Aus der
Neurologischen Universitätsklinik Tübingen
Abteilung Neurologie mit Schwerpunkt Neurodegenerative Erkrankungen

Isoform specific Interactome Analysis of Spastin

Inauguraldissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard-Karls-Universität
zu Tübingen

vorgelegt von

Melanie Wayand

2022

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Privatdozentin Dr. R. Schüle-Freyer
2. Berichterstatter: Professor Dr. O. Rieß

Tag der Disputation: 18.01.2022

A scientist in his laboratory is not a mere technician: he is also a child confronting natural phenomena that impress him as though they were fairy tales. – Marie Curie

To all the amazing people, that took part in my discovery of this fairy tale.

List of abbreviations

Abbreviations	Description
ABC	Ammonium bicarbonate
ATP	Adenosine triphosphate
BDNF	brain derived neurotrophic factor
BP	Biological processes
BSA	Bovine serum albumin
CC	Cellular compartments
CHMP	charged multivesicular protein
CMV	cytomegalovirus
CNS	Central nervous system
DAPI	4',6-diamidino-2-phenylindole
ddH_2O	double distilled H_2O
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DsRed	Discosoma red fluorescent protein
DTT	dithiothreitol
E. coli	Escherichia coli
E4	Exon 4
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic
ER	endoplasmic reticulum
FBS	Fetal Bovine Serum
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	Green fluorescent protein
GOI	gene of interest
HA	hemagglutinin
HRP	Horse radish peroxidase
HSP	Hereditary Spastic Paraplegia
IAA	Iodacidamid
IP	Immunoprecipitation
IPS	Induced pluripotent stem cells
kDa	Kilodalton
LB	Lysogeny broth
MF	Molecular function
MIT	Microtubule interacting and trafficking domain
MOPS	3-(N-morpholino) propanesulfonic acid
MTBD	Microtubule binding domain
NaCl	Sodium chloride
NES	Nuclear export signal
NLS	Nuclear localization signal
NP40	Nonidet P40
NSE	Neuron specific enolase
NUP	Nucleoporin
PBS	phosphate buffered saline

PCR	Polymerase chain reaction
PI	proteinase inhibitor
RA	retinoic acid
RT	Room temperature
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
SOC	Super Optimal broth with Catabolite repression
SPG	spastic paraplegia gene
TEMED	Tetramethylethylenediamine
TFA	Trifluoracetic acid
UCSC	University of California Santa Cruz

List of figures

Figure 1: Schematic of the human voluntary motoric nervous system (Blackstone, 2012) 4

Figure 2: Schematic of known Spastic Paraplegia Genes (SPGs) and their mode of inheritance.....6

Figure 3: Schematic of a neuron, clustering of HSP genes in different pathways 8

Figure 4: Schematic of the functional structure of the M1 and M87 isoforms of the SPG4 protein spastin (Solowska and Baas, 2015).....10

Figure 5: Schematic of the four endogenously expressed spastin isoforms 11

Figure 6: Schematic of the cloning strategy used in this work 21

Figure 7: Scheme of the integration of a pcDNATM5/FRT/TO expression vector into a pcDNATM5/FRT/TO Host Line 30

Figure 8: Workflow of the Interactome analysis using a label free mass spectrometry approach 38

Figure 9: Schematic of the affinity purification using the Flag tag.....39

Figure 10: Scheme of the filtering procedure of the mass spectrometry results 42

Figure 11: Representative sequence of mutated nucleotides before and after mutagenesis.....46

Figure 12: Protein expression and localization of the spastin isoforms in Hek293T cells 47

Figure 13: Agarose gel electrophoresis of the four SPAST isoforms.....48

Figure 14: Validation of the cloning result exemplary shown for the M87 spastin isoform after cloning into the pcDNATM5/FRT/TO vector 49

Figure 15: Western blot analysis of the tetracycline-dependent expression of spastin in a Flp-InTM T-RexTM expression cell line 51

Figure 16: Western Blot analysis of the expression of the M1 spastin isoform after the addition of different doses of MG132 52

Figure 17: Immunofluorescence staining of SH-SY5Y cells before, during and after undergoing a differentiation protocol with retinoic acid and brain derived neurotrophic factor 53

Figure 18: Perseus workflow of the statistical analysis of the mass spectrometry interactome results 55

Figure 19: Volcano Blot of the LC/MS results of the M1 (A) and M87 (B) Isoform of spastin 61

Figure 20: Functional analysis of statistically significant interaction candidates of M1 spastin (DAVID platform, version 6.8) 62

Figure 21: Functional analysis of statistically significant interaction candidates of M87 spastin (DAVID platform, version 6.8) 63

Figure 22: Isoform specific protein-protein network of the M1 (A) and M87 (B) isoform of spastin 65

Figure 23: Western blot analysis of the performed co-immunoprecipitations 67

List of tables

Table 1: SPAST Isoforms....16

Table 2: Composition of the Agar (pH=7)....17

Table 3: Composition LB medium (pH=7)....17

Table 4: Primer SPAST sequencing....18

Table 5: Sequencing PCR reaction....18

Table 6: Sequencing PCR protocol....19

Table 7: Primers mutagenesis....20

Table 8: Composition agarose gel (2%)....22

Table 9: SPAST amplification reaction....22

Table 10: Primer SPAST amplification with restriction sites....23

Table 11: PCR conditions SPAST amplification....23

Table 12: A-tailing reaction....24

Table 13: Ligation reaction....24

Table 14: Restriction Enzymes used for cloning into pENTRTM/D-TOPO® 25

Table 15: Enzymatic plasmid digestion....25

Table 16: Ligation reaction pENTRTM/D-TOPO® 26

Table 17: Composition LR reaction....27

Table 18: Composition transfection mixture....29

Table 19: Composition Host Line Medium....29

Table 20: Transfection mixture for the transfection of a Flp-In™ T-Rex™ host cell line in a 10cm petri dish....31

Table 21: Medium Flp-In™ T-Rex™ expression cell line....32

Table 22: Protein preparation for Western Blotting....34

Table 23: Composition Bis-Tris gel....35

Table 24: Antibodies used for Western Blotting....36

Table 25: Composition of all solutions used for western blotting....36

Table 26: Composition of all buffers used for the affinity purification 40

Table 27: Extract of the final matrix showing all proteins tested significant ($p<0.05$) in the performed two sample test, as well in the Significance B test for the M1 isoform (including exon 4) 56

Table 28: Extract of the final matrix showing all proteins tested significant ($p<0.05$) in the performed two sample test, as well in the Significance B test for the M87 isoform (including exon 4) 57

Content

LIST OF ABBREVIATIONS ..... **IV**

LIST OF FIGURES ..... **VI**

LIST OF TABLES ..... **VIII**

1. INTRODUCTION **4**

1.1 HEREDITARY SPASTIC PARAPLEGIAS (HSPS) AND THE VOLUNTARY MOTORIC NERVOUS SYSTEM **4**

1.2 GENETICS OF HEREDITARY SPASTIC PARAPLEGIAS (HSPS) **5**

1.3 CLINIC OF HEREDITARY SPASTIC PARAPLEGIA (HSP) **6**

1.4 CELLULAR PATHWAYS AFFECTED IN HSP **7**

1.5 THE SPASTIC PARAPLEGIA GENE 4 (SPG4) - *SPAST* **9**

1.5.1 SPASTIN AND ITS FOUR ISOFORMS 9

1.5.2 PATHOPHYSIOLOGICAL HYPOTHESES FOR SPG4 12

1.6 THE USE OF SH-SY5Y CELL LINES AS A NEURONAL *IN VITRO* CELL MODEL **13**

1.7 THE USE OF PROTEOMICS IN THE ANALYSIS OF PROTEIN-PROTEIN INTERACTIONS **14**

1.8 OBJECTIVE OF THIS THESIS **15**

2. MATERIALS AND METHODS **16**

2.1 AMPLIFICATION AND SEQUENCING OF *SPAST* – ISOFORMS **16**

2.1.1 AMPLIFICATION OF *SPAST*-PLASMIDS BY HEAT SHOCK TRANSFORMATION OF CHEMICALLY POTENT ESCHERICHIA COLI BACTERIA 16

2.1.2 SANGER SEQUENCING OF THE *SPAST* ISOFORMS 18

2.2 CORRECTION OF BASE CHANGES BY SITE-DIRECTED MUTAGENESIS **19**

2.3 CLONING OF *SPAST* CDNA INTO A PCDNA™5/FRT/TO VECTOR **20**

2.3.1 DETECTION OF PCR PRODUCTS BY GEL ELECTROPHORESIS 21

2.3.2 CLONING OF THE *SPAST* ISOFORMS INTO A PGEM®-T EASY VECTOR 22

2.3.3 CLONING INTO PENTR™/D-TOPO® VECTOR 25

2.3.4 CLONING INTO PCDNA™5/FRT/TO VECTOR BY GATEWAY® CLONING 26

2.4 CULTIVATION OF HEK293T AND SH-SY5Y CELLS **27**

2.4.1 CULTIVATION OF HEK293T CELLS 28

2.4.2 Transfection of Hek293T cells with fluorescently labeled *SPAST* and HA tagged *SPAST* ... 28
2.4.3 Cultivation of a SH-SY5Y Flp-In™ T-Rex™ host cell line ... 29
2.4.4 Cotransfection of pcDNA™5/FRT/TO-*SPAST* and pOG44 into a Flp-In™ T-Rex™ host line ... 30
2.4.5 Selection of Colonies by Hygromycin ... 31
2.4.6 Induction of Flag-spastin expression by doxycycline ... 32
2.4.7 Differentiation of SH-SY5Y cells into a neuronal phenotype by retinoic acid and brain derived neuronal factor ... 33
2.5 Protein isolation and detection by Western Blotting ... 33
2.5.1 Cell lysis and protein isolation ... 33
2.5.2 Quantitative Analysis of cellular protein ... 34
2.5.3 Detection of spastin through a Western Blot ... 34
2.6 Increase of the expression of the M1 isoform by inhibiting protein degradation 37
2.7 Interactome analysis of spastin through a label free mass spectrometry approach ... 38
2.7.1 Extraction of Flag tagged spastin by Flag Affinity Purification ... 39
2.7.2 Detection of interaction partners by mass spectrometry ... 41
2.7.3 Qualitative and quantitative analysis of the mass spectrometric data ... 42
2.8 Validation of spastin interactions ... 44

3. RESULTS ... 45

3.1 Generation of a SH-SY5Y cell line with an inducible, stable spastin overexpression ... 45
3.1.1 Correction of variants of unknown significance in the *SPAST* sequence ... 45
3.1.2 Validation of the functionality of the SPAST plasmids by transient overexpression in Hek293T cells ... 46
3.1.3 Cloning of the *SPAST* isoforms into pcDNA™5/FRT/TO vectors ... 48
3.1.4 Induction and analysis of the spastin expression in SH-SY5Y overexpression cell lines ... 49
3.1.5 Investigation of the weak expression of M1 spastin ... 51
3.1.6 SH-SY5Y as a neuron-like cell model ... 52
3.2 Detection of spastin-protein interactions by mass spectrometry ... 53
3.2.1 Analysis of the mass spectrometricinteractome data ... 54
3.2.2 Functional annotation of significant interaction candidates ... 61

3.3 VALIDATION OF INTERACTION PARTNERS 65
3.3.1 SELECTION OF PROMISING SPASTIN INTERACTION CANDIDATES FOR FURTHER VALIDATION 66
3.3.2 CO-IMMUNOPRECIPITATION OF SPASTIN AND IDENTIFIED SPASTIN INTERACTION CANDIDATES 66

4. DISCUSSION 68

4.1 FLP-IN™ T-REX™ SH-SY5Y EXPRESSION CELL LINES AS NEURON-LIKE CELL MODEL FOR SPG4 68
4.2 WEAK EXPRESSION OF M1 SPASTIN IN THE OVEREXPRESSION CELL MODEL 69
4.3 CONFIRMATION OF KNOWN SPASTIN INTERACTION PARTNERS IN THE MASS SPECTROMETRY-BASED INTERACTOME ANALYSIS 70
4.4 DIFFERENCES IN THE IDENTIFIED PROTEIN INTERACTIONS BETWEEN THE TWO SPASTIN ISOFORMS 71
4.5 POSSIBLE INTERACTION OF M1 SPASTIN WITH THE NUP107-160 COMPLEX AND ITS POSSIBLE RELEVANCE 71
4.6 POSSIBLE INTERACTION OF M1 SPASTIN WITH MITOCHONDRIAL PROTEINS AND ITS POSSIBLE RELEVANCE 73
4.7 LIMITATIONS OF THE CURRENT APPROACH 74
4.8 CONCLUSION AND OUTLOOK 75

5. ABSTRACT 76

6. ZUSAMMENFASSUNG 78

7. REFERENCES 80

8. ERKLÄRUNG ZUM EIGENANTEIL 85

9. ACKNOWLEDGEMENTS 86

1. Introduction

1.1 Hereditary Spastic Paraplegias (HSPs) and the voluntary motoric nervous system

Hereditary Spastic Paraplegias (HSPs) are a heterogeneous group of monogenetically inherited, neurodegenerative disorders. A prevalence of 1.8-12/100,000 (Ruano et al., 2014, Braschinsky et al., 2009, Erichsen et al., 2009) defines them as rare diseases. Mutations in a wide range of over 80 spastic paraplegia genes (SPGs) (Novarino et al., 2014, Schule and Schols, 2011) lead to a length-dependent axonopathy of the upper motor neurons in the corticospinal tract. Clinically, HSP manifests with a progressive spasticity and weakness of the lower limbs as core symptom.

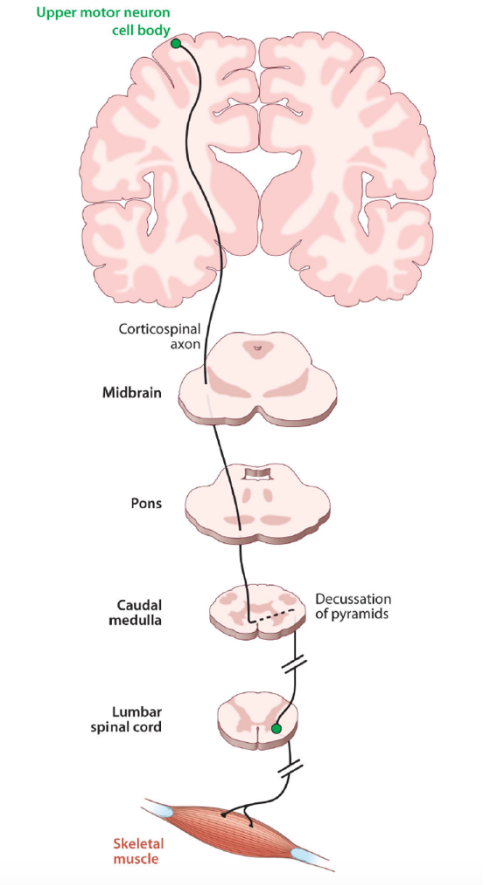


Figure 1: Schematic of the human voluntary motoric nervous system (Blackstone, 2012); it displays the division into an upper motor neuron, that originated in layer V of the motoric cortex and then follows the corticospinal tract, which mostly decussates in the caudal medulla and continues to the spinal level, where the transmission to a lower motor neuron takes place. This second motor neuron leads to its individual muscle.

The pyramidal motor nervous system is required to perform any kinds of voluntarily controlled movements. As shown in figure 1, it can be divided into two levels, first the corticospinal tract (upper motor neurons) which is directly or by intermediate neurons connected to the second level, the lower motor neurons, that transfer action potentials from the anterior horns of the spinal cord to the muscle.

The upper motor neurons originate mostly in layer V of the motor cortex and project their axonal processes along the pyramidal tract of the spinal cord. It either decussates in the caudal medulla and follows the lateral tract or follows the anterior tract, which decussates on the spinal level where it synapses either directly or by intermediate interneurons to lower motor neurons. These second motor neurons lead to their individual muscles and end with a neuromuscular synapsis.

With a length of about one meter, the upper motor neurons are among the longest nerves of the human body. This extreme polarization is necessary for their fast transmission but is also considered the main factor for their vulnerability that leads to the development of neurodegenerative disorders.

1.2 Genetics of Hereditary Spastic Paraplegias (HSPs)

HSPs can be inherited by various modes of inheritance, including autosomal dominant, autosomal recessive and X-linked modes. While HSP used to be mostly a diagnosis of exclusion for a long time, advances in human molecular genetics over the past decades have led to the discovery of numerous genes causing HSP. About 80 SPGs have been identified to date (Schule and Schols, 2011, Novarino et al., 2014). These distinct genes are numerically termed SPG 1- ~80. Despite these advances, about 40% of HSP patients currently still remain without a genetic diagnosis (Schule et al., 2016). Further genetic heterogeneity is therefore to be expected.

As the phenotype of HSP patients poorly correlates with the underlying mutation and the symptoms of the different genetic subtypes overlap, the only way to reach a final diagnosis is genetically.

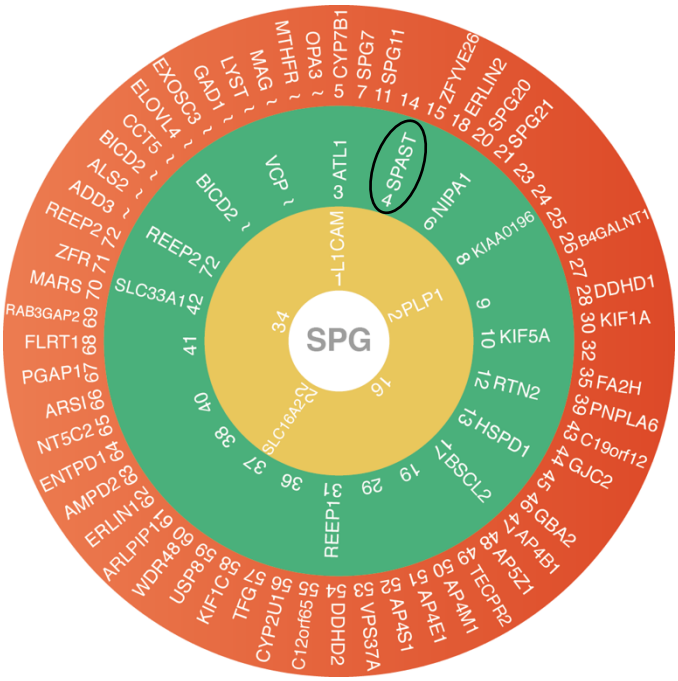


Figure 2: Schematic of known spastic paraplegia genes (SPGs) and their mode of inheritance

All spastic paraplegia genes are shown with their numeric label (SPG 1, 2…) according to the order of discovery. Additionally, the genes are sorted by their mode of inheritance: X-chromosomal (yellow, inner circle), autosomal dominant (green, middle circle) and autosomal recessive (orange, outer circle). The SPG4 gene SPAST, that is further investigated in this work, is marked by a black circle (adapted from © Schüle, R., Springer Verlag).

1.3 Clinic of Hereditary Spastic Paraplegia (HSP)

The core symptom of HSPs is a progressive spasticity and weakness of the lower limbs, resulting in a spastic gait disorder, caused by an axonopathy of the upper motor neurons.

Additionally, pure and complicated forms of HSP can be distinguished. Whereas pure forms are mainly characterized by spastic paraparesis, variably accompanied by urinary urgency or mild deficits of vibration sense, complicated cases are characterized by a wide variety of additional symptoms, such as cognitive deficits, peripheral neuropathy, cerebellar ataxia or seizures.

Symptoms typically progress continuously over decades and can clinically be monitored by the Spastic Paraplegia Rating Scale (SPRS) (Schule et al., 2006). Generally, most patients are considered to have a normal life span.

Partially due to the diversity in causative mutations, a wide spectrum of ages of onset and disease manifestations can be observed, varying from a mild, slowly progressing spasticity, barely affecting the patients' quality of life, to an early inability to walk or dependence on a wheelchair caused by complex neurologic disability.

The ages of onset are spread anywhere between early childhood and late adulthood and correlate in part with the underlying mutation (Schule et al., 2016, Klebe et al., 2015).

The immense clinical and genetic heterogeneity of HSPs hamper clinical counseling and reliable prediction of disease course and prognosis.

There are no causal treatment strategies available for any of the HSP subtypes today. Available treatment strategies are currently restricted to management of symptoms, like the use of antispastic drugs to relieve spasticity and pain, or medical treatment of urinary urgency and the implementation of continuous physical therapy.

1.4 Cellular pathways affected in HSP

Due to the high number of genes and associated proteins implicated in HSPs, a variety of affected cellular pathways have been discovered to be involved in HSP pathogenesis.

To structure these proteins, encoded by SPGs, Blackstone (2012) introduced four functional groups: proteins involved in membrane trafficking, mitochondrial proteins, proteins that play a role in myelination and miscellaneous.

The biggest group consists of proteins, that participate in membrane trafficking, for example by influencing endoplasmic reticulum (ER) morphology or endosomal trafficking, which therefore seems to play an important role in the pathogenesis of HSP. Miscellaneous, less frequently affected pathways include the synthesis of lipids and their metabolisms, axon pathfinding in CNS development and cell adhesion (Blackstone, 2012).

Many HSP genes and implicated proteins take part in more than one of the listed pathways.

The distributions and key functions of some of the more common SPGs in a neuronal cell are displayed in figure 3. It highlights the breadth of involved cellular functions as well as the enriched involvement of SPG proteins in some pathways. For instance, the proteins spastin (SPG4), atlastin (SPG3), REEP1 (SPG31), erlin2 (SPG18) and reticulon2 (SPG12) all share a common function in the shaping of the ER morphology.

Furthermore, proteins like for example spastin (SPG4), spartin (SPG20), spatacsin (SPG11) and strumpellin (SPG8) are involved in endosomal trafficking and thereby membrane trafficking.

As the phenotype of different HSP forms can be very similar, it is hypothesized, that many, if not all spastic paraplegia genes may be connected by one common pathway, which ultimately leads to the axonal degeneration of upper motor neurons and the manifestation of the disease.

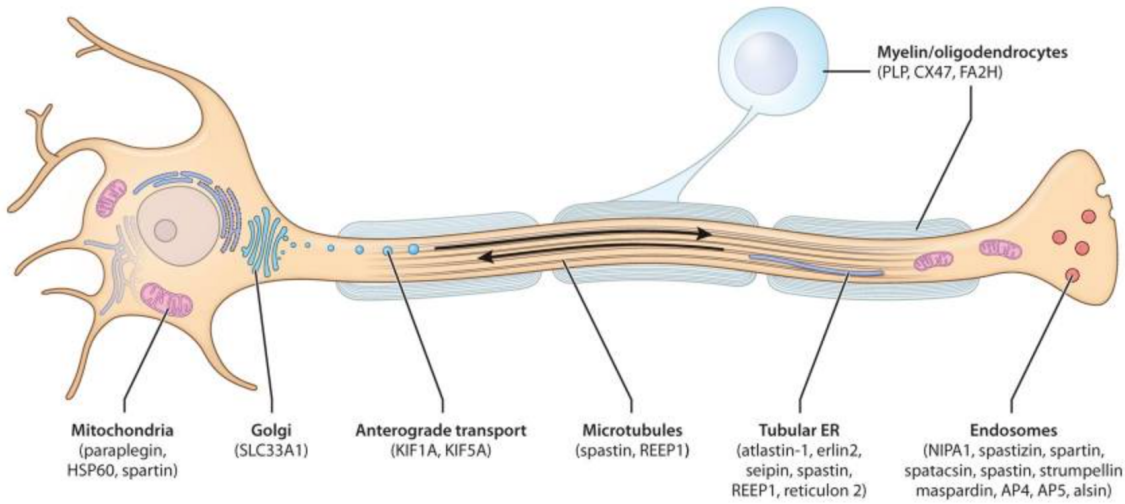


Figure 3: Schematic of a neuron, clustering of HSP genes in different pathways
A selection of SPG proteins is shown with their cellular localization and some of their cellular tasks. Several HSP genes cluster in cellular compartments (i.e. atlastin-1, erlin2, seipin, spastin, REEP1 and reticulon2 together play a role in the shaping of the ER; spastin and REEP1 are involved in the severing of microtubules)

1.5 The Spastic Paraplegia Gene 4 (SPG4) - *SPAST*

The most common form of autosomal dominant inherited HSP is caused by mutations in the SPG4 gene, *SPAST*, coding for the protein spastin.

Disease causing mutations in der *SPAST* gene were first described in 1999 by Hazan et al., which led to the genetic diagnosis of a great number of HSP cases. In almost 50% of autosomal dominant inherited index cases, a mutation in der *SPAST* gene could be detected (Schule and Schols, 2011, Bürger et al., 2000). Mutations include missense, nonsense or splice site mutations, small insertions or deletions as well as genomic deletions often comprising several exons (Beetz et al., 2006, Depienne et al., 2007).

The phenotype of SPG4 comprises the characteristic spastic gait disorder; additionally 50% of the patients suffer from an impaired sense of vibration and about one third of cases report sphincter disturbances (Solowska and Baas, 2015).

The main age of onset is in young adulthood, although there is a huge age of onset variability between the first and the last years of life. Even an intrafamilial variation in symptom expression can be observed.

1.5.1 Spastin and its four isoforms

Since SPG4 and its encoded protein spastin were discovered as a cause for HSP in 1999, its role in the disease mechanism was intensively investigated.

Spastin is, as its homolog katanin, an AAA ATPase, with the main function to sever microtubules. Interestingly, spastin was shown to play a role in a lot of pathways disrupted in HSP, including intracellular transport, membrane shaping or cytoskeleton dynamics.

Two neighboring domains are required to fulfill this function, an enzymatically active AAA domain and a microtubule binding domain (MTBD). For the severing of microtubules, in the presence of ATP, the protein assembles into hexamers with a central pore, through which a microtubule is threaded and subsequently severed (White et al., 2007, Roll-Mecak and Vale, 2008).

The N-terminus of spastin has been reported to mediate interaction with other proteins and target the protein to specific subcellular compartments and is not

required for microtubule severing. It contains a microtubule-interacting and trafficking domain (MIT), a domain which is found in several endosomal proteins and typically mediates interaction with proteins of the CHMP (charged multivesicular body protein) family. Hereby, it plays a role in endosomal trafficking (Reid et al., 2005). The N-terminus furthermore carries two nuclear localization signals as well as two nuclear export signals (NESs) that were shown to be active *in vitro* (Beetz et al., 2004).

Additionally, an N-terminal interaction with other spastic paraplegia proteins, for example atlastin-1, REEP 1 and reticulons (Solowska and Baas, 2015) was shown.

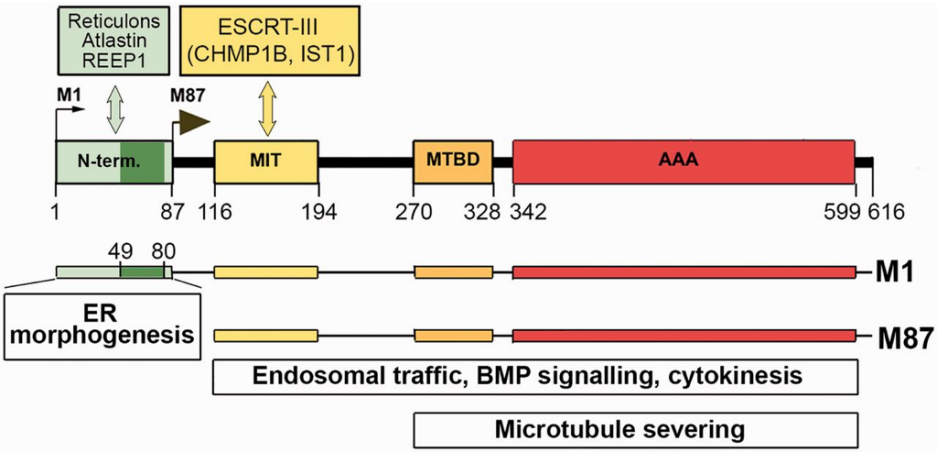


Figure 4: Schematic of the functional structure of the M1 and M87 isoforms of the SPG4 protein spastin (Solowska and Baas, 2015)

The most important functional domains of spastin are displayed. A C-terminal AAA domain and a linked microtubule binding domain (MTBD) are required for microtubule severing. Additionally, the interaction of spastin with proteins of the ESCRT-III complex by a microtubule-interacting and trafficking domain is shown, resulting in an influence on endosomal trafficking, BMP signaling and cytokinesis. The N-terminus of M1 spastin is shown to interact with REEP1, atlastin and reticulons and plays a part in the ER morphogenesis.

In humans, four isoforms of spastin are endogenously expressed. Two AUG start codons in the spastin mRNA lead to the translation of a long M1 isoform (68kDa) and an 86 amino acid shorter M87 isoform (60kDa).
Additionally, alternative splicing results in the facultative skipping of exon 4 (E4).

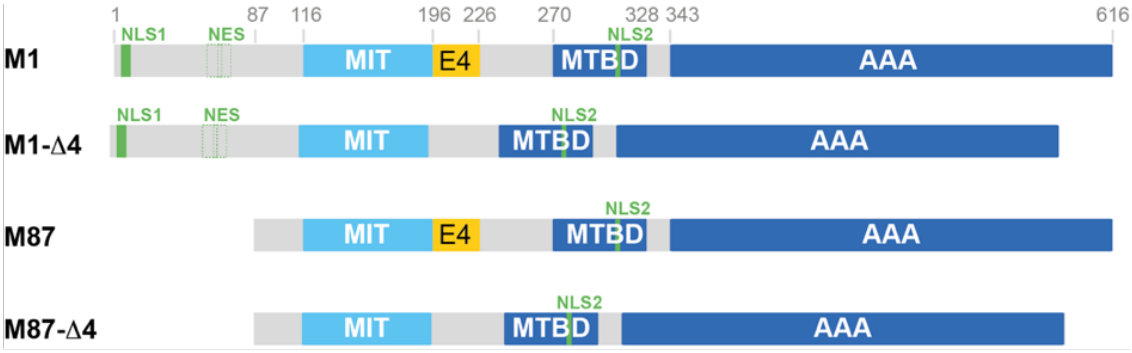


Figure 5: Schematic of the four endogenously expressed spastin isoforms
Two AUG start codons in the mRNA of spastin lead to the transcription of a long (M1) and a short (M87) isoform; due to alternative splicing there are two more isoforms lacking exon 4 (E4). Furthermore, the rough functional structure of the spastin protein is shown (according to figure 4). Nuclear localization sequences (NLS1,2) and the Nuclear export sequence (NES) are mostly located at the N-terminus, which leads to the shorter isoforms lacking the NLS1 and NES resulting in the nucleoplasmic localization of the M87 isoforms.

These four isoforms of human spastin seem to play different roles in cellular pathways. It was shown that the shorter M87 isoform is widely expressed in the human central nervous system including cortical pyramidal neurons, hippocampus and cerebellum, whereas the M1 isoform shows a relevant expression exclusively in the spinal cord (Wharton et al., 2003).
Spastin isoforms also differ regarding their subcellular localization. While M87 spastin is distributed throughout the cytoplasm and additionally – mediated by its single nuclear localization sequence (NLS2) – localizes in the nucleus, the M1 isoform co-localizes with specific areas of the ER complex, the three-way junctions. A hydrophobic region in the first 86 amino acids of M1 spastin forms a hairpin, allowing the integration of the M1 isoform into the membrane of the

endoplasmic reticulum. Here it is known to interact with the proteins atlastin (SPG3) and REEP1 (SPG31), both also known HSP-associated proteins and plays a role in the shaping of the ER (Reid et al., 2005, Sanderson et al., 2006). The nuclear export signal (NES), which is located in the first amino acids and therefore only contained in the longer M1 isoform leads to an effective nuclear export of this isoform. Consequently, M1 is not detectable in the nucleoplasm (Claudiani et al., 2005).

Remarkably, the M87 isoform is endogenously much higher expressed than the longer M1 isoform of spastin. As a reason, a weak Kozak sequence preceding the first AUG start codon and an upstream open reading frame (ORF) partially overlapping the first AUG start codon are discussed (Kozak, 2002, Claudiani et al., 2005).

1.5.2 **Pathophysiological hypotheses for SPG4**

The discovery of mutations in the *SPAST* gene as a common cause for autosomal dominant HSP led to various speculations regarding the pathomechanism of the disease (Solowska et al., 2010).

With a mutation spectrum mostly consisting of truncating mutations, such as nonsense or splice site mutations, the most postulated theory in the beginning was a loss of function of the protein spastin as explanation for the disease (Evans et al., 2005). In this scenario, due to the inactivation of one allele of the *SPAST* gene, the axons of the upper motor neurons degenerate as a consequence of insufficient microtubule severing resulting in deficient intracellular transport.

This theory is supported by a variety of experiments in different cell and animal models.

For instance, the knock-out of spastin in zebrafish led, by a lack of microtubule severing, to massive disturbances in the outgrowth of axons, which is suggested to play a role in childhood onset HSP cases (Wood et al., 2006).

Spastin-null drosophila showed low eclosion rates of about 20% and severe motor deficits as they couldn’t jump or fly (Sherwood et al., 2004). Additionally, axons were not able to regenerate in drosophila lacking one allele of spastin (Stone et al., 2012).

Neuronally differentiated pluripotent stem cells, derived from fibroblasts of SPG4 patients, showed decreased spastin levels, axonal swellings with loosely arranged microtubules and a lower number as well as shorter, less branched neurites (Havlicek et al., 2014).

The loss of function theory provides a plausible explanation for the disease mechanism for a major part of the *SPAST* mutations, located in the AAA domain of the protein, leading to the impairment of the microtubule severing activity. However, there is a number of mutations reported outside this area, not affecting the function of the AAA domain. Therefore, not all of SPG4 cases can be explained by the loss of function theory.

Furthermore, Solowska et al. (2010) investigated the microtubule severing function in cells with some selected SPG4-associated mutations of the *SPAST* gene located outside the AAA ATPase domain and didn't detect any impairment in microtubule severing.

Recent studies used a different approach and examined the potential cellular toxicity of truncated M1 isoforms of spastin. It was shown, that some mutations in the *SPAST* gene lead to the expression of a truncated, non-functional M1 isoform of spastin, which accumulates in the cell cytoplasm suggesting a toxic gain-of-function leading to the degeneration of the cell (Solowska et al., 2017).

In conclusion, many experiments in different animal and cell models support the explanation of a loss of function as pathomechanism in SPG4, yet not all disease-causing mutations can be explained by this theory. Therefore, at least one other disease mechanism of SPG4 is to be assumed.

1.6 The use of SH-SY5Y cell lines as a neuronal *in vitro* cell model

The selection of neuronal cell models is highly limited by the fact that terminally differentiated neurons are post mitotic and cannot be propagated further.

Patient-derived induced pluripotent stem cells (iPSCs) that can be differentiated into cortical neurons (Hauser et al., 2016) allow to study disease pathophysiology in a human neuronal model system. As this is an expensive and time-consuming process, however, we decided to use the SH-SY5Y neuroblastoma cell line in

this work, with the possibility of confirming promising results subsequently in an IPS derived neuronal cell system.

SH-SY5Y is a subclone of the SK-N-SH cell line that was originally derived from a biopsy taken from a bone marrow metastasis of a neuroblastoma (Kovalevich and Langford, 2013).

As SH-SY5Y cells express several human neuronal markers, they are often used as an easy to handle, low cost *in vitro* model for neurological diseases. At the same time, as a commercially obtainable cell line, they allow the work with human mammalian cells without ethical concerns.

Additionally, by small changes in the cell medium, SH-SY5Y cells can be differentiated into a more mature neuronal phenotype with axon-like structures and express several neuronal markers including ßIII – tubulin, microtubule-associated protein (MAP2) and neuron specific enolase (NSE) (Kovalevich and Langford, 2013).

1.7 The use of proteomics in the analysis of protein-protein interactions

A mass spectrometry-based proteomics approach was used in this work to identify new protein-protein interactions affecting the protein encoded by SPG4, spastin. Many proteins perform their functions by forming macromolecular complexes and thereby fulfill their role in the cell metabolism, for example through the activation or depression of pathways or the transmission of information. Therefore, the investigation of protein-protein interactions is a crucial step in the understanding of the molecular function of a protein.

There are various ways to investigate protein-protein interactions, such as the yeast two hybrid system, mass spectrometry or different versions of immunoprecipitations (Braun and Gingras, 2012).

Unbiased proteomic approaches, including yeast two hybrid or mass spectrometry, have been successfully used in the past to uncover previously unknown protein functions. The discovery of atlastin, the protein affected in the SPG3 subtype of hereditary spastic paraplegias, as a binding partner of spastin (Sanderson et al., 2006) or the interaction of spastin with Reticulon 1 and

Reticulon 3 (Mannan et al., 2006) are only two example for protein interactions identified by a yeast-two-hybrid screen.

Mass spectrometry is unbiased, highly sensitive and can be performed in mammalian, or even neuronal, cells. This is of great importance as most protein interactions are cell type specific.

In this work we performed a mass spectrometry-based interactome analysis that enabled the construction of a protein-protein interaction network of spastin, a so called interactome, and identified novel interaction partners of spastin.

1.8 Objective of this thesis

Although there are several hypotheses on how mutations in der *SPAST* gene may lead to upper motor neuron degeneration and ultimately Hereditary Spastic Paraplegia, the exact pathomechanism of the disease still remains unsolved.

In particular, the differential functions of the M1- and the M87 isoforms of spastin and their relative contribution to the pathophysiology remain unclear. Another open question is what role exon 4, only present in two of the endogenously expressed isoforms of spastin, plays in the proteins function.

The aim of this work was to identify interaction partners specific to the four main spastin isoforms. Therefore, SH-SY5Y neuroblastoma cell lines with a stable, inducible overexpression of the four isoforms of spastin were established. These cells were used to perform a proteomics based interactome analysis to identify new interaction partners of spastin and thereby reveal novel functions of the protein.

The results give an inside into the role the protein plays in different pathways and allow speculations about the underlying disease mechanism in SPG4.

2. Materials and Methods

2.1 Amplification and Sequencing of *SPAST* – Isoforms

The four *SPAST* isoforms (Table 1) were used in this work to generate inducible, stable, spastin overexpressing SH-SY5Y cell lines.

In a first step, plasmids containing the *SPAST* gene were amplified and the accuracy of the *SPAST* cDNA was confirmed by sanger sequencing. The plasmids were then further used for cloning. The following methods (amplification of plasmid DNA and sanger sequencing) were used in several steps of the following cloning procedure.

As a generous gift of Christian Beetz (Jena), we received the following four isoforms of the *SPAST* gene in transport vectors containing a resistance for kanamycin:

Table 1: SPAST Isoforms

SPAST 1-616	M1 Isoform	pDsRed-backbone
SPAST 87-616	M87 Isoform	pDsRed-backbone
SPAST 1-616 del Ex4	M1 Isoform lacking Exon 4	pEGFP-backbone
SPAST 87-616 del Ex4	M87 Isoform lacking Exon 4	pEGFP-backbone

The plasmids were solved in 150µl H_2O (ddH_2O, Merck) and for amplification transformed into chemically competent Escherichia coli bacteria (2.1.1).

2.1.1 Amplification of *SPAST*-plasmids by heat shock transformation of chemically potent Escherichia coli bacteria

The insertion of plasmid DNA in chemically competent Escherichia coli bacteria (E. coli) for amplification by heat shock transformation is a standard procedure in molecular biology and was used in various steps in this work.

To this end an aliquot of 10µl E. coli bacteria (Escherichia coli, New England Biolabs, Ipswich, MA/USA) was thawed on ice and 2µl of the plasmid solution were added. After an incubation on ice for 30 minutes a heat shock transformation was performed at 42°C for 30 seconds (Thermomixer comfort, Eppendorf AG,

Hamburg). Next, 75µl of SOC-Medium (New England Biolabs, Ipswich, MA/USA) were added and the mixture was shaken (300rpm) at 37°C for one hour.
Afterwards the solution was distributed on pre-warmed, Kanamycin (25µg/ml) containing agar plates (table 2) and incubated at 37°C over night.

Table 2: Composition of the Agar (pH=7)

Contents			Amount
Agar powder	Tryptone	10g/L	20g
	Yeast Extract	5g/L	
	NaCl	10g/L	
	Agar	15g/L	
H_2O			ad 500ml

The next day colonies were picked and put in 3ml of LB medium (table 3) containing Kanamycin (25µg/ml) to grow over night at 37°C while they were shaken (180rpm, Ecotron, Infors AG, Bottmingen/Basel, CH).

Table 3: Composition LB medium (pH=7)

Contents			Amount
LB powder	Tryptone	10g/L	20g
	Yeast Extract	5g/L	
	NaCl	5g/L	
H_2O			ad 1l

The next day, the bacteria suspension was centrifuged (4000rpm) for 10 minutes and the DNA isolated using the QIAprep Spin Miniprep Kit (Quiagen, Hilden) following the manufacturer's instructions. The DNA concentration in the eluted sample was measured at 260nm using a spectrophotometer (NanoDrop 1000, Thermo Scientific, Waltham, MA/USA).
When higher concentrations of plasmid DNA were needed (i.e. for transient transfections of mammalian cells), plasmids were inoculated in 100ml LB medium

(table 4) and the DNA was isolated using the EndoFree Plasmid Maxi Kit (Quiagen, Hilden) following the manufacturer's instructions.

2.1.2 **Sanger sequencing of the *SPAST* Isoforms**

Sanger sequencing was used to analyze DNA sequences on nucleotide level. This is a standard method, based on the incorporation of labeled di-deoxynucleotide triphosphates leading to a chain termination in a polymerase chain reaction.

For the complete sequencing of the *SPAST* isoforms, the following primers (table 4), labelled after their nucleotide position in the M1 isoform, were used in a sequencing PCR (Table 5). The standard PCR protocol used for sanger sequencing is shown below (Table 6).

Table 4: Primer SPAST sequencing

Primer	Sequence (5'-3')
782 reverse	TAGGTGCTCTATGGTGGCCT
1144 reverse	CTTCCCATTCCCAGGTGGAC
1360 forward	AGCACGATGCTAGTAGACGC
1523 reverse	GTAGTCTTGTCTCCTCATTTGGT

Table 5: Sequencing PCR reaction

Components	Volume (µl)
5x Buffer (Thermo Scientific, Waltham, MA/USA)	1.65
Primer 10µM	1
BigDye	0.7
ddH_2O	4.65
DNA sample	2
Total	10

ddH_2O: double distilled water (Merck)

Table 6: Sequencing PCR protocol

Step	Temperature	Time
1	94°C	1 minute
2	94°C	10 seconds
3	50°C	5 seconds
4	60°C	4 minutes
5	Go to step 2 / 29x	
6	4°C	forever

The PCR product was purified by adding 30µl NaAc/EtOH (3M NaAc pH 5.2 1:25 in 100% EtOH) and centrifuged at 3220g for 45 minutes at RT. After discarding the supernatant, the samples were washed with 100µl 70% EtOH and centrifuged for another 15 minutes at 3220g. After discarding the supernatant, the tubes were centrifuged for 1min at 600g upside down on a paper towel and the dried PCR product was diluted in 15µl ddH_2O and vortexed for 30min at room temperature. The diluted PCR product (7µl) was mixed with 15µl formamide for further sequencing.

The analysis of the base sequence was carried out on a Genetic Analyzer 3130xl (Applied Biosystems, Foster City, USA), the results were processed using the SeqA6 Sample Manager (Applied Biosystems) and Pregap4 and viewed in Gap4 (both Staden package, Open source; http://staden.sourceforge.net). The DNA sequences were compared to the *SPAST* cDNA sequence included in the UCSC database (http://genome.ucsc.edu) and identified variations checked for known single nucleotide polymorphisms (SNP) in the Ensemble database (http://www.ensembl.org/index.html).

2.2 Correction of base changes by site-directed Mutagenesis

After the analysis of base changes in the investigated *SPAST* genes in comparison to the *SPAST* cDNA (UCSC database), two variations didn't match a known SNP. As a functional impact of these variants couldn't be ruled out, they were corrected by site-directed mutagenesis. This method is based on a polymerase chain reaction (PCR) using a wide range Q5 polymerase and primers

containing the corrected nucleotide sequence. It thereby allows the exchange of single bases as well as small insertions or deletions in plasmid DNA.
The two variants in the *SPAST* isoforms (position c.158 and c.1850, positions counted in the M1 isoform (NM_014946)) were corrected by a mutagenesis reaction using a Q5 site directed Mutagenesis Kit (New England Biolabs, Ipswich, MA/USA). Mutagenesis was performed according to the manufacturer's protocol. Adequate primers were designed (Table 7) by using the NEBase Changer (New England Biolabs) and contained the corrected nucleotide sequence.

Table 7: Primers mutagenesis

Primer	Sequence (5'-3')
c158 C > T fw	CTGTACTATT ***t**** CTCCTACCCGCTG
c158 C > T rv	GTTCCGCTTATGCGGCG
c1850 T > A fw	ACCACTGTTT ***a**** AGGATCCACC
c1850 T > A rv	ATCTCCAAAGTCCTTGTTC

**the corrected nucleotides are marked by small, bold letters*

The result was amplified by heat shock transformation using chemically potent E. coli as described in 2.1.1 and was checked by sanger sequencing of the changed region and subsequent sequencing of the entire gene.
All sequenced plasmids, as well as the plasmids of the following intermediate cloning steps were stored long term in glycerol stocks (800µl E. coli - plasmid solution, 200µl glycerol) at -80°C.

2.3 Cloning of *SPAST* cDNA into a pcDNA™5/FRT/TO vector

The aim of this step was to clone the isoforms of the *SPAST* gene, into a pcDNA™5/FRT/TO vector (Karsten Boldt, medical proteome center, Tübingen) with an integrated Strep/Flag Tag (Gloeckner et al., 2009b) and a FRT site (Flp-In™ T-Rex™ system). This vector is part of the Flp-In™ T-Rex™ system and can be stably integrated into a Flp-In™ T-Rex™ host line. This allows the generation of an inducible overexpression cell line.

Since the direct cloning of the gene into a pENTR™/D-TOPO® vector by following the instructions of the pENTR™ Directional TOPO Cloning Kit (Invitrogen, Carlsbad, CA/USA) wasn't successful after several attempts, the *SPAST* isoforms were first cloned it into the pGEM®-T Easy vector (Promega GmBH, Mannheim). The complete cloning strategy is shown in figure 6. The separate steps are further described in the following sections.

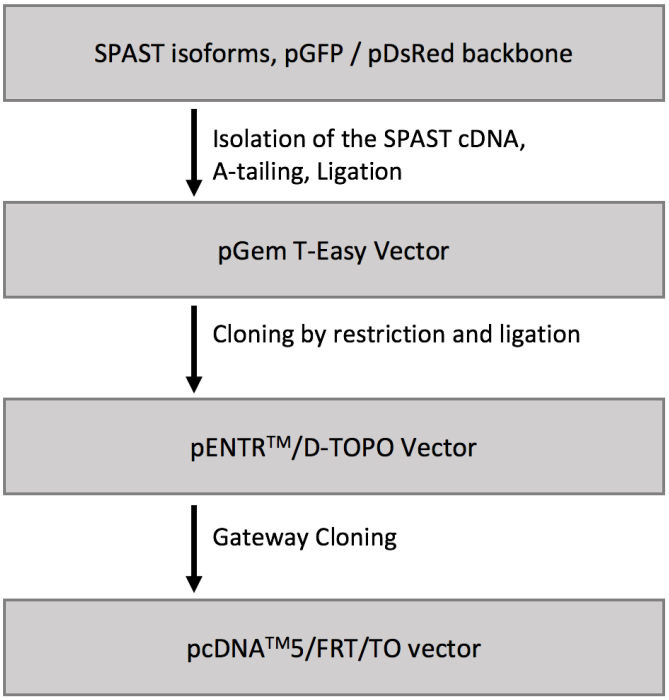


Figure 6: Schematic of the cloning strategy used in this work
First, the cDNA was isolated by a polymerase chain reaction and integrated into a pGEM®-T Easy vector; next the cDNA was transferred to a pENTR™/D-TOPO® vector by digestion with restriction enzymes and subsequent ligation; finally, the cloning into the pcDNA™5/FRT/TO vector took place by Gateway® cloning.

2.3.1 **Detection of PCR products by gel electrophoresis**

The ability to separate DNA fragments by their length by gel electrophoresis was used in various steps in this work to check results of PCRs or enzymatic plasmid digests. Therefore, 6X DNA Loading Dye (Thermo Fisher Scientific, Waltham, USA) was added to a DNA sample (ratio 1:5) and the sample was loaded on a

2% agarose gel (Table 8). The Gene Ruler DNA Ladder Mix (Life Technologies (Thermo Fisher Scientific), Waltham, USA) was used as a size marker.

Table 8: Composition agarose gel (2%)

Components	Amount
1/2x TBE buffer	200ml
SeaKem®LE Agarose (Lonza, Basel, Switzerland)	2,5g
Midori Green Advance (Biozym, Vienna, Austria)	8µl

**1/2x TBE buffer (45mM Tris Base, 45mM Boric Acid and 10mM EDTA in H_2O) and agarose were boiled up together. After a short cool down Midori Green was added and the gel was poured.*

Gel electrophorese was carried out at 150V for about 30 minutes. Afterwards, bands were detected in a UV gel documentation chamber (V029119, Vilber Lourmat, Eberhardzell, Germany).

2.3.2 Cloning of the *SPAST* isoforms into a pGEM®-T Easy vector

The four *SPAST*-Isoforms were amplified from the transport vector by a polymerase chain reaction. For further cloning by enzymatic restriction a Not I restriction site was created at the N-terminus and an Asc I (Sgs I) restriction site at the C-terminus. The primers designed for this step are listed in Table 11.
In this step, the Q5 Hot Start High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA/USA) was used in the following reaction (Table 9):

Table 9: SPAST amplification reaction

Components	Volume (µl)
Q5 Hot Start High-Fidelity 2X Master Mix	12,5
10µM Forward Primer	1,25
10µM Reverse Primer	1,25
Template DNA (10ng/µl)	1
Nuclease-free water	9,0
Total	25

Table 10: Primer SPAST amplification with restriction sites

Primer	Sequence (5'-3')
M1 forward (Not I restriction site)	***GCGGCCGC**** ATGAATTCTCCGGGTGG
M87 forward (Not I restriction site)	***GCGGCCGC**** ATGGCAGCCAAGAGG
Reverse (Asc I/Sgs I restriction site)	***GGCGCGCC**** TTAAACAGTGGTATCTCCAA

*Restriction sites (*Not I restriction site, ** Asc I (Sgs I) restriction site) are marked in bold letters.*

Table 11: PCR conditions SPAST amplification

Step	Temperature	Time
1, Initial Denaturation	98°C	30 seconds
2, Denaturation	98°C	10 seconds
3, Annealing	63°C	2:20 minutes
4, Elongation	72°C	3 minutes
5	Go to step 2 / 29x	
6, Final Extension	72°C	20 minutes
7, Hold	4°C	forever

The result was checked by gel electrophoresis as described in 2.3.1 and remaining PCR mix was purified using the QIAquick PCR purification kit (Quiagen, Hilden) following the manufacturers protocol. The DNA concentration was measured with a spectrophotometer (NanoDrop 1000, Thermo Scientific, Waltham, MA/USA) at 260nm.

As the pGEM®-T Easy is an open vector with a single thymine (T) overhang, an adenosine (A)-tail on the desired DNA insert is required for cloning into the vector. An A-tail was added by performing the following reaction (table 12) which incubated for 30 minutes at 70°C.

Table 12: A-tailing reaction

Components	Volume (µl)
5X Colorless GoTaq Reaction Buffer (Promega, Madison, USA)	2
10 mM dNTPs (Thermo Fisher Scientific, Waltham, USA)	0,2
G2 GoTaq polymerase (Promega, Madison, USA)	0,1
PCR product	amount calculated as described
ddH_2O	Ad 10

The amount of PCR product needed for the ligation reaction was calculated with a “Molar Ratio of Insert : Vector” calculator (Promega) and according to that, the appropriate amount was used in the Adenosine-tailing reaction.
The ligation-reaction mix (table 13) incubated at 4°C over night.

Table 13: Ligation reaction

Components	Volume (µl)
PCR product	2
T4 DNA Ligase (New England Biolabs, Ipswich, MA/USA)	1
10X T4 DNA ligase reaction buffer (New England Biolabs, Ipswich, MA/USA)	1
pGEM®-T Easy (Promega GmBH, Mannheim)	0,5
ddH_2O	ad 10µl

The next day the resulting plasmid was amplified by transformation into chemically potent bacteria (Escherichia coli) as described above (2.1).
The correct integration of the gene into the vector was checked by sequencing of the gene overlapping into the vector as described in 2.1.2 (Primer: 782rv, 1360fw; table 4).

2.3.3 Cloning into pENTR™/D-TOPO® vector

Other than the pGEM®-T Easy vector, the pENTR™/D-TOPO® vector is a closed DNA ring. For cloning into this vector, a small part has to be cut out by enzymatic digestion resulting in the formation of two sticky endings. An insert that was digested with the same enzymes, and therefore has the same endings enabling ligation.

For this cloning step, the restriction sites added in 2.3.1 were used. Both, the pGEM®-T Easy vector containing the inserted *SPAST* gene as well as the pENTR™/D-TOPO® Vector were enzymatically digested with the following restriction enzymes (table 14).

Table 14: Restriction Enzymes used for cloning into pENTRTM/D-TOPO®

Enzyme	
Not I	Thermo Scientific (Waltham, MA/USA)
Asc I, (Sgs I)	Thermo Scientific (Waltham, MA/USA)

The conditions for the combined digestion were determined by the DoubleDigest calculator (www.thermofisher.com). According to that, both the pGEM®-T Easy vector containing the *SPAST* gene and the empty pENTR™/D-TOPO® were first digested with Asc I (SgsI) in the following reaction (table 15).

Table 15: Enzymatic plasmid digestion

Contents	Amount
10X Tango Buffer (Thermo Scientific, Waltham, MA/USA)	1µl
AscI (Thermo Scientific, Waltham, MA/USA)	0,5µl
Vector	1µg
ddH_2O	ad 10µl

Next, 2 µl of the enzyme Not I (Thermo Scientific, Waltham, MA/USA) and 1,25µl of Tango buffer were added and the mixture and incubated again for one hour at 37°C. The result was checked by gel electrophoresis and the desired bands were

cut out for gel purification using the QIAquick Gel Extraction Kit (Quiagen, Hilden, Germany), following the manufacturer's protocol. The DNA amount was measured using a spectrophotometer (NanoDrop 1000, Thermo Scientific, Waltham, MA/USA).

The ligation reaction was performed as described below (table 16). The amount needed for ligation was calculated using the Promega ligations calculator (http://www.promega.com/a/apps/biomath/) and an Insert:Vector ratio of 1:1.

Table 16: Ligation reaction pENTR™/D-TOPO®

Contents	Volume (µl)
T4 DNA Ligase (New England Biolabs, Ipswich, MA/USA)	1
10X T4 DNA ligase reaction buffer (New England Biolabs, Ipswich, MA/USA)	1
Digested *SPAST* insert	calculated amount
pENTR™/D-TOPO® vector (Thermo Scientific; Waltham, MA/USA)	14ng
ddH_2O	Ad 10µl

2.3.4 Cloning into pcDNA™5/FRT/TO Vector by Gateway® Cloning

In a final cloning step, the *SPAST* gene was transferred to the destination vector (pcDNA™5/FRT/TO; medical proteome center, Tübingen) using the Gateway® system. This system takes use of the ability of phage λ to integrate its genome into E. coli bacteria by site-specific recombination. Specific recombination sites in both organisms (attP in phage λ and attB in E. coli) allow an integration process called lysogeny. Two enzymes are needed for this event, an integrase encoded in the genome of phage λ and the integration host factor of E. coli. Upon the recombination of the recombination sites, new attR and attL sites are generated. As the excision can be catalyzed by the same enzymes, this recombination process is reversible.

Gateway® cloning is an in vitro use of this system: An attR site in the pENTR™/D-TOPO® vector (Thermo Scientific; Waltham, MA/USA) and an attL site in the

pcDNA™5/FRT/TO vector allows an efficient recombination of the two vectors resulting in the transfer of an insert from the entry vector to the destination vector.

Table 17: Composition LR reaction

Contents	Amount
Destination vector (pcDNA™5/FRT/TO)	500ng
pENTR™/D-TOPO® vector + *SPAST*	100ng
Clonase II Enzyme Mix (Thermo Scientific; Waltham, MA/USA)	2µl
Tris-EDTA (TE) buffer, pH 8.0 (Quiagen, Hilden)	Ad 10µl

The listed LR mixture (table 17) incubated at 25°C for one hour. Next, 1µl of Proteinase K (2µg/µl, Thermo Scientific; Waltham, MA/USA) was added and the reaction incubated at 37°C for 10 minutes. The result was then transformed in chemically potent bacteria (Escherichia coli) which were distributed on Ampicillin (50µg/ml; Thermo Scientific; Waltham, MA/USA) containing agar plates overnight. Transformants were checked by enzymatic digestion with Bam HI (Thermo Scientific (Waltham, MA/USA)) as described in 2.3.3. A BamHI/Lsp buffer (Thermo Scientific, Waltham, MA/USA) was used.
At the end, all plasmids were sequenced to check the DNA sequence and the correct integration of the *SPAST* gene behind the Strep/Flag Tag as described in 2.1.

2.4 Cultivation of Hek293T and SH-SY5Y cells

All cells were cultured at 37°C, 5% CO_2 and 100% relative humidity (Incubator Heracell 240, Heraeus, Hanau). Media were changed every 3 to 4 days. All media and phosphate buffered saline (Dulbecco's PBS, Sigma life science, St. Louis, MA/USA) were prewarmed to 37°C in a water bath. All cell culture work was performed under sterile conditions. As some of the used antibiotics (hygromycin, zeocin, doxycycline) are very light sensitive all steps involving those substances were carried out in the dark and the media were protected from light.

When cells reached a confluency of 80-90%, they were passaged and split. To this end, they were first washed once with PBS (Dulbecco's phosphate buffered saline, Sigma life science, St. Louis, MA/USA) and detached from the surface by 1% trypsin (Biochrom GmbH, Berlin) in PBS. After an incubation time of 3-5 minutes at 37°C the reaction was stopped by adding the same amount of medium containing 15% fetal bovine serum (FBS, Thermo Scientific (Waltham, MA/USA)) and the cells were carefully distributed on new plates.

2.4.1 Cultivation of Hek293T cells

Hek293T is a human embryonic kidney cell line that was created by transforming human kidney cells with an adeno virus 5. The cell line distinguishes itself by its easy handling and high transfection efficiency and is therefore popular as a cell model in a lot of experiments. In this work transfected Hek293T cells were used for transfection with fluorescent tags.

The cells were cultured in DMEM (Dulbecco's Modified Eagle Medium; Thermo Scientific, Waltham, MA/USA) with 15% FBS (Fetal bovine serum, Thermo Scientific, Waltham, MA/USA) as and were passaged when 80% confluent.

2.4.2 Transfection of Hek293T cells with fluorescently labeled *SPAST* and HA tagged *SPAST*

The high transfection efficiency of Hek293T cells was used at the beginning of this work. The original *SPAST* plasmids with fluorescent backbone (table 1; Christian Beetz, Jena) were transfected into Hek293T cells to check the functionality of the plasmid. The cells where either used for western blotting to validate the spastin expression on protein level or they were fixated, and the fluorescence label detected by fluorescence microscopy.

The transfection of cells by Lipofectamine takes advantage of the ability of lipid droplets to pass the plasma membrane and reach the cytoplasm. This process is used by packing plasmids with the integrated cDNA of the desired protein into lipid droplets and treating cultured cells with this mixture. The cDNA then is transiently integrated in the cell and the desired protein is expressed.

For transfection, at day 0, Hek293T cells were seeded in 6 wells plates at an appropriate density of 50-60%. On day 1 the medium was changed and the transfection mixture (table 18) was carefully, dropwise added to the cells.

Table 18: Composition transfection mixture

Contents	Amount
Lipofectamine 2000 (Thermo Scientific, Waltham, MA/USA)	20µl
desired plasmid DNA for transfection	3µg
OptiMem (Thermo Scientific, Waltham, MA/USA)	500µl

After an incubation of 48 hours, on day 3, cells were lysed or fixated and used for further experiments.

2.4.3 **Cultivation of a SH-SY5Y Flp-In™ T-Rex™ host cell line**

As a generous gift, we received a SH-SY5Y Flp-In™ T-Rex™ host cell line (Karsten Boldt, medical proteome center Tübingen) generated following the instructions of the Flp-In™ T-Rex™ core kit (Invitrogen, Carlsbad, CA/USA).
The Flp-In™ T-Rex™ host cell line is characterized by an included FRT site, which is needed to stably integrate genes in the genome, combined with a tetracycline-dependent CMV/$TetO_2$ promotor which is later used for the tetracycline dependent transcription and expression of the desired protein.
Due to selection steps used in the generation of this host cell line, the cells carry resistances to the antibiotics blasticidin and zeocin and were therefore cultivated in the following medium (table 19).

Table 19: Composition Host Line Medium

Contents	
DMEM (Dulbecco's Modified Eagle Medium) (DMEM GlutaMAX™, Thermo Scientific, Waltham, MA/USA)	Ad 500ml
FBS (Fetal Bovine Serum) (Thermo Scientific, Waltham, MA/USA)	15% (v/v)
Penicillin/ Streptomycin (Biochrom GmbH, Berlin)	1% (v/v)

Blasticidin (InvivoGen, Toulouse/FR)	5µg/ml
Zeocin (InvivoGen, Toulouse/ FR)	100µg/ml

2.4.4 Cotransfection of pcDNA™5/FRT/TO-*SPAST* and pOG44 into a Flp-In™ T-Rex™ host line

For the generation of a Flp-In™ T-Rex™ expression cell line the host line cells were co-transfected with two plasmids: the destination vector containing a *SPAST* isoform and a pOG44 plasmid coding for a recombinase that is necessary for the rearrangement of the two FRT sites and by that the stably integration of the plasmid into the host line's genome. The destination vector carries an antibiotic resistance to hygromycin, while the zeocin resistance of the host line gets lost during recombination.

For further selection, two negative controls were created, that where treated similar: one transfection with just one plasmid (pOG44) and the host line itself.

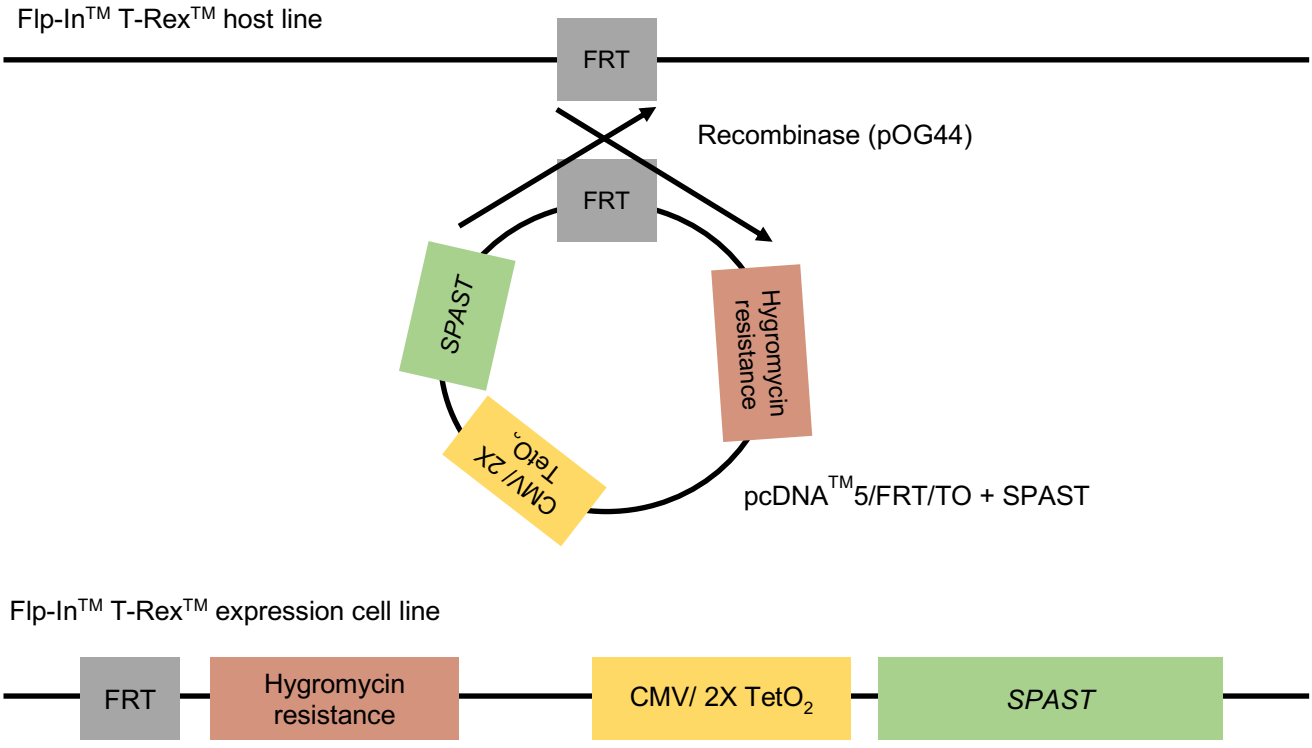


Figure 7: Scheme of the integration of a pcDNATM5/FRT/TO expression vector into a pcDNATM5/FRT/TO Host Line

the recombination of both FRT sites is shown, the directly neighbored zeocin resistance of the Host Line is lost during this process, the hygromycin resistance of the expression vector is integrated into the cell's genome. (adapted from pcDNA™5/FRT/TO user manual, Invitrogen Carlsbad, CA/USA)

Before transfection, cells were counted using a Neubauer chamber with a depth of 0.1mm. By a trypan blue stain, dead cells were marked and excluded in the count. The determination of the cell count was performed according to standard protocols.

On day 0, 5 x 10^6 cells were distributed on each of six 10cm petri dishes. For the transfection of the Flp-In™ T-Rex™ host line, Lipofectamine 2000 (Thermo Scientific, Waltham, MA/USA) and a ratio pcDNA™5/FRT/TO: pOG44 (w/w) of 1:4 was chosen. A total of 8µg of plasmid was used in each approach (table 23). After removing the medium, the cells were washed once with DMEM containing 15% FBS, which was also used as transfection medium.

The transfection mixture (table 20) incubated for 20min at RT and was then added dropwise to the cells.

Table 20: Transfection mixture for the transfection of a Flp-In™ T-Rex™ host cell line in a 10cm petri dish

Contents	Amount
DMEM (Dulbecco's Modified Eagle Medium)	3 ml
Lipofectamine 2000	20µl
pcDNA™5/FRT/TO containing a *SPAST* isoform	1,6µg
pOG44	6,4µg

On day 2 the medium was changed to DMEM containing 15%FBS and 1% Penicillin/Streptomycin (Biochrom GmbH, Berlin).

2.4.5 **Selection of Colonies by Hygromycin**

Two days after transfection, on day 3, the selection of successfully co-transfected cells carrying a resistance for hygromycin started by passaging and splitting the

cells 1:4 as described in 2.4.1. and changing the medium to the final selection medium, containing hygromycin instead of zeocin.

Table 21: Medium Flp-In™ T-Rex™ expression cell line

Contents	
DMEM (Dulbecco's Modified Eagle Medium) (DMEM GlutaMAX™, Thermo Scientific, Waltham, MA/USA)	Ad 500ml
FBS (Fetal Bovine Serum) (Thermo Scientific, Waltham, MA/USA)	15% (v/v)
Penicillin/ Streptomycin (Biochrom GmbH, Berlin)	1% (v/v)
Blasticidin (InvivoGen, Toulouse/FR)	5µg/ml
Hygromycin (InvivoGen, Toulouse/ FR)	100µg/ml

In the following weeks, the medium was changed twice a week until no attached, living cells remained in the negative, non-transfected, controls. After 2 to 3 weeks first colonies were visible, got marked and were picked 4-5 weeks after transfection into 24 well plates.
When confluent, they were passaged to 6 wells and further screened for the desired expression of Flag-spastin.

2.4.6 **Induction of Flag-spastin expression by doxycycline**

Due to the tetracycline-regulated CMV/TetO_2 promotor linked to the *SPAST* gene in the Flp-In™ T-Rex™ expression cell lines, the spastin overexpression could be induced by the addition of this antibiotic to the cell culture medium. To validate the influence of tetracycline on the transcription, the protein expression after treatment with different concentrations of doxycycline (0µg/ml, 0.5µg/ml, 1µg/ml, 5µg/ml, 10µg/ml) was examined.
For the doxycycline treatment, clones were split on 6 wells. When they were about 80% confluent doxycycline was added.
After 48h cells were lysed for further experiments.

2.4.7 Differentiation of SH-SY5Y cells into a neuronal phenotype by retinoic acid and brain derived neuronal factor

SH-SY5Y cells were chosen for this experiment due to their resemblance to human neuronal tissue by their molecular neuronal characteristics and the capability to differentiate into a neuron-like phenotype in the presence of retinoic acid (RA).

The differentiation of SH-SY5Y cells was performed according to Kovalevich and Langford (2013).

Cells were cultured on 6well plates until ca. 60% confluent. The cells incubated in DMEM medium with 15% FCS and 10µM retinoic acid (RA, Sigma-Aldrich) for 6 days. The medium was changed to serum free DMEM F12 medium on day 7 and 50ng/ml BDNF (brain derived neurotrophic factor, Peprotech) was added, in which the cells incubated for another 6 days. Finally, on day 13 day the cells were fixated for immune fluorescent staining.

2.5 Protein isolation and detection by Western Blotting

2.5.1 Cell lysis and protein isolation

The isolation of proteins from cultured cells was needed to perform western blotting or immunoprecipitation (IP). Lysis was performed using a buffer containing the detergence Nonidet P40 (NP40), that mildly destroys the cell's membrane structure and releases intracellular protein.

For protein isolation cells were washed once with PBS. Afterwards they were scratched (Cell-scraper, Sarstedt, Nürnbrecht) in an appropriate amount of lysis buffer (50mM Tris-base (pH7.4); 150mM NaCl, 0.5% (v/v) NP40 Igepal (Sigma-Aldrich), 2% proteinase inhibitor cocktail (cOmplete mini, Roche) in ddH_2O) on ice and transferred to a 1.5ml tube. After an incubation of 10 minutes on ice the samples were centrifuged at 14.000rpm and 4°C for 30minutes.

The supernatant containing the cellular proteins was transferred to another tube, shock frozen in liquid N_2 and stored at -80°C long term. A small amount was removed and diluted 1:10 with PBS for the quantitative protein analysis by a BCA (bicinchoninic acid) protein assay to determine the protein concentration

2.5.2 Quantitative Analysis of cellular protein

The amount of protein in the sample was determined by performing a bicinchoninic acid (BCA) assay. Therefore, the Pierce BCA Protein Assay Kit (Thermo Scientific, Waltham, MA/USA) was used. It was carried out following the manufacturers protocol in a 96 well plate, using triplets of 10µl of diluted proteins (1:10 in PBS) and 200µl of the BCA reagent mixture (reagent A:B – 50:1) per well.

After an incubation of 30 minutes at 37°C the absorbance was measured at a wavelength of 595nm. A cascade of bovine serum albumin (BSA) dilutions was used to create a standard curve to calculate the sample protein concentrations.

2.5.3 Detection of spastin through a Western Blot

With the aim of validating the expression of Flag-tagged spastin in the Flp-In™ T-Rex™ expression cell line western blots were performed. This method was also used for co-immunoprecipitations later in this work.

Western blotting is a standard technique to identify and quantify proteins in a cell lysate. Three steps can be distinguished: First the proteins are separated by their molecular weight through gel electrophoresis, then they are transferred to a membrane where in a last step they are visualized by a protein-specific primary and secondary antibody combination.

All solutions needed for western blotting are listed in table 25.

In a first step the protein samples needed to be denatured to form their secondary structure, that was stabilized by addition of dithiothreitol (DTT; Thermo Scientific, Waltham, MA/USA). To this end the following mixture (table 22) was incubated at 96°C for 10min. Based on the BCA assay results an appropriate amount of isolated protein was calculated. A total of 25-80µg protein was used in each western blot sample. After a cool down to room temperature the samples were loaded on the Bis-Tris gel (table 26). As a protein marker, the Precision Plus Protein™ Dual Color standard (Biorad, Munich) was used.

Table 22: Protein preparation for Western Blotting

Contents	
25-80µg isolated protein	calculated volume (25-80µg)
5X Pierce™ Lane Marker Reducing Sample Buffer (Thermo Scientific, Waltham, MA/USA)	20% (v/v)
10X DTT (Thermo Scientific, Waltham, MA/USA)	10% (v/v)
ddH_2O	Ad 25-70µl

For gel electrophoresis, an acrylamide gel composed of a stacking (4%) and a separating (8%) gel was poured. Gel electrophoresis was carried out in an electrophoresis chamber (Biorad, Munich) filled with running buffer (table 25) at 120V, 100mA for ca. 1h.

Table 23: Composition Bis-Tris gel

Contents	Volume (ml) Stacking gel (4%)	Volume (ml) Separating gel (8%)
3.5X Bis-Tris buffer (1 M Bis-Tris in ddH_2O, pH 6.5)	0.86	2.29
40% Acrylamide (v/v) (AppliChem, Darmstadt)	0.3	1.6
ddH_2O	1.81	4.03
10% (v/v) APS (0.4M APS in ddH_2O)	0.03	0.08
TEMED	0.003 (3µl)	0.008 (8µl)
total of	3	8

**APS: ammonium persulfate; TEMED: tetramethylethylenediamine*

Next, the separated samples were transferred to a nitrocellulose membrane (polyvinyl difluoride) in a transfer buffer (table 25) at 4°C, 25V and 50mA overnight.

To check the transfer, the protein on the membrane was stained with a Ponceau S solution (Sigma Aldrich, Chemie GmbH, Munich) for 15min at RT. The

membrane was then washed three times with TBS-T (tris buffer saline + 0,1% (v/v) Tween 20) and blocked with milk (5% (w/v) nonfat dried milk powder (Sigma-Adrich) in TBS-T) for 1h at room temperature.
A specific primary antibody (table 24) was added and incubated overnight at 4°C. The housekeeping gene GAPDH was used as a loading control. All antibodies were diluted in 5ml milk (5% (w/v) nonfat dried milk powder in TBS-T).

Table 24: Antibodies used for Western Blotting

Primary Antibody	Dilution	Company (Cat.No.)
Spastin (mouse)	1:5,000	Santa Cruz Biotechnology (sc-81624)
Flag (mouse)	1:10,000	Sigma-Aldrich (F3165)
GAPDH (mouse)	1:10,000	Meridian Life Science (H86504M)
NUP43 (rabbit)	1:5,000	Sigma-Aldrich (PA5-55514)
ATP5A (mouse)	1:1,000	Abcam (ab14748)
Atlastin (mouse)	1:1,000	Sigma-Aldrich (WH0051062M3)
HRP-anti-mouse	1:10,000	Jackson ImmunoResearch, Cambridgeshire, UK (115-035-003)
HRP-anti-rabbit	1:10,000	Jackson ImmunoResearch, Cambridgeshire, UK (111-035-003)

**GAPDH: glyceraldehyde 3-phosphate dehydrogenase; HRP: horse reddish peroxidase; all antibodies were diluted in 5% (w/v) nonfat dried milk powder in TBS-T*

The next day, the membrane was washed three times with TBS-T for 10min and subsequently incubated with a secondary antibody (table 27) conjugated with a horseradish peroxidase (HRP) for 1h at room temperature. After three washing steps with TBS-T for 10min an enhanced chemiluminescence (ECL; Pierce™ ECL Western Blotting Substrate, Thermo Scientific) solution was added and the signal resulting of the activity of the horse reddish peroxidase linked to the secondary antibody was detected through a ChemiDoc XRS+ imaging system (Biorad, Munich) and analyzed with the Image Lab 5.0 software.

Table 25: Composition of all solutions used for western blotting

Western Blot solutions	Compositions

1X TBS (tris buffered saline)	50mM Tris-base (pH7.4); 150mM NaCl in ddH_2O
1X TBS-T	50mM Tris-base (pH7.4); 150mM NaCl, 0.1% Tween20 (v/v) in ddH_2O
1X running buffer	1M MOPS, 1M Tris-Base, 70mM SDS, 20mM EDTA in ddH_2O
1X transfer buffer	0.25M Bicine, 0,25M Bis-Tris, 20mM EDTA, 10% (v/v) methanol in ddH_2O

**NaCl: sodium chloride, MOPS: 3-(N-morpholino) propanesulfonic acid; SDS: sodium dodecyl sulfate; EDTA: ethylenediaminetatraacetic acid, ddH_2O: double-distilled H_2O*

2.6 Increase of the expression of the M1 isoform by inhibiting protein degradation

Since the protein expression of the M1 isoform of spastin in the Flp-In™ T-Rex™ expression cell line was extremely low and could barely be detected by western blotting, an increased degradation of the M1 isoform on protein level was considered as a possible cause.

To increase the expression of M1 spastin, the degradation of cellular protein was inhibited by limiting the proteasomal activity by addition of the potent cell-permeable proteasome inhibitor MG132 (Sigma-Aldrich).

The Flp-In™ T-Rex™ expression cell line was induced for 24h with 1ng/µl doxycycline (as described in 2.4.6) and as additional supplement two concentrations of MG132 (5µM and 10µM) were examined.

Cells were lysed after an incubation of 24h and a western blot was performed.

2.7 Interactome analysis of spastin through a label free mass spectrometry approach

In the following two sections the generated Flp-In™ T-Rex™ SH-SY5Y spastin expression cell lines were used to perform a mass spectrometry-based interactome analysis. An overview of these methods is illustrated in figure 8.

Induced, as well as non-induced cells (as negative control) were lysed and an affinity purification performed to isolate spastin in complex with bound interaction partners. The eluted protein mixture was digested into peptides through which the contained proteins were identified by mass spectrometry. After the qualitative and quantitative analysis of the mass spectrometry data a functional annotation was performed, allowing the formation of isoform specific spastin interaction networks, so called interactomes.

All methods of the following two sections (affinity purification and mass spectrometry) were carried out in cooperation with the medical proteome center Tübingen (Karsten Boldt).

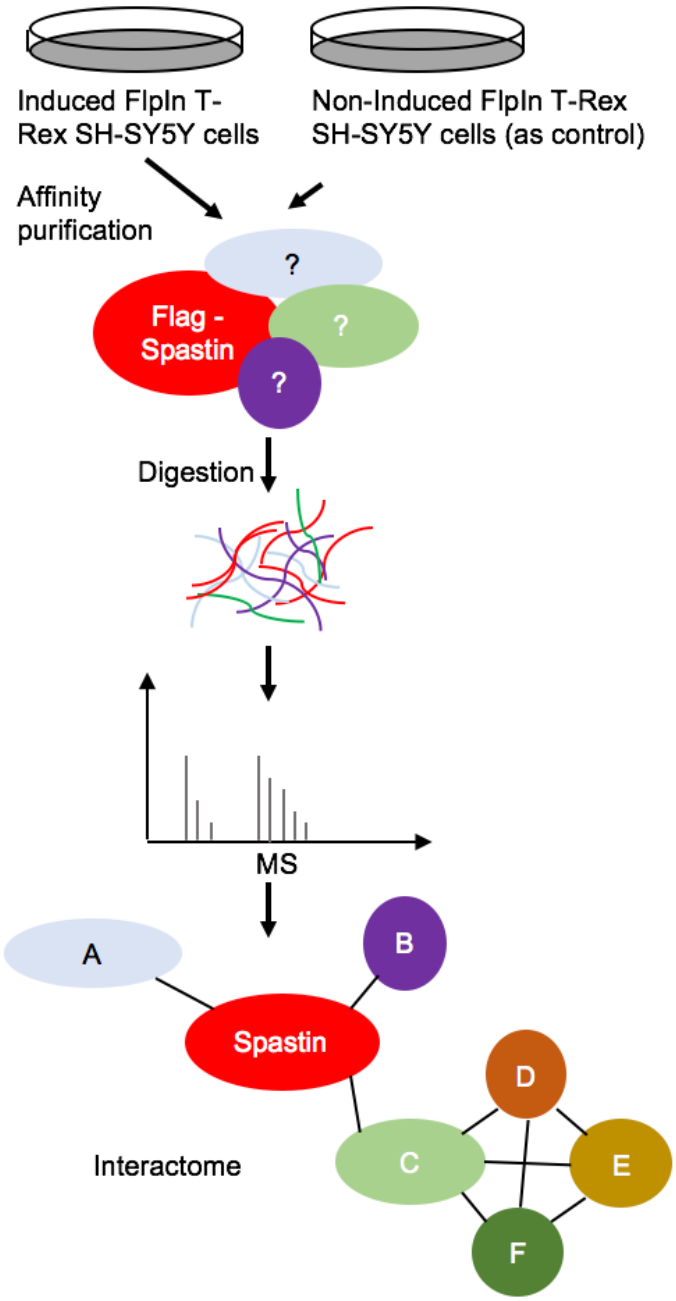


Figure 8: Workflow of the Interactome analysis using a label free mass spectrometry approach

2.7.1 **Extraction of Flag tagged spastin by Flag Affinity Purification**

To perform an interactome analysis of the spastin isoforms, the first step was to purify the overexpressed, tagged spastin in complex with all bound interaction partners from the SH-SY5Y cell lysate.

Affinity purification was carried out as published by Gloekner et al (2007). However, other than described in the publication, we used a one-step purification (Flag-tag) instead of the tandem affinity purification (Strep/Flag) to avoid an decreased amount of protein resulting in false negatives. All buffers needed for affinity purification are listed in table 30.

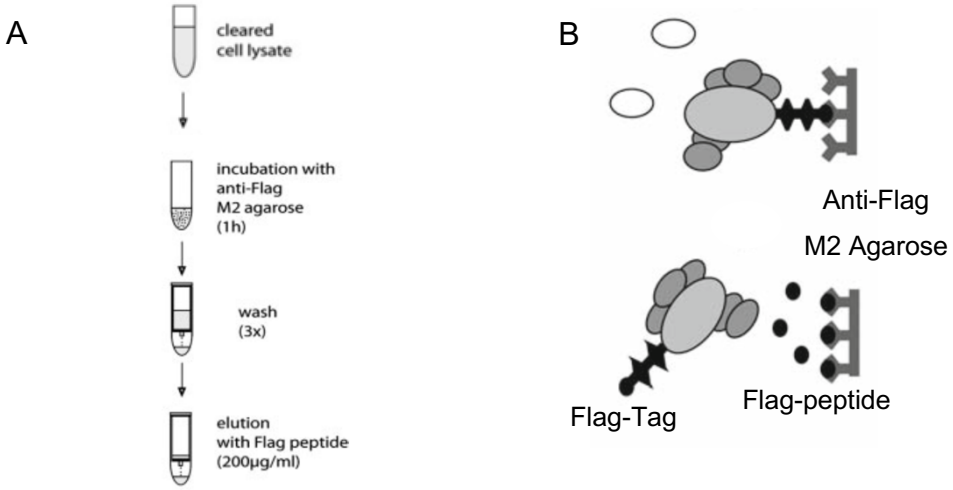


Figure 9: Schematic of the affinity purification using the Flag tag
A: the cell lysate incubated with anti-Flag beads and the mixture incubated on an anti-Flag M2 agarose. After three washing steps the desired protein was obtained by elution with the Flag peptide. B: the binding of a Flag tagged protein to anti Flag beads and its displacement by a Flag peptide is shown. Adapted from Gloeckner, Boldt, and Schumacher (2007)

For each spastin isoform, four 150cm cell culture flasks of each condition (induced for 48h with 1ng/µl Doxycycline and non-induced) were lysed in 6ml lysis buffer (table 30). The cell lysis was performed as described in a previous section (1.6.3).

The Flag beads (anti-Flag-M2-agarose-beads, mouse, SIGMA; A 2220; 150µl of the bead-mixture for each sample) were prepared by solving them in 600µl TBS

and three washing steps, once with 500µl lysis buffer and twice with 500µl wash buffer (table 27). After each washing step the beads were centrifuged (1min at 5000rpm) and the supernatant was discarded.
The washed beads were solved in 300µl TBS, added to the protein lysate and the mixture incubated in an end-over-end shaker at 4°C for one hour.
Next, the beads were isolated by centrifuging at 4°C and 5000 rpm for one minute. The supernatant was discarded, and the beads transferred to an anti-Flag M2 agarose column.
The column was washed 3 times by adding 500µl wash buffer and discarding the flow through. After the last washing step, the columns were centrifuged until dry (1000rpm, 10sec) and 200µl flag peptide were added on the dry column for elution. After an incubation of 10min the columns were centrifuged (1000rpm, 1min). The eluate was stored at -20°C for up to two weeks.

Table 26: Composition of all buffers used for the affinity purification

	Lysis Buffer	Wash Buffer
10X TBS	1ml	1ml
Nonidet P40	50µl	10µl
Protease inhibitor	200µl	-
ddH_2O	Ad 10ml	Ad 10ml

**TBS: tris buffered saline (table 28)*

In preparation for a mass spectrometric analysis, the affinity purified protein samples were precipitated out of the eluate. To this end, 800µl methanol (100%, Merck), 200µl chloroform (100%, Merck) and 600µl ddH_2O were added to 200µl protein eluate. After a centrifugation step the upper phase was discarded and another 600µl methanol (100%, Merck) added. After centrifuging for 2min at 16,000g the supernatant was discarded, and the protein pellet was dried under laminar flow for 10 minutes. The protein pellets were stored at -80°C for up to one week.

2.7.2 Detection of interaction partners by mass spectrometry

Mass spectrometry works by separating molecules, in this experiment peptides, by their individual mass. This process requires the ionization of the molecules. The gaseous ions are then characterized by their mass to charge (m/z) ratio and thereby allow conclusion regarding the original peptide. Each protein has its specific peptide pattern and can thus be identified. The following steps were carried out as published by Gloeckner et al. (2009b).

Five biological replicates of every spastin isoform were measured. The interactome analysis was performed by a label free quantification using non-induced SH-SY5Y cells as negative control.

In preparation for the mass spectrometric measurements, the precipitated proteins were first digested into peptides by performing an in-solution tryptic digest. The protein pellet was therefore first solved in 30µl ammonium bicarbonate (ABC, 50nM, Sigma Aldrich). Next, 4µl RapiGest SF Surfactant (Waters, Eschborn) were added. After adding 1µl Dithioreitol Solution (DTT, 100mM, Merck) to stabilize the secondary protein structure the solution incubated on a thermo block at 60°C for 10 minutes. After a short cool down 1µl iodacetamide solution (IAA, 300mM, Merck) was added and the mixture incubated for 30 minutes in complete darkness.

For the digestion step 2µl trypsin (trypsin from porcine pancreas, proteomics grade, Sigma) were added and incubated at 37°C over night. The next day, the sample was centrifuged at 9,000g for one minute and the enzymatic reaction was stopped by adding 1.7µl trifluoracetic acid (TFA, final concentration of 5%, Fluka). The acidified solution was transferred into a tube with polypropylene insert (200µl, Supelco).

After another incubation for 10 minutes the tube was centrifuged at 16,000g for 15 minutes and the clear middle phase, containing the digested proteins transferred to a new tube.

The samples were purified with StageTips (Thermo Scientific), following the manufacturer's protocol.

All mass spectrometric measurements were performed on a NanoRSLC3000 HPLC system (Diodex, Idstein, Germany), coupled to an Orbitrap Fusion™

Tribrid™ Mass Spectrometer (Thermo Scientific, Waltham, MA/USA) by a nano spray ion source. The digested peptide samples were automatically injected and loaded at a flow rate of 6µl/min in 98% buffer C (0.1% trifluoroacetic acid in HPLC-grade water) and 2% buffer B (80% acetonitrile and 0.08% formic acid in HPLC-grade water) onto a nano trap column (75µm i.d. x 25cm, Acclaim PepMap100 C18, 3µm, 100Å; Dionex).

2.7.3 Qualitative and quantitative analysis of the mass spectrometric data

The analysis of the raw mass spectrometric data leading to the identification of proteins was carried out as described in Gloeckner et al. (2009b).

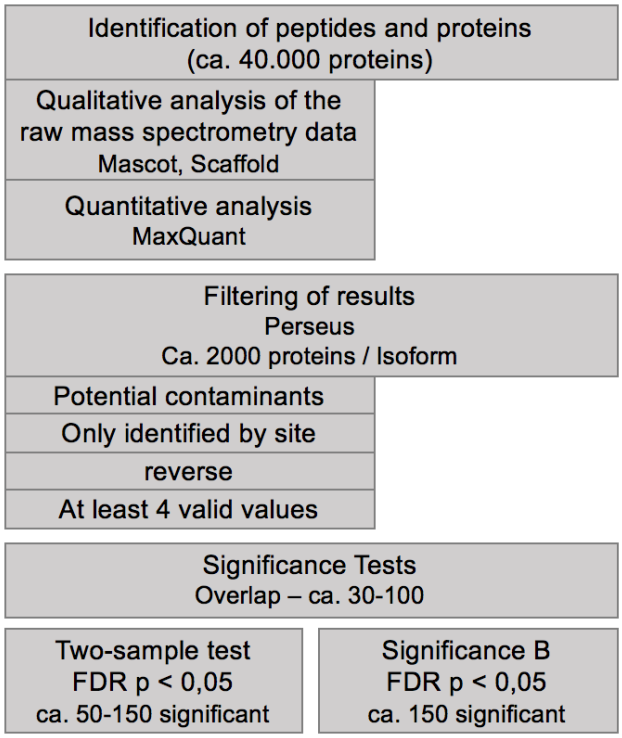


Figure 10: Scheme of the filtering procedure of the mass spectrometry results through a qualitative and quantitative analysis proteins were identified in the raw mass spectrometric data. The results were further filtered, and statistical testing was performed.

The qualitative analysis of the raw mass spectrometric data was performed by Mascot (Matrix Science, Boston, USA; version 2.4.0) and Scaffold (version

Scaffold_4.0.3, Proteome Software Inc. Portland, USA). For the analysis of all MS/MS samples, Mascot was set up to search the Swiss-Prot_2012_05 database (selected for Homo sapiens, 2012_05, 20245 entries). Trypsin was set as digestion enzyme. Mascot was searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 10.0 PPM.

Scaffold was used for the validation of MS/MS based peptide and protein identifications. Peptide identifications with a probability greater than 95.0% by the Peptide prophet algorithm with Scaffold delta-mass correction were accepted. Protein identifications were accepted with a probability greater than 99.0% and at least two identified peptides.

The mass spectrometric raw data was processed by quantitative analysis using the MaxQuant software (version 1.3.0.5) assuming trypsin as cleavage enzyme, allowing two missed cleavages per peptide. Peptide and protein false discovery rates were set to 1%. For peptide and protein identification the human subset of the human proteome reference set provided by SwissProt (Release 2012_01 534,242 entries) was used.

The MaxQuant contaminant database search was enabled to automatically detect known contaminants, including e.g. keratins. To perform protein quantification, a minimum number of two unique peptides with a minimum length of seven amino acids was needed for detection; only unique peptides were selected for quantification.

The further statistical analysis of the spastin interactome was carried out with the Perseus software (version 1.6.1.3). To narrow the list down to the significant hits, commonly used filters were applied. The results were filtered for potential contaminants (such as keratins and other common proteins), proteins that were only identified by side (identification by post-translational modifications) and reverse recognitions (reverse peptide sequence). Additionally, only proteins, that were identified at least in four samples out of the five repetitions were analyzed further. Two statistical tests were performed to explore the significance of possible spastin interactions. In a two sample test, the difference of the mean log2 transformed values of the two conditions (induced sample and the non-

induced one) was calculated for the quintuplicates. A p-value limit of 0.05 in the permutation based FDR was chosen.

$$t = \frac{(x_1 - x_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

x: mean of the log2 transformed values
s: standard deviation
n: sample size

Another approach on the significancy testing was to calculate a significance score for log protein ratios induced/non-induced (significance A; Cox and Mann (2008)). A p-value limit of 0.05 was chosen after a permutation based FDR correction for multiple testing.

$$Significance\ A = \frac{1}{2} erfc\left(\frac{z}{\sqrt{2}}\right) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-t^2/2} dt$$

As the bulk distribution of log protein ratios depends on the signal intensity, it was taken into account through a significance B testing, similar to significance A but dependent on the signal intensity.

After the analysis of the interactome mass spectrometric data with the Perseus software, significant hits ($p < 0,05$, permutation-based FDR) in either t-test or significance B were further evaluated by a functional analysis testing for common cellular pathways/molecular functions/cellular compartments. To this end the DAVID online platform (version 6.8) was used. Homo sapiens was set at host.

2.8 Validation of spastin interactions

Subsequent to the functional annotation of the interactome mass spectrometric data, a small number of promising, possible spastin interactors were chosen for further validation.

Those selected protein-protein interactions were investigated by performing co-immunoprecipitations (Co-IPs). To this end the spastin overexpression in the Flp-In™ T-Rex™ SH-SY5Y expression cell line was induced by addition of doxycycline and MG132 was added to the M1 isoform cell line to increase the protein expression. The affinity purification was performed as described in 2.8.1 and spastin and its interaction partners detected by western blotting (2.5).

3. Results

3.1 Generation of a SH-SY5Y cell line with an inducible, stable spastin overexpression

In this work, a Flp-In™ T-Rex™ SH-SY5Y expression cell line was generated to overexpress the four endogenously existing isoforms of spastin, the protein affected in the SPG4 subtype of hereditary spastic paraplegias (HSP). To this end, the four isoforms of the *SPAST* gene (M1, M87, M1 del Ex 4 and M87 del Ex 4) were cloned into pcDNA™5/FRT/TO vectors and subsequently integrated in a Flp-In™ T-Rex™ SH-SY5Y host cell line. Finally, the tetracycline dependent expression of spastin was validated by western blotting.

3.1.1 Correction of variants of unknown significance in the *SPAST* sequence

In a first step, the four isoforms of the *SPAST* gene were analyzed on nucleotide level by sanger sequencing and the result was compared to the *SPAST* cDNA registered in the UCSC database. Deviations in the base sequence were checked for known single nucleotide polymorphisms (SNPs; ENSEMBLE database). As a result, two base changes in position c.158 T>C and c.1850 A>T (labelled after their position in the M1 *SPAST* isoform) emerged. As the benign nature of the first variant, a missense mutation leading to a single amino acid change, couldn't be assumed and the second variant consisted of a base exchange resulting in the deletion of the stop codon, both variants were corrected by performing a mutagenesis. The successful correction was again validated by sanger sequencing (figure 11).

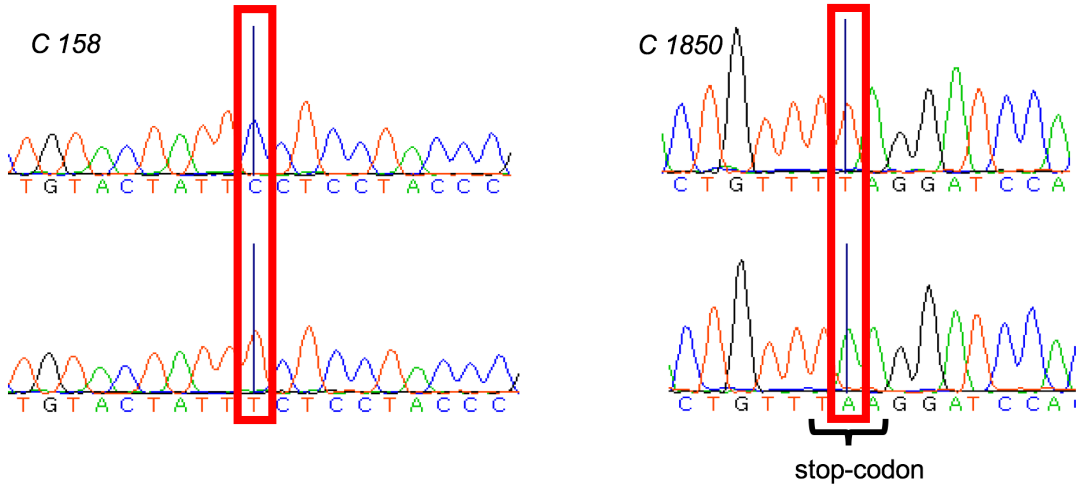


Figure 11: Representative sequence of mutated nucleotides before and after mutagenesis

Shown are the mutated nucleotides in position c.158T>C (A) and c.1850A>T (B) before (upper row) and after (lower row) correction by mutagenesis. The corrected nucleotide is marked by a red box.

3.1.2 Validation of the functionality of the SPAST plasmids by transient overexpression in Hek293T cells

To ensure the functionality of the *SPAST* DNA sequences, Hek293T cells were transfected with the corrected fluorescence tagged *SPAST* plasmids resulting in a transient overexpression of the spastin isoforms.

As a result, after the transfection of Hek293T cells all isoforms of spastin were expressed and could be detected by their fluorescent DsRed, respectively GFP tag as well as on protein level by western blotting (figure 12).

As an additional finding, the different isoforms of spastin differed in their cellular localization, as well as in their expression levels, in transfected Hek293T cells. While the two M87 isoforms were homogenously distributed in the entire cytoplasm and were even slightly present in the nucleoplasm, the M1 isoforms showed a punctate pattern in specific intracellular locations. This finding matches with the previously described localization of the isoforms (Reid et al., 2005, Solowska and Baas, 2015, Claudiani et al., 2005).

In the following western blot, the expression of all four spastin isoforms in transfected Hek293T cells is shown. As the attached GFP (~ 27kDa), respectively

DsRed (~ 27kDa) tag has a higher molecular weight then the Strep-Flag tag (~ 13kDa), used in later experiments, the detection height of the spastin isoforms differs from the western blots shown in the following chapters.

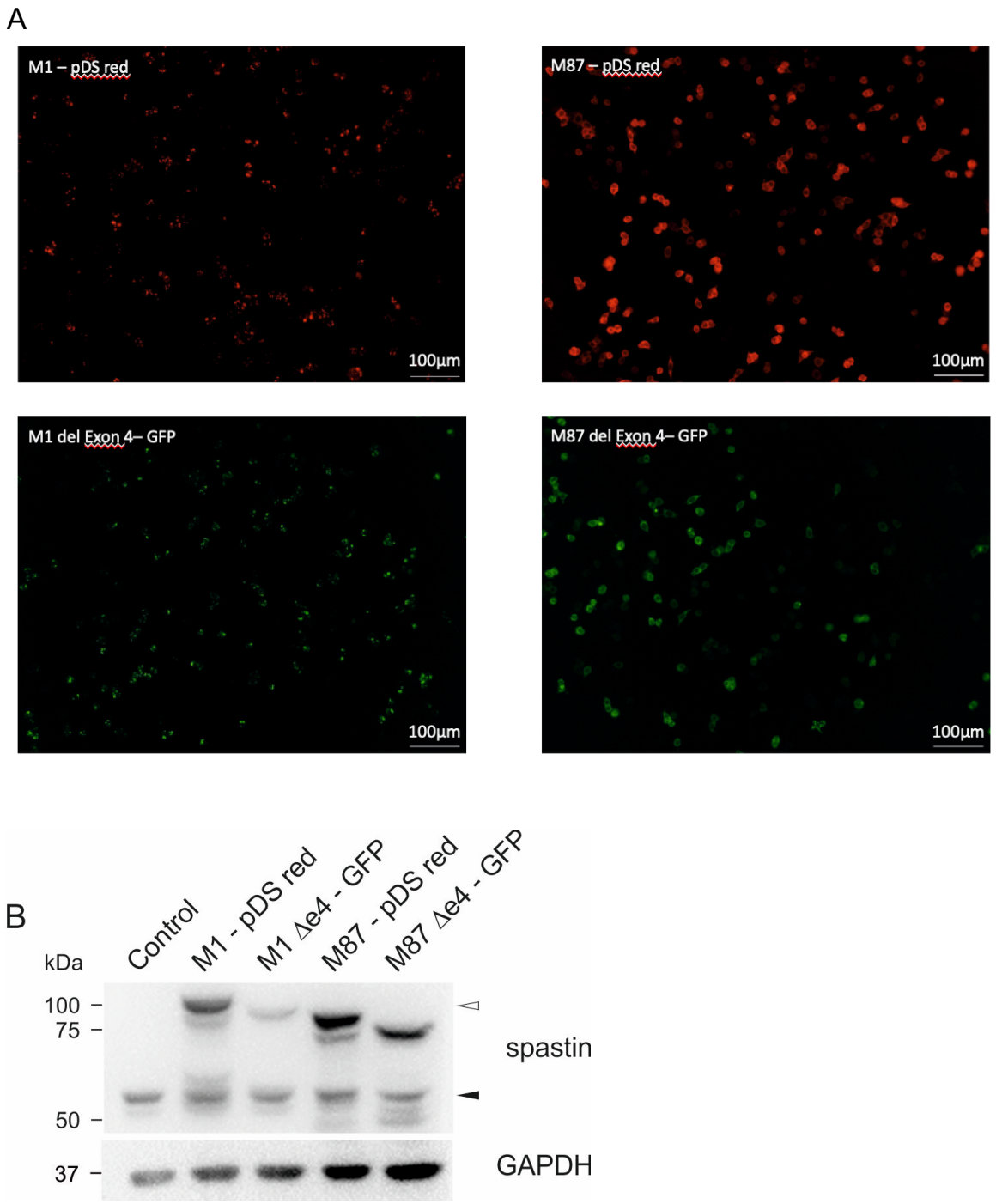


Figure 12: Protein expression and localization of the spastin isoforms in Hek293T cells

A: Fluorescence microscopy of transfected Hek293T cells. Successfully transfected cells emerge by a fluorescent signal originating from the fluorescent tagged spastin (GFP/DsRed). Notably, the M1 isoform (with and without deletion of exon 4) shows a very punctate specific cellular localization, while the M87 isoforms are homogenously distributed in the cytoplasm. Zeiss Axiovert 200M, 20X; B: Western blot analysis of the four overexpressed isoforms of spastin after transfection of Hek293T cells. The black arrowhead marks the endogenously expressed spastin, the white arrowhead indicated the overexpressed, tagged isoforms of spastin. All spastin isoforms could be detected on protein level. Non-transfected Hek293T cells served as negative control, the house keeping protein GAPDH was used as loading control. Primary antibodies: mouse anti-spastin (Santa Cruz) 1:5,000; mouse anti-GAPDH (Meridian) 1:10,000; secondary antibody: HRP coupled anti-mouse (Dianova) 1:10,000

3.1.3 **Cloning of the *SPAST* isoforms into pcDNA™5/FRT/TO vectors**

The corrected *SPAST* open reading frames were then cloned into pcDNA™5/FRT/TO vectors. As part of the Flp-In™ T-Rex™ system they allow the stable integration of the plasmid into a Flp-In™ T-Rex™ host cell line and thereby enable the inducible, stable overexpression of a gene of interest. The *SPAST* isoforms were therefore first isolated by PCR and cloned into a pGEM®-T Easy vector.

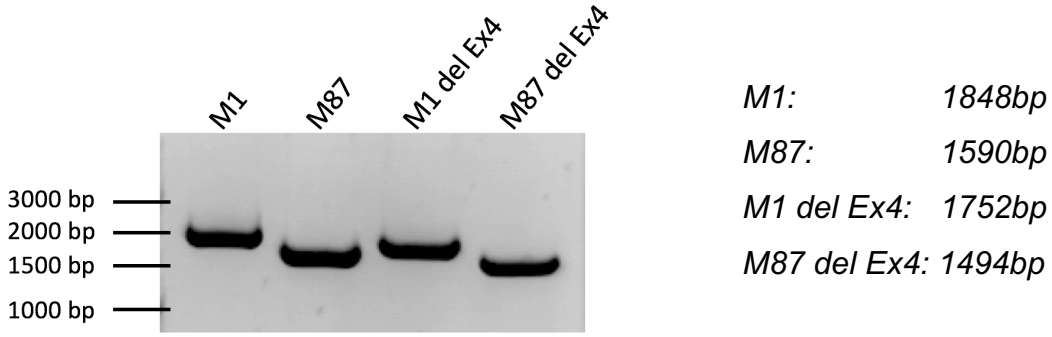


Figure 13: Agarose gel electrophoresis of the four SPAST isoforms amplified by PCR; Marker: GeneRuler DNA Ladder Mix bp: base pairs

Through enzymatic digestion and ligation, the *SPAST* gene was then transferred to a pENTR™/D-TOPO® vector and finally cloned into a pcDNA™5/FRT/TO vector containing a Strep-Flag tag (provided by Karsten Boldt, medical proteome center, Tübingen) by Gateway® cloning. Each cloning step was validated by an enzymatic digestion and sanger sequencing. Finally, the correct integration of the *SPAST* isoform attached to a N-terminal Strep-Flag tag was validated by sanger sequencing and is exemplary shown for the M87 isoform (figure 14).

In the following steps, the Strep-Flag tag was used to perform an affinity purification and to specifically detect the induced protein expression by western blotting.

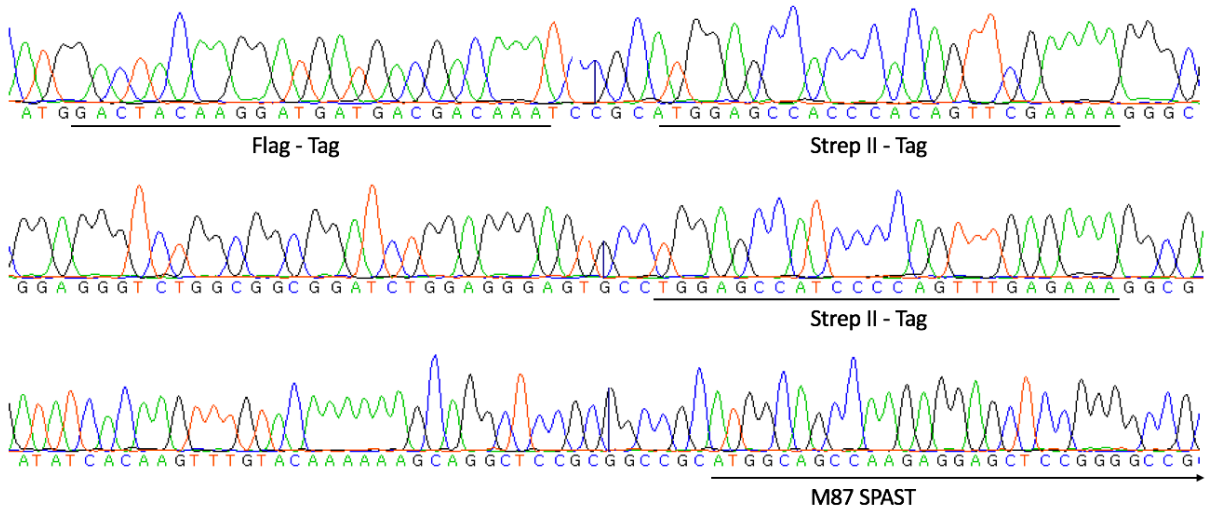


Figure 14: Validation of the cloning result exemplary shown for the M87 spastin isoform after cloning into the pcDNATM5/FRT/TO vector

Sanger sequencing showing the correct integration of the M87 SPAST isoform with an upstream, N-terminal Strep-Flag tag consisting of two Strep II tags (WSHPQFEK) and a Flag tag (DYKDDDDK). The tag sequences and the first nucleotides of the M87 spastin isoform are underlined and labeled in black.

3.1.4 **Induction and analysis of the spastin expression in SH-SY5Y overexpression cell lines**

To create Flp-In™ T-Rex™ expression cell lines, *SPAST* - pcDNA™5/FRT/TO vectors and a pOG44 plasmid, coding for a recombinase, were co-transfected

into a Flp-In™ T-Rex™ SH-SY5Y host cell line (medical proteome center, Tübingen).

After antibiotic selection, the Flp-In™ T-Rex™ SH-SY5Y clones were treated with doxycycline to activate the tetracycline dependent $TetO_2$-promotor and thereby induce the overexpression of the spastin isoforms. The tetracycline dependent expression of spastin was validated by western blotting (exemplary shown for the M87 isoform in figure 15, B).

During all western blot experiments, as control for the sufficient transfer of the proteins from an acrylamide gel to the nitrocellulose membrane, a protein staining with Ponceau S was performed. This staining is exemplary shown for the M87 induction western blot (figure 15, A).

Surprisingly, the overexpression of the M1 spastin isoforms (with and without exon 4 deletion) resulted in a very low expression of the protein after induction with doxycycline, with a protein amount, that was not sufficient for a definite detection on protein level through western blotting. After an extremely long exposure time (~300s), a very weak signal could be detected.

The isoform specific expression of all four spastin isoforms is shown in a western blot (figure 15, C).

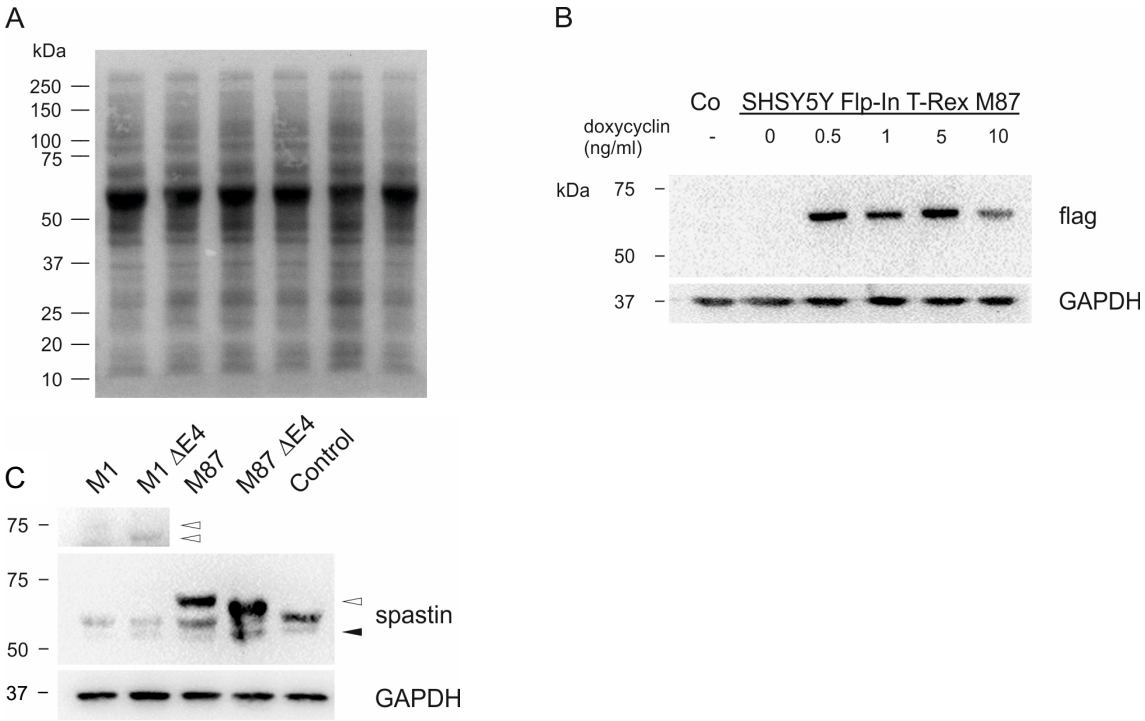


Figure 15: Western blot analysis of the tetracycline-dependent expression of spastin in a Flp-In™ T-Rex™ expression cell line
A: Ponceau staining of the nitrocellulose membrane after the blotting procedure, showing the sufficient protein transfer. B: Western blot analysis of the Flag-M87-spastin expression after treatment with different doses of doxycycline. The tetracycline-dependent expression of Flag-spastin is shown. The non-transfected Flp-In™ T-Rex™ host cell line served as negative control (not shown). C: Western Blot analysis of the expression of the four spastin isoforms in induced Flp-In™ T-Rex™ overexpression cell lines. For the induction, the cells were treated with 5µg/ml doxycycline for 48h. The endogenous spastin expression is marked by a black, the induced spastin expression by a white arrowhead. To detect M1-spastin a long exposure time of 300s was required. Primary antibodies: mouse Anti-Flag (Sigma-Aldrich) 1:10,000; mouse anti-spastin (Santa Cruz) 1:5,000; mouse anti-GAPDH (Meridian) 1:10,000; secondary antibody: HRP coupled anti-mouse (Dianova) 1:10,000; protein marker: Precision Plus Protein™ Dual Color standard

3.1.5 **Investigation of the weak expression of M1 spastin**

As shown in the previous section, the M1 isoforms of spastin showed a surprisingly low expression in the Flp-In™ T-Rex™ overexpression cell model. As a possible cause, degradation of M1 spastin on protein level was considered. To test this hypothesis and with the aim of reaching higher protein concentrations for later experiments, an additional experiment was performed, in which the expression of M1 spastin was induced in the Flp-In™ T-Rex™ overexpression cell line, while inhibiting the protein degradation with different concentrations of the potent proteasome inhibitor MG132. As a result, a significantly increased amount of M1 spastin was detected, while higher concentrations and a longer exposure to the proteasome inhibitor resulted in increased cell death.
Fortunately, all mass spectrometric experiments were sensitive enough to detect M1 spastin without the addition of MG132, so that the proteasome inhibitor was not used in the following interactome analysis.

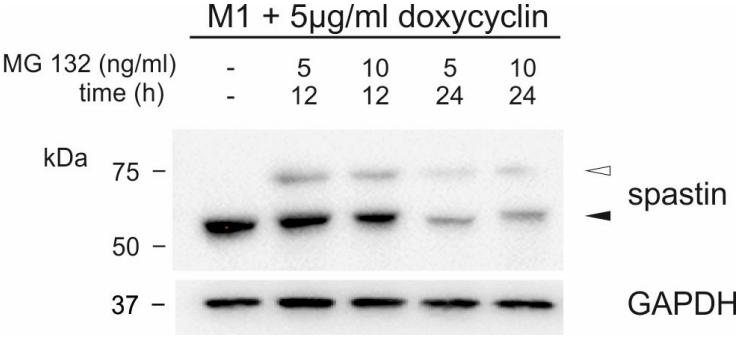


Figure 16: Western Blot analysis of the expression of the M1 spastin isoform after the addition of different doses of MG132

The protein expression was induced with 5µg/ml doxycycline for either 12 or 24 hours. Additionally, the cells were treated with different concentrations the proteasome inhibitor MG132. The black arrow marks the endogenously expressed spastin (M1 + M87 isoform), the white arrow the induced M1 spastin. A significantly increased expression of M1 spastin is shown, detectable by western blotting. Primary antibodies: mouse anti-spastin (Santa Cruz) 1:5,000; mouse anti-GAPDH (Meridian) 1:10,000; secondary antibody: HRP coupled anti-mouse (Dianova) 1:10,000

3.1.6 **SH-SY5Y as a neuron-like cell model**

One reason to choose SH-SY5Y cells as model for neurological diseases is their ability to differentiate in a more neuron-like phenotype in the presence of retinoic acid (RA). A differentiation protocol was established in our lab, according to Kovalevich and Langford (2013). Cells were treated with retinoic acid (RA) for 6 days followed by exposure to brain derived neurotrophic factor (BDNF) for another 6 days while undergoing a serum starvation.

To demonstrate the neuronal phenotype, immunofluorescence staining of neuronal factors such as ßIII-tubulin and neuron specific enolase (NSE) was performed in different stages of the differentiation. As demonstrated in figure 17 the cells undergo a phenotypic change as they build dendrite like structures. Additionally, the expression of neuronal markers increases with the increasing degree of differentiation.

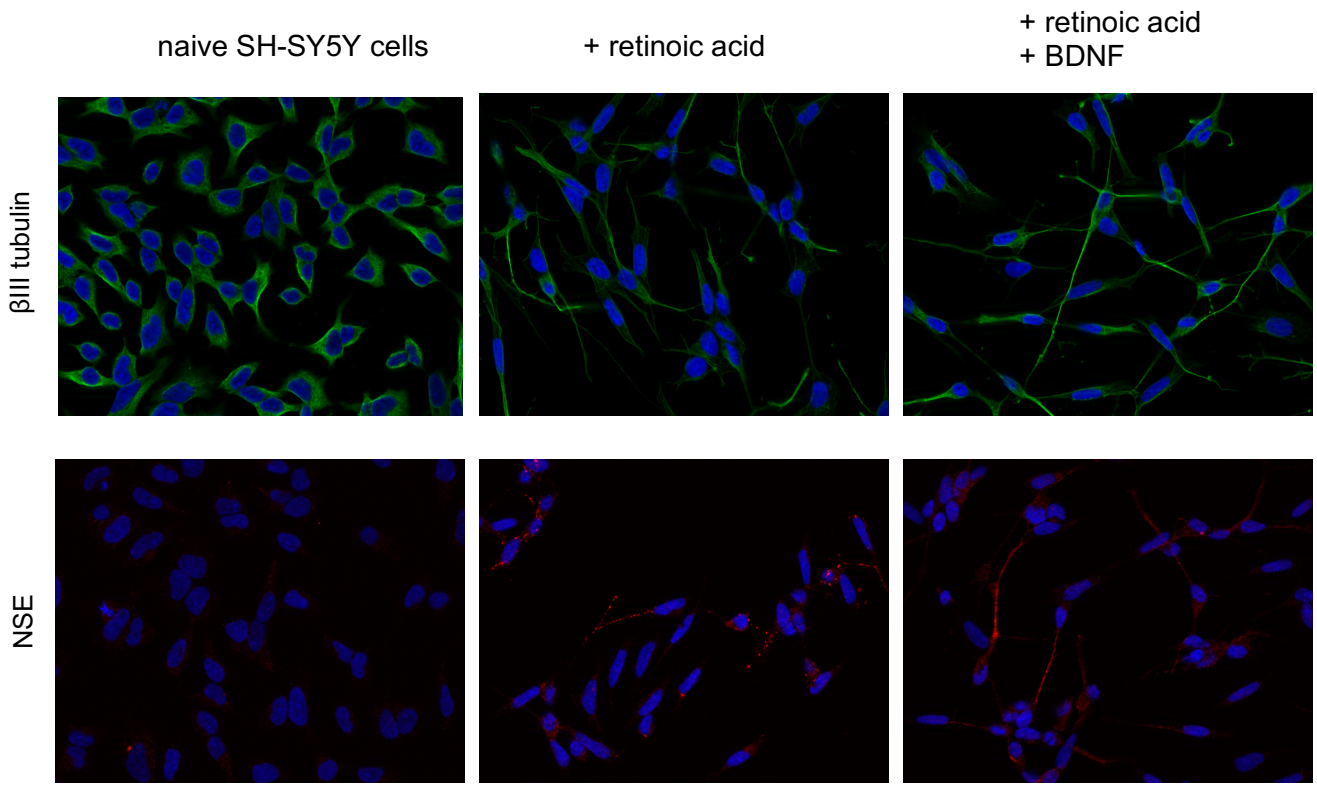


Figure 17: Immunofluorescence staining of SH-SY5Y cells before, during and after undergoing a differentiation protocol with retinoic acid and brain derived neurotrophic factor

Primary antibodies: anti-β III-tubulin (1:1,000, mouse, anti-NSE (1:250) rabbit; secondary antibodies: Alexa 488 anti-mouse (1:1,000), Alexa 568 anti-rabbit (1:1,000); nuclear staining with DAPI; 40X magnification with oil. NSE: neuron specific enolase; BDNF: brain derived neurotrophic factor

3.2 Detection of spastin-protein interactions by mass spectrometry

The overexpressed, tagged spastin isoforms were used to perform an isoform specific mass spectrometry-based interactome analysis.

To this end, the Strep-Flag-spastin overexpression was induced and in a first step, the tagged protein in complex with bound interaction partners was isolated through a Flag affinity purification (as described in Gloeckner et al. (2009a), described in detail in 2.8.1). The proteins contained in the isolated sample were then identified through a mass spectrometric analysis.

Five biological replicates of each isoform were examined, non-induced Flp-In™ T-Rex™ SH-SY5Y expression cell lines served as a negative control. The mass spectrometric raw data was analyzed and the potential hits were filtered (shown in figure 10). Then, possible spastin interactors were tested for their statistical significance. Candidate proteins were then clustered into functional groups, which finally allowed the formation of isoform specific protein-protein interaction networks.

As the Flp-In™ T-Rex™ SH-SY5Y expression clone of the M87 del exon 4 spastin isoform was unfortunately discovered to be contaminated with another cell line during the time of the data analysis, the data of the isoforms lacking exon 4 was left out in the further steps. Additionally, as the results of the M1 spastin isoform lacking exon 4 did not show any relevant differences to the full-length M1 isoform, the following analysis focusses on both full-length isoforms, M1 and M87 spastin.

3.2.1 **Analysis of the mass spectrometricinteractome data**

To perform a mass spectrometry-based interactome analysis, the purified Strep-Flag-spastin complexes underwent an in-solution trypsin digest. The resulting peptides were then detected and identified by their molecular weight to mass ratio (detailed description in 2.8). After the qualitative (Mascot, Scaffold) and quantitative (MaxQuant) analysis of each sample, the resulting list of detected proteins was further filtered, and each hit tested for its statistical significance. To this end the following workflow (figure 18) was established using the Perseus software (Version 1.6.1.3).

Two statistical tests were performed in this analysis: A two sample test was used to calculate the significance of the difference between the mean log2 transformed values of the doxycycline induced versus the non-induced sample. In another approach the significancy was tested through a significance B score for log protein ratios induced/non-induced depending on the signal intensity (Cox and Mann (2008)). The two statistical tests are explained in detail in the Material and

Methods section (2.8.3). Proteins showing significant values ($p<0.05$) in both statistical tests were then further investigated by performing a functional annotation.

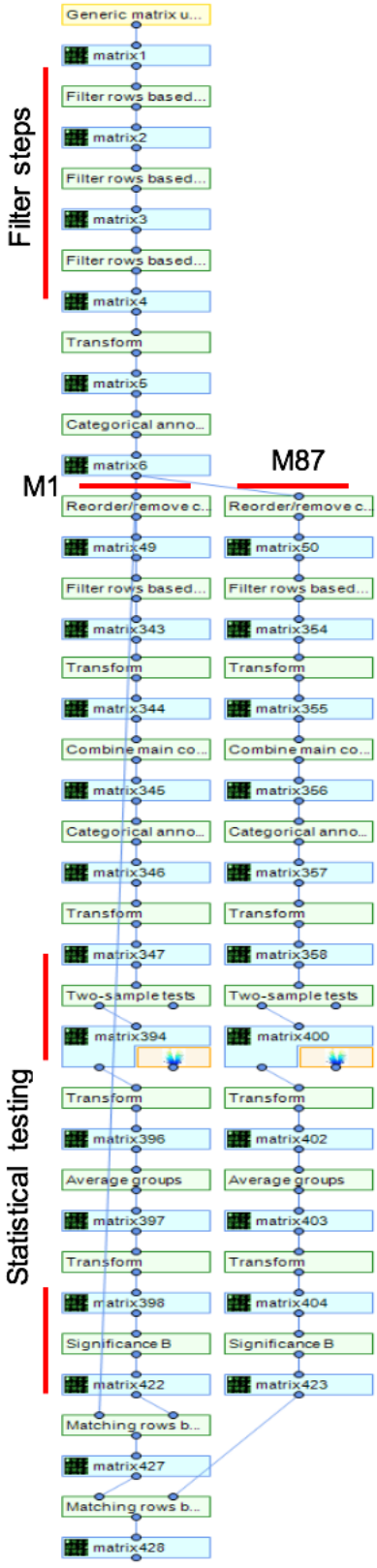


Figure 18: Perseus workflow of the statistical analysis of the mass spectrometry interactome results

The identified proteins were filtered for potential contaminants, proteins that were at least identified in four of the five repetitions, proteins that were only identified by site and reverse recognitions. Statistical testing (two-sample test and significance B) was separately performed for each isoform.

In the end the matrices of both isoforms were reunited in one final matrix.

Perseus software, Version 1.6.1.3

Selected columns of the final matrix (matrix 428 of the Perseus workflow) resulting from the previous workflow demonstrated in figure 18 are shown in the following table.

The listed proteins were tested significant ($p<0.05$) in both performed significance tests. The number of unique peptides identified and the resulting coverage of the unique protein sequence, as well as the significance levels of the performed significance tests are included.

Table 27: Extract of the final matrix showing all proteins tested significant ($p<0.05$) in the performed two sample test, as well in the Significance B test for the M1 isoform (including exon 4)

Gene names	Unique peptides	Unique sequence coverage [%]	-Log Student's T-test p-value LFQ intensity M1-Doxy_LFQ intensity M1-Co	LFQ intensity M1-Doxy_x/y_LFQ intensity M1-Co Significance B
SPAST	90	869	6,31308	1,06E-206
NUP85	27	471	4,61984	6,93E-140
NUP43	14	495	6,34497	2,12E-112
HM13	8	236	1,58943	3,71E-73
NUP160	42	393	3,85072	3,15E-60
NUP37	11	408	4,21251	1,13E-40
NUP107	29	45	1,98925	8,30E-25
NUP133	31	38	2,19166	7,61E-20
IST1	14	533	1,96702	1,49E-19
SLC25A6	5	128	1,36184	3,75E-17
NUP98	41	336	1,59491	9,66E-16
ATP5O	10	601	1,43117	2,67E-05
ATP5B	26	673	1,47818	3,48E-05
ATP5C1	13	383	1,4105	8,40E-05
ATP5A1	36	617	1,38319	0,000220264
TGIF2LX	6	349	1,98216	0,00102021
GNAI3	5	22	1,3531	0,00214705
MPHOSPH10	11	248	1,34384	0,00242638
ICT1	4	238	1,37286	0,00735574
PICALM	4	121	1,92985	0,0105455
MRPL43	9	53	1,76339	0,0115177
CKAP4	29	608	1,61763	0,0162243
PTBP2	11	365	1,65279	0,0213292

LRRFIP2	8	171	1,34629	0,0238565
PSAT1	9	314	1,57936	0,039708
THSD7A	12	124	1,4217	0,0412066
MRPS28	6	38	1,53121	0,0444897
MRPL27	6	432	1,57635	0,0448211
MRPL2	5	275	1,49525	0,0453638

Table 28: Extract of the final matrix showing all proteins tested significant (p<0.05) in the performed two sample test, as well in the Significance B test for the M87 isoform (including exon 4)

Gene names	Unique peptides	Unique sequence coverage [%]	-Log Student's T-test p-value LFQ intensity M87-Doxy_LFQ intensity M87-Co	LFQ intensity M87-Doxy_x/y_LFQ intensity M87-Co Significance B
SPAST	91	869	5,53576	3,50E-26
IST1	14	533	3,05516	9,38E-23
CHMP1A	6	224	4,42422	1,52E-10
CHMP4B	6	304	1,95791	1,69E-07
CHMP2A	5	189	3,65565	3,05E-07
MPHOSPH10	11	248	2,33436	3,48E-06
STAU1	30	605	3,43238	1,78E-05
HSPA6	14	107	1,85523	2,36E-05
ZNF629	16	299	2,23982	3,75E-05
RPL13A	17	552	2,85684	6,29E-05
RPL37A	7	598	2,09571	7,58E-05
RPL15	21	618	2,45676	9,94E-05
RPL10	17	547	2,07126	0,00017128
PRKRA	18	642	2,09938	0,000198118
RPL8	23	673	2,87696	0,000199851
RPL3	30	536	2,59783	0,000231614
RPL32	13	556	1,95503	0,000239668
RPL18A	14	597	2,93536	0,000244038
BMS1	43	427	1,66819	0,000339619
GTPBP4	24	446	1,79173	0,000346226
DDX50	22	358	2,41285	0,000435237
SPATA5	25	439	1,75712	0,000548325
RPL21	15	569	2,40783	0,000595
GNL3	20	461	1,61346	0,000630953
RPL31	10	544	2,43975	0,000667348
DDX18	22	455	1,49112	0,00074779

RPL26	17	614	2,68078	0,000879622
RPL18	11	484	2,32802	0,000892111
RPL17	15	614	2,59124	0,000894321
GRWD1	17	536	1,31255	0,000907535
C7orf50	12	521	3,08533	0,00104057
RFC1	18	244	3,87374	0,00111024
RPL30	12	80	2,14936	0,00132473
SRPK2	18	265	2,08421	0,00148322
RPL12	12	733	1,77638	0,0017135
RPL6	27	601	3,03884	0,00173286
RPL10A	17	594	1,98687	0,00176393
POLRMT	20	24	2,02042	0,0020945
EXOSC6	7	29	1,62794	0,00222745
RBM28	27	354	1,38684	0,00228534
RPL36	7	362	1,83322	0,00233073
NAT10	21	245	1,38252	0,00242295
RSBN1L	17	358	1,47422	0,00290275
CHMP5	6	42	2,14342	0,00310099
RPL27	15	713	2,11966	0,00337122
RPL7	28	665	1,92818	0,00372695
RPL19	16	50	2,55134	0,00387419
RPL5	30	70	2,36	0,00390899
PWP1	10	319	1,529	0,00411184
HP1BP3	16	26	1,57829	0,00438295
RCL1	10	316	2,09454	0,00444335
RPL13	22	611	2,66397	0,00451241
TRMT1L	26	591	1,81364	0,00458354
RPL4	34	571	2,13129	0,00468539
RPL7A	24	50	1,45531	0,00512268
EPB41L5	20	462	2,54364	0,00556815
RPL14	10	40	1,40807	0,00612962
NO66	14	359	1,43172	0,00618637
DDX56	12	303	1,48363	0,00621671
ZNF512	18	487	1,81499	0,00667868
EXOSC3	7	451	4,13132	0,00740053
EXOSC9	10	301	3,0803	0,00758961
KNOP1	11	358	2,11758	0,00776049
NOC3L	8	139	1,59124	0,00830227
EXOSC2	10	396	2,804	0,00924714
DDX27	12	175	1,9315	0,00944566
SRPK1	19	357	2,25374	0,0106434

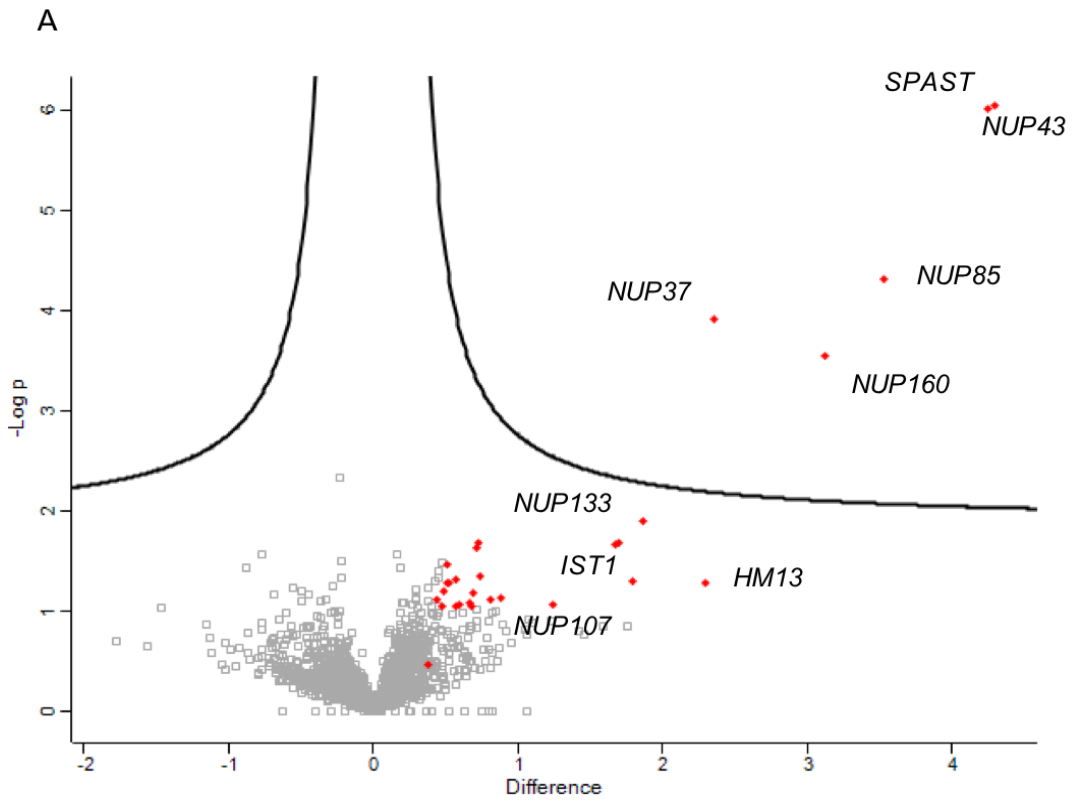
RPL29	3	208	3,00492	0,0117764
RPL36A	10	425	2,37227	0,0119509
NIFK	11	461	2,45955	0,0126128
RPL9	17	719	1,79057	0,0144036
SPATS2L	9	233	1,64652	0,014413
SRP14	11	507	2,54998	0,0148716
NOL10	10	233	1,91011	0,0149105
RPL23A	16	551	1,66215	0,0160973
EXOSC5	8	557	1,67508	0,0167731
MRTO4	11	515	1,43625	0,0169153
GTF3C3	12	198	1,83413	0,0170359
ZNF768	6	117	2,32035	0,0171453
RPL26L1	14	117	2,60575	0,018477
RPL22	8	641	2,34186	0,0185675
TFB1M	7	231	1,46417	0,0188106
EXOSC4	7	38	2,06687	0,0189883
ZNF316	16	256	2,02126	0,0191133
ISG20L2	6	241	1,35092	0,0194607
SRP9	6	616	1,76666	0,0194997
EXOSC1	8	523	2,34163	0,020287
EXOSC7	10	577	1,42394	0,0215725
TOP2B	19	175	1,37088	0,0221593
TSPYL5	13	379	2,26449	0,0235283
ZCCHC6	16	152	1,6608	0,0261301
GPATCH4	9	258	2,78736	0,0271471
RPL24	11	446	2,58466	0,0301678
HSPA5	49	599	2,02232	0,0306356
GTF3C5	20	484	1,67333	0,0321941
WDR36	15	24	1,30658	0,0323938
RPF2	5	258	2,34685	0,034781
RPL28	17	664	2,04176	0,0348355
GTF3C4	26	46	1,67927	0,0369325
ZNF787	7	258	2,7291	0,0381441
MYEF2	17	378	1,68071	0,0395757
TOP1	30	409	1,84138	0,0424373
RPL11	16	68	1,5531	0,0448072
PATZ1	7	14	1,90683	0,0477127
H1F0	7	284	1,38017	0,0486711
NEMF	14	175	1,48725	0,0491373

The results of the two-sample test are illustrated in form of volcano blots (figure 19) displaying the difference of the protein detection intensity in the doxycycline induced samples versus the non-induced negative control against the -log p-value. The black line marks the significance level with a permutation-based FDR <0,05 and s=0.01. Therefore, all proteins above the line are tested significant after correction for multiplicity testing.

The results show, that as result of the induced overexpression, spastin was significantly enriched in both samples. Known interaction partners as *IST1* or the proteins of the *ESCRTIII* complex (*CHMP1A, CHMP2A, CHMP4, CHMP5)* appear for both isoforms.

Noticeably, multiple proteins of the nuclear pore complex (*NUP 37, 43, 85, 160*) were detected in the M1 spastin sample.

The analysis of the M87 spastin isoform revealed a higher overall number of hits.



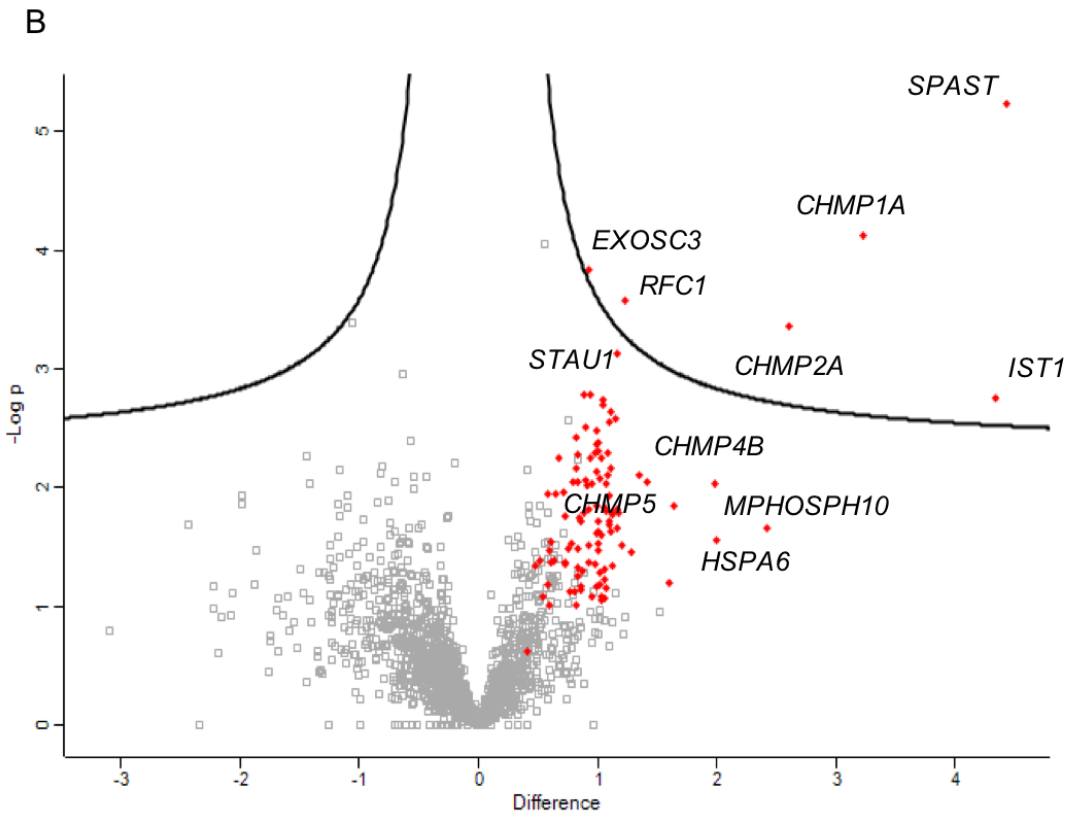


Figure 19: Volcano Blot of the LC/MS results of the M1 (A) and M87 (B) Isoform of spastin

The black line marks the permutation-based FDR = 0,05, s0 = 0,1 significance level, all proteins above the line reached significant values in a two-sample test after correction for multiplicity testing. Proteins, that reached a p-value < 0,05 in a two-sample test and an additional Significance B test are marked in red (Perseus Version 1.6.1.3).

3.2.2 **Functional annotation of significant interaction candidates**

Spastin interaction candidates that reached a p-value <0.05 in the performed two-sample test as well as in the significance B test were further investigated and organized into groups by performing a functional annotation.

The DAVID open source platform was used to divide potential spastin interactors into functional groups, formed after classes defined by the gene ontology consortium, the GO terms. All proteins were categorized after their localization in cellular compartments (CC), their molecular function (MF) and their involvement in biological processes (BP).

	Category	Protein count	p-value	Benjamini
GO-Term_CC	nuclear pore outer ring	5	1,70E-10	8,70E-09
	kinetochore	5	1,80E-06	3,10E-05
	mitochondrial proton-transporting ATP synthase complex, catalytic core F(1)	3	1,50E-05	1,90E-04
	mitochondrial large ribosomal subunit	3	6,60E-04	6,70E-03
	myelin sheath	4	1,20E-03	1,00E-02
	nuclear membrane	3	1,40E-02	8,90E-02
	extracellular exosome	9	1,80E-02	1,00E-01
	nuclear periphery	2	2,50E-02	1,20E-01
	centrosome	3	8,30E-02	3,30E-01
GO-Term_MF	nucleocytoplasmic transporter activity	3	3,50E-04	4,30E-03
	poly(A) RNA binding	8	3,60E-04	3,00E-03
	proton-transporting ATPase activity, rotational mechanism	3	5,30E-04	3,30E-03
	structural constituent of nuclear pore	2	2,80E-02	1,30E-01
	structural constituent of ribosome	3	4,70E-02	1,80E-01
GO-Term_BP	ATP synthesis coupled proton transport	3	4,00E-04	2,90E-02
	nuclear pore complex assembly	2	1,20E-02	2,60E-01
	mRNA export from nucleus	2	4,00E-02	5,30E-01
	ATP hydrolysis coupled proton transport	2	4,40E-02	4,90E-01
	protein import into nucleus	2	5,50E-02	5,10E-01
	regulation of transcription, DNA-templated	4	6,10E-02	4,90E-01

Figure 20: Functional analysis of statistically significant interaction candidates of M1 spastin (DAVID platform, version 6.8)

GO: gene ontology; CC: cellular compartment; MF: molecular function; BP: biological pathway

	Category	Protein count	p-value	Benjamini
GO-Term_CC	ribosome	35	3,40E-39	1,30E-37
	intracellular non-membrane-bounded organelle	71	4,90E-36	1,30E-34
	ribonucleoprotein complex	41	5,70E-34	1,00E-32
	cytosolic part	25	2,40E-27	3,30E-26
	nucleolus	33	1,20E-19	1,50E-18
	nuclear lumen	39	2,90E-15	2,90E-14
	intracellular organelle lumen	42	1,10E-14	9,90E-14
	exosome (RNase complex)	8	2,90E-13	2,10E-12
	signal recognition particle, endoplasmic reticulum targeting	2	3,80E-02	2,10E-01
GO-Term_MF	RNA binding	44	9,00E-32	5,80E-30
	structural molecule activity	35	6,00E-23	2,60E-21
	3'-5'-exoribonuclease activity	7	5,90E-11	1,90E-09
	ribonuclease activity	7	4,40E-06	6,30E-05
	nuclease activity	9	8,80E-06	1,10E-04
	rRNA binding	5	3,60E-05	4,20E-04
	RNA polymerase III transcription factor activity	3	1,80E-03	1,80E-02
	DNA topoisomerase (ATP-hydrolyzing) activity	2	2,60E-02	2,20E-01
	ATP-dependent helicase activity	4	2,60E-02	2,10E-01
	purine NTP-dependent helicase activity	4	2,60E-02	2,10E-01
	ATPase activity, coupled	6	3,30E-02	2,40E-01
GO-Term_BP	translational elongation	35	1,90E-51	6,30E-49
	ribosome biogenesis	24	5,60E-28	6,10E-26
	ribonucleoprotein complex biogenesis	25	2,20E-25	1,80E-23
	rRNA processing	17	5,10E-19	3,30E-17
	ncRNA metabolic process	17	1,40E-12	5,10E-11
	tRNA transcription	3	8,60E-04	2,30E-02
	rRNA transcription	3	6,00E-03	1,40E-01
	negative regulation of translational elongation	2	1,30E-02	2,50E-01
	SRP-dependent cotranslational protein targeting to membrane	2	3,80E-02	5,50E-01
	RNA biosynthetic process	6	4,40E-02	5,40E-01
	DNA topological change	2	4,50E-02	5,30E-01
	cellular macromolecular complex assembly	6	5,60E-02	5,90E-01
	protein targeting to ER	2	5,70E-02	5,80E-01
	cotranslational protein targeting to membrane	2	7,50E-02	6,40E-01

Figure 21: Functional analysis of statistically significant interaction candidates of M87 spastin (DAVID platform, version 6.8)

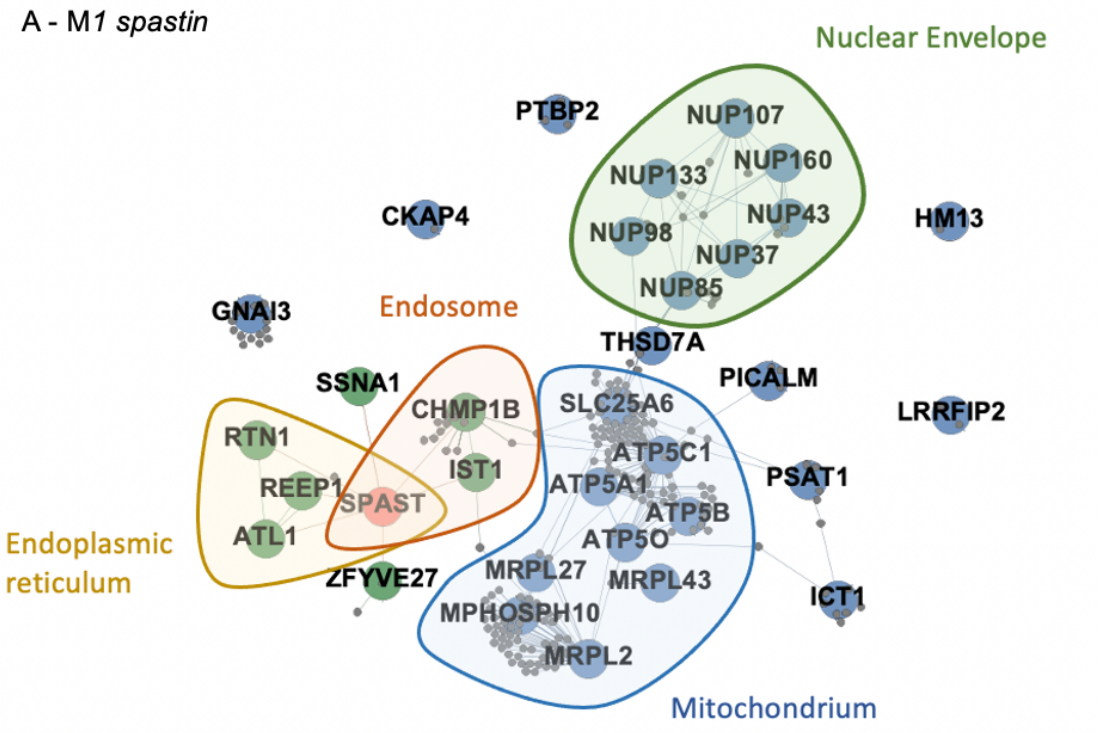
GO: gene ontology; CC: cellular compartment; MF: molecular function; BP: biological pathway

The results of the functional annotation were then used to create isoform specific protein-protein interaction networks (figure 22). Proteins that could be assigned to a common cellular pathway, cellular localization or protein complex are clustered by colored frames.

Noticeably, additionally to the previously known location of M1 spastin at the ER and its interaction with endosomal proteins (CHMP1B, IST1), multiple nucleoporins (NUP) showed up in the analysis. As another finding, several proteins of the mitochondrial membrane, such as SLC25A6, ATPC1 or ATP5A appeared in the analysis.

Overall, a much higher number of identified proteins resulted from the investigation of the M87 spastin isoform. Besides the known endosomal interaction partners of the protein, several ribosomal proteins can be noted.

These hits, considered as possible novel spastin interaction partners, are further examined and their relevance discussed in the following chapters.



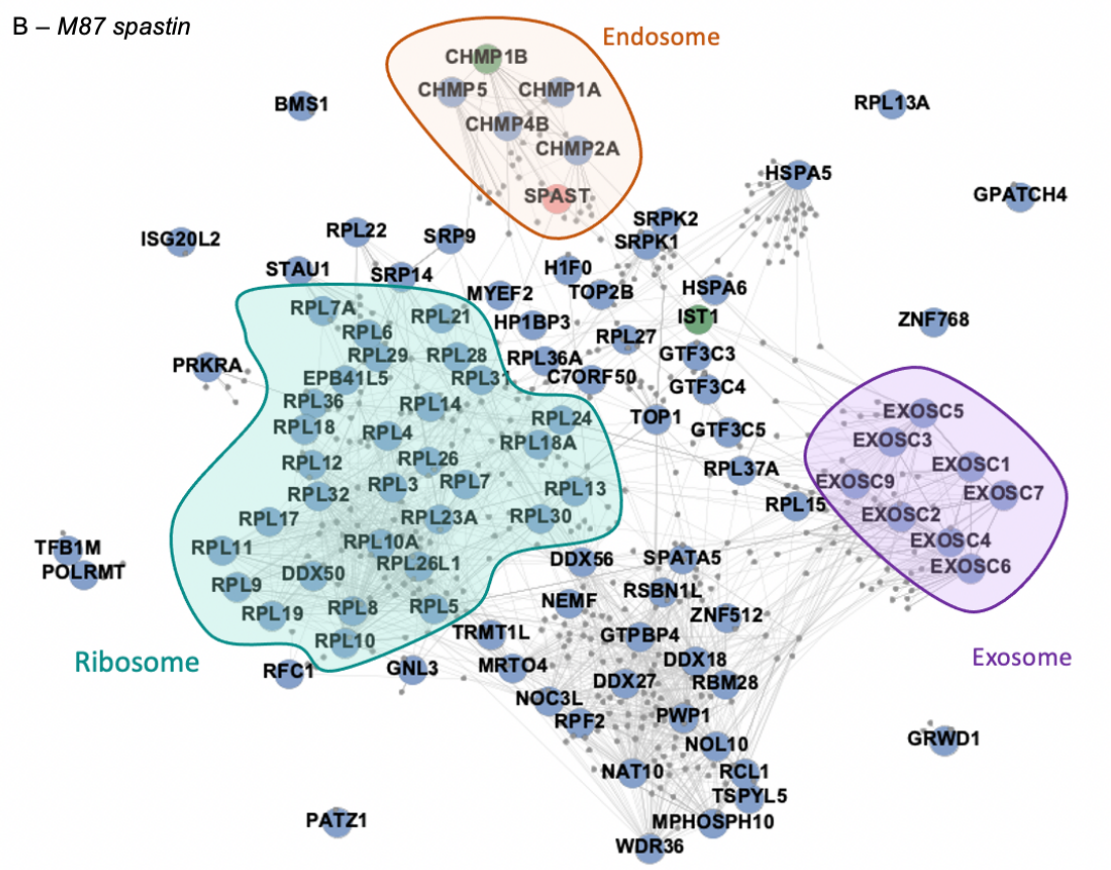


Figure 22: Isoform specific protein-protein network of the M1 (A) and M87 (B) isoform of spastin

each node represents one protein, interacting proteins are connected by a grey line; spastin is marked in red, known interaction partners are represented in green (not all of them were identified in the performed interactome analysis; ATL1, REEP1, RTN1, SSNA1 and ZFYVE27 were added as known interaction partners), proteins that appear as novel spastin interacting partners in the performed interactome analysis are colored blue; functional groups were selected after common GO-terms (figure 20/21) and are marked by a colored frame (Gephi Version 0.9.1.).

3.3 Validation of interaction partners

As a next step, the spastin interactions identified by mass spectrometry were further validated by an independent method. Therefore, the protein-protein interaction was confirmed via a co-immunoprecipitation. To this end, spastin was

isolated by an affinity purification and both, spastin and its considered interaction partner were detected in the eluate by western blotting.

3.3.1 Selection of promising spastin interaction candidates for further validation

After the analysis of the interactome data, various proteins emerged as spastin interaction candidates. To validate those interactions, promising hits where chosen for further validation.

When looking of the functional clusters represented in the M1 spastin interactome analysis (3.2.2), two functional groupings stood out:

Firstly, various proteins of the nuclear pore complex (NUP 37, 43, 85, 98, 107, 133, 160) emerged as possible M1 spastin interactors. As those proteins are known to form the subcomplex NUP107-160 of the nuclear pore complex (Walther et al., 2003) and an interaction of spastin with proteins of the nuclear envelope hasn't been postulated to far, NUP43 was selected as representative for this group for further validation.

Secondly, an entire cluster of mitochondrial proteins (e.g. ATP5A, SLC25A6, MRPL34) appeared in the analysis. As the impairment of mitochondrial processes was previously shown to be impaired in some types of HSPs (e.g. SPG 7, 13; (Blackstone, 2012, Solowska and Baas, 2015)) and there have recently been discussions on the role of Mitochondria-ER-contact sites and its relevance in the pathogenesis of HSPs (Paillusson et al., 2016, Phillips and Voeltz, 2016), ATP5A was chosen as a representative of this cluster for further validation.

3.3.2 Co-Immunoprecipitation of spastin and identified spastin interaction candidates

To validate spastin-protein interactions identified in the mass spectrometry-based interactome analysis co-immunoprecipitation experiments were performed. To this end, spastin complexes were isolated by a flag affinity purification (as described in 3.3.1). Spastin and its considered novel interaction partners were then detected in the eluate by a western blot. The expression of the investigated

proteins in SH-SY5Y cells was confirmed in the input sample. As negative control, a regular SH-SY5Y cell line was used.

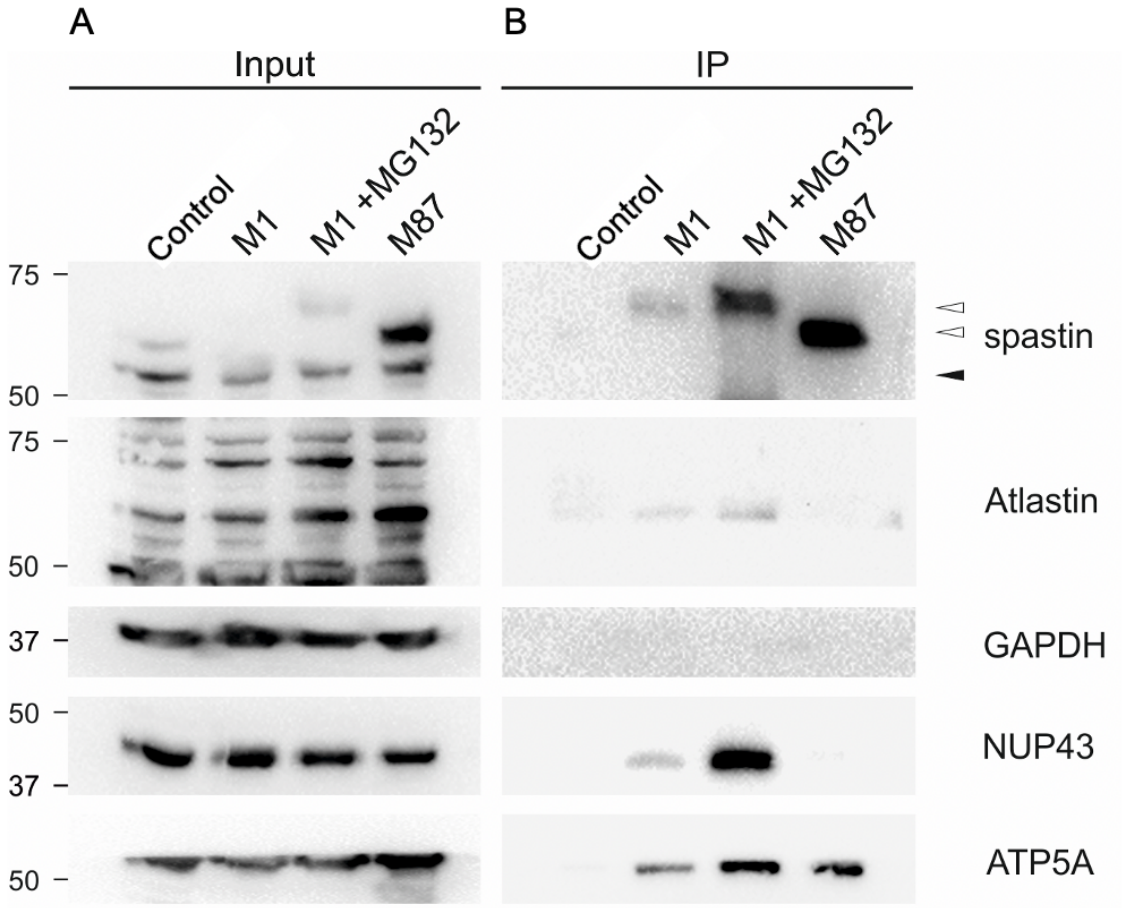


Figure 23: Western blot analysis of the performed co-immunoprecipitations
A: western blot showing the expression of all investigated proteins in the whole cell lysate (input) of Flp-In™ T-Rex™ SH-SY5Y expression cells; B: protein expression in the co-immunoprecipitated (co-IP) sample. The black arrowhead marks the endogenously expressed spastin in the input sample. The white arrowheads indicate the overexpressed, tagged isoforms of spastin.
Primary antibodies: mouse anti-spastin (Santa Cruz) 1:5,000; mouse anti-GAPDH (Meridian) 1:10,000; secondary antibody: HRP coupled anti-mouse (Dianova) 1:10,000, HRP coupled anti-rabbit (Dianova) 1:10,000; protein marker: Precision Plus Protein™ Dual Color standard

4. Discussion

Mutations in the spastic paraplegia gene 4 (SPG4), *SPAST*, are by far the most common cause for autosomal dominant inherited HSP (Bürger et al., 2000, Solowska and Baas, 2015). The gene encodes the protein spastin, an AAA ATPase with the main known function of severing microtubules (Lumb et al., 2012, Evans et al., 2005). Additionally, functions in endosomal trafficking and ER morphogenesis, and thereby membrane modelling were previously described (Park et al., 2010, Allison et al., 2013, Blackstone et al., 2011). The majority of prior investigations on the disease-causing pathomechanism in SPG4 implicate a loss of function of the truncated protein or a toxicity of the accumulating M1 protein isoform (Solowska et al., 2010, Solowska et al., 2017, Qiang et al., 2019). Interestingly, it was shown that not all disease-causing spastin mutations can be explained by these theories, so that the overall pathomechanism of SPG4 still remains unsolved.

In this work, possible novel isoform specific spastin functions were revealed through the investigation of protein interactions by a mass spectrometry-based interactome analysis. Therefore, isoform specific Flp-In™ T-Rex™ SH-SY5Y overexpression cell lines were generated.

Identified possible spastin interactors were subsequently further validated through co-immunoprecipitation studies.

4.1 Flp-In™ T-Rex™ SH-SY5Y expression cell lines as neuron-like cell model for SPG4

A great challenge in the investigation of diseases affecting the human nervous system is the establishment of a compatible model to simulate endogenous processes. The upper motor neurons, which are the cells mainly affected in HSPs (Solowska and Baas, 2015), can't be obtained and propagated further in cell cultures. Therefore, alternative cell lines are used in scientific experiments to imitate the human nervous system.

One option, that has been established in the research of a variety of neurological diseases such as Parkinson's disease (Xicoy et al., 2017) is the use of the

mammalian SH-SY5Y cell line, derived from a metastasis of a human neuroblastoma (Kovalevich and Langford, 2013).

In this work, the Flp-In™ T-Rex™ system was used to create an overexpression SH-SY5Y cell model for each of the four endogenously expressed isoforms of the SPG4 protein spastin (Spitzer et al., 2013). This neuron-like cell model allows the stable, inducible overexpression of a tagged protein without ethical concerns. As the cells are fairly easy to maintain it allows the establishment of a reliable, reproducible system, enabling a broad investigation on cellular protein interactions.

On the downside, protein-protein interaction dynamics are very vulnerable and are easily affected by changes of environmental conditions, such as protein expression levels or the composition of proteins in the cell (Xing et al., 2016). The overexpression of a protein could thereby trigger unspecific bindings or lead through the affection of post-translational modifications and an impaired protein lifetime to an atypical protein behavior for example in terms of cellular localization and protein-protein interaction dynamics.

To summarize, SH-SY5Y cells serve in this work as a mammalian, neuron-like, easy to obtain cell model and allow a first investigation of the isoform specific spastin interactome. Nevertheless, the possible influence on spastin interactions resulting from the spastin overexpression, as well as the use of a simplified cell model has to be taken into account in the further evaluation of identified spastin interaction partners.

4.2 Weak expression of M1 spastin in the overexpression cell model

Under endogenous conditions, the isoforms of spastin were previously shown to differ not only in their distribution and traceability in different human tissues (Solowska et al., 2008), but also in their expression levels in neuronal cells. The longer M1 isoform is known to be expressed at much lower levels than the M87 isoform. A poor Kozak sequence (Kozak, 2002), leading to a leaky transcription of this isoform and the N-terminal overlap of another gene were implicated as causes for this observation (Claudiani et al., 2005, Solowska et al., 2017).

It was also shown, that the expression of M1 spastin can be enhanced by an improvement of the Kozak sequence in HeLa, respectively iPS cells (Claudiani et al., 2005, Solowska et al., 2014).
Surprisingly, the expression levels of the isoforms of spastin distinctly differed in the overexpression cell model in this work (shown in 3.1.5). As the two isoforms were tagged at the N-terminus with the exact same sequence, the endogenous Kozak sequence was abolished and the transcription triggered by an identical doxycycline dependent promotor. Still, the M1 isoform of spastin, as well as the M1 isoform leaking exon 4, both showed a lower expression than both of the M87 isoforms.
As this finding cannot be explained with the theories postulated so far, we propose active degradation of the M1 isoform via the ubiquitin proteasome system as an additional mechanism to keep M1 levels low as high concentrations of M1 spastin were shown to have toxic effects on neuronal cells (Solowska et al., 2017). I accordance with this hypothesis we were able to increase the protein levels of M1 spastin by inhibition of the proteasomal protein degradation system.

4.3 Confirmation of known spastin interaction partners in the mass spectrometry-based interactome analysis

The identification of known spastin interaction partners played a crucial role in the validation of the experimental strategy used in this work.
As demonstrated in the isoform-specific interactome networks (shown in 3.2.2.) the results of the mass spectrometry-based interactome analysis did confirm the known interaction between both isoforms of spastin with IST1 and various components of the ESCRTIII-complex like CHMP1B. These interactions with the MIT domain of spastin have previously been shown to play an important role in membrane shaping processes (Allison et al., 2013).
Other known interaction partners like atlastin (SPG3), that is known to interact only with the M1 isoform of spastin, did not reach significant values, which could be a result of its low expression levels in SH-SY5Y cells or other limitations of this approach (as discussed in 4.6).

4.4 Differences in the identified protein interactions between the two spastin isoforms

As indicated in the protein-interaction networks (3.2.2), a much higher number of overall hits resulted from the mass-spectrometry-based interactome analysis of the M87 isoform than for the M1 isoform of spastin. One explanation for this finding is the major difference of expression levels of the isoforms, resulting in a higher protein input after immunoprecipitation and possibly leading to a higher number of false positives. Especially the high number of identified ribosomal proteins in the interactome analysis of the M87 isoform of spastin is highly suspicious for false positives. As both isoforms are known to have similar interaction partners involving the MIT and MTBD protein domains and cellular localization does not play a role after cell lysis, as performed in this experimental strategy, M87 specific protein interactions seem unlikely.

4.5 Possible interaction of M1 spastin with the NUP107-160 complex and its possible relevance

One striking finding in the interactome analysis of the different isoforms of spastin was the significant appearance of protein belonging to the Nucleoporin 107-160 complex. These interactions were exclusively identified in the M1 interactome (shown in 3.3.2).

The possible interaction of M1 spastin with this complex was exemplary confirmed by co-immunoprecipitation of M1 spastin but not M87 spastin for the protein NUP43 (shown in 3.4.1).

The NUP 107-160 complex, consisting of the proteins NUP 37, 43, 85, 96, 107 and 160, forms a subunit of the nuclear pore (Walther et al., 2003). During mitosis is was shown, that the complex locates to the kinetochores (Orjalo et al., 2006, Zuccolo et al., 2007). Depletion of the NUP 107-160 complex leads to a complete lack of nuclear pore complexes (NPC) and a delay in mitosis with reduced kinetochore tensions and a poor kinetochore-microtubule attachment. Those findings lead to the hypothesis, that the NUP 107-160 complex plays a major role

in the assembly of the nuclear pore complex and it is presumed to promote the spindle assembly during mitosis (Orjalo et al., 2006, Nakano et al., 2011, Zuccolo et al., 2007).

Considering the known functions of the NUP 107-160 complex and in regard of what we already know about the function of M1 spastin, we propose that there are at least two possible explanations for this interaction:

Firstly, M1 spastin is known to form a hydrophobic hairpin structure at the proteins N-terminus, that integrates in specific locations of the ER membrane (Fowler and O'Sullivan, 2016, Solowska and Baas, 2015). As the nuclear envelope is part of the ER and the two cellular structures are directly connected, a co-localization of the proteins at the nuclear pores is imaginable. Following this hypothesis M1 spastin could play a role, for example in the regulation of nuclear import and export processes or just in the formation of the nuclear pores.

A second hypothesis assumes an interaction between the two proteins during mitosis. While the NUP 107-160 complex is known to play a role at the attachment of microtubules at the kinetochores and thereby the assembly of the spindle structure, spastin has an AAA domain, used for the severing of microtubules. Furthermore, spastin and its known interaction partners of the ESCRT-III complex were previously assumed to play a role in the disassembly of the mitotic spindle formation and resemblance of the nuclear envelope in anaphase (Vietri et al., 2015). Hence, an interaction in this context is also to consider.

This leads to the question if an interaction of M1 spastin with the NUP 107-160 complex only takes place in mitotic cells, such as the SH-SY5Y cell model used in this work or if it also plays a role in post-mitotic neurons and thus may have relevance for the pathogenesis of SPG4.

The interactome-analysis, as well as the validation by co-immunoprecipitation shows, that the interaction with the NUP 107-160 complex is exclusive for the M1 isoform, but not the M87 isoform of spastin. This finding is highly suggestive for an interaction between that NUP proteins and a domain of spastin localized within the N-terminal 86 amino acids of M1 spastin. Regarding the functional impact of this interaction, one possible scenario could be the recruitment of M1 spastin, for

example to the spindle apparatus, possibly to play a role in processes requiring microtubule severing.
Nevertheless, the significance of this interaction in human in vivo neurons and in the pathogenesis of SPG4 remains uncertain and requires further investigations.

4.6 Possible interaction of M1 spastin with mitochondrial proteins and its possible relevance

Previous investigations on the pathomechanism of HSP showed, that an impaired mitochondrial function can lead to hereditary forms of spastic paraplegia, for example as in SPG7 or SPG13 (Blackstone, 2012, Solowska and Baas, 2015). An involvement of mitochondrial processes could so far be detected in various forms of hereditary spastic paraplegia, but not SPG4 (McDermott et al., 2003).
Interestingly, a number of proteins that appeared in the interactome analysis of M1 spastin are part of the F1 subunit of the mitochondrial ATP synthase. The proteins ATP5A1, ATP5O, ATP5C1 and ATP5B all showed significantly enriched values in the induced M1 spastin sample. These mitochondrial proteins did not reach significant levels in the interactome analysis of the M87 isoform, nevertheless, the results of the co-immunoprecipitation experiments did show an interaction between ATP5A and the M87 isoforms of spastin.
The ATP synthase located in the mitochondrial inner membrane uses the electrochemical gradient of protons between the mitochondrial matrix and the cytoplasm to produce ATP out of ADP. The complex consists of two major domains: the F0 complex, a proton channel integrated in the inner mitochondrial membrane and a F1 complex, containing the catalytic core. Both protein complexes are linked by a rotary mechanism (Junge and Nelson, 2015, Ruhle and Leister, 2015).
Mitochondria-ER contact sites have lately been under suspicion to play a substantial role in the pathogenesis of neurodegenerative diseases (Paillusson et al., 2016, Phillips and Voeltz, 2016). Furthermore, REEP1, the protein mutated in SPG31, which is additionally known as an interaction partner of spastin and is known to co-localize with spastin at three-way junctions of the endoplasmic reticulum, was shown to facilitate ER-mitochondria interactions. Disease-causing

mutations in REEP1 lead to the diminishment of this interaction and defects in neurite growth resulting in neuronal degeneration (Lim et al., 2015, Park et al., 2010).

The ensuing question, whether these processes play a role in the pathogenesis of SPG4 in particular, will require further investigations on the co-localization of spastin with the inner mitochondrial membrane in vital neuronal cells and the impact on the function of the ATP synthase in cells with mutated or depleted spastin.

4.7 Limitations of the current approach

Protein-protein interactions are known to be very sensitive to changes in cellular conditions such as protein expression levels. The experimental structure in this work limits the estimation of the significance of the investigated novel M1 spastin-protein interactions in SPG4. Further investigation is therefore required.

As discussed in a previous chapter, the overexpression SH-SY5Y cell model is a simple and commonly used model in the investigation of neurodegenerative diseases. Nevertheless, the overexpression and the fact, that other than neurons, SH-SY5Y is a mitotic cell line, there might be differences in protein interaction dynamics, as well as protein behavior in terms of localization or expression levels.

Another consideration is, if the investigated protein-protein-interactions actually take place in vital cells, or if the interaction was just triggered be the disintegration of the cellular compartments through the cell lysis in this experiment.

Therefore, these findings cannot be automatically applied for processes in the human nervous system in the pathogenesis of SPG4.

Summarizing, this approach only allows speculations on protein interactions, that have to be further validated through co-localization experiments and require investigations in a second model, for example neuronal cells derived from pluripotent stem cells.

4.8 Conclusion and outlook

This thesis and the affiliated experiments were designed to further investigate possible pathogenic processes in the most common form of hereditary spastic paraplegia, SPG4, caused by mutations in the SPAST gene.

We were able to identify the two novel interaction partners NUP43 of the M1 isoform of spastin and ATP5A of both spastin isoforms in a mass spectrometry-based interactome analysis using an overexpression SH-SY5Y cell line. Those possible spastin interactions were further validated in immunoprecipitation experiments, in which the structural ability of the proteins to interact was verified. To evaluate the relevance of those interactions in human neuronal cells further investigation and validation will be required. Ensuing this work, co-localization experiments will be performed to test the required co-localization of these interaction in vital, neuronal cells.

Furthermore, their relevance in the pathogenesis of SPG4 has to be examined on endogenous expression levels in post-mitotic neuronal cells.

5. Abstract

Hereditary Spastic Paraplegias (HSPs) are a heterogeneous group of inherited neurodegenerative disorders, that are distinguished by an axonopathy of the upper motor neurons and therefore clinically present with a spasticity and weakness of the lower limbs. Complicated forms of the disease can include additional symptoms such as cognitive impairment, ataxia or myopathy (Klebe et al., 2015).

HSPs can be inherited in an autosomal recessive, dominant or X-chromosomal manner. The most common form of autosomal dominant HSP is caused by a mutation in the Spastic Paraplegia Gene 4 (SPG4), encoding for the protein spastin, which was first described in 1999 (Bürger et al., 2000, Solowska and Baas, 2015). Four isoforms of the protein, that are shown to differ in their cellular expression levels and localization (Claudiani et al., 2005, Solowska and Baas, 2015), are known to be endogenously expressed.

Furthermore, spastin integrates several key pathways of HSP pathogenesis, including membrane shaping, cytoskeleton dynamics as well as intracellular transport and is known to interact with a number of HSP-associated proteins such as Atlastin or REEP1 (Evans et al., 2006, Park et al., 2010).

While there have been many postulations about the disease mechanism of SPG4, such as a loss of function of the protein (Solowska et al., 2010) or the toxicity of the truncated M1 spastin isoform (Solowska et al., 2017), it was shown that not all disease-causing mutations can be explained by these theories. Therefore, up to today, the pathomechanism remains uncertain.

In this work, a mass spectrometry-based approach was chosen to perform an isoform-specific interactome analysis of spastin. The Flp-In™ T-Rex™ system was used to create stable SH-SY5Y overexpression cell lines for the four endogenously expressed spastin isoforms. The tagged protein was then isolated by immunoprecipitation and bound interaction partners were identified by mass spectrometry.

Promising interaction candidates were subsequently confirmed in co-immunoprecipitation studies.

Resulting from this workflow, we were able to reveal the two novel protein-protein interaction partners of the SPG4 protein spastin, NUP43 and ATP5A. Our findings indicate an interaction of the longer M1 isoform of spastin with the NUP107-160 complex, a subunit of the nuclear pore complex, that is known to play a major role in the assembly of the nuclear pore complex and is presumed to promote the spindle assembly during mitosis.

Surprisingly, another finding was the interaction of spastin with proteins of the F1 subunit of the mitochondrial ATP synthase, such as ATP5A. As an impairment of mitochondrial functions was previously shown for other forms of HSP (e.g. SPG7,13), an affection seems possible.

As those novel M1 spastin interactions were identified in a simplified cell model after cell lysis, they will need to be confirmed a second in vivo cell model, for example through co-localization studies. Furthermore, the relevance of these possible spastin interactions in post-mitotic neuronal cells requires further investigation.

A better understanding of the spastin function in health and disease will hopefully bring us closer to revealing the disease mechanism in SPG4 and the development of treatment options.

6. Zusammenfassung

Hereditäre spastische Paraplegien (HSP) sind eine heterogene Gruppe neurodegenerativer Erkrankungen, die sich durch eine Axonopathie des ersten Motoneurons auszeichnen und folglich klinisch mit einer spastischen Parese der unteren Extremitäten einhergehen. Komplizierte Formen der Erkrankung können zusätzliche Symptome wie eine kognitive Beeinträchtigung, Ataxie oder Myopathie umfassen (Klebe et al., 2015).

HSPs können autosomal rezessiv, dominant oder X-chromosomal vererbt werden. Die häufigste Form der autosomal-dominanten HSP wird durch eine Mutation im SPAST-Gen (SPG4) verursacht, das für das Protein Spastin kodiert, das 1999 erstmals beschrieben wurde (Bürger et al., 2000, Solowska und Baas, 2015).

Das codierte Protein Spastin wird endogen in Form von vier Isoformen exprimiert, die sich sowohl in ihren Expressionsleveln als auch ihrer zellulären Lokalisation unterscheiden (Claudiani et al., 2005, Solowska und Baas, 2015).

Darüber hinaus spielt Spastin eine wichtige Rolle in mehreren Schlüsselpfade der HSP-Pathogenese, wie unter anderem Membranstoffwechel, Zytoskelett-Dynamik sowie intrazellulären Transport und es ist bekannt, dass Spastin mit einer Reihe von HSP-assoziierten Proteinen wie Atlastin (SPG3) oder REEP1 (SPG31) interagiert (Evans et al., 2006, Park et al., 2010).

Obwohl viele Hypothesen zur Pathogenese von SPG4, wie z.B. ein Funktionsverlust des Proteins (Solowska et al., 2010) oder die Toxizität der verkürzten Spastin-Isoform M1 (Solowska et al., 2017) gestellt wurden, konnten bislang nicht alle krankheitsverursachenden Mutationen durch diese Theorien erklärt werden. Der Pathomechanismus der SPG4 ist bis heute ungeklärt.

In dieser Arbeit wurde ein Massenspektrometrie-basierter Ansatz gewählt, um eine Interaktom-Analyse der verschiedenen Isoformen von Spastin durchzuführen. Das Flp-In™ T-Rex™ System wurde verwendet, um SH-SY5Y Überexpressionszelllinien für die verschiedenen endogen exprimierten Spastin-Isoformen zu erzeugen. Das Protein wurde dann durch eine Immunpräzipitation isoliert und gebundene Interaktionspartner durch massenspektrometrisch

identifiziert. Mögliche Spastin-Interaktoren wurden anschließend in Ko-Immunopräzipitationsstudien bestätigt.

Als Resultat dieser Arbeit konnten wir zwei neue Interaktionspartner von Spastin, NUP43 und ATP5A identifizieren. Unsere Ergebnisse zeigen eine Interaktion von der längeren M1-Isoform von Spastin mit dem NUP107-160-Komplex, einer Untereinheit des Kernporenkomplexes, der eine wichtige Rolle beim Zusammenbau des Kernporenkomplexes spielt und vermutlich am Aufbau des Spindelapparats während der Mitose beteiligt ist.

Überraschenderweise zeigten sich zudem Hinweise auf eine Interaktion zwischen M1-Spastin und Proteinen der F1-Untereinheit der mitochondrialen ATP-Synthase, wie z.B. ATP5A. Da eine Affektion der mitochondrialen Funktionen bereits für andere Formen der HSP gezeigt wurde, scheint eine Beteiligung an dieser Stelle denkbar.

Da die genannten Interaktionen in einem vereinfachten SH-SY5Y Zellmodell nach Lyse der Zellen identifiziert wurden, ist eine weitere Bestätigung in einem zweiten in vivo Zellmodell nötig, denkbar sind beispielsweise Co-Lokalisationsstudien. Außerdem muss die Relevanz dieser möglichen Spastin-Interaktionen für die Pathogenese der HSP in post-mitotischen neuronalen Zellen weiter untersucht werden.

Wir hoffen, dass ein besseres Verständnis der Funktion von Spastin und helfen kann die Pathogenese der SPG4 besser zu verstehen und Ansätze für mögliche Behandlungsstrategien zu entwickeln.

7. References

ALLISON, R., LUMB, J. H., FASSIER, C., CONNELL, J. W., TEN MARTIN, D., SEAMAN, M. N., HAZAN, J. & REID, E. 2013. An ESCRT-spastin interaction promotes fission of recycling tubules from the endosome. *J Cell Biol,* 202**,** 527-43.

BEETZ, C., BRODHUN, M., MOUTZOURIS, K., KIEHNTOPF, M., BERNDT, A., LEHNERT, D., DEUFEL, T., BASTMEYER, M. & SCHICKEL, J. 2004. Identification of nuclear localisation sequences in spastin (SPG4) using a novel Tetra-GFP reporter system. *Biochem Biophys Res Commun,* 318**,** 1079-84.

BEETZ, C., NYGREN, A. O., SCHICKEL, J., AUER-GRUMBACH, M., BURK, K., HEIDE, G., KASSUBEK, J., KLIMPE, S., KLOPSTOCK, T., KREUZ, F., OTTO, S., SCHULE, R., SCHOLS, L., SPERFELD, A. D., WITTE, O. W. & DEUFEL, T. 2006. High frequency of partial SPAST deletions in autosomal dominant hereditary spastic paraplegia. *Neurology,* 67**,** 1926-30.

BLACKSTONE, C. 2012. Cellular pathways of hereditary spastic paraplegia. *Annu Rev Neurosci,* 35**,** 25-47.

BLACKSTONE, C., O'KANE, C. J. & REID, E. 2011. Hereditary spastic paraplegias: membrane traffic and the motor pathway. *Nat Rev Neurosci,* 12**,** 31-42.

BRASCHINSKY, M., LUUS, S. M., GROSS-PAJU, K. & HALDRE, S. 2009. The prevalence of hereditary spastic paraplegia and the occurrence of SPG4 mutations in Estonia. *Neuroepidemiology,* 32**,** 89-93.

BRAUN, P. & GINGRAS, A. C. 2012. History of protein-protein interactions: from egg-white to complex networks. *Proteomics,* 12**,** 1478-98.

BÜRGER, J., FONKNECHTEN, N., HOELTZENBEIN, M., NEUMANN, L., ELFRIEDE, B., HAZAN, J. & REIS, A. 2000. Hereditary spastic paraplegia caused by mutations in the SPG4 gene. *European Journal of Human Genetics*.

CLAUDIANI, P., RIANO, E., ERRICO, A., ANDOLFI, G. & RUGARLI, E. I. 2005. Spastin subcellular localization is regulated through usage of different translation start sites and active export from the nucleus. *Exp Cell Res,* 309**,** 358-69.

COX, J. & MANN, M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol,* 26**,** 1367-72.

DEPIENNE, C., FEDIRKO, E., FORLANI, S., CAZENEUVE, C., RIBAI, P., FEKI, I., TALLAKSEN, C., NGUYEN, K., STANKOFF, B., RUBERG, M., STEVANIN, G., DURR, A. & BRICE, A. 2007. Exon deletions of SPG4 are a frequent cause of hereditary spastic paraplegia. *J Med Genet,* 44**,** 281-4.

ERICHSEN, A. K., KOHT, J., STRAY-PEDERSEN, A., ABDELNOOR, M. & TALLAKSEN, C. M. 2009. Prevalence of hereditary ataxia and spastic paraplegia in southeast Norway: a population-based study. *Brain,* 132**,** 1577-88.

EVANS, K., KELLER, C., PAVUR, K., GLASGOW, K., CONN, B. & LAURING, B. 2006. Interaction of two hereditary spastic paraplegia gene products, spastin and atlastin, suggests a common pathway for axonal maintenance. *Proc Natl Acad Sci U S A,* 103**,** 10666-71.

EVANS, K. J., GOMES, E. R., REISENWEBER, S. M., GUNDERSEN, G. G. & LAURING, B. P. 2005. Linking axonal degeneration to microtubule remodeling by Spastin-mediated microtubule severing. *J Cell Biol,* 168**,** 599-606.

FOWLER, P. C. & O'SULLIVAN, N. C. 2016. ER-shaping proteins are required for ER and mitochondrial network organization in motor neurons. *Hum Mol Genet,* 25**,** 2827-2837.

GLOECKNER, C. J., BOLDT, K., SCHUMACHER, A. & UEFFING, M. 2009a. Tandem affinity purification of protein complexes from mammalian cells by the Strep/FLAG (SF)-TAP tag. *Methods Mol Biol,* 564**,** 359-72.

GLOECKNER, C. J., BOLDT, K. & UEFFING, M. 2009b. Strep/FLAG tandem affinity purification (SF-TAP) to study protein interactions. *Curr Protoc Protein Sci,* Chapter 19**,** Unit19 20.

HAUSER, S., ERZLER, M., THEURER, Y., SCHUSTER, S., SCHULE, R. & SCHOLS, L. 2016. Establishment of SPAST mutant induced pluripotent stem cells (iPSCs) from a hereditary spastic paraplegia (HSP) patient. *Stem Cell Res,* 17**,** 485-488.

HAVLICEK, S., KOHL, Z., MISHRA, H. K., PROTS, I., EBERHARDT, E., DENGUIR, N., WEND, H., PLOTZ, S., BOYER, L., MARCHETTO, M. C., AIGNER, S., STICHT, H., GROEMER, T. W., HEHR, U., LAMPERT, A., SCHLOTZER-SCHREHARDT, U., WINKLER, J., GAGE, F. H. & WINNER, B. 2014. Gene dosage-dependent rescue of HSP neurite defects in SPG4 patients' neurons. *Hum Mol Genet,* 23**,** 2527-41.

JUNGE, W. & NELSON, N. 2015. ATP synthase. *Annu Rev Biochem,* 84**,** 631-57.

KLEBE, S., STEVANIN, G. & DEPIENNE, C. 2015. Clinical and genetic heterogeneity in hereditary spastic paraplegias: from SPG1 to SPG72 and still counting. *Rev Neurol (Paris),* 171**,** 505-30.

KOVALEVICH, J. & LANGFORD, D. 2013. Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. *Methods Mol Biol,* 1078**,** 9-21.

KOZAK, M. 2002. Pushing the limits of the scanning mechanism for initiation of translation. *Gene,* 299**,** 1-34.

LIM, Y., CHO, I. T., SCHOEL, L. J., CHO, G. & GOLDEN, J. A. 2015. Hereditary spastic paraplegia-linked REEP1 modulates endoplasmic reticulum/mitochondria contacts. *Ann Neurol,* 78**,** 679-96.

LUMB, J. H., CONNELL, J. W., ALLISON, R. & REID, E. 2012. The AAA ATPase spastin links microtubule severing to membrane modelling. *Biochim Biophys Acta,* 1823**,** 192-7.

MANNAN, A. U., BOEHM, J., SAUTER, S. M., RAUBER, A., BYRNE, P. C., NEESEN, J. & ENGEL, W. 2006. Spastin, the most commonly mutated protein in hereditary spastic paraplegia interacts with Reticulon 1 an endoplasmic reticulum protein. *Neurogenetics,* 7**,** 93-103.

MCDERMOTT, C. J., TAYLOR, R. W., HAYES, C., JOHNSON, M., BUSHBY, K. M., TURNBULL, D. M. & SHAW, P. J. 2003. Investigation of mitochondrial function in hereditary spastic paraparesis. *Neuroreport,* 14**,** 485-8.

NAKANO, H., WANG, W., HASHIZUME, C., FUNASAKA, T., SATO, H. & WONG, R. W. 2011. Unexpected role of nucleoporins in coordination of cell cycle progression. *Cell Cycle,* 10**,** 425-33.

NOVARINO, G., FENSTERMAKER, A. G., ZAKI, M. S., HOFREE, M., SILHAVY, J. L., HEIBERG, A. D., ABDELLATEEF, M., ROSTI, B., SCOTT, E., MANSOUR, L., MASRI, A., KAYSERILI, H., AL-AAMA, J. Y., ABDEL-SALAM, G. M. H., KARMINEJAD, A., KARA, M., KARA, B., BOZORGMEHRI, B., BEN-OMRAN, T., MOJAHEDI, F., EL DIN MAHMOUD, I. G., BOUSLAM, N., BOUHOUCHE, A., BENOMAR, A., HANEIN, S., RAYMOND, L., FORLANI, S., MASCARO, M., SELIM, L., SHEHATA, N., AL-ALLAWI, N., BINDU, P. S., AZAM, M., GUNEL, M., CAGLAYAN, A., BILGUVAR, K., TOLUN, A., ISSA, M. Y., SCHROTH, J., SPENCER, E. G., ROSTI, R. O., AKIZU, N., VAUX, K. K., JOHANSEN, A., KOH, A. A., MEGAHED, H., DURR, A., BRICE, A., STEVANIN, G., GABRIEL, S. B., IDEKER, T. & GLEESON, J. G. 2014. Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. *Science,* 343**,** 506-511.

ORJALO, A. V., ARNAOUTOV, A., SHEN, Z., BOYARCHUK, Y., ZEITLIN, S. G., FONTOURA, B., BRIGGS, S., DASSO, M. & FORBES, D. J. 2006. The Nup107-160 nucleoporin complex is required for correct bipolar spindle assembly. *Mol Biol Cell,* 17**,** 3806-18.

PAILLUSSON, S., STOICA, R., GOMEZ-SUAGA, P., LAU, D. H., MUELLER, S., MILLER, T. & MILLER, C. C. 2016. There's Something Wrong with my MAM; the ER-Mitochondria Axis and Neurodegenerative Diseases. *Trends Neurosci,* 39**,** 146-57.

PARK, S. H., ZHU, P. P., PARKER, R. L. & BLACKSTONE, C. 2010. Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network. *J Clin Invest,* 120**,** 1097-110.

PHILLIPS, M. J. & VOELTZ, G. K. 2016. Structure and function of ER membrane contact sites with other organelles. *Nat Rev Mol Cell Biol,* 17**,** 69-82.

QIANG, L., PIERMARINI, E., MURALIDHARAN, H., YU, W., LEO, L., HENNESSY, L. E., FERNANDES, S., CONNORS, T., YATES, P. L., SWIFT, M., ZHOLUDEVA, L. V., LANE, M. A., MORFINI, G., ALEXANDER, G. M., HEIMAN-PATTERSON, T. D. & BAAS, P. W. 2019. Hereditary spastic paraplegia: gain-of-function mechanisms revealed by new transgenic mouse. *Hum Mol Genet,* 28**,** 1136-1152.

REID, E., CONNELL, J., EDWARDS, T. L., DULEY, S., BROWN, S. E. & SANDERSON, C. M. 2005. The hereditary spastic paraplegia protein spastin interacts with the ESCRT-III complex-associated endosomal protein CHMP1B. *Hum Mol Genet,* 14**,** 19-38.

ROLL-MECAK, A. & VALE, R. D. 2008. Structural basis of microtubule severing by the hereditary spastic paraplegia protein spastin. *Nature,* 451**,** 363-7.

RUANO, L., MELO, C., SILVA, M. C. & COUTINHO, P. 2014. The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies. *Neuroepidemiology,* 42**,** 174-83.

RUHLE, T. & LEISTER, D. 2015. Assembly of F1F0-ATP synthases. *Biochim Biophys Acta,* 1847**,** 849-60.

SANDERSON, C. M., CONNELL, J. W., EDWARDS, T. L., BRIGHT, N. A., DULEY, S., THOMPSON, A., LUZIO, J. P. & REID, E. 2006. Spastin and atlastin, two proteins mutated in autosomal-dominant hereditary spastic paraplegia, are binding partners. *Hum Mol Genet,* 15**,** 307-18.

SCHULE, R., HOLLAND-LETZ, T., KLIMPE, S., KASSUBEK, J., KLOPSTOCK, T., MALL, V., OTTO, S., WINNER, B. & SCHOLS, L. 2006. The Spastic Paraplegia Rating Scale (SPRS): a reliable and valid measure of disease severity. *Neurology,* 67**,** 430-4.

SCHULE, R. & SCHOLS, L. 2011. Genetics of hereditary spastic paraplegias. *Semin Neurol,* 31**,** 484-93.

SCHULE, R., WIETHOFF, S., MARTUS, P., KARLE, K. N., OTTO, S., KLEBE, S., KLIMPE, S., GALLENMULLER, C., KURZWELLY, D., HENKEL, D., RIMMELE, F., STOLZE, H., KOHL, Z., KASSUBEK, J., KLOCKGETHER, T., VIELHABER, S., KAMM, C., KLOPSTOCK, T., BAUER, P., ZUCHNER, S., LIEPELT-SCARFONE, I. & SCHOLS, L. 2016. Hereditary spastic paraplegia: Clinicogenetic lessons from 608 patients. *Ann Neurol,* 79**,** 646-58.

SHERWOOD, N. T., SUN, Q., XUE, M., ZHANG, B. & ZINN, K. 2004. Drosophila spastin regulates synaptic microtubule networks and is required for normal motor function. *PLoS Biol,* 2**,** e429.

SOLOWSKA, J. M. & BAAS, P. W. 2015. Hereditary spastic paraplegia SPG4: what is known and not known about the disease. *Brain,* 138**,** 2471-84.

SOLOWSKA, J. M., D'ROZARIO, M., JEAN, D. C., DAVIDSON, M. W., MARENDA, D. R. & BAAS, P. W. 2014. Pathogenic mutation of spastin has gain-of-function effects on microtubule dynamics. *J Neurosci,* 34**,** 1856-67.

SOLOWSKA, J. M., GARBERN, J. Y. & BAAS, P. W. 2010. Evaluation of loss of function as an explanation for SPG4-based hereditary spastic paraplegia. *Hum Mol Genet,* 19**,** 2767-79.

SOLOWSKA, J. M., MORFINI, G., FALNIKAR, A., HIMES, B. T., BRADY, S. T., HUANG, D. & BAAS, P. W. 2008. Quantitative and functional analyses of spastin in the nervous system: implications for hereditary spastic paraplegia. *J Neurosci,* 28**,** 2147-57.

SOLOWSKA, J. M., RAO, A. N. & BAAS, P. W. 2017. Truncating mutations of SPAST associated with hereditary spastic paraplegia indicate greater accumulation and toxicity of the M1 isoform of spastin. *Mol Biol Cell,* 28**,** 1728-1737.

SPITZER, J., LANDTHALER, M. & TUSCHL, T. 2013. Rapid creation of stable mammalian cell lines for regulated expression of proteins using the Gateway(R) recombination cloning technology and Flp-In T-REx(R) lines. *Methods Enzymol,* 529**,** 99-124.

STONE, M. C., RAO, K., GHERES, K. W., KIM, S., TAO, J., LA ROCHELLE, C., FOLKER, C. T., SHERWOOD, N. T. & ROLLS, M. M. 2012. Normal

spastin gene dosage is specifically required for axon regeneration. *Cell Rep,* 2**,** 1340-50.

VIETRI, M., SCHINK, K. O., CAMPSTEIJN, C., WEGNER, C. S., SCHULTZ, S. W., CHRIST, L., THORESEN, S. B., BRECH, A., RAIBORG, C. & STENMARK, H. 2015. Spastin and ESCRT-III coordinate mitotic spindle disassembly and nuclear envelope sealing. *Nature,* 522**,** 231-5.

WALTHER, T. C., ALVES, A., PICKERSGILL, H., LOIODICE, I., HETZER, M., GALY, V., HULSMANN, B. B., KOCHER, T., WILM, M., ALLEN, T., MATTAJ, I. W. & DOYE, V. 2003. The conserved Nup107-160 complex is critical for nuclear pore complex assembly. *Cell,* 113**,** 195-206.

WHARTON, S. B., MCDERMOTT, C. J., GRIERSON, A. J., WOOD, J. D., GELSTHORPE, C., INCE, P. G. & SHAW, P. J. 2003. The cellular and molecular pathology of the motor system in hereditary spastic paraparesis due to mutation of the spastin gene. *J Neuropathol Exp Neurol,* 62**,** 1166-77.

WHITE, S. R., EVANS, K. J., LARY, J., COLE, J. L. & LAURING, B. 2007. Recognition of C-terminal amino acids in tubulin by pore loops in Spastin is important for microtubule severing. *J Cell Biol,* 176**,** 995-1005.

WOOD, J. D., LANDERS, J. A., BINGLEY, M., MCDERMOTT, C. J., THOMAS-MCARTHUR, V., GLEADALL, L. J., SHAW, P. J. & CUNLIFFE, V. T. 2006. The microtubule-severing protein Spastin is essential for axon outgrowth in the zebrafish embryo. *Hum Mol Genet,* 15**,** 2763-71.

XICOY, H., WIERINGA, B. & MARTENS, G. J. 2017. The SH-SY5Y cell line in Parkinson's disease research: a systematic review. *Mol Neurodegener,* 12**,** 10.

XING, S., WALLMEROTH, N., BERENDZEN, K. W. & GREFEN, C. 2016. Techniques for the Analysis of Protein-Protein Interactions in Vivo. *Plant Physiol,* 171**,** 727-58.

ZUCCOLO, M., ALVES, A., GALY, V., BOLHY, S., FORMSTECHER, E., RACINE, V., SIBARITA, J. B., FUKAGAWA, T., SHIEKHATTAR, R., YEN, T. & DOYE, V. 2007. The human Nup107-160 nuclear pore subcomplex contributes to proper kinetochore functions. *EMBO J,* 26**,** 1853-64.

8. Erklärung zum Eigenanteil

Die Arbeit wurde im Hertie-Institut für klinische Hirnforschung in Tübingen unter Betreuung von PD Dr. Rebecca Schüle-Freyer durchgeführt.

Die Konzeption der Studie erfolgte durch PD Dr. Rebecca Schüle-Freyer.

Sämtliche Versuche wurden (nach Einarbeitung durch Labormitglieder Ulrike Ulmer, Andrés Caballero und Jennifer Reichbauer) von mir eigenständig durchgeführt. Alle Massenspektrometrie-Versuche wurden in Zusammenarbeit mit Karsten Boldt und Nicola Horn im Medical Proteome Center Tübingen durchgeführt.

Die statistische Auswertung der Interaktom-Ergebnisse erfolgte in Zusammenarbeit mit Karsten Boldt.

Ich versichere, das Manuskript selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

Tübingen, den

9. Acknowledgements

I would like to express my sincere thank you to Rebecca Schüle for the opportunity to work in her lab, the generous support over the last years and especially for the welcoming and warm atmosphere she creates in her lab.

Further I would like to thank all members of the Schüle Lab for their strong support during the time of this work. Especially Ulrike Ulmer, who taught me most of the techniques in my first months, Jennifer Reichbauer, who not only always contributed good advices, but also helped out whenever she could, Maike Nagel, who especially supported me during finishing this work, Selina Reich, who always invested time to forward her tips and tricks and Ina Gehweiler, who provided the best mental support anyone could ever get. Thank you all for the time and effort you invested and the great times we spend together!

I would also like to thank Karsten Boldt and Nicola Horn for the great cooperation with the medical proteome center Tübingen. Thank you for sharing your expertise, for the time you invested in my project and for answering all my questions!

Last but not least a very warm thank you to my amazing family and Dennis, for their unconditional love and support, for always believing in me and cheering me up whenever needed. I wouldn’t be where I am now without you!