MLLT10	myeloid/lymphoid or mixed-lineage leukemia translocated to 10
Unique	IgA	NDC80	NDC80 homolog, kinetochore complex component (S. cerevisiae)
Unique	IgA	NDUFA5	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5, 13kDa
Unique	IgA	NOP14	NOP14 nucleolar protein homolog (yeast)
Unique	IgA	NUP133	nucleoporin 133kDa
Unique	IgA	PFDN5	prefoldin subunit 5
Unique	IgA	POLDIP3	polymerase (DNA-directed), delta interacting protein 3
Unique	IgA	POMT1	protein-O-mannosyltransferase 1
Unique	IgA	PSMC5	proteasome (prosome, macropain) 26S subunit, ATPase, 5
Unique	IgA	PSMD11	proteasome (prosome, macropain) 26S subunit, non-ATPase, 11
Unique	IgA	RABEPK	Rab9 effector protein with kelch motifs

Unique	IgA	RABGAP1L	RAB GTPase activating protein 1-like
Unique	IgA	RBBP6	retinoblastoma binding protein 6
Unique	IgA	RDBP	Negative elongation factor E (NELF-E)(RNA-binding protein RD)
Unique	IgA	RPL5	small nucleolar RNA, C/D box 21
Unique	IgA	RPS3	ribosomal protein S3
Unique	IgA	RSBN1L	round spermatid basic protein 1-like
Unique	IgA	SCFD1	sec1 family domain containing 1
Unique	IgA	SH3GL3	SH3-domain GRB2-like 3
Unique	IgA	SLC39A7	Zinc transporter SLC39A7 (Solute carrier family 39 member 7)
Unique	IgA	SPEF2	sperm flagellar 2
Unique	IgA	SUV420H1	suppressor of variegation 4-20 homolog 1 (Drosophila)
Unique	IgA	TMPO	thymopoietin
Unique	IgA	TPM4	tropomyosin 4
Unique	IgA	TRBC2	T cell receptor beta constant 2
Unique	IgA	UBE4A	ubiquitination factor E4A (UFD2 homolog, yeast)
Unique	IgA	VRK1	vaccinia related kinase 1
Unique	IgA	ZFP36	zinc finger protein 36, C3H type, homolog (mouse)
Unique	IgA	ZKSCAN2	zinc finger with KRAB and SCAN domains 2
Unique	IgA	ZNF346	zinc finger protein 346
Potentially specific	IgA	MAZ	MYC-associated zinc finger protein
Potentially specific	IgA	ENSA	endosulfine alpha
Potentially specific	IgA	RPS12	ribosomal protein S12
Potentially specific	IgA	ZNF695	zinc finger protein 695
Potentially specific	IgA	EBNA1BP2	EBNA1 binding protein 2
Potentially specific	IgA	ADD1	adducin 1 (alpha)
Potentially specific	IgA	FAM32A	family with sequence similarity 32, member A
Potentially specific	IgA	NUMA1	nuclear mitotic apparatus protein 1
Potentially specific	IgA	SH3GL1	SH3-domain GRB2-like 1

**Table 16. Positive clones from Multiple Sclerosis screening**

Priority	Autoantibody	Gene ID	Gene Description
Highly specific and unique	IgA	DEAF1	deformed epidermal autoregulatory factor 1 (Drosophila)
Highly specific and unique	IgA	H2AFY2	H2A histone family, member Y2
Highly specific and unique	IgA	IMPACT	Impact homolog (mouse)
Highly specific and unique	IgA	NOP56	small nucleolar RNA, C/D box 86
Highly specific and unique	IgA	RNF157	ring finger protein 157
Highly specific and unique	IgA	RRP1	ribosomal RNA processing 1 homolog (S. cerevisiae)
Highly specific and unique	IgA	HMG2	high-mobility group nucleosomal binding domain 2
Highly specific and unique	IgA	SFRS16	splicing factor, arginine/serine-rich 16

Highly specific and unique	IgA	ZFP106	zinc finger protein 106 homolog (mouse)
Highly specific and unique	IgA	GTF3C1	general transcription factor IIC, polypeptide 1, alpha 220kDa
Highly specific and unique	IgA	ANKHD1	ankyrin repeat and KH domain containing 1
Highly specific and unique	IgA	ATP5D	ATP synthase, H <sup>+</sup> transporting, mitochondrial F1 complex, delta subunit
Highly specific and unique	IgA	C11orf2	Protein fat-free homolog (Another new gene 2 protein)
Highly specific and unique	IgG	NOP56	small nucleolar RNA, C/D box 86
Highly specific and unique	IgG	NOP58	NOP58 ribonucleoprotein homolog (yeast)
Highly specific and unique	IgG	RPL24	ribosomal protein L24
Unique	IgG	AC100771	60S ribosomal protein L17 (60S ribosomal protein L23)(PD-1)
Unique	IgG	ADCY9	adenylate cyclase 9
Unique	IgG	ALDOA	aldolase A, fructose-bisphosphate
Unique	IgG	ARFGAP1	ADP-ribosylation factor GTPase activating protein 1
Unique	IgG	ATXN7L3	ataxin 7-like 3
Unique	IgG	CACNA1E	calcium channel, voltage-dependent, R type, alpha 1E subunit
Unique	IgG	CNOT1	CCR4-NOT transcription complex, subunit 1
Unique	IgG	CTSK	cathepsin K
Unique	IgG	DCX	doublecortin
Unique	IgG	DPPA4	developmental pluripotency associated 4
Unique	IgG	EIF2A	eukaryotic translation initiation factor 2A, 65kDa
Unique	IgG	FAM192A	family with sequence similarity 192, member A
Unique	IgG	GRIA1	glutamate receptor, ionotropic, AMPA 1
Unique	IgG	HES5	hairy and enhancer of split 5 (Drosophila)
Unique	IgG	HIST1H2AC	histone cluster 1, H2ac
Unique	IgG	IL11RA	interleukin 11 receptor, alpha
Unique	IgG	LAMA4	laminin, alpha 4
Unique	IgG	ME3	malic enzyme 3, NADP(+)-dependent, mitochondrial
Unique	IgG	MLL3	B melanoma antigen family, member 3
Unique	IgG	NMT1	N-myristoyltransferase 1
Unique	IgG	OTUD5	OTU domain containing 5
Unique	IgG	POLR3E	polymerase (RNA) III (DNA directed) polypeptide E (80kD)
Unique	IgG	PPP1R8	protein phosphatase 1, regulatory (inhibitor) subunit 8
Unique	IgG	PSME1	proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)
Unique	IgG	RPL13A	small nucleolar RNA, C/D box 32A
Unique	IgG	RPL18	ribosomal protein L18
Unique	IgG	RPL7A	small nucleolar RNA, C/D box 24
Unique	IgG	SFRS7	serine/arginine-rich splicing factor 7
Unique	IgG	SKIV2L2	superkiller viralicidic activity 2-like 2 (S. cerevisiae)
Unique	IgG	TALDO1	transaldolase 1
Unique	IgG	TRIO	triple functional domain (PTPRF interacting)
Unique	IgG	USP1	ubiquitin specific peptidase 1
Unique	IgG	ZNF238	zinc finger protein 238
Unique	IgG	ZNF354A	zinc finger protein 354A
Unique	IgG	ZNF681	zinc finger protein 681
Unique	IgG	ZNF721	ATP-binding cassette, sub-family A (ABC1), member 11

Unique	IgG	ZNHIT1	zinc finger, HIT type 1
Unique	IgA	AHSA1	AHA1, activator of heat shock 90kDa protein ATPase homolog 1
Unique	IgA	DCTN1	dynactin 1
Unique	IgA	DHX8	DEAH (Asp-Glu-Ala-His) box polypeptide 8
Unique	IgA	GABARAP	GABA(A) receptor-associated protein
Unique	IgA	GNL2	guanine nucleotide binding protein-like 2 (nucleolar)
Unique	IgA	GPATCH1	G patch domain containing 1
Unique	IgA	KIF22	kinesin family member 22
Unique	IgA	LAMB2	laminin, beta 2 (laminin S)
Unique	IgA	MAST2	microtubule associated serine/threonine kinase 2
Unique	IgA	MPP1	membrane protein, palmitoylated 1, 55kDa
Unique	IgA	NAP1L4	nucleosome assembly protein 1-like 4
Unique	IgA	NCSTN	nicastrin
Unique	IgA	NIPSNAP	nipsnap homolog 1 (C. elegans)
Unique	IgA	NKRF	NFKB repressing factor
Unique	IgA	PODXL2	podocalyxin-like 2
Unique	IgA	PTK2B	PTK2B protein tyrosine kinase 2 beta
Unique	IgA	SCYL1	SCY1-like 1 (S. cerevisiae)
Unique	IgA	SFRS4	serine/arginine-rich splicing factor 4
Unique	IgA	WDR3	WD repeat domain 3
Unique	IgA	WDR73	WD repeat domain 73
Unique	IgA	WHSC1	Wolf-Hirschhorn syndrome candidate 1
Unique	IgA	ZNF711	zinc finger protein 711
Potentially specific	IgG	AC004081	60S ribosomal protein L6
Potentially specific	IgG	EBNA1BP2	EBNA1 binding protein 2
Potentially specific	IgG	FTH1	ferritin, heavy polypeptide 1
Potentially specific	IgG	MAP1B	microtubule-associated protein 1B
Potentially specific	IgG	PHIP	pleckstrin homology domain interacting protein
Potentially specific	IgG	PTN	Pleiotrophin
Potentially specific	IgG	RPL14	ribosomal protein L14 [
Potentially specific	IgG	RPL29	ribosomal protein L29
Potentially specific	IgG	RSL1D1	ribosomal L1 domain containing 1
Potentially specific	IgG	SAMD14	sterile alpha motif domain containing 14
Potentially specific	IgG	STMN4	stathmin-like 4
Potentially specific	IgG	VRK3	vaccinia related kinase 3
Potentially specific	IgG	ZNF44	zinc finger protein 44
Potentially specific	IgA	CIRBP	cold inducible RNA binding protein
Potentially specific	IgA	MLL5	myeloid/lymphoid or mixed-lineage leukemia 5
Potentially specific	IgA	RPL8	ribosomal protein L8
Potentially specific	IgA	RPS12	ribosomal protein S12
Potentially specific	IgA	RPS8	small nucleolar RNA, C/D box 55
Potentially specific	IgA	RPS6	ribosomal protein S6
Potentially specific	IgA	RSL1D1	ribosomal L1 domain containing 1
Potentially specific	IgA	FTH1	ferritin, heavy polypeptide 1

Potentially specific	IgA	RPS25	ribosomal protein S25
----------------------	-----	-------	-----------------------

**Table 17. Positive clones from screening Healthy vs. Alzheimer's Disease**

Priority	Antibody	Gene ID	Gene Description
Highly specific and unique	IgA	RPS15	ribosomal protein S15
Highly specific and unique	IgA	RPS23	ribosomal protein S23
Unique	IgG	C14orf153	kinesin light chain 1
Unique	IgG	HARS	histidyl-tRNA synthetase
Unique	IgG	MBD3	methyl-CpG binding domain protein 3
Unique	IgA	FAM50A	family with sequence similarity 50, member A
Unique	IgA	HMGXB4	HMG box domain containing 4
Unique	IgA	PDCD6	programmed cell death 6
Unique	IgA	RPS2	ribosomal protein S2
Unique	IgA	PSMB5	proteasome (prosome, macropain) subunit, beta type, 5
Unique	IgA	PSMA1	proteasome (prosome, macropain) subunit, alpha type, 1
Potentially specific	IgG	AZGP1	alpha-2-glycoprotein 1, zinc-binding
Potentially specific	IgG	BMS1	BMS1 homolog, ribosome assembly protein (yeast)
Potentially specific	IgG	CD320	CD320 molecule
Potentially specific	IgG	CHMP1A	chromatin modifying protein 1A
Potentially specific	IgG	ICA1	islet cell autoantigen 1, 69kDa
Potentially specific	IgG	MORF4L1	mortality factor 4
Potentially specific	IgG	MRPS24	mitochondrial ribosomal protein S24
Potentially specific	IgG	RASD1	RAS, dexamethasone-induced 1
Potentially specific	IgG	SARNP	SAP domain containing ribonucleoprotein
Potentially specific	IgA	PRAP1	proline-rich acidic protein 1
Potentially specific	IgA	PIP5K1C	phosphatidylinositol-4-phosphate 5-kinase, type I, gamma
Potentially specific	IgA	RPS25	ribosomal protein S25
Potentially specific	IgA	AC004086.1	60S ribosomal protein L6 (Tax-responsive enhancer element-binding protein 107)
Potentially specific	IgA	AP3D1	adaptor-related protein complex 3, delta 1 subunit
Potentially specific	IgA	BCCIP	BRCA2 and CDKN1A interacting protein
Potentially specific	IgA	CAP1	CAP, adenylate cyclase-associated protein 1 (yeast)
Potentially specific	IgA	DHX9	DEAH (Asp-Glu-Ala-His) box polypeptide 9
Potentially specific	IgA	GDAP1L1	ganglioside-induced differentiation-associated protein 1-like 1
Potentially specific	IgA	ING4	inhibitor of growth family, member 4
Potentially specific	IgA	LRRC47	leucine rich repeat containing 47
Potentially specific	IgA	LUC7L	LUC7-like ( <i>S. cerevisiae</i> )
Potentially specific	IgA	MAP2	microtubule-associated protein 2
Potentially specific	IgA	RPL23A	small nucleolar RNA, C/D box 4A
Potentially specific	IgA	RPL29	ribosomal protein L29
Potentially specific	IgA	RPL37A	ribosomal protein L37a
Potentially specific	IgA	RPS8	small nucleolar RNA, C/D box 55
Potentially specific	IgA	SRGAP1	SLIT-ROBO Rho GTPase activating protein 1
Potentially specific	IgA	TTC3	tetratricopeptide repeat domain 3
Potentially specific	IgA	BAZ1A	bromodomain adjacent to zinc finger domain, 1A

Potentially specific	IgA	RPS14	ribosomal protein S14
----------------------	-----	-------	-----------------------

**Table 18. Positive clones from screening Healthy vs. Multiple Sclerosis**

Priority	Antibody	Gene ID	Gene Description
Highly specific and unique	IgA	FEN1	flap structure-specific endonuclease 1
Unique	IgA	ACSL3	acyl-CoA synthetase long-chain family member 3
Unique	IgA	LCORL	ligand dependent nuclear receptor corepressor-like
Unique	IgA	TARBP2	TAR (HIV-1) RNA binding protein 2
Unique	IgA	LUC7L	LUC7-like ( <i>S. cerevisiae</i> )
Potentially specific	IgG	ATP9A	ATPase, class II, type 9A
Potentially specific	IgG	AZGP1	alpha-2-glycoprotein 1, zinc-binding
Potentially specific	IgG	BMS1	BMS1 homolog, ribosome assembly protein (yeast)
Potentially specific	IgG	CCDC125	coiled-coil domain containing 125
Potentially specific	IgG	CD320	CD320 molecule
Potentially specific	IgG	CHMP1A	chromatin modifying protein 1A
Potentially specific	IgG	KDM3B	lysine (K)-specific demethylase 3B
Potentially specific	IgG	MORF4L1	mortality factor 4
Potentially specific	IgG	MRPS24	mitochondrial ribosomal protein S24
Potentially specific	IgG	RASD1	RAS, dexamethasone-induced 1
Potentially specific	IgG	RNASEN	ribonuclease type III, nuclear
Potentially specific	IgG	SARNP	SAP domain containing ribonucleoprotein
Potentially specific	IgG	UFC1	ubiquitin-fold modifier conjugating enzyme 1
Potentially specific	IgA	PRAP1	proline-rich acidic protein 1
Potentially specific	IgA	CAP1	CAP, adenylate cyclase-associated protein 1 (yeast)
Potentially specific	IgA	CCDC125	coiled-coil domain containing 125
Potentially specific	IgA	EHD1	EH-domain containing 1
Potentially specific	IgA	LRRC47	leucine rich repeat containing 47
Potentially specific	IgA	MYOZ3	myozenin 3
Potentially specific	IgA	RNF216	ring finger protein 216
Potentially specific	IgA	RPL23A	small nucleolar RNA, C/D box 4A
Potentially specific	IgA	SRGAP1	SLIT-ROBO Rho GTPase activating protein 1
Potentially specific	IgA	TTC3	tetratricopeptide repeat domain 3
Potentially specific	IgA	STMN4	stathmin-like 4
Potentially specific	IgA	PIN4	protein (peptidylprolyl cis/trans isomerase) NIMA-interacting, 4

## 10.2. Lists of autoantigens from phage display screenings

“Enrichment” represents the difference in the read numbers per gene between the initial library and the accordant library from the 4<sup>th</sup> selection round. “Sera” represents the number of sera, in which the accordant gene was enriched.

**Table 19. Enriched uniques from Alzheimer's Disease screening**

Enrichement	Sera	Antibody	Gene ID	Gene Description
8,302,857	1	IgG & IgA	SLC12A3	solute carrier family 12 (sodium/chloride transporters), 3
320,592	1	IgG & IgA	LDB3	LIM domain binding 3
194,463	2	IgA	SULT1C3	sulfotransferase family, cytosolic, 1C, member 3
192,966	1	IgG & IgA	FOXH1	forkhead box H1
174,828	1	IgG & IgA	WNT3	wingless-type MMTV integration site family, member 3
143,765	1	IgG & IgA	CYP11B2	cytochrome P450, family 11, subfamily B, polypeptide 2
134,114	1	IgG & IgA	TMC2	transmembrane channel-like 2
81,693	1	IgG & IgA	SHH	sonic hedgehog
75,409	1	IgG & IgA	OPRD1	opioid receptor, delta 1
64,185	1	IgG & IgA	KCNQ1	potassium voltage-gated channel, KQT-like subfamily, 1
63,231	1	IgG & IgA	HRK	harakiri, BCL2 interacting protein (contains only BH3)
59,898	2	IgA	NDRG4	NDRG family member 4
58,798	2	IgA	DTNBP1	dystrobrevin binding protein 1
49,394	1	IgA	CYTH3	cytohesin 3
48,514	1	IgG & IgA	PER2	period homolog 2 (Drosophila)
44,940	1	IgG & IgA	PGGT1B	protein geranylgeranyltransferase type I, beta subunit
42,449	1	IgG & IgA	APC	adenomatous polyposis coli
40,093	1	IgG & IgA	IL1RAPL1	interleukin 1 receptor accessory protein-like 1
39,680	1	IgG & IgA	GLP1R	glucagon-like peptide 1 receptor
34,182	1	IgG & IgA	EN1	engrailed homeobox 1 spleen focus forming virus (SFFV) proviral integration oncogene spi1
29,576	1	IgG & IgA	SPI1	
28,982	1	IgG & IgA	CDY2A	chromodomain protein, Y-linked, 2A
26,902	1	IgG & IgA	CHRNA9	cholinergic receptor, nicotinic, alpha 9
24,881	1	IgG & IgA	MAP3K10	mitogen-activated protein kinase kinase kinase 10
24,446	1	IgG & IgA	BRDT	bromodomain, testis-specific
23,732	1	IgG & IgA	SIM1	single-minded homolog 1 (Drosophila)
19,908	1	IgA	PAX6	paired box 6
18,034	1	IgG & IgA	PRKG2	protein kinase, cGMP-dependent, type II [
16,929	1	IgG & IgA	ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif, 2
12,804	1	IgG & IgA	ULK1	unc-51-like kinase 1 (C. elegans)
12,779	1	IgG & IgA	SOLH	small optic lobes homolog (Drosophila)
10,229	1	IgG & IgA	NOS1AP	nitric oxide synthase 1 (neuronal) adaptor protein
9,807	1	IgG & IgA	MALL	mal, T-cell differentiation protein-like
8,109	1	IgG & IgA	SOX21	SRY (sex determining region Y)-box 21
6,888	1	IgG & IgA	NCAN	neurocan
6,313	1	IgG & IgA	EPX	eosinophil peroxidase
5,943	1	IgG & IgA	PDE6C	phosphodiesterase 6C, cGMP-specific, cone, alpha prime
5,849	1	IgG & IgA	ZIC3	Zic family member 3 (odd-paired homolog, Drosophila)
5,751	1	IgG	AS3MT	arsenic (+3 oxidation state) methyltransferase
5,687	1	IgG & IgA	ASB4	ankyrin repeat and SOCS box-containing 4
5,676	1	IgG & IgA	RHO	rhodopsin
5,221	2	IgG & IgA	MAGOH	mago-nashi homolog, proliferation-associated (Drosophila)

5,164	1	IgG & IgA	HCRT	hypocretin (orexin) neuropeptide precursor
4,948	1	IgG & IgA	GAGE1	G antigen 3
4,628	1	IgG & IgA	HCRTR2	hypocretin (orexin) receptor 2
4,625	1	IgA	ANXA2	annexin A2
3,867	1	IgG & IgA	PPP1R3A	protein phosphatase 1, regulatory (inhibitor) subunit 3A
3,675	1	IgG & IgA	FOXC2	forkhead box C2 (MFH-1, mesenchyme forkhead 1)
3,357	1	IgG & IgA	DMD	dystrophin
3,019	1	IgG & IgA	TLL2	tolloid-like 2
2,381	2	IgA	GDI1	GDP dissociation inhibitor 1
2,365	1	IgG & IgA	EVC2	Ellis van Creveld syndrome 2
2,365	1	IgG & IgA	F5	coagulation factor V (proaccelerin, labile factor)
2,300	1	IgA	TCEB1	transcription elongation factor B (SIII), polypeptide 1
2,212	1	IgG & IgA	FOSB	FBJ murine osteosarcoma viral oncogene homolog B chronic lymphocytic leukemia up-regulated 1 opposite strand
2,019	1	IgG & IgA	CLLU1OS	
1,944	1	IgG	C14orf1	Probable ergosterol biosynthetic protein 28
1,929	1	IgG	DEFB104A	defensin, beta 104A
1,916	1	IgA	FOLR3	folate receptor 3 (gamma)
1,890	2	IgG & IgA	PRRC1	proline-rich coiled-coil 1
1,700	1	IgG & IgA	RTN4	reticulon 4
1,690	2	IgG	SMCP	sperm mitochondria-associated cysteine-rich protein
1,662	1	IgA	MND1	meiotic nuclear divisions 1 homolog
1,530	1	IgG	DEC1	deleted in esophageal cancer 1
1,475	1	IgG & IgA	PDX1	pancreatic and duodenal homeobox 1
1,403	1	IgG & IgA	FZD8	frizzled homolog 8 (Drosophila)
1,392	1	IgG & IgA	RNF39	Ring finger protein 39
1,263	1	IgG & IgA	ZDHHC3	zinc finger, DHHC-type containing 3
1,258	1	IgG	PMP22	peripheral myelin protein 22
1,227	1	IgG	LYRM2	LYR motif containing 2
1,185	1	IgG	PSME3	proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki)
1,177	1	IgA	RAP2A	RAP2A, member of RAS oncogene family
1,165	1	IgG	IGF1	insulin-like growth factor 1 (somatomedin C)
1,031	1	IgG	CPNE8	copine VIII

**Table 20. Enriched uniques from Multiple Sclerosis screening**

Enrichment	Sera	Antibody	Gene ID	Gene Description
272,328	2	IgA	TOB1	transducer of ERBB2, 1
34,619	1	IgA	THAP10	THAP domain containing 10
9,106	2	IgG & IgA	CCDC56	coiled-coil domain containing 56
6,568	2	IgG & IgA	GNG3	guanine nucleotide binding protein (G protein), gamma 3
4,899	2	IgA	COPE	coatomer protein complex, subunit epsilon
4,580	1	IgA	DHFR	dihydrofolate reductase pseudogene 1
3,354	2	IgG & IgA	COX7A1	cytochrome c oxidase subunit VIIa polypeptide 1
3,194	2	IgA	DUSP18	dual specificity phosphatase 18
2,038	1	IgG	SFXN2	sideroflexin 2

2,019	1	IgA	FXYD6	FXYD domain containing ion transport regulator 6
1,827	1	IgG	EPB49	erythrocyte membrane protein band 4.9 (dematin)
1,750	1	IgG	HIGD1A	HIG1 hypoxia inducible domain family, member 1A

**Table 21. Enriched uniques from Healthy screening**

Enrichement	Sera	Antibody	Gene ID	Gene Description
1,802,764	1	IgA	CASP4	caspase 4, apoptosis-related cysteine peptidase
3,809	1	IgG & IgA	TFG	TRK-fused gene
2,537	1	IgG & IgA	MIF	macrophage migration inhibitory factor
2,179	1	IgG	NDUFB7	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 7
1,689	1	IgG & IgA	C22orf39	UPF0545 protein C22orf39
1,418	1	IgG	AKR7A2	aldo-keto reductase family 7, member A2
1,268	1	IgG	NDUFA3	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3
1,222	1	IgG	MCEE	methylmalonyl CoA epimerase
1,194	1	IgG	UBE2A	ubiquitin-conjugating enzyme E2A (RAD6 homolog)

## 11. Appendix

### 11.1. List of figures

Figure 1: Correlation between the temporal course of AD development and changing biomarker values .....	20
Figure 2: Hypothetical intersection of auto-antigenicity profiles between AD, MS, healthy individuals and other neurodegenerative disorders.....	28
Figure 3: Overview of the protein macroarray production and usage .....	29
Figure 4: Workflow of the phage display bio-panning process.....	32
Figure 5: Overview of the phage display selection process .....	33
Figure 6: Schematic representation of bacteriophage M13 and different monovalent display types.....	35
Figure 7: Workflow of the Illumina Genome Analyzer .....	38
Figure 8: Overall experimental workflow .....	41
Figure 9: LR recombination .....	60
Figure 10: Coupling principle of tosylactivated magnetic beads .....	70
Figure 11: KingFisher Flex magnetic particle processor .....	72
Figure 12: MTP loading schema during selection on the KingFisher particle processor.....	73
Figure 13: MTP loading schema for ELISA on the KingFisher particle processor .....	75
Figure 14: Total IgG and IgA titer in donors' blood sera.....	78
Figure 15: Example of a hybridized high-density protein macroarray .....	79
Figure 16: Summery plots of macroarray screening results.....	79
Figure 17: IPA analyses of the results from macroarray screening.....	83
Figure 18: Cloning strategy for the pYG-full-ORF phage vectors.....	85
Figure 19: PCR for cloning the pYG-fusion-pelB vector.....	86
Figure 20: PCR for cloning the pYG-fusion-DsbA vector .....	86
Figure 21: PCR for cloning the pYG-LZ-pelB vector .....	87
Figure 22: Sequent PCR for cloning the pYG-LZ-DsbA/TorA vectors.....	87
Figure 23: Results of ccdB tests with pYG destination vectors .....	88
Figure 24: Full-ORF inserts size distributions .....	90
Figure 25: Gene coverages in initial pYG/OC libraries.....	91
Figure 26: Full-ORF insert size (bp) distributions in all pYG/OC initial libraries .....	92
Figure 27: Distribution of predicted native subcellular localizations of the full-ORF proteins	93
Figure 28: Western blot analysis of EGFP expression in <i>E. coli</i> periplasm.....	94
Figure 29: Fluorescence measurements of EGFP periplasm fractions.....	94
Figure 30: Western blot analysis of EGFP presenting pYG phages .....	95

---

Figure 31: ELISA analysis of EGFP presenting pYG phages .....	96
Figure 32: Fluorescence measurements of pYG/EGFP phage suspensions.....	97
Figure 33: Schematic representation of the bait-target construct during bio-panning.....	98
Figure 34: Western blot analysis of beads coating efficiency with primary antibody.....	99
Figure 35: Western blot analysis of coating beads with human serum autoantibodies and cross-reactivity test of the primary $\alpha$ -human antibodies .....	100
Figure 36: Polyclonal phage ELISA results of bio-panning rounds with pYG-LZ vectors and serum IgG autoantibodies .....	101
Figure 37: Polyclonal phage ELISA results of bio-panning rounds with pYG-LZ vectors and serum IgA autoantibodies .....	102
Figure 38: Polyclonal phage ELISA results of bio-panning rounds with pYG-fusion vectors and serum IgG/IgA autoantibodies.....	103
Figure 39: Preliminary tests for GAPDH enrichment validation. A. Colony PCR with Fos_seq x GAPDH_back primers .....	104
Figure 40: Validation of the GAPDH enrichment process .....	105
Figure 41: DNA sample preparation for Illumina sequencing.....	106
Figure 42: PRDX1 expression and affinity purification .....	111
Figure 43: Signal distribution and statistical analysis of antigen ELISA with PRDX1.....	112
Figure 44: TANK expression and affinity purification .....	114
Figure 45: TRAF4 expression and affinity purification.....	115
Figure 46: ANXA2 expression and affinity purification .....	116
Figure 47: Signal distribution and statistical analysis of antigen ELISA with ANXA2.....	117
Figure 48: DTNBP1 expression and affinity purification .....	119
Figure 49: Signal distribution and statistical analysis of antigen ELISA with DTNBP1 .....	120
Figure 50: GDI1 expression and affinity purification.....	122
Figure 51: NDRG4 expression and affinity purification .....	123
Figure 52: Signal distribution and statistical analysis of antigen ELISA with NDRG4 .....	124
Figure 53: PAX6 expression and affinity purification.....	126
Figure 54: ANKHD1 expression and affinity purification .....	127
Figure 55: Signal distribution and statistical analysis of antigen ELISA with ANKHD1 .....	128
Figure 56: DEAF1 expression and affinity purification .....	130
Figure 57: Signal distribution and statistical analysis of antigen ELISA with DEAF1 .....	131
Figure 58: Test ELISA with purified Ig-proteins and HRP-conjugated $\alpha$ -human antibodies. 133	
Figure 59: Test ELISA with serum Ig-proteins and HRP-conjugated $\alpha$ -human antibodies... 134	
Figure 60: Intersection between shared genes spotted on the protein macroarrays and genes represented in the human full-ORF library .....	148

## 11.2. List of tables

Table 1: Comparison of the most widely used AD diagnostic tools.....	21
Table 2: Annotations of human AD sera .....	49
Table 3: Annotations of human MS sera .....	49
Table 4: Annotations of human Healthy sera .....	50
Table 5: Human sera, screened on protein macroarrays .....	66
Table 6: Human sera, screened with phage display .....	71
Table 7: Program protocol for the 1 <sup>st</sup> selection round on the KingFisher magnetic particle processor.....	72
Table 8: Program protocol for ELISA on the KingFisher magnetic particle processor .....	74
Table 9: Positive clones analysis from macroarray screening .....	81
Table 10: Hit genes analysis from macroarray screening .....	81
Table 11: Amplification of pYG/OC libraries in DH10B cells .....	89
Table 12: pYG/OC libraries amplification in XL1 Blue/TG1 cells.....	90
Table 13: Analysis of enriched genes in phage display .....	107
Table 14: Potential biomarker candidates, chosen for re-validation .....	108
Table 15: Positive hits from Alzheimer's Disease screening .....	169
Table 16: Positive clones from Multiple Sclerosis screening .....	171
Table 17: Positive clones from screening Healthy vs. Alzheimer's Disease .....	174
Table 18: Positive clones from screening Healthy vs. Multiple Sclerosis.....	175
Table 19: Enriched uniques from Alzheimer's Disease screening .....	176
Table 20: Enriched uniques from Multiple Sclerosis screening .....	177
Table 21: Enriched uniques from Healthy screening .....	178