universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Purification and characterization of a potential therapeutic protein for spinal muscular atrophy“

verfasst von / submitted by

Zsófia Kormányos

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 830
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Molekulare Mikrobiologie, Mikrobielle Ökologie und Immunbiologie
Betreut von / Supervisor:	Assoc.Prof. Priv.-Doz. Dr.med.univ. Franco Laccone

Acknowledgements

I would first like to thank my thesis supervisor Assoc.Prof. Priv.-Doz. Dr.med.univ. Franco Laccone for giving me the opportunity to perform my master thesis in his laboratory. I am grateful for the knowledge and experience I gained during the short period I spent there.

I would like to emphasis my gratitude to the members of the Laccone group: Mag. Dr. Hannes Steinkellner for his supervising and guidance in the laboratory, Dr. Alex Beribisky for motivating me to search for solutions and on the way to have a deeper understanding of proteins and Anna Huber for her advice but mostly for her friendship.

My sincere thanks go to the Department of Medical Genetics of the Medical University of Vienna for providing a space for me and all the employees for either helping me, cheering me up and making it a fun and welcoming place for me.

I would like to thank the Vienna Biocenter Core Facilities (VBCF) for performing the Thermofluor, dynamic light scattering and mass spectrometry experiments.

Last but not least I would like to thank my family and friends for their support and encouragement.

Table of contents

Acknowledgements **3**
Table of contents **7**
1. Introduction **9**
1.1. Spinal muscular atrophy (SMA) 9
1.2. Molecular genetics 10
1.3. Screening 12
1.4. Treatment 12
1.5. Therapeutic protein expression and purification 14
2. Materials and methods **18**
2.1. Cultivation of cell lines 18
2.1.1. Mouse Motor Neuron-like Hybrid cell line - NSC-34 18
2.1.2. Human Schwan cell line - HSC 18
2.1.3. Mouse Embryonic Fibroblast cell line - NIH3T3 18
2.2 Knockdown of SMN1 in HSC cells 19
2.3. Knockdown of SMN1 in NSC-34cells 19
2.4. Casting and running of SDS-PAGE 19
2.5. Cloning strategy of pET28-His-SMN1-Strep and pET28-His-TAT-SMN1-Strep 20
2.6. Expression of TAT-SMN1 and SMN1 21
2.6.1. Expression method 1 21
2.6.2. Expression method 2 21
2.6.3. Expression method 3 22
2.6.4. Expression method 4 22
2.6.5. Expression method 5 22
2.6.6. Expression method 6 22
2.7. Lysis methods of TAT-SMN1 and SMN1 overexpressed bacteria 23
2.8. Protein purification 25
3.8.1. His-Gravi Trap Sepharose Column 25
3.8.2. Strep Tactin Column 25
2.8.3. Ion exchange chromatography 26
2.8.4. PD10 desalting column 27
2.8.5. Dyalisis 28

2.9. Protein concentration with 10kD MWCO 28
2.10. Storage of purified proteins 29
2.11. Purified protein concentration determination 30
2.12. Protein stability test 30
2.13. Mass spectrometry, Dynamic light scattering and Thermofluor 30
2.14. Immunofluorescence 30
2.15. Transduction of different types of TAT-SMN1 and SMN1 and cell lysis for western-blot 33
2.16. Western-blot 33
3. Results 35
3.1. Lysis in native lysis buffer of pellets produced by various expression methods 35
3.2. Different lysis methods of TAT-SMN1 pellets produced at 20°C in LB-rich medium 37
3.3. Denaturing purification of TAT-SMN1 and SMN1 on Strep column 41
3.4. Purifications with dialysis 43
3.5. TAT-SMN1 purification with cation exchange column 43
3.6. Lysis with 5 % n-propanol and 2 M urea and purification with His column 48
3.7. Lysis with 5 % DMSO and 2 M urea and refolding on the Strep column 49
3.8. Lysis with 5 % n-propanol and 2 M urea and purification with Strep column and PD-10 columns 49
3.9. Mass spectrometry, dynamic light scattering, Thermofluor 51
4. Discussion 54
5. Conclusion 60
6. References 62
7. Appendix 66
List of Figures: 66
List of tables: 67
Abstract 68
Zusammenfassung 69

1. Introduction

1.1. Spinal muscular atrophy (SMA): SMA is an autosomal recessive genetic disorder characterized by the degeneration of spinal cord motor neurons, atrophy of skeletal muscles and generalized weakness. It carries a relatively high incidence rate of one per every 10.000 live births with one in every 50 people being a carrier. SMA is the second most common lethal autosomal genetic disorder after cystic fibrosis and the leading genetic cause of infant death. The disease was first described by Guido Werdnig (Werdnig, 1891) and by Johann Hoffman (Hoffmann, 1893). SMA can be categorized into 4 groups of severity based on the age of onset and the highest motor function achieved. Type I (Werdning-Hoffmann disease) is the most severe, type II which is intermediately severe, type III (Kugelberg-Welander disease) is quite mild and type IV in which the affected only start showing symptoms in their 20s or 30s. In the following table the different types of severity are described by the onset and the highest function achieved (Lunn & Wang, 2008) (**Table 1**).

	OMIM number	Age at onset	Highest function achieved	Natural age of death
Type I (severe, Werdnig-Hoffmann disease)	253300	0–6 months	Never sit	<2 years
Type II (intermediate)	253550	7–18 months	Sit, never stand	>2 years
Type III (mild, Kugelberg-Welander disease)	253400	>18 months	Stand and walk	Adult
Type IV (adult)	271150	Second or third decade	Walk during adulthood	Adult

Table 1. Types of spinal muscular atrophy and their defining characteristics (Lunn & Wang, 2008).

Compared to a healthy individual, only one gene is mutated but it causes various differences on many levels. If the *Survival motor neuron 1* (*SMN1)* gene is absent or mutat-

ed, the SMN1 protein produced from it is unable to carry out its function, which in turn leads to motor neuron dysfunction or death (S. Lefebvre et al., 1997). Specifically, motor neurons cannot communicate with the muscles and consequentially the latter lose strength and start to shrink (atrophy). In the end the movements (walking, swallowing and breathing) are restricted (**Figure 1**).

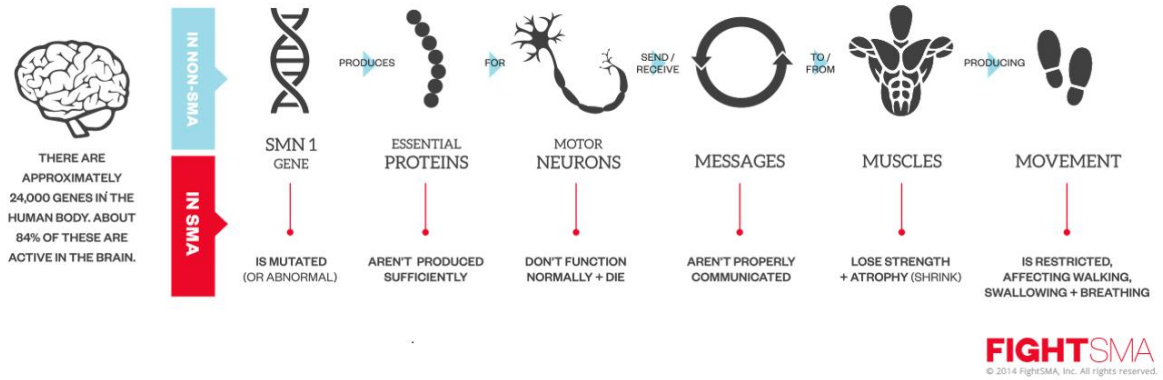

Figure 1. Effects of SMA on different levels compared to a healthy individual (http://fightsma.org/what-sma).

1.2. Molecular genetics: In 98 % of the cases SMA is caused by the homozygous disruption in *SMN1* which was identified as the causative gene of SMA in 1995. *SMN* is an approximately 20kb gene coding for a 294-amino acid protein which is located on the 5q13 chromosome (Brzustowicz et al., 1990). In affected individuals, this region is either missing or interrupted while the patients retaining this gene have a point mutation or short deletion (Suzie Lefebvre et al., 1995). SMN is present in multiple copies in the human body: *SMN1* has a telomeric copy while *SMN2* possesses multiple centromeric copies. The *SMN1* gene produces 90 % of full length 38kDa SMN1 protein. *SMN2*, on the other hand, is only present in humans and only 10 % of the produced protein is full length and functional, with the other 90% being truncated (**Figure 2**). A cytosine to thymine substitution (850C>T) in the *SMN2* gene promotes exon 7 skipping

which results in the translation of *SMN2* rather than *SMN1*, as the aforementioned thymine residue is not a component of the exonic splicing enhancer complex (Butchbach, 2016).

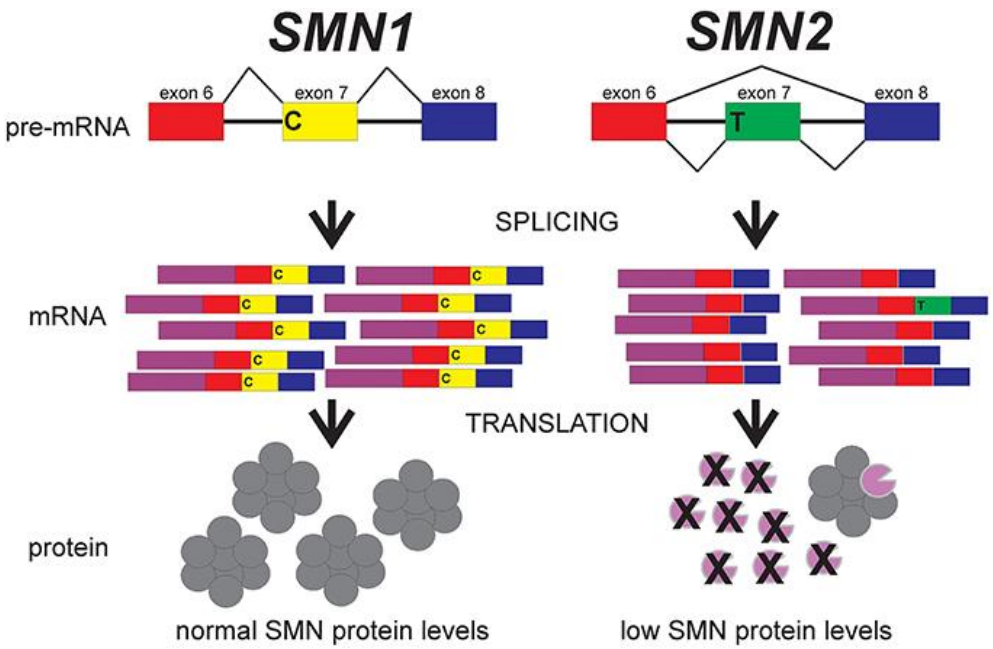

Figure 2. Different **s**plicing of *SMN1* and *SMN2* resulting in different protein levels (Butchbach, 2016).

The severity of SMA depends on the *SMN2* copy number in the affected individual since *SMN2* also produces full length protein. The higher the *SMN2* copy number, the higher motor function is achieved (Butchbach, 2016).

SMN1 is a ubiquitously expressed protein located in the cytoplasm, neuronal growth cones, neuronal extensions and in the nucleus. In the nucleus, the SMN1 forms subnuclear bodies called gems which are near coiled coil bodies. SMN1 possesses multiple functions depending on its location. It has a role in RNA metabolism (snRNP's), in actin cytoskeleton dynamics, in mRNA transport, in ubiquitin homeostasis, in bioenergetics pathway and in synaptic vesicle release. Although it has many functions, none of the above mentioned is fully responsible for SMA (**Figure 3**) (Bowerman et al., 2017).

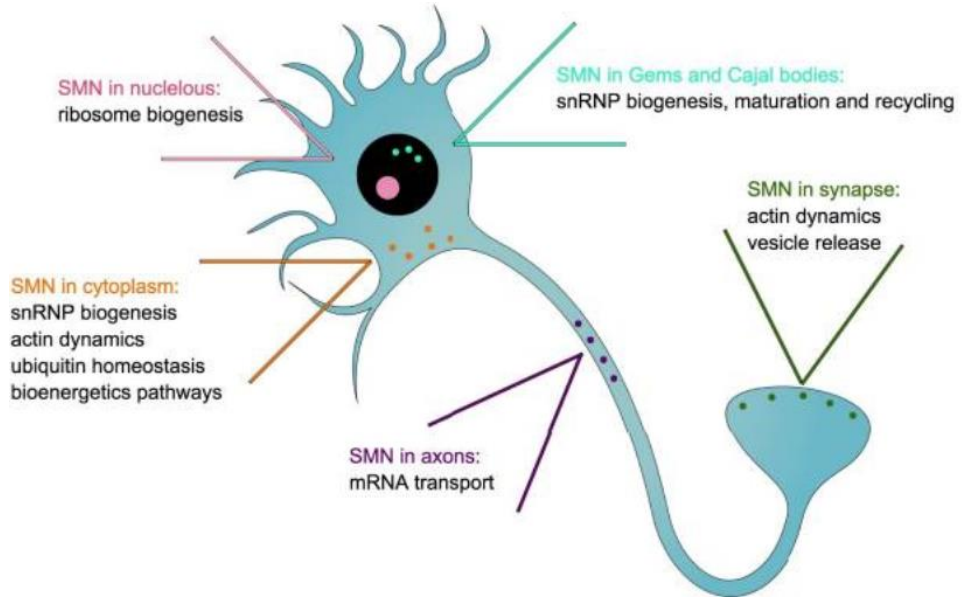

Figure 3. SMN1 localization in a neuronal cell and its functions (Bowerman et al., 2017).

1.3. Screening: Due to the high carrier frequency, testing of parents and siblings of an affected child is important in order to obtain information for later reproductive planning. The siblings of an affected child should be tested if they are carriers then their partners should be tested as well. A couple, who already had a child with SMA, has a 25% chance for reoccurrence. In any event, families are offered a prenatal diagnosis. Prognosis is determined by the phenotypic severity (from type I – high mortality within the first year of life to type IV – symptoms only appear in adult life). Screening for SMA could be one way to avert this disease (D'Amico et al., 2011).

1.4. Treatment: The first FDA-approved drug for SMA treatment is an intronic splicing silencer (ISS-N1) called Spinraza™ (nusinersen). Spinraza™ is an antisense oligonucleotide-mediated splicing correction of SMN2 (Ottesen, 2017). There are several other approaches to tackle this serious illness: gene therapy with viral vectors to deliver SMN1 directly, *trans*-splicing RNA's, neurotrophic agents and motor neurons derived from stem cells (**Figure 4**) (Bowerman et al., 2017).

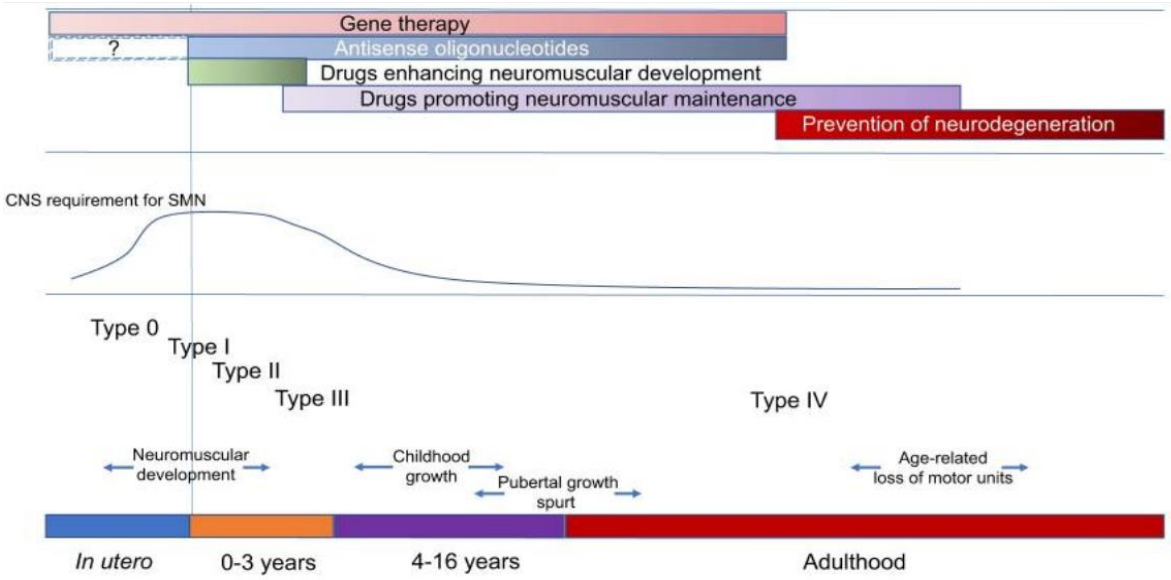

Figure 4. Different therapeutic approaches based on the time of disease progression (Bowerman et al., 2017).

Our approach to treat this debilitating disease was the development of a protein replacement therapy. Our aim was to produce functional and stable SMN1 protein which is able to go through the cell membrane. To achieve the transduction in *in vitro* models, we used amongst others NSC-34 cells (a cellular model for SMA), and we created a TAT (transactivator of transcription)-SMN1 fusion protein. TAT is the trans-activating transcriptional activator of HIV1 virus and it can deliver different cargos through the blood-brain barrier (BBB) and the cell membrane. This method was first described in 1988 and since then it is continuously improved (Frankel & Pabo, 1988). TAT is able to deliver full length proteins into primary or transduced cells within less than 5 minutes. TAT fusion proteins regardless of size or function can be added to mammalian cell culture or injected into mice. Other method also exists for transduction like: transfection of expression vector, microinjection, or diffusion of peptidyl mimetics. These methods are more or less successful but cannot be strictly regulated. Therefore, cell penetrating peptides are a better option for transduction of proteins (Becker-Hapak,

McAllister, & Dowdy, 2001). The arginine-rich cell penetrating peptide TAT is one of the most widely used arginine-rich cell penetrating peptide (**Figure 5**) (Nakase et al., 2017).

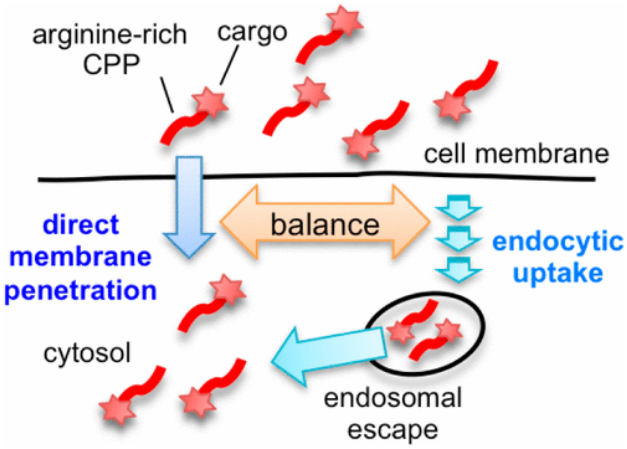

Figure 5. TAT directed transport mechanisms of cargo into the cytosol (Nakase et al., 2017).

1.5. Therapeutic protein expression and purification: Expressing eukaryotic proteins in bacteria could be challenging because the bacteria can accumulate the protein in inclusion bodies. There are many factors which contribute to inclusion body formation such as high temperature or lack of eukaryotic chaperons. Inclusion bodies barely contain any host proteins; they are mainly composed of the expressed recombinant protein. There are two types of inclusion bodies: classical inclusion bodies and non-classical inclusion bodies. In general, non-classical ones are more fragile and easier to lyse. Conventionally, the solubilisation of the classical inclusion bodies happens with high concentration chaotrophes (urea, guanidium chloride (GdnHCl)). This method involves the complete denaturation of the protein which often makes the refolding processes challenging, sometimes leading to formation of protein aggregates (Singh et al., 2015). There are many techniques to tackle protein aggregation but there exists no universal solution because of various nature of different proteins. One commonly used

additive in the storage buffer which could suppress aggregation is tris (2-carboxyethyl phosphine (TCEP). TCEP is a reducing agent similar to β-mercaptoethanol (ß-ME) (Wingfield, 2016).

Non-classical inclusion bodies usually could be lysed with mild lysis methods (extreme pH, low concentration of detergent (N-laurylsarcosine) or organic solvents (β-ME, n-propanol)) combined with low amounts of urea (2-3 M) (A. Singh et al., 2015). Non-classical inclusion bodies contain proteins in native secondary structure therefor if you lyse them with milder reagents; there is a lower chance of aggregation or no need for refolding (**Figure 6**).

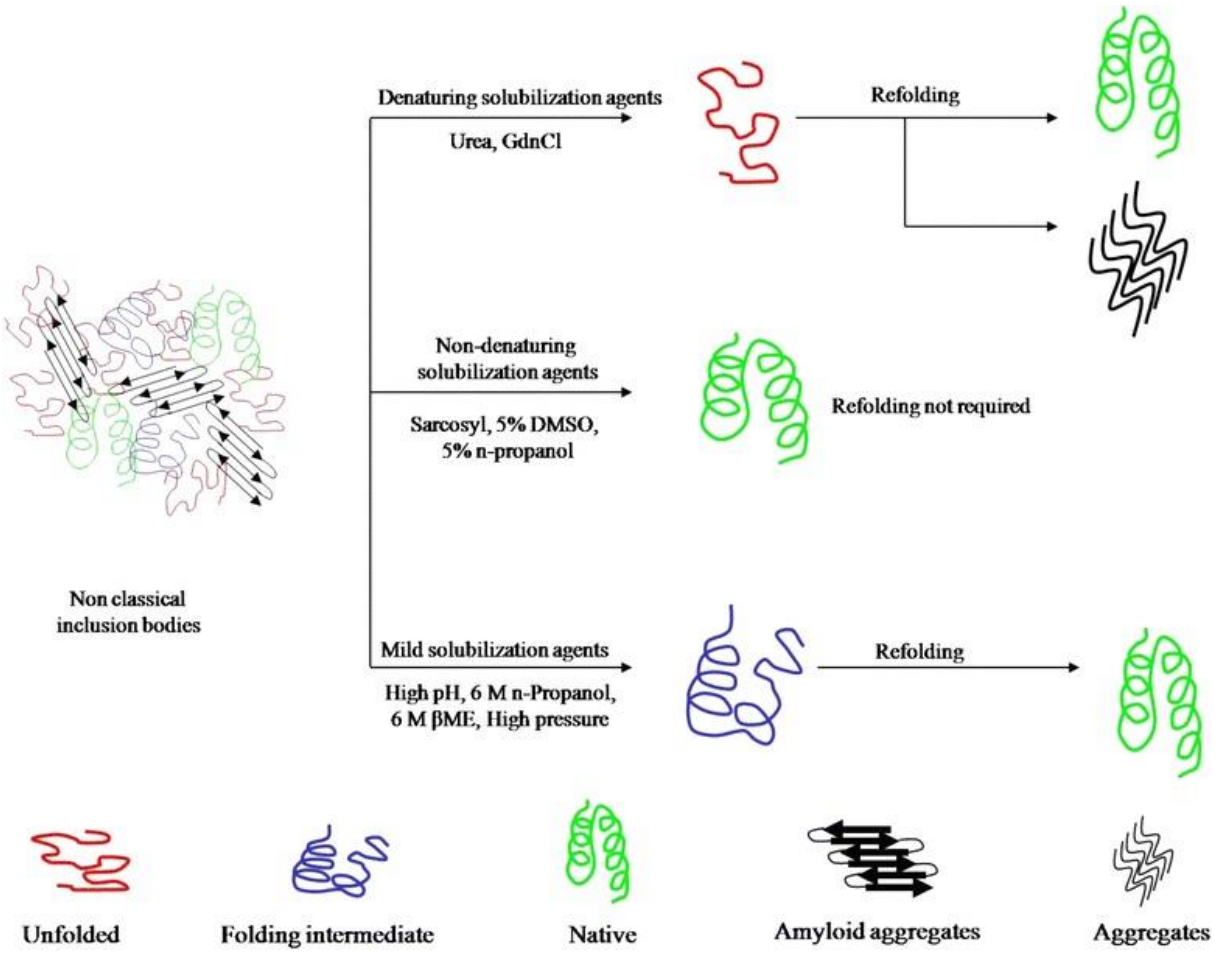

Figure 6. Different solubilisation methods and consequences in protein folding (A. Singh et al., 2015).

Another TAT fusion protein which has therapeutic potential (TAT-Hsp70) was the most efficiently transduced and had the highest biological activity when purified in native conditions (Nagel et al. 2008). There are expression methods used several methods to supress inclusion body formation and produce soluble protein. Expressing proteins at low temperature (15-25ºC) is one of them. At the same time, lower temperature causes slower growth and lower protein yield. Expressing proteins in other strains like *Origami*^{TM} DE3 (Novagen) can also serve as a solution for insoluble proteins. This strain is a thioredoxin reductase mutant, so the disulphide bond formation in the cyto-

plasm is enhanced (Rosano & Ceccarelli, 2014). Overexpressing molecular chaperons together with the protein of interest could help in correct protein folding and to get correct tertiary structure (Nishihara, Kanemori, & Yanagi, 2000).

2. Materials and methods

2.1. Cultivation of cell lines

2.1.1. Mouse Motor Neuron-like Hybrid cell line - NSC-34: Cells were thawed and diluted 1:5 in NSC-34 media consisting of DMEM (41965-29, Gibco), 10 % FBS (F9665, Sigma) and 1 % penicillin – streptomycin (#15140122, Gibco). As the cells settled on the bottom of the flask (4-6 hours) media was changed to NSC-34 media in order to remove dimethyl sulfoxide (DMSO, #D2438-50ML, Sigma). Cells were split twice a week in 1:6 ratio with 0,25 % trypsin-EDTA.

2.1.2. Human Schwan cell line - HSC: Cells were thawed and transferred into a falcon containing prewarmed HSC media. HSC media contained DMEM (#11965-092, Gibco), 0,2 % glucose (#15023-021, Sigma), 2 mM L-glutamine (#25030-081, Sigma), 10 % FBS (#F4135, Sigma), 2 µM forskolin (#F6886, Sigma) and 1 % penicillin – streptomycin (#15140122, Gibco). Then the cells were centrifuged for 10 minutes at 1000 xg on RT. The supernatant was removed from the cell pellet. Then the cell pellet was re-suspended in HSC media and transferred into a cell culture flask. Cells were split by trypsinization (0,05 %) twice a week in 1:4 ratio.

2.1.3. Mouse Embryonic Fibroblast cell line - NIH3T3: After thawing, the cells were treated as HSC. The media for NIH3T3 cells contained DMEM (41966-39, Gibco), 10 % FBS (#F4135, Sigma) and 1 % penicillin – streptomycin (#15140122, Gibco). Cells were split 1:5 thrice a week.

All cell lines were handled among sterile environment and kept in a 37ºC incubator with 5 % CO_2.

2.2 Knockdown of SMN1 in HSC cells

Cells were plated in a 12-well plate and grown for 24 hours. Cells were 40 % confluent when the transfection mixture was added. The mixture contained 2 ml serum free media, 8 µl INTERFERin (#409-10, Polyplus) and 1 nM siRNA (#S104950925, Qiagen) per well. The reagents were vortexed and spun down and then incubated for 10 minutes at room temperature before being added to the wells. As one negative control scrambled siRNA (#D-001810-01-05, Thermo Scientific Dharmacon) was used in 1 nM concentration in combination with INTERFERin and as another negative control only INTERFERin was added to the well. 24 hours after the transfection media was changed to fresh HSC media and another 48 hours later, the cells were lysed in RIPA buffer.

2.3. Knockdown of SMN1 in NSC-34 cells

In order to establish the cell count needed for transfection, different cell counts ($6x10^5$, $5x10^5$, $4\ x10^5$, $3\ x10^5$, $2\ x10^5$) were plated in a 12 well plate and grown for 24 hours. Wells were transfected with serum free media, INTERFERin and 50 nM s74016 siRNA (#4390815, Ambion). Media was changed 24 hours after transfection and cells were lysed after 72 hours of transfection. Based on western-blot results, the best cell count was used from here on. To determine the best INTERFERin and siRNA concentration, $4\ x\ 10^5$ cells were plated and grown for 24 hours. 6 µl of INTERFERin were combined with 5 nM and 10 nM siRNA and the same was carried out with 8 µl of INTERFERin. The rest of the procedure was identical as described above.

2.4. Casting and running of SDS-PAGE

12 % sodium dodecyl sulphate polyacrylamide (SDS-gels) gels were made using ProSieve Gel solution (#50618, Lonza). All gels were 1,5 mm thick with 10 pockets for

loading. Samples for SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) were prepared as follows: 1x Sample Buffer (7 ml TRIS-HCl (1 M pH=6.8), 6 ml glycerol, 2 g SDS, 2,4 mg bromphenol blue,7 ml ddH2O) and 5 mM β-ME were added then vortexed and heated for 5 minutes at 95°C. After loading, the gels were run in an electrophoresis tank filled with electrophoresis buffer in the refrigerator for 50 minutes at 180 V. After running, the gels were either used for western-blot or stained with Coomassie staining solution (0,1 % Coomasie Blue R250, 10 % acetic acid, 50 % methanol, 40 % ddH_2O) and then destained by destaining solution (10 % acetic acid, 50 % methanol, 40 % ddH_2O).

2.5. Cloning strategy of *pET28-His-SMN1-Strep* and *pET28-His-TAT-SMN1-Strep*

His-TAT-SMN1-Strep sequence was amplified by rtPCR from *pJexpress411-64109-TAT-SMN1-StrepMOD* vector and the *pBluescript* plasmid was digested with EcorV (#R0195S, NEB). The products had been loaded on an agarose gel and were then extracted by QIAquick Extraction Kit (#28706, Qiagen). The vector and the insert (1:4) were ligated and transformed into ElectroMAX DH10B (#18290015, ThermoFisher Scientific) electrocompetent cells. The transformation was grown on LB-agar plates supplemented with ampicillin and X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) for positive clone selection. The positive clones were subsequently used for plasmid extraction with GeneElute Plasmid Miniprep Kit (#PLN350-1KT, Sigma) and which was analyzed by restriction endonuclease digestion on agarose gel. The clones which were shown to be positive were sequenced. Minipreps and pET28_TAT_SMN1_Strep (#17) vectors were cut with NcoI-HF (#3193, NEB) and BsrGI-HF (#R3575, NEB). Gel elutions of the restriction products were performed and a ligation reaction was set up. The ligated vector and insert was transformed into Elec-

troMAX DH10B cells and grown on LB-agar-kanamycin plates. Miniprep was carried out and the clones were test restricted and the positive ones were grown for midiprep (Quiagen Plasmid Plus Midi Kit (#12943, Qiagen)). The plasmids were sequenced resulting in *pET28-His-SMN1-Strep*. The *pET28-His-TAT-SMN1-Stre*p was provided (midiprep from 04.08.2011, 530 ng/µl).

2.6. Expression of TAT-SMN1 and SMN1

2.6.1. Expression method 1: 100 ng of *pET28-TAT-SMN1-Strep (#17)* and pET28_His_SMN1_Strep were transformed into *E. cloni* EXPRESS BL21 (DE3) (#60300-1, Lucigen) electrocompetent cells respectively by electroporation (25 µF, 200 Ω, 2.5 kV). The transformation was plated on LB-agar-kanamycin plates. A single colony was picked for inoculation of the overnight culture. The overnight culture was diluted 1:25 and grown for 5 hours, 37°C, 220 rpm, then incubated for 1 hour on ice. This culture was then further diluted 1:10 in LB-rich-kanamycin media and protein expression was induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After induction, the cultures were grown for 24 hours, at 220 rpm and 30°C. Cells were harvested by centrifugation (4°C, 4000 xg, and 20 minutes).

2.6.2. Expression method 2: The transformation and plating followed the protocol described above. A single colony was picked and grown overnight at 37°C and 220 rpm in LB-kanamycin media. This culture was centrifuged for 20 minutes at 4°C and 4000 xg then re-suspended in LB-rich-kanamycin media. The re-suspended bacteria were diluted 1:20 in LB-rich media and grown at 20°C until it reached an OD of 0,8 and it was induced by 0,5 mM IPTG. The induced culture was grown overnight at 20°C and 220 rpm then cells were collected by centrifugation.

2.6.3. Expression method 3: The protocol follows the same steps as in expression method 2 with the exception that the expression media was supplemented with 3 % ethanol, 5 % glucose and 0,4 M sucrose respectively.

2.6.4. Expression method 4: TAT-SMN1 was co-transformed into *E. cloni* BL21 (DE3) with *pG-KJE8* plasmid (#3340, Takara), which codes for 5 different chaperons. The induction of the chaperons with 0,5 mg/ml L-arabinose and 0,5 ng/ml tetracycline happened when the cultures were diluted into the expression media. Four different expression media were used: LB-rich, LB-rich +3 % ethanol, LB-rich +5 % glucose and LB-rich+0,4 M sucrose.

2.6.5. Expression method 5: TAT-SMN1 was transformed and co-transformed with *pG-JKE8* into *Arctic express* DE3 chemically competent cells (#230192, Agilent) respectively. Transformation was carried out according to manual; protein expression was performed at 12°C.

2.6.6. Expression method 6: TAT-SMN1 and *pG-JKE8* were co-transformed and TAT-SMN1 was transformed into *Origami*^{TM} DE3 competent cells (#71345-3, Novagen). All steps and recommendations were followed in the manual. The protein was expressed at 20ºC.

The expression methods are summarized below (**Table 2**).

strain	plasmids	growth temperature	time of induction	expression temperature	growth media
E. cloni BL21	TAT-SMN1	37°C	after 5 hours	30°C	LB-rich
E. cloni BL21	TAT-SMN1	37°C	OD=0.8	20°C	LB-rich
E. cloni BL21	TAT-SMN1	37°C	OD=0.8	20°C	LB-rich+3% ethanol
E. cloni BL21	TAT-SMN1	37°C	OD=0.8	20°C	LB-rich+0.4M sucrose
E. cloni BL21	TAT-SMN1	37°C	OD=0.8	20°C	LB-rich+5% glucose
E. cloni BL21	TAT-SMN1+pG-KJE8	37°C	OD=0.8	20°C	LB-rich
E. cloni BL21	TAT-SMN1+pG-KJE9	37°C	OD=0.8	20°C	LB-rich+3% ethanol
E. cloni BL21	TAT-SMN1+pG-KJE10	37°C	OD=0.8	20°C	LB-rich+0.4M sucrose
E. cloni BL21	TAT-SMN1+pG-KJE11	37°C	OD=0.8	20°C	LB-rich+5% glucose
Arctic express	TAT-SMN1	37°C	OD=0.8	12°C	LB-rich
Arctic express	TAT-SMN1+pG-KJE11	37°C	OD=0.8	12°C	LB-rich
Origami DE3	TAT-SMN1	37°C	OD=0.8	20°C	LB-rich
Origami DE3	TAT-SMN1+pG-KJE11	37°C	OD=0.8	20°C	LB-rich

Table 2. Combination of strains of bacteria with expressed plasmids, expression temperature and growth media used for protein expression.

2.7. Lysis methods of TAT-SMN1 and SMN1 overexpressed bacteria

Numerous different lysis buffers were screened to lyse inclusion bodies in the mildest way possible. All lysis buffers were pH=8.0 and sterile filtered. Before lysis the following compounds were freshly added: 5 mM β-mercaptoethanol, 50 µg/ml lysozyme, 25 U/ml benzoase nuclease, 100 µg/ml PMSF and 1 X protease inhibitor cocktail.

Denaturing lysis buffers:

- 8 M urea, 100 mM Tris, 500 mM NaCl, 10 % glycerol, 1 mM EDTA
- 6 M GCl, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 4 M GCl, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA

Mild lysis buffers:

- 5 % DMSO, 2 M urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % n-propanol, 2 M urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA

- 0,2 % lauryl sarcosine, 2 M urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 6 M n-propanol, 2 M urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % DMSO, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % n-propanol, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % DMSO, 250 mM urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % DMSO, 500 mM urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % DMSO, 750 mM urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 5 % DMSO, 1 M urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 2 M GCl, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA
- 1 M GCl, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA

Non-denaturing lysis buffer:

500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA

In case of non-denaturing lysis, the pellets were thawed on ice and re-suspended by pipetting and vortexing with the lysis buffer. Incubation followed for 30 minutes on the shaker on ice. Then the cell suspension was sonicated 6 X 45 seconds with 2 minutes of cooling break. Sonicator (102C (CE), Branson) was set to output 2 and duty cycle 60 %. Lysed cells were transferred into centrifuge tubes and centrifuged for 30 minutes, at 4°C and 17.000 xg. The supernatant was used for the subsequent protein purification. All pellets from different expressions were lysed in non-denaturing buffer.

Denaturing purification was performed the same manner except for temperature. Incubation and centrifugation steps were at room temperature.

Mild lysis methods involve two major steps. First step is exactly follows as non-denaturing lysis. After the first step, the supernatant was discarded and the pellet (inclusion bodies) are re-suspended in the denaturing buffer, incubated for 30 minutes on RT and centrifuged for 50 minutes, RT and 17.000g. The supernatant was further purified. The denaturing and mild lysis methods were only applied to the *E. cloni* pellets containing just the TAT-SMN1 plasmid expressed at 20°C in LB-rich medium.

2.8. Protein purification

3.8.1. His-Gravi Trap Sepharose Column: Bacterial cultures were lysed in 8 M urea, 100 mM Tris, 500 mM NaCl, 10 % glycerol, 1 mM EDTA, pH=8.0 and 5 % n-propanol, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA pH=8.0 buffers as described above respectively. The denaturing purification was performed at RT and the semi-denaturing at 4ºC. The workflow of the two purifications was the same with the only exception of the different lysis buffers. First step was to drain the column of the storage buffer then add 5 ml of lysis buffer and 500 mM imidazole and 10ml of lysis buffer+20 mM imidazole. The column was then equilibrated with lysis buffer and 20 mM imidazole, 5 mM β-ME. Lysate was applied and the flow through was collected. Excess and unbound material was washed away with lysis buffer, 20 mM imidazole and 5 mM β-ME. Wash fractions were collected to be able to examine the quality of the washing. Protein was finally eluted with lysis buffer+500 mM imidazole+ 5 mM β-ME in the smallest volume possible to have a pure but highly concentrated protein.

3.8.2. Strep Tactin Column: The lysate was diluted from 8 M urea to 2 M urea for column usage. Storage buffer was drained and the column was equilibrated with 20 ml

of lysis buffer. Filtered lysate was applied to the column and flow through was collected. Column was washed with 50 ml of lysis buffer and 5mM β-mercaptoethanol and two wash fractions were collected. Protein was eluted with 30ml of lysis buffer and 5 mM β-ME and 50 mM biotin (after biotin was added pH was re-adjusted to 8.0). Three elution fractions were collected, and then all fractions were run on SDS-gel. The column was regenerated with 10 mM NaOH and then washed with storage buffer.

2.8.3. Ion exchange chromatography: Before the Enrich S Cation Exchange Chromatography, TAT-SMN1 was buffer exchanged on a PD10 column into a low salt and high urea containing buffer (6 M urea, 20 mM Tris, 200 mM NaCl, 10 % glycerol, pH=7.0) then concentrated to 2 ml and loaded on the cation exchange column. Before loading the protein, column was washed with water and with low salt buffer. For purification, the EnrichSVaryFlows program was used. Elution was performed with an increasing salt concentration by loading the mixture of a low and high salt buffer. Low salt buffer contained: 20 mM Tris, 200 mM NaCl, 6 M Urea, 10 % glycerol pH = 7.0 and high salt buffer contained: 20 mM Tris, 1 M NaCl, 6 M Urea, 10 % glycerol pH = 7.0. TAT-SMN1 has eluted at a NaCl concentration of approximately 680 mM. After several attempts, the concentration step to 2ml was discarded and the sample was loaded with a 50 ml superloop. In the case of SMN1, an anion exchange column was used and the low salt buffer contained 6 M urea, 10 % glycerol, 20 mM Tris, 80 mM NaCl, 0,4 mM TCEP (#C4706, Sigma) pH=9.0. The high salt buffer's components were 6 M urea, 10 % glycerol, 20 mM Tris, 1 M NaCl, 0,4 mM TCEP pH=9.0. SMN1 has eluted from the anion exchange column at a NaCl concentration of approximately 320 mM. After the proteins were purified on ion exchange column they were used for buffer exchange to the storage buffers.

2.8.4. PD-10 desalting column: All steps were carried out 4ºC. Column was drained then washed with 10ml of sterile H_2O, and with 15 ml of 20 mM Tris buffer with the same pH as the storage buffer. It was equilibrated with 10 ml of storage buffer. A maximum volume of 2,4 ml of the protein was applied to the column and buffer exchanged into 10 ml of storage buffer. 1 ml fractions were collected. Protein containing fractions were determined with Bradford protein assay and SDS-gel.

Storage buffers of TAT-SMN1:

- 200 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 0,5 mM EDTA pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 1 mM TCEP pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 300 mM Arginine pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 1 mM TCEP, 0,05 % CHAPS pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 2 mM TCEP pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 1 % bovie serum albumin (BSA, #A9674-50G, Sigma) pH=7.0
- 400 mM NaCl, 20 mM Tris, 30 % glycerol pH=7.0
- 400 mM NaCl, 2 M urea, 20 mM Tris, 10 % glycerol pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 1 M prolin pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 400 mM lysine pH=7.0

- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 400 mM Arginine pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 400 mM Arginine+1 mM TCEP pH=7.0

Storage buffers of SMN1:

- 200 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0
- 400 mM NaCl, 20 mM Tris, 10 % glycerol, 1 mM TCEP pH=8.0

2.8.5. Dyalisis: Both His -Gravi Trap and Strep TActin column elution of TAT-SMN1 was dyalised with Slide A lyser cassette against 200 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0, 400 mM NaCl, 20 mM Tris, 10 % glycerol, pH=7.0 respectively.

2.9. Protein concentration with 10kD MWCO

The benchtop centrifuge was pre-cooled to 4ºC and the membrane of the concentrator was pre-washed with the same buffer in which the protein was. The concentrator was spun in the centrifuge for 10 minutes then the buffer was discarded. The protein was transferred into the concentrator and was centrifuged at 2000 xg until final volume was reached. Flow through was discarded.

Summarized purification methods are in table 3 below (**Table 3**).

Lysis buffer	Strep	His	Ion-exchange	PD10	Dyalisis	MWCO	Storage buffer
8M urea		X			X		200mM NaCl
8M urea		X		X		X	200mM NaCl
8M urea	X			X		X	200mM NaCl
8M urea	X		X	X		X	200mM NaCl
8M urea	X		X	X		X	400mM NaCl
8M urea	X		X	X		X	0,5mM EDTA
8M urea	X		X	X		X	1mM TCEP
8M urea	X		X	X		X	300mM Arginine
8M urea	X		X (superloop)	X			1mM TCEP, 0,05% CHAPS
8M urea	X		X (superloop)	X			1mM TCEP
5% DMSO, 2M urea	X			X			400mM NaCl
5% DMSO, 2M urea	X			X			1mM TCEP
5% DMSO, 2M urea	X			X			1% BSA
5% DMSO, 2M urea	X			X			30% glycerol
5% DMSO, 2M urea	X			X			2M urea
5% DMSO, 2M urea	X						refolded on strep column
5% n-propanol, 2M urea		X		X			2mM TCEP
5% n-propanol, 2M urea		X		X			400mM NaCl
5% n-propanol, 2M urea	X			X			400mM NaCl
5% n-propanol, 2M urea	X			X			1mM TCEP
5% n-propanol, 2M urea	X			X			2M urea
5% n-propanol, 2M urea	X			X			400mM lysine
5% n-propanol, 2M urea	X			X			1M proline
5% n-propanol, 2M urea	X				X		400mM NaCl
4M GuCl	X			X			400mM NaCl
4M GuCl	X			X			1mM TCEP
4M GuCl	X			X			1mM TCEP, 400mM Arginine
4M GuCl	X			X			400mM Arginine
4M GuCl	X			X			1M GuCl

Table 3. Combination of applied lysis methods, columns and storage buffers for purification of TAT-SMN1. Except the aforementioned components of the lysis buffer they contained: 500 mM NaCl, 100 mM Tris, 1 mM EDTA and 10 % glycerol, in the case of storage buffer the ingredients not shown here were: 20 mM Tris and 10 % glycerol. If salt concentration of the storage buffer is not specified, then it was 400 mM and the pH=7.0.

2.10. Storage of purified proteins

Different storage conditions were tested. Samples from the same purification were stored for 24 hours at room temperature, 4°C, -20°C, -80°C, snap-freezing in liquid ni-

trogen then -80°C and adding proteinase inhibitor cocktail and then snap-freezing and stored at -80°C. Next day samples were thawed and run on SDS-gel to determine the most efficient storage condition. Later on, all proteins were snap-frozen and stored at -80°C.

2.11. Purified protein concentration determination

Purified protein concentration after thawing was determined with Bradford Protein Assay according to manual.

2.12. Protein stability test

The examined proteins were incubated in a 37°C incubator for 8 hours. Samples were taken at 0, 60, 120, 240, 360 and 480 incubation time and then s. run on 12 % SDS-gel.

2.13. Mass spectrometry, Dynamic light scattering and Thermofluor

TAT-SMN1 was purified through Strep-Tactin column. For Thermofluor the purified protein in 2 M urea buffer was submitted. The same sample was used for the dynamic light scattering (DLS) experiments. For mass spectrometry TAT-SMN1 was loaded on a SDS-gel. The gel had been stained after running and then was de-stained overnight. The band of interest was cut out in sterile environment. All buffers used for the entire procedure were sterile filtered. The band was submitted for mass spectrometry. All three techniques were carried out by Vienna Bio Center (VBC).

2.14. Immunofluorescence

In case of NIH3T3 cell line, $1x10^5$ cells were plated in a 4 chamber labtech chamber and grown overnight or until it reached around 90-100 % confluency. Media was changed into a media which had the appropriate concentration of the various previously purified proteins. This media was on the cells for 5-6 hours (**Figure 7**).

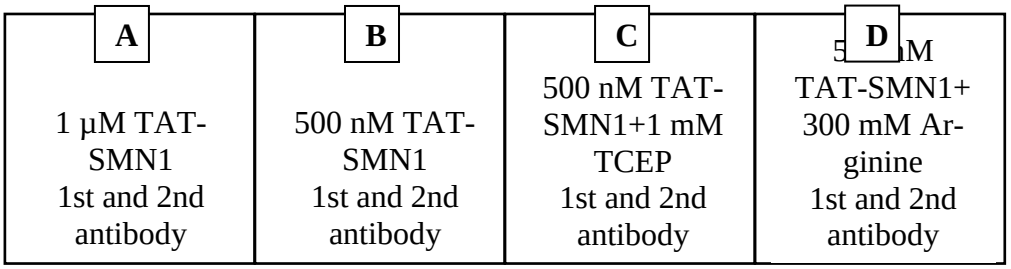

Figure 7. Labtech chamber seeded with NIH3T3 cells and incubated with **A** 1µM TAT-SMN1. **B.** 500 nM TAT-SMN1. **C.** 500 nM cation-exchange purified TAT-SMN1 with 1 mM TCEP in the storage buffer. **D.** 500 nM cation-exchange purified TAT-SMN1 with 300 mM Arginine in the storage buffer

To be able to cultivate NSC-34 cells on glass surface, the chambers had to be coated with 5 µg/cm^2 fibronectin (#F1141, Sigma). 7x10^5 cells were seeded and grown for 24 hours then incubated with the correct concentration of the proteins of interest for 5-6 hours. As a negative control, the same conditions were applied to two separate wells with the exception that one was only incubated with the secondary antibody (**Figure 8**).

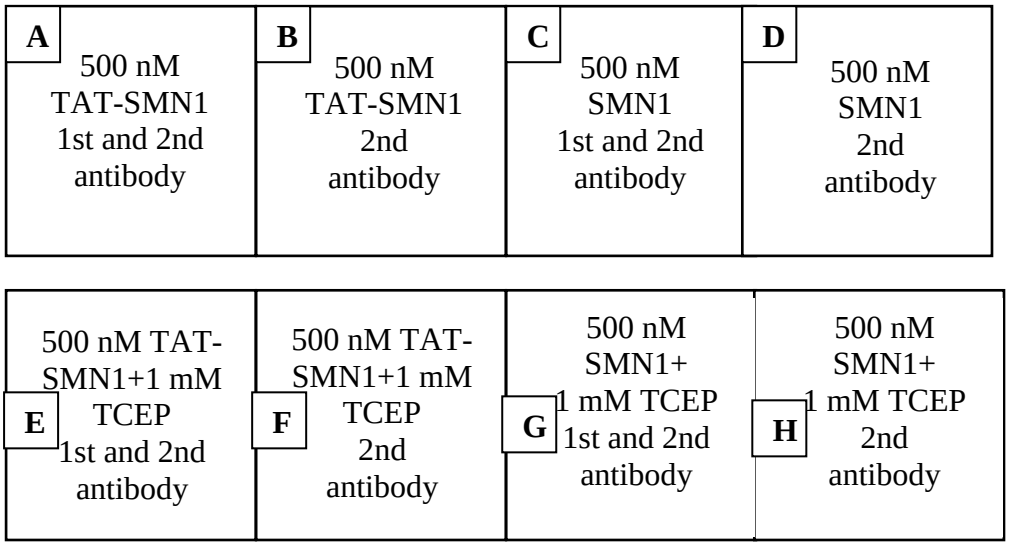

Figure 8. Labtech chambers seeded with NSC-34 cells. All proteins were used in 500nM concentration. **A.** TAT-SMN1. **B.** TAT-SMN1 as negative control. **C.** SMN1. **D.** SMN1 and only secondary antibody. **E.** TAT-SMN1 ion exchanged in TCEP supplemented storage buffer. D Same as E but only with secondary antibody. **G.** Incubation with SMN1 ion exchanged with TCEP in the storage buffer. **H.** negative control of G.

Media was removed from the wells and immediately after they were washed with ice-cold DPBS. Cells were fixed with 4 % formaldehyde for 15 minutes. To remove the formaldehyde, wells were washed 3x then the cells were permeabilized with 100 % ice-cold ethanol for 10 minutes. Ethanol was washed away and the cells were blocked for 1 hour at RT with 1 % BSA and 5 % normal goat serum (NGS). Then the blocking solution was removed and primary antibody was added: Strep MAB classic diluted 1:200 in 1 % BSA. Incubation with the primary antibody was at 4°C overnight. Following this incubation, solution was removed and cell layer was washed and then secondary antibody was added: Alexa goat antimouse 488 diluted 1:200. Incubation lasted for 1 hour, on RT, on shaker and in the dark. Incubation with secondary antibody was followed by

washing step and counterstaining with 4',6-diamidino-2-phenylindole (DAPI) diluted 1:500 for 5 minutes, on the shaker, in the dark. A final wash was performed and the chambers were then removed. The slides were covered with Fluoroprep and a coverslip and were let to dry and cool down. Images were prepared using confocal microscope.

2.15. Transduction of different types of TAT-SMN1 and SMN1 and cell lysis for western-blot

HSC and NSC-34 cells were seeded and knocked down as previously described. They were incubated with various protein samples for 8 hours. To remove proteins stuck on the membrane cells were treated with trypsin. NSC34 was treated with 0,004 % trypsin and HSC with 0,00025 % trypsin. After trypsinization, cells were washed with cold PBS and were lysed in RIPA buffer (50 mM Tris, 150 mM NaCl, 1 % SDS, 0,5 % Na-deoxycholat, 1 % NP40) supplemented with proteinase inhibitor cocktail. Cells were removed from the wells by cell scraper and transferred into an Eppendorf tube and incubated for 30 minutes on ice. Lysates were centrifuged in a 4°C benchtop centrifuge at 10.000g for 10 minutes. Supernatant was carefully pipetted into a fresh, precooled Eppendorf tube. Cell lysates were stored for long term at -80°C.

2.16. Western-blot

Cell lysates mixed with Sample Buffer were run on a 12 % SDS-gel as described above. The gel was blotted onto a nitrocellulose membrane 35 minutes 0,5 A in a blotting tank with blotting buffer in it at 4°C. After blotting, if the blot needed to be cut then it was stained with Poncaeu red, cut and de-stained with water. The blot was then blocked in 5 % milk powder in PBS-Tween. For SMN1 detection, purified mouse anti-SMN1 (BD biosciences, 610647) primary antibody was used in a dilution of 1:5000 in

5 % milk powder. As loading control, tubulin was used and to detect monoclonal anti-β-tubulin (#T4026, Sigma) was diluted 1:2000 in 5 % milk powder. With both antibodies blot was incubated overnight at 4°C. Before incubation with the secondary antibodies, the blot was washed and incubated with secondary antibody affinity purified AB peroxidase labelled goat-anti-mouse (#04-18-06, KPL) in both cases diluted 1:10.000. Incubation lasted 1 hour at RT. When the incubation was finished, the blot was washed again and then Chemiglow stable peroxidase buffer and luminol (#60-12596-00, Protein simple) was mixed 1:1 and pipetted on the blot. After 5 minutes of incubation, bands were detected with chemi imager. If the blot needed to be re-probed then it was stripped. The first step was washing the blot 3 times with PBS then rolling it into a flask with stripping buffer with β-ME in it. The flask was incubated in a 56°C water bath with a shaking bottom for 20 minutes. Afterwards the blot was washed again thrice and blocked again.

3. Results

3.1. Lysis in native lysis buffer of pellets produced by various expression methods

SMN1 pellets expressed at 20°C were lysed in native buffer and purified with the Strep-column. Samples were taken throughout the purification and run on SDS-gel. The protein was only present in inclusion bodies (**Figure 9**).

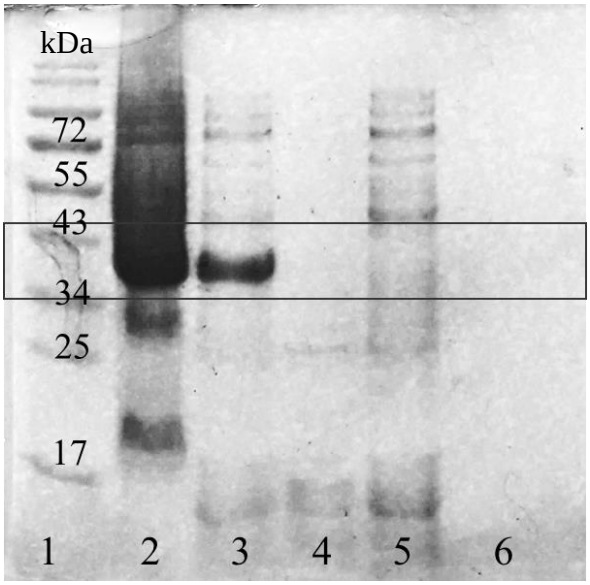

Figure 9. Lysis of SMN1 pellet produced at 20°C in LB-rich media. Lane 1: Marker. Lane 2: Pellet after centrifugation of the lysate. Lane 3: Lysate before centrifugation. Lane 4: Second wash fraction of the Strep column. Lane 5: First wash fraction of the Strep column. Lane 6: Elution fraction 2 (from 5ml to 17ml) of the Strep column.

TAT-SMN1 was expressed at 20°C in LB-rich media and 0,4 M sucrose, LB-rich media and 5 % glucose and LB-rich media and 2 % ethanol respectively. TAT-SMN1 was also co-expressed with protein chaperones at the same temperature and in the same media with LB-rich media additionally. These pellets were lysed in native lysis buffer. After lysis, one could determine that the protein is still located in inclusion bodies in all cases (**Figure 10**).

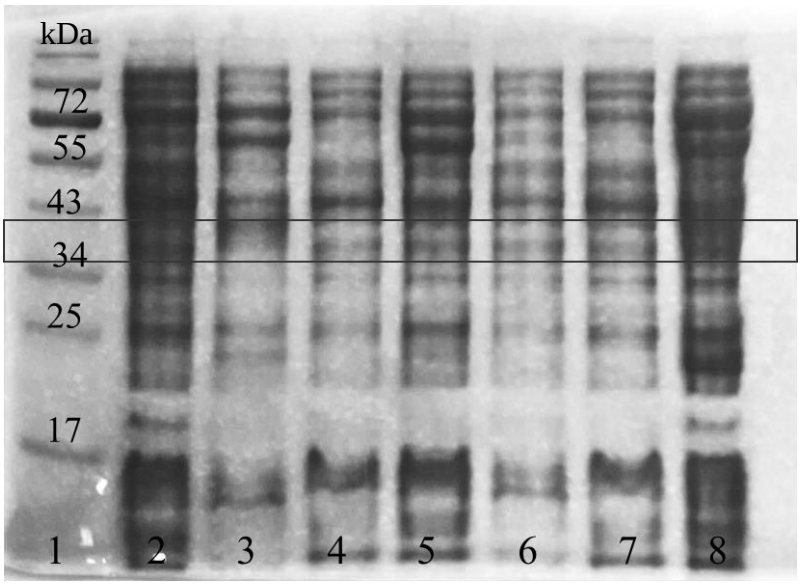

Figure 10. Lyses in native buffer of TAT-SMN1 and TAT-SMN1+*pG-KJE8* produced in different LB-rich media. Lane 1: Marker. Lane 2: TAT-SMN1 produced in LB-rich + 0,4 M Sucrose. Lane 3: TAT-SMN1 + protein chaperones produced in LB-rich + 0,4 M sucrose. Lane 4: TAT-SMN1 produced in LB-rich + 5 % glucose. Lane 5: TAT-SMN1 + *pG-KJE8* produced in LB-rich + 5 % glucose. Lane 6: TAT-SMN1 produced in LB-rich + 2 % ethanol. Lane 7: TAT-SMN1 + protein chaperone produced in LB-rich + 2 %ethanol. Lane 8: TAT-SMN1 + protein chaperones produced in LB-rich medium.

TAT-SMN1 and TAT-SMN1 + protein chaperones were also expressed in *Arctic express* DE3 cells at 12°C and the pellets were lysed in native lysis buffer. TAT-SMN1 was still only present in inclusion bodies. TAT-SMN1+*pG-KJE8* was co-expressed in *Origami*^{TM} DE3 in order to produce native protein. The bacteria were lysed in native buffer. We were not able to detect TAT-SMN1 in the soluble fraction of the lysate (**Figure 11**).

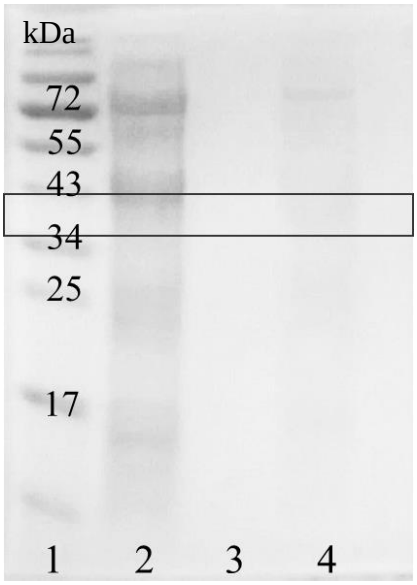

Figure 11. Native lysis of TAT-SMN1 and *pGKJ-E8* produced in *Origami*^{TM} DE3. Lane 1: Marker. Lane 2: Supernatant of lysate after centrifugation. Lane 3: Empty. Lane 4: Induced bacteria culture before centrifugation.

3.2. Different lysis methods of TAT-SMN1 pellets produced at 20°C in LB-rich medium

It was previously established that the protein could be extracted from inclusion bodies with 8 M urea. In order to avoid problems occurring in the later purification steps, we tried out different lysis methods. Another denaturing agent instead of urea was used called guanidium-chloride. Various amounts were used: 6 M, 4 M, 2 M and 1 M. All four concentrations were sufficient for inclusion body lysis (**Figure 12**).

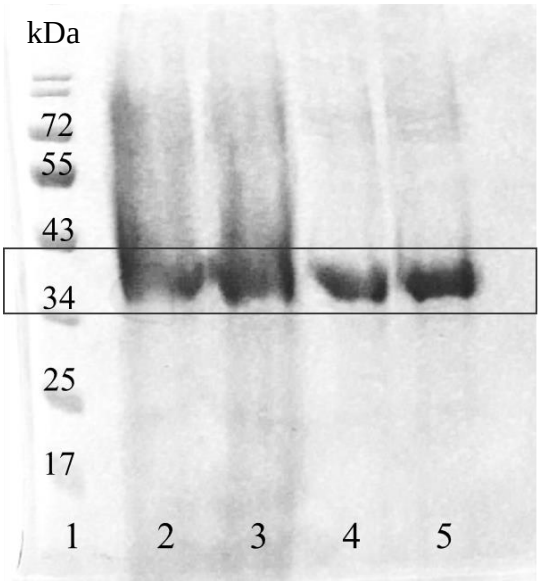

Figure 12. Lysis of TAT-SMN1 pellet produced at 20°C in LB-rich media in different concentration of Guanidium-chloride. Lane 1: Marker. Lane 2: Supernatant of lysate in 6 M GuCl. Lane 3: Supernatant of lysate in 4 M GuCl. Lane 4: Supernatant of lysate in 2 M GuCl. Lane 5: Supernatant of lysate in 1 M GuCl.

Also, partially denaturing lysis methods were studied. All buffers contained 2M urea next to additional substances like DMSO, n-propanol (5 % and 6 M) and lauryl sarcosine. 5 % DMSO and 5 % n-propanol was successful but 6 M n-propanol and lauryl sarcosine possibly lysed the inclusion bodies and also degraded the protein (**Figure 13**).

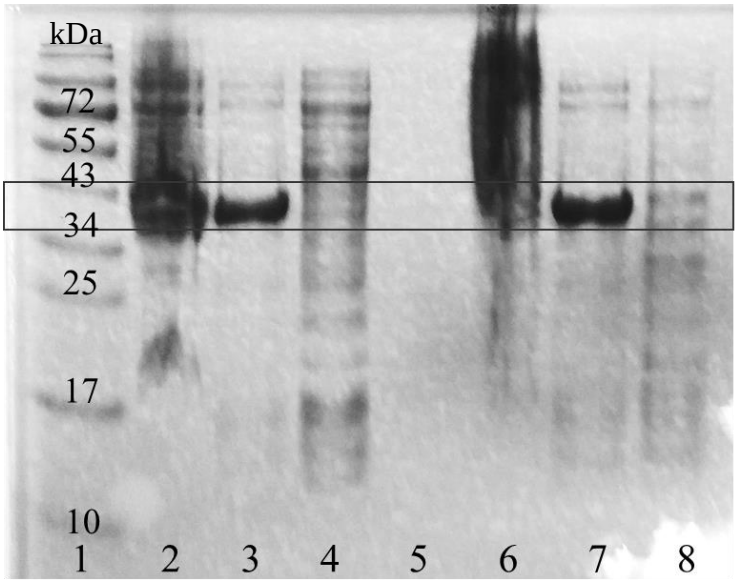

Figure 13. Lysis of TAT-SMN1 pellet produced at 20°C in LB-rich media in various mild lysis buffers with 2 M urea. Lane 1: Marker. Lane 2: Pellet of lysate lysed in 5 % DMSO after centrifugation. Lane 3: Supernatant of lysate lysed in 5 % DMSO after centrifugation. Lane 4: Supernatant of lysate lysed in 6 M n-propanol after centrifugation. Lane 5: Empty. Lane 6: Pellet of lysate lysed in 5 % n-propanol after centrifugation. Lane 7: Supernatant of lysate lysed in 5 % n-propanol after centrifugation. Lane 8: Supernatant of lysate lysed in 0,2 % lauryl sarcosine.

Based in the previous results, 5 % DMSO and 5 % n-propanol buffer without urea was used on both TAT-SMN1 and SMN1 pellets. In case of TAT-SMN1, no protein was detected. SMN1 pellets were originally bigger in size resulting in more protein. This would explain the low amount of protein from the TAT-SMN1 pellets and the higher amount from SMN1 (**Figure 14**).

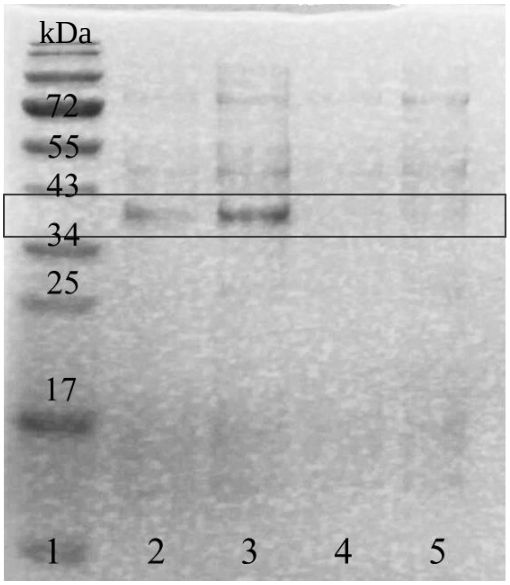

Figure 14. Lysis of SMN1 and TAT-SMN1 produced at 20°C in LB-rich media in 5 % DMSO and 5 % n-propanol. Lane 1: Marker. Lane 2: SMN1 lysed with 5 % DMSO. Lane 3: SMN1 lysed in 5 % n-propanol. Lane 4: TAT-SMN1 lysed in 5 % DMSO. Lane 5: TAT-SMN1 lysed in 5 % n-propanol

After the lysis with urea-free buffers was not successful, increasing concentration of urea was added to the 5 % DMSO buffer. Slightly more protein was present in the supernatant but not sufficient amount for further purifications (**Figure 15**). Therefore, further purifications were performed with 8 M urea, 5 % DMSO with 2 M urea, 5 % n-propanol with 2 M urea and 4 M guanidium-chloride.

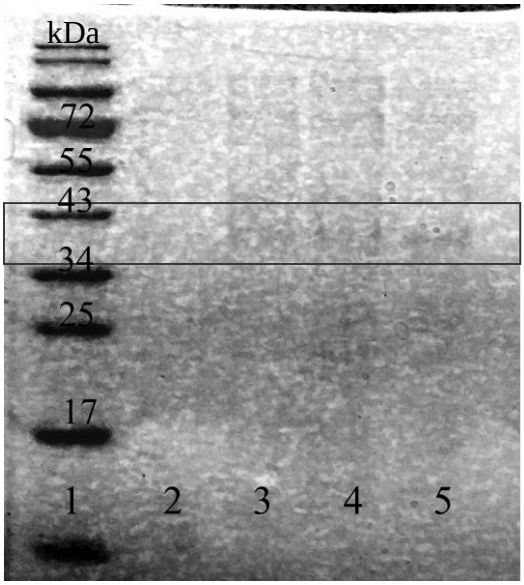

Figure 15. Lyses of TAT-SMN1 pellet produced at 20°C in LB-rich media with various concentrations of urea and 5 % DMSO. Lane 1: Marker. Lane 2: Supernatant after centrifugation lysed with 250 mM urea. Lane 3: Supernatant after centrifugation lysed with 500 mM urea. Lane 4: Supernatant after centrifugation lysed with 750 mM urea. Lane 5: Supernatant after centrifugation lysed with 1 M urea.

3.3. Denaturing purification of TAT-SMN1 and SMN1 on Strep column

All denaturing purifications with 8 M urea gave rise to final protein samples with varying concentrations. In case of TAT-SMN1, the highest final concentration after Strep column, PD-10 columns was 962 ng/µl and with the same procedure SMN1's concentration was 632 ng/µl. Protein concentration in all cases were determined with Bradford Protein Assay (#5000006, BioRad). TAT-SMN1 came off from the PD10 column mainly in the fractions from 5 ml to 7 ml (**Figure 16**) but on the other hand SMN1 came off from fraction 2 to 6 (**Figure 17**). Even though the concentration of the proteins were sufficient for downstream experiments, however, both Western blot and immunofluorescence did not show TAT-SMN1 transduction (**Figure 18**).

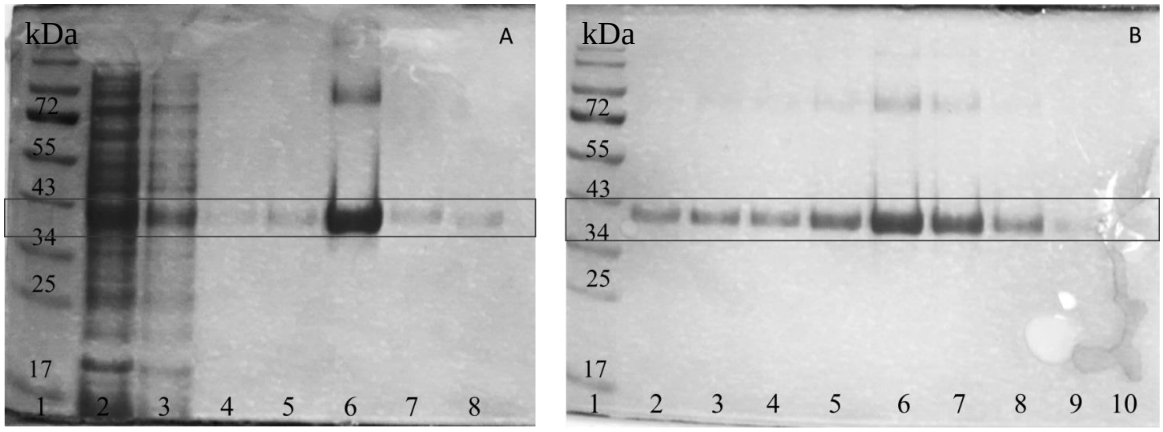

Figure 16. Denaturing purification of TAT-SMN1 with 8 M urea on Strep column and PD-10. **A.** fractions of the Strep column. Lane 1: Marker. Lane 2: Flow-through. Lane 3: First wash fraction (40 ml). Lane 4: Second wash fraction (10 ml). Lane 5: First elution (4 ml). Lane 6: Second elution (12 ml). Lane 7: third elution (14 ml). Lane 8: First elution of PD-10. **B.** Elution fractions of PD-10. Lane 1: Marker. Lanes 2-10: from second to tenth elution fractions of PD-10.

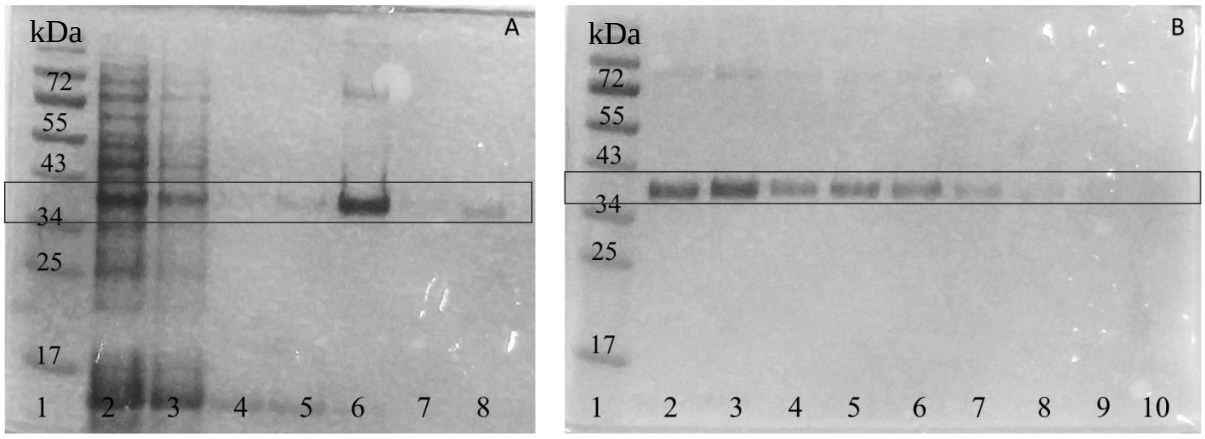

Figure 17. Denaturing purification of SMN1 with 8M urea on Strep column and PD-10. A. fractions of the Strep column. Lane 1: Marker. Lane 2: Flow-through. Lane 3: First wash fraction (40 ml). Lane 4: Second wash fraction (10 ml). Lane 5: First elution (4 ml). Lane 6: Second elution (12 ml). Lane 7: Third elution (14 ml). Lane 8: First elution of PD-10. **B.** Elution fractions of PD-10. Lane 1: Marker. Lanes 2-10: From second to tenth elution fractions of PD10.

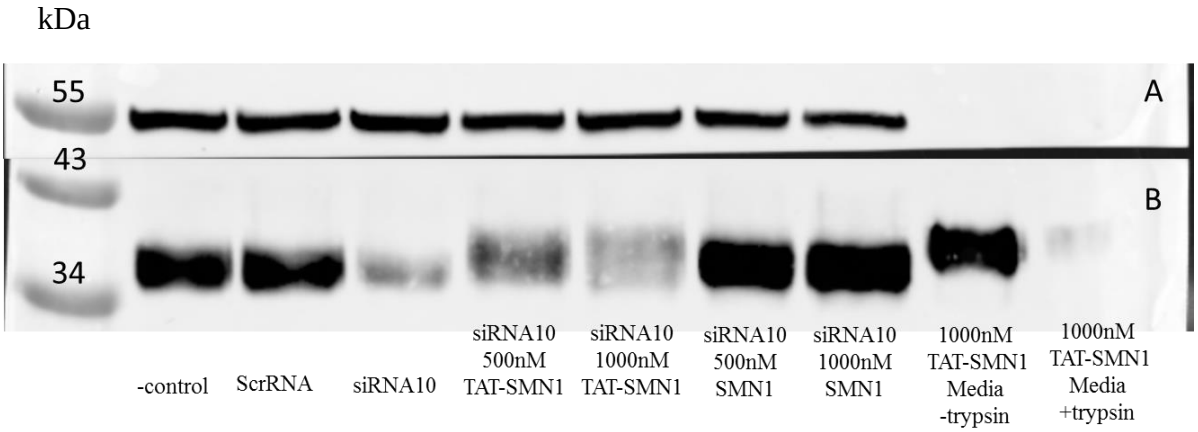

Figure 18. Western blot of SMN1 knockdown HSC cells incubated with TAT-SMN1 and SMN1. A. anti ß-tubulin.**B.** anti SMN1.

3.4. Purifications with dialysis

Dialysis on TAT-SMN1 was attempted two times. The first time, the pellets were lysed in 8 M urea, 100 mM Tris, 500 mM NaCl, 10 % glycerol, 1 mM EDTA, purified using the His column and then dialyzed against 200 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0. The second time, the pellets were lysed in 5 % n-propanol, 2 M urea, 500 mM NaCl, 100 mM Tris, 10 % glycerol, 1 mM EDTA, purified through Strep column and then dialyzed against 400 mM NaCl, 20 mM Tris, 10 % glycerol pH=7.0. In both cases after dialysis visible precipitation could be observed in the dialysis cassette. The concentration of the protein could not be measured and no visible protein signal was seen on a SDS gel.

3.5. TAT-SMN1 purification with cation exchange column

TAT-SMN1 was purified on the Strep column with 8M urea always successfully. It was followed by buffer exchange and concentration then the ion exchange column. Finally, TAT-SMN was buffer-exchanged into the storage buffer. After several very large decreases in protein concentration after freezing and thawing, we tried different

storage conditions. We concluded that higher salt concentration is more beneficial for the protein and our regular storage method is the best option (**Figure 19**).

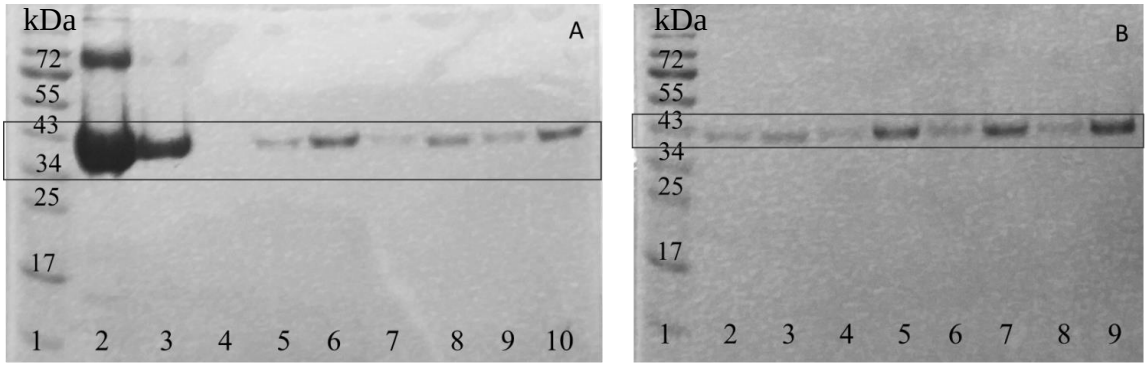

Figure 19. Different storage conditions of TAT-SMN1. A. Lane 1: Marker. Lane 2: Concentrated protein loaded on the ion exchange column. Lane 3: Cation exchange column elution. Lane 4: Cation exchange column flow-through. Lane 5: 200 mM NaCl, before freezing. Lane 6: 400 mM NaCl, before freezing. Lane 7: 200 mM NaCl, RT. Lane 8: 400 mM NaCl, RT. Lane 9: 200 mM NaCl, 4°C. Lane 10: 400 mM NaCl, 4°C. **B.** Lane 1: marker. Lane 2: 200 mM NaCl, -20°C. Lane 3: 400 mM NaCl, -20°C. Lane 4: 200 mM NaCl, -80°C. Lane 5: 400 mM NaCl, -80°C. Lane 6: 200 mM NaCl, snap-freezing, -80°C. Lane 7: 400 mM NaCl, snap freezing, -80°C. Lane 8: 200 mM NaCl and Proteinase inhibitor cocktail (PIC), snap freezing, -80°C. Lane 9: 400 mM NaCl and PIC, snap freezing, -80°C.

Based on previous work, we suspected that the problem could be aggregation. There are several available compounds that can be used to suppress protein aggregation. We tried arginine, TCEP, EDTA and all three combined in the final storage buffer. The final buffer without additives contained: 400 mM NaCl, 20 mM Tris, 10 % glycerol pH=7. The final concentrations we reached are shown in the table below (**Table 4**) (**Figure 20**).

storage buffers	final concentration
no addatives	~100ng/µl
1mM TCEP	318ng/µl
0.5mM EDTA	~100ng/µl
300mM arginine	341ng/µl
1mM TCEP, 0.5mM EDTA,300mM arginine	159ng/µl

Table 4. The achieved final concentration of TAT-SMN1 in different storage buffers.

Arginine and TCEP reduced aggregation to a similar extent. However, taking into consideration that high arginine amounts may affect the cell culture experiments, and the fact the Thermofluor experiment also showed that TCEP is our best option we chose to proceed with TCEP.

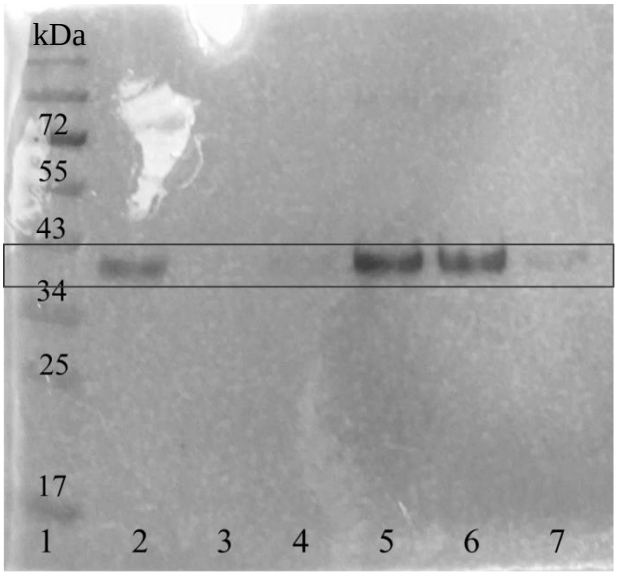

Figure 20. Lane 1: marker. Lane 2: Cation exchange column elution. Lane 3: Final protein with no additives (400 mM NaCl, 20 mM tris, 10 % glycerol pH=7.0). Lane 4: Final protein and 0,5 mM EDTA. Lane 5: final protein and 1 mM TCEP. Lane 6: Final protein and 300 mM arginine. Lane 7: final protein, 0,5 mM EDTA, 1 mM TCEP and 300 mM arginine.

To further increase the concentration of both proteins, both concentration steps during purification were removed. This resulted in a much higher yield of protein: 879 ng/µl and based on the SDS-PAGE no protein was lost due to freezing and thawing (**Figure 21**). The same procedure was applied to SMN1 purification resulting in 632 ng/µl final concentration.

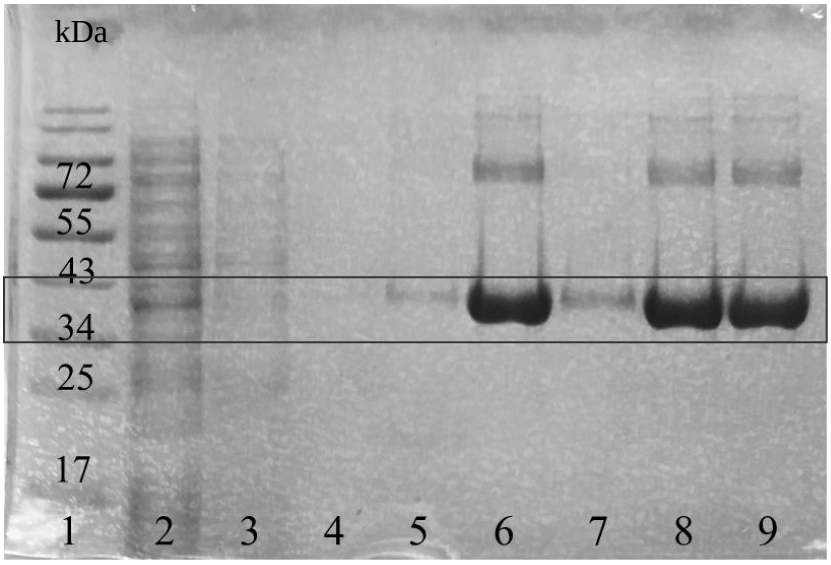

Figure 21. TAT-SMN1 purification with cation exchange column, without concentration and with 1 mM TCEP in the storage buffer. Lane 1: Marker. Lane2: Flow-through of Strep column. Lane 3: First wash fraction of Strep column (40 ml). Lane 4: Second wash fraction of Strep column (10 ml). Lane5: First elution (4 ml). Lane 6: Second elution (12 ml). Lane7: Third elution (9 ml). Lane 8: Final TAT-SMN1 sample before freezing. Lane 9: Final TAT-SMN1 sample after thawing.

Proteins previously purified without ion exchange as well as their cation exchange-purified counterparts were both subjected to a protein stability test. The test showed that none of the proteins are stable for 8 hours at 37°C. In all cases around there is a decrease in about 50 % of the band intensity in the SDS-PAGE (**Figure 22**). The downstream experiments were still performed and confirmed the conclusion about aggregation. Based on the western blot we could not detect transduction of TAT-SMN1 (**Fig-**

ure 23). Immunofluorescence showed that TAT-SMN1 is less prone to aggregation than SMN1 (**Figure 24**). Also, the introduction of the ion exchange column made the aggregation problem more prevalent. To overcome this obstacle mild lysis methods were introduced in the following purifications.

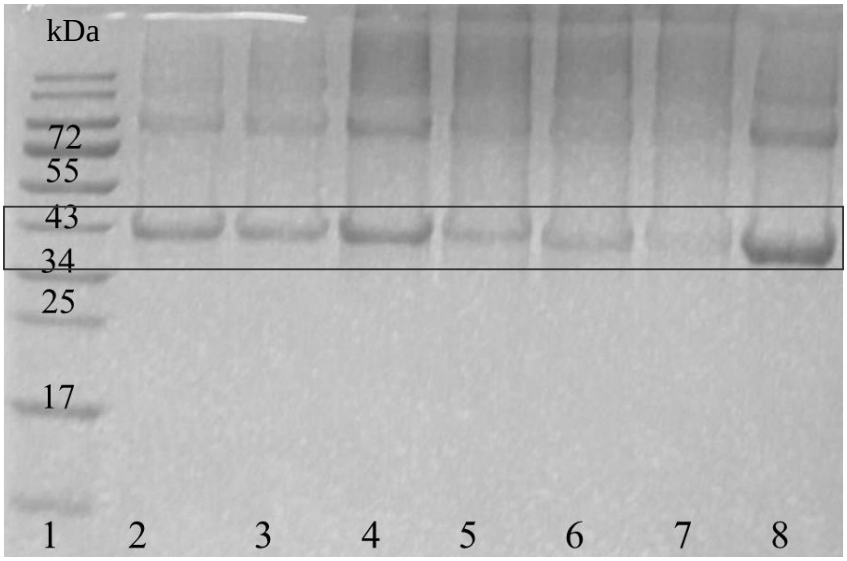

Figure 22. Stability test of TAT-SMN1, SMN1, ion exchanged TAT-SMN1 and ion exchanged SMN1. Lane 1: Marker. Lane 2: TAT-SMN1 0 hours. Lane 3: TAT-SMN1 8 hours. Lane 4: ion exchanged TAT-SMN1 0 hours. Lane 5: ion exchanged TAT-SMN1 8 hours. Lane 6: SMN1 0 hours. Lane 7: SMN1 8hours. Lane 8: ion exchanged SMN1 0 hours.

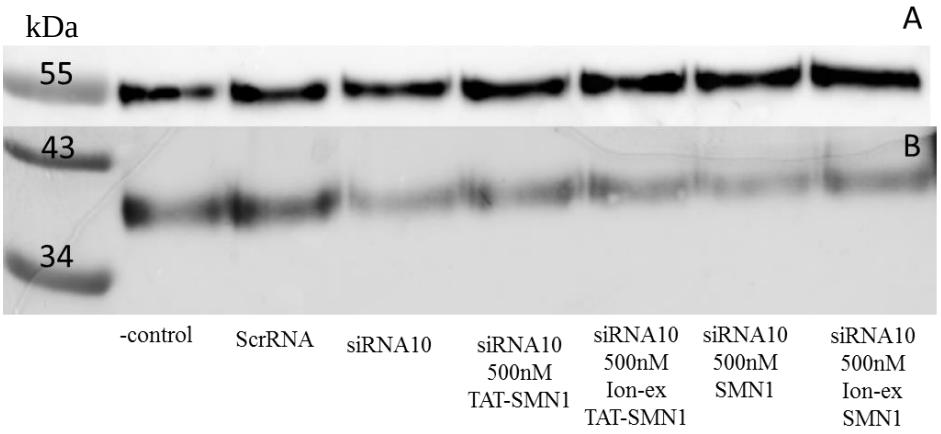

Figure 23: Western blot of knocked down NSC-34 incubated with different proteins. A. anti β-tubulin. **B.** anti SMN1.

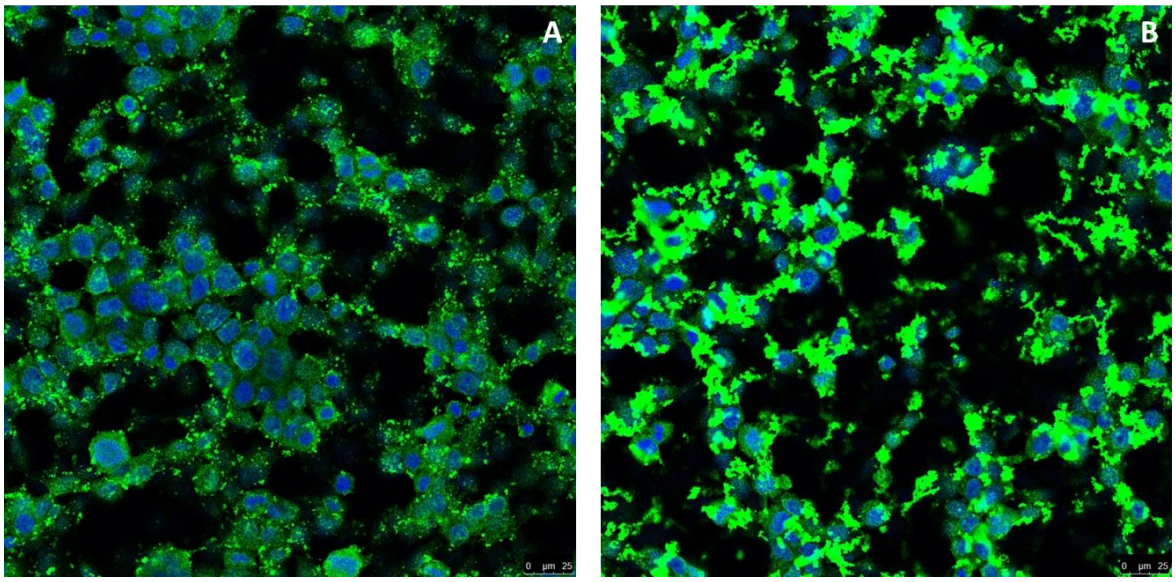

Figure 24. A. NSC-34 cells incubated with 500 nM ion exchanged TAT-SMN1 for 4 hours. **B.** NSC-34 cells incubated with 500 nM ion exchanged SMN1 for 4 hours.

3.6. Lysis with 5 % n-propanol and 2 M urea and purification with His column

There were several problems associated with purification TAT-SMN1 with the His column. There was always a big amount of protein in the flow through meaning that a lot of protein was not bound to the column and hence lost. Also, based on SDS gel an approximately 20 kDa band was always present when the protein was purified through

His column. This probably occurs because of the way the column interacts with the protein. From this point on only, Strep column was used for purification.

3.7. Lysis with 5 % DMSO and 2 M urea and refolding on the Strep column

Refolding of the denatured TAT-SMN1 happened on the Strep column. This was largely an unsuccessful experiment. The protein eluting from the Strep column was almost undetectable on an SDS gel.

3.8. Lysis with 5 % n-propanol and 2 M urea and purification with Strep column and PD-10 columns

We attempted to add various additives to the final storage buffer to see which would be the best in suppressing aggregation. 2 M urea to the storage buffer served as a positive control. Surprisingly the storage buffer without any aggregation suppressant was the best option but still some aggregation occurred compared to the positive control. This protein also failed to maintain its concentration throughout the protein stability test. This sample (1 mM TCEP which helped to reduce aggregation in case of denaturing purification) has shown the largest amount of aggregation in this case. Additives tried were: 1 mM TCEP, 1 % bovine serum albumin (BSA), 30 % glycerol, 1 M proline, 400 mM lysine (**Figure 25, Figure 26**).

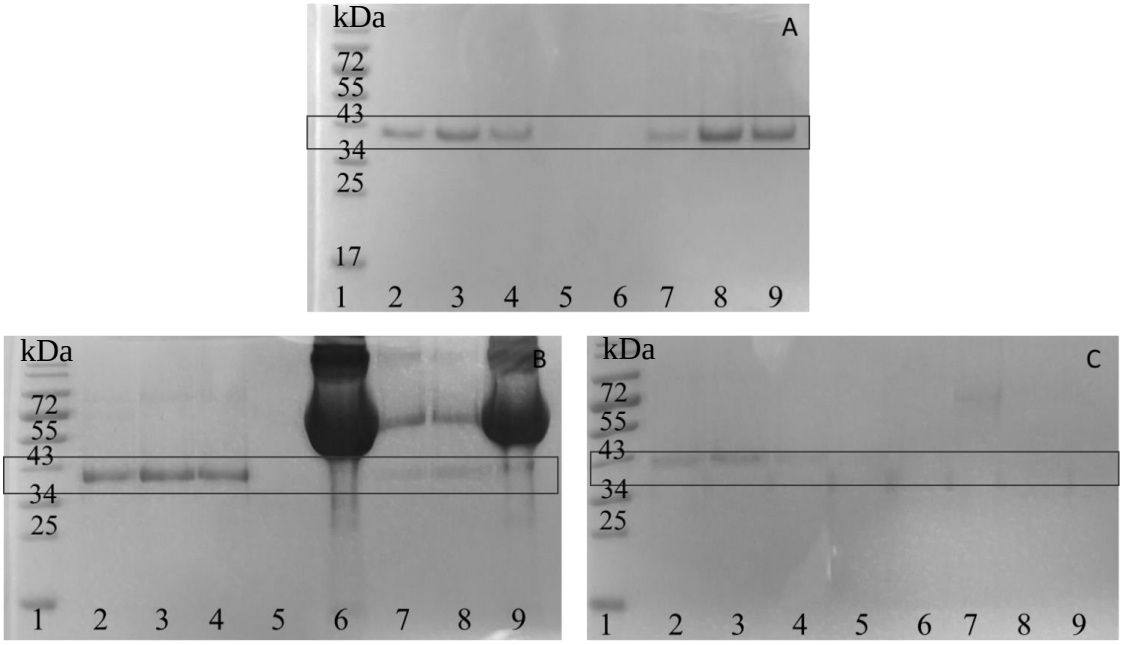

Figure 25. A. Lane 1: Marker. Lanes 2-4: PD-10 fractions with 1 mM TCEP. Lanes 5-6: Empty. Lanes 7-9: PD-10 fraction without additives (400 mM NaCl, 20 mM tris, 10 % glycerol pH=7.0). **B.** Lane 1: Marker. Lanes 2-4: PD-10 fractions with 2 M urea. Lane 5: empty. Lane 6-9: PD10 fractions with 1 % BSA. **C.** Lane 1: Marker. Lanes 2-9: PD-10 fraction with 30 % glycerol.

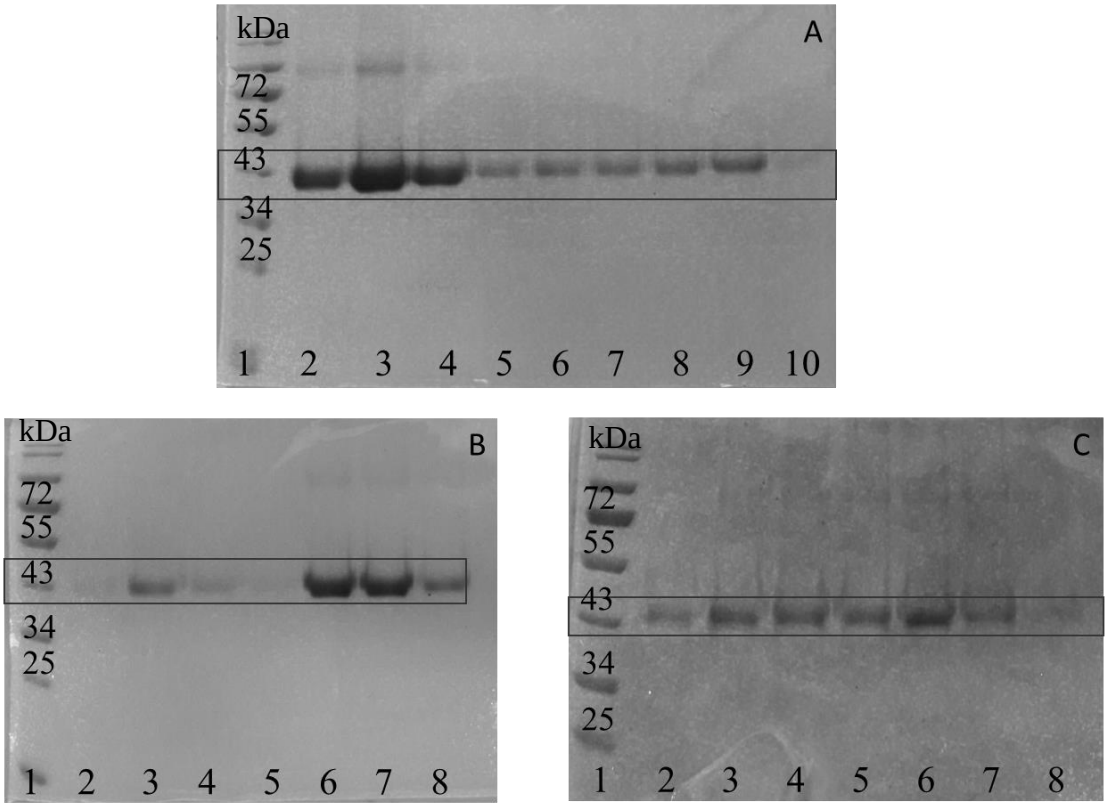

Figure 26. **A.** Lane 1: Marker. Lane 2-4: PD-10 fractions with 2 M urea. Lanes 5-10: PD-10 fractions with 1 M proline. **B.** Lane 1: marker. Lanes 2-5: PD-10 fractions with 1 mM TCEP. Lanes 6-8: PD-10 fraction without additives (400 mM NaCl, 20 mM tris, 10 % glycerol pH=7.0). **C.** Lane 1: Marker. Lanes 2-8: PD-10 fractions with 400 mM lysine.

3.9. Mass spectrometry, dynamic light scattering, Thermofluor

There is an upper band appearing on the gels in cases of both TAT-SMN1 and SMN1. The TAT-SMN1 upper band was submitted for mass spectrometry (**Figure 27**). The band was around the size of 80 kDa and was identified as an oligomer of TAT-SMN1 (**Table 5**).

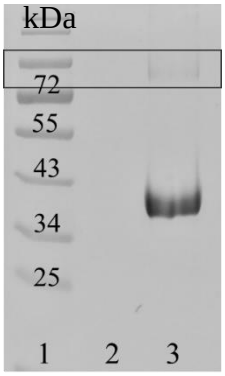

Figure 27. Band submitted for mass spectrometry. Lane 1: Marker. Lane 2: Empty. Lane 3: TAT-SMN1.

Protein IDs	Majority protein IDs	Protein names	Gene names	Number of proteins	Peptides	Razor + unique peptides	Unique peptides	Sequence coverage [%]	Mol. weight [kDa]	Sequence length	Q-value	Score	Intensity	iBAQ	MS/MS count
Q16637-X	Q16637-X	Survival motor neuron protein	SMN1	1	22	22	22	56.3	35.5	325	0	323.3	8.36E+10	9.29E+09	463
P0A6Y8	P0A6Y8	Chaperone protein DnaK	dnaK	1	34	34	34	49.1	69.1	638	0	323.3	2.42E+09	6.90E+07	94
P0A6M8	P0A6M8	Elongation factor G	fusA	1	28	28	28	51.7	77.6	704	0	323.3	7.15E+08	2.04E+07	55
P27302;P3	P27302	Transketolase 1	tktA	2	19	19	19	28.4	72.2	663	0	323.3	4.64E+08	1.66E+07	34
P76558	P76558	NADP-dependent malic enzyme	maeB	1	21	21	21	29.0	82.4	759	0	278.1	2.28E+08	5.56E+06	28
P0AC41	P0AC41	Succinate dehydrogenase flavop	sdhA	1	15	15	15	30.1	64.4	588	0	323.3	2.66E+08	8.58E+06	22
P0A853	P0A853	Tryptophanase	tnaA	1	15	15	15	34.2	52.8	471	0	178.1	1.44E+08	6.25E+06	16
P77398	P77398	Bifunctional polymyxin resistan	arnA	1	15	15	15	22.9	74.3	660	0	158.3	1.04E+08	2.74E+06	16

Table 5. Mass spectrometry analysis of the upper SDS-PAGE band of TAT-SMN1. First eight entries best protein matches are listed.

The DLS results have clearly demonstrated the protein aggregation. All the buffer conditions which were tested showed the same results. The majority of the protein (98,8 % of peak intensity) forms large aggregates which are around 550 nm and that correlates to 8755220 kDa. Only a small amount (1,2 % peak intensity) is identified as non-aggregating TAT-SMN1 with an approximate 50kDa (**Figure 28**).

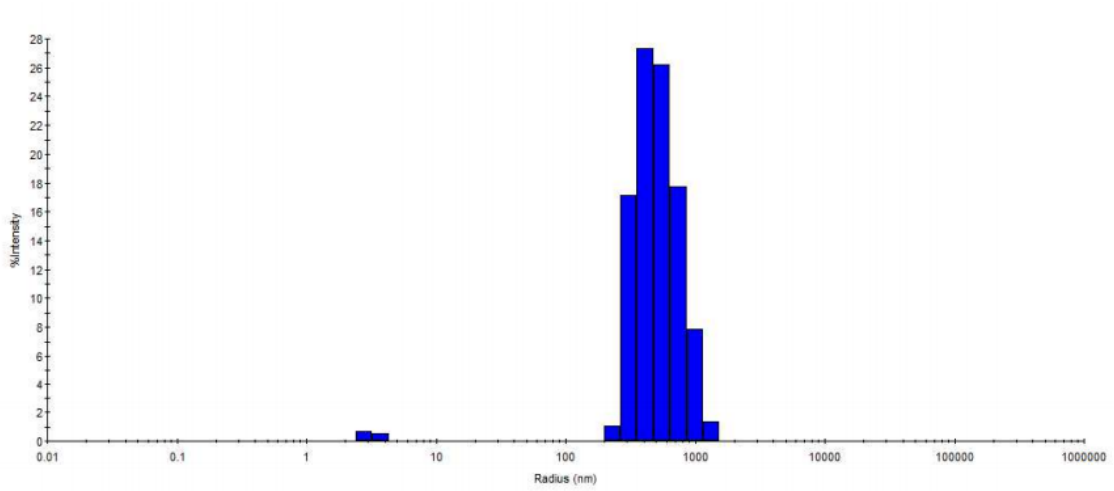

Figure 28. DLS results of TAT-SMN1. Smaller peak: monomeric TAT-SMN1 (~3,2 nm), bigger peak: aggregated TAT-SMN1 (~550 nm).

4. Discussion

The main goal of this study was to purify TAT-SMN1 and SMN1 and transduce them into various cell lines, such as NSC-34, an *in-vitro* model used for SMA. SMN1 protein is largely unstructured (as many proteins involved in neurodegenerative diseases) with hydrophobic residues and it is likely that this makes it prone to aggregation. Apart from the unstructured regions it contains a Tudor domain in the middle of the protein. The Tudor domain is composed largely of β-sheets and it is highly conserved just as SMN1 (Bühler et al., 1999). Proteins possessing the Tudor domain are binding methylated lysine and arginine on their binding partners in order to promote the assembly of macromolecular complexes. These complexes are involved in RNA metabolism, DNA damage response and chromatin modification (Pek, Anand, & Kai, 2012). SMN1 binds SMN-interacting protein 1 (SIP1) and spliceosomal Sm protein through its Tudor domain. These are components of the spliceosomal uridine-rich small ribonucleoprotein (U-snRNP) (Bühler et al., 1999).

The proteins which have both unstructured and structured regions are often referred to as hybrid proteins or intrinsically disordered proteins (IDPs). These proteins possess scarce tertiary and/or secondary structure elements. They may have multiple binding partners because of their structural plasticity but also tend to aggregate under certain conditions (Uversky, 2015). We faced the problem of protein aggregation in the beginning of our experiments and from this point; the main aim was to purify stable, non-aggregating protein for downstream experiments. The challenge herein was that SMN1 expresses into inclusion bodies, which can already facilitate aggregate formation. First, we tried to decrease the expression temperature but we were unable to inhibit inclusion body formation. To eliminate the problem of aggregation of purified SMN1 and TAT-SMN1 proteins, we tried a number of expression variants. The protein

constructs were produced in *E.cloni* BL21 DE3, *Origami*^{TM} DE3 and *Arctic express* DE3 strains. The *Arctic express* strain allowed us to carry out the expression at 12ºC, however the protein was still present in inclusion bodies and not in the bacterial supernatant following cell disruption.

The next tested strain of bacteria was the *Origami*^{TM} DE3 which has a mutation in both, the thiredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which in turn facilitates correct disulphide bond formation. Expressing in *Origami*^{TM} DE3 did not solve the solubility problem of the protein (Sørensen & Mortensen, 2005). In all three *E. coli* strains we tried to co-express TAT-SMN1 with a *pG-KJE8* vector, which poses chaperons GroEL, GroES, DnaK, DnaJ and GrpE. There is evidence that overexpression of molecular chaperons with the protein of interest could help in correct protein folding and to result in a correct tertiary structure (de Marco et al., 2005). However, we were not able to detect soluble TAT-SMN1.

In another effort to make TAT-SMN1 soluble, different expression media additives were used. We added 0.4M sucrose, 5% glucose and 2% ethanol to LB-rich media respectively and combine them by co-expression with *pGKJE8* plasmid (Chhetri, Kalita, & Tripathi, 2015). Then we lysed the bacterial pellet with native lysis buffer and run the lysate on SDS-Page. However, we could not identify TAT-SMN1. We have therefore concluded that changing the expression host, decreasing the temperature or adding additives to the expression media could not eliminate the inclusion bodies. Another would have been co-expressing TAT-SMN1 and SMN1 together with a potential binding partners such as SIP1, Sm protein (Bühler et al., 1999) , spliceosomal snRNP proteins, HuD (Akten et al., 2011), hnRNP-R and gry-rbp/hnRNP-Q (Rossoll et al., 2002). These are just a few known binding partners that could have been co-expressed with TAT-SMN1 and SMN1 but for protein replacement therapy this is not an option.

Hence, we then turned our attention to the employment of lysis methods. We have already have observed in former experiments that TAT-SMN1 could be lysed from inclusion bodies with buffers containing 8 M urea. However, this method leads to complete denaturation of the protein and the consequentially causing refolding problems including aggregation. We also tested the transduction of TAT-SMN1 and SMN1 (lysed in 8 M urea, purified on Strep column and buffer exchanged on PD-10) but our results were not conclusive. In Western blot analysis we can assume a transduction through cell membrane of 500 nM and 1000 nM TAT-SMN1. However, we are not fully convinced that TAT-SMN1-fusion protein was inside the cell membrane or just adhering to it. As a negative control we did the same incubation with SMN1 (without TAT) and we detected a stronger transduction compared to TAT-SMN1. This in theory would be impossible but based on the immunofluorescence pictures we could say that the signal on the Western blot is only severe aggregation on the cell membrane and not transduction.

From now on we only worked with protein expressed at 20ºC in LB-rich medium. There is another denaturing agent, guanidium chloride, which we tried at different concentrations (1 M, 2 M, 4 M and 6 M) in the hope of that the denaturation is complete or different in nature therefore the renaturation would happen without aggregation (Monera, Kay, & Hodges, 1994). All four concentrations of GuCl could lyse the inclusion bodies but the samples which contained TAT-SMN1 purified with GuCl were impossible to load on an SDS-Page quantitatively. Even after two buffer exchanges, the GuCl was still present in a concentration which caused the sample buffer and the protein to form a precipitate thus we were unable to analyse these purifications.

We tried to use so-called mild lysis methods which do not fully denature the protein. Therefore, one would expect that it will be easier to refold without aggregation.

There are buffers which can lyse non-classical inclusion bodies: 6 M n-propanol and 0,2 % lauryl sarcosine (combined with 2 M urea) (S. M. Singh et al., 2012). Both of these buffers lysed the TAT-SMN inclusion bodies but also degraded the protein based on as the gel pictures indicate (**Figure 13**). Buffers able to lyse non-classical inclusion bodies are 5 % DMSO and 5 % n-propanol combined with 2 M urea (A. Singh et al.,2015). These buffers were able to lyse the inclusion bodies without degrading the protein. From this we could conclude that the protein is located in non-classical inclusion bodies. Based on only the concentration of the purified proteins, 5 % n-propanol was superior to 5 % DMSO, but a higher percentage of biologically active protein could have been extracted with DMSO. This was the case with human granulocyte-colony stimulating factor (hG-CSF). hG-CSF was expressed into non-classical inclusion bodies and lysed with 5 % n-propanol and 5 % DMSO. However, in that case, these non-detergent lysis methods were not employed in combination with urea as in our case (Jevševar et al., 2008). We also attempted to change the buffer conditions to as close to native ones as possible and tried 5 % DMSO with different urea concentrations: 0 mM, 250 mM, 500 mM and 1 M but only 2 M additional urea is sufficient for the lysis, all the other concentrations showed no success (**Figure 15**).

DMSO and n-propanol did not affect the biological activity of hG-CSF, but in our case this was not an option (Jevševar et al., 2008). TAT-SMN1 was indeed stable in the lysis buffer but due to the fact that we would have had to use the protein for downstream experiments on NSC-34 cells, we needed to remove the urea, DMSO or n-propanol from the buffer. Therefore, even though we managed to lyse the inclusion bodies, we still needed to find an additive in the storage buffer which can supress aggregation but does not harm the cells during the downstream transduction experiments.

Finally, there are also major differences in the secondary structure of the two proteins (hG-CSF and SMN1). SMN1, as stated earlier, is made up of largely unstructured regions and β-sheets. The rigidity of β-sheets and the hydrophobic nature of this protein could contribute to the aggregation (Abdalla, 2016). On the other hand, hG-CSF it is mainly composed of α-helices and these structural differences can account for the differences in the proteins' aggregation propensities.

The last items we could screen for were the refolding methods, the storage conditions and the storage buffers. We mainly tested denaturing purifications (8 M), denaturing purification with ion-exchange column and mild lysis with 5 % n-propanol. First of all, we identified the best storage condition for the ion-exchanged protein by leaving protein from the same purification in different environments overnight. The next day we performed SDS-PAGE to identify the condition where the least amount of TAT-SMN1 was aggregated. The best way for TAT-SMN1 to be stored is 400 mM NaCl, snap freezing, and storage at -80°C.

Dialysis is another refolding method. We tried to dialyse two TAT-SMN1 proteins, one which was purified with 8M urea and another which was purified with 5 % n-propanol. In both cases almost the entire amount of protein precipitated.

On-column refolding with decreasing amount of denaturing agent is another option for avoid aggregation in some cases. We have used this method on an ion-exchange (denaturing purification) and on Strep columns (5 % n-propanol) (Lemercier, Bakalara, & Santarelli, 2003). In our case this method did not work, it even increased the aggregation rather than suppressing it.

Various storage buffers were tested on ion-exchange purified and 5 % n-propanol purified proteins. For the ion-exchange purified protein, we selected 1 mM TCEP (Willis et al., 2005), 300 mM arginine and 0,5 mM EDTA as aggregation sup-

pressants. We saw increases in TAT-SMN1 concentration in two additives: TCEP and arginine. Without concentration and always keeping the protein cold we managed to obtain 879 ng/µl of protein. This protein was stable during freezing and thawing but it was not stable at 37ºC. Therefore, we could not use it for testing transduction.

In order to determine the best storage buffer which would inhibit aggregation, we submitted a sample of this protein for Thermofluor experiment (screens 96 different buffers). Thermoflour monitors protein stability over increasing temperature as a function of dye binding (Reinhard et al., 2013). This experiment has shown that TAT-SMN1 aggregates in all of these 96 conditions; however, it also identified the TCEP containing buffer as one in which aggregation is limited.

To complement Thermoflour, TAT-SMN1 was also subjected to DLS. DLS is a dye-free technique measures the diameter of particles in solution by firing a laser beam at the solution and quantifying dynamic fluctuation of light scattering (Goldburg, 1999). In our case the TAT-SMN1 sample was diluted in different buffers to see which one has the lowest amount aggregates and the highest amount of monomeric protein. All the buffers which were screened showed the same results. A low intensity peak of monomeric TAT-SMN1 and a high intensity peak of aggregated TAT-SMN1. This experiment validated our theory of aggregation.

We compared the stability of TAT-SMN1 purified in denaturing conditions (8M urea, without ion-exchange) with and without ion-exchange by incubating them at 37ºC and we saw an approximately 50 % decrease in protein concentration after 8 hours. We performed the downstream experiments which would serve to determine whether transduction had occurred and both the Western blot and the immunofluorescence showed no transduction or were not conclusive. On the immunofluorescence pictures, we saw strong signal from the protein attached to the cell, probably to the cell membrane.

This led us to try the new storage buffers in combination with the mild lyses method. In case of the 5 % n-propanol purification, we first tried more additives: 1 mM TCEP, 1 % BSA, 30 % glycerol, 1 M proline and 400 mM lysine (Kim, Yan, & Zhou, 2006). No additives and 2 M urea in the buffer served as controls. Surprisingly none of the additives had a major impact on the aggregation, but the protein in the buffer without additives looked the most concentrated after the 2 M urea one (positive control). This does not mean that that there is no aggregation in the additive-free buffer, but it raises the question, what would be the best storage buffer for TAT-SMN1 purified using the mild lysis methods. Because of time limitation, we could not submit another sample for Thermofluor.

On the SDS-PAGE every time we ran the protein an upper band around 80-90 kDa was visible. Our initial guess was bacterial chaperon. We submitted that band for mass spectrometry and the analyses showed that it mainly contains SMN1 protein. SMN1 possesses a self-oligomerization domain on exon 6 between residues 249-278 (Lorson et al., 1998). The upper band could be an oligomer of the protein, as the ones that form gems in the nucleus (S. Lefebvre et al., 1997).

5. Conclusion

This work is focusing on expression and purification of fusion protein, TAT-SMN1 and SMN1. We knew the proteins are expressed into inclusion bodies therefor we lysed them with denaturing lysis buffer. We saw that aggregation was happening which we tried to supress by various additives. TCEP proved to be the best aggregation suppressant in this case but aggregation still occurred. After several trials we turned our interest to suppressing inclusion body formation. We changed expression hosts (*Origami*^{TM} DE3and *Arctic express*), decreased expression temperature, co-expressed them

with different chaperons (GroEL, GroES, DnaK, DnaJ and GrpE) and add glucose, sucrose and ethanol to the expression medium. Even though, we could not produce a soluble protein, we managed to identify mild lysis methods which could solubilize the inclusion bodies. This made us come to the conclusion that the protein is expressed into non-classical inclusion bodies. Our findings indicate that stable TAT-SMN1 fusion protein can be purified, however additional efforts are required to completely tackle its aggregtion.

We could not make a final conclusion about the transduction of TAT-SMN1 and SMN1 because it is possible that the protein is aggregating on the cell membrane and that can give false positive results.

6. References

Abdalla, M. (2016). Important Factors Influencing Protein Crystallization. *Global Journal of Biotechnology and Biomaterial Science*, *2*(1), 025–028. https://doi.org/10.17352/gjbbs.000008

Akten, B., Kye, M. J., Hao, L. T., Wertz, M. H., Singh, S., Nie, D., … Sahin, M. (2011). Interaction of survival of motor neuron (SMN) and HuD proteins with mRNA cpg15 rescues motor neuron axonal deficits. *Proceedings of the National Academy of Sciences*, *108*(25), 10337–10342. https://doi.org/10.1073/pnas.1104928108

Becker-Hapak, M., McAllister, S. S., & Dowdy, S. F. (2001). TAT-mediated protein transduction into mammalian cells. *Methods*, *24*(3), 247–256. https://doi.org/10.1006/meth.2001.1186

Bowerman, M., Becker, C. G., Yáñez-Muñoz, R. J., Ning, K., Wood, M. J. A., Gillingwater, T. H., & Talbot, K. (2017). Therapeutic strategies for spinal muscular atrophy: SMN and beyond. *Disease Models & Mechanisms*, *10*(8), 943–954. https://doi.org/10.1242/dmm.030148

Brzustowicz, L. M., Lehner, T., Castilla, L. H., Penchaszadeh, G. K., Wilhelmsen, K. C., Daniels, R., … Gilliam, T. C. (1990). Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q1 1.2–13.3. *Nature*, *344*, 540. Retrieved from http://dx.doi.org/10.1038/344540a0

Bühler, D., Raker, V., Lührmann, R., & Fischer, U. (1999). Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: Implications for spinal muscular atrophy. *Human Molecular Genetics*, *8*(13), 2351–2357. https://doi.org/10.1093/hmg/8.13.2351

Butchbach, M. E. R. (2016). Copy Number Variations in the Survival Motor Neuron Genes: Implications for Spinal Muscular Atrophy and Other Neurodegenerative Diseases. *Frontiers in Molecular Biosciences*, *3*(March), 1–10. https://doi.org/10.3389/fmolb.2016.00007

Chhetri, G., Kalita, P., & Tripathi, T. (2015). An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX*, *2*, 385–391. https://doi.org/10.1016/j.mex.2015.09.005

D'Amico, A., Mercuri, E., Tiziano, F. D., & Bertini, E. (2011). Spinal muscular atrophy. *Orphanet Journal of Rare Diseases*, *6*(1), 71. https://doi.org/10.1186/1750-1172-6-71

de Marco, A., Vigh, L., Diamant, S., & Goloubinoff, P. (2005). Native folding of aggregation-prone recombinant proteins in Escherichia coliby osmolytes, plasmid- or benzyl alcohol over-expressed molecular chaperones. *Cell Stress Chaperones*, *10*, 329–339. https://doi.org/10.1379/CSC-139R.1

Frankel, A. D., & Pabo, C. O. (1988). Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, *55*(6), 1189–1193. https://doi.org/10.1016/0092-8674(88)90263-2

Goldburg, W. I. (1999). Dynamic light scattering. *American Journal of Physics*, *67*(12), 1152–1160. https://doi.org/10.1119/1.19101

Hoffmann, J. (1893). Ueber chronische spinale Muskelatrophie im Kindesalter, auf familiärer Basis. *Deutsche Zeitschrift Für Nervenheilkunde*, *3*(6), 427–470. https://doi.org/10.1007/BF01668496

Jevševar, S., Gaberc-Porekar, V., Fonda, I., Podobnik, B., Grdadolnik, J., & Menart, V. (2008). Production of Nonclassical Inclusion Bodies from Which Correctly Folded Protein Can Be Extracted. *Biotechnology Progress*, *21*(2), 632–639. https://doi.org/10.1021/bp0497839

Kim, S.-H., Yan, Y.-B., & Zhou, H.-M. (2006). Role of osmolytes as chemical chaperones during the refolding of aminoacylase. *Biochemistry and Cell Biology = Biochimie et Biologie Cellulaire*, *84*(1), 30–38. https://doi.org/10.1139/o05-148

Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., … Melki, J. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, *80*(1), 155–165. https://doi.org/10.1016/0092-8674(95)90460-3

Lefebvre, S., Burlet, P., Liu, Q., Bertrandy, S., Clermont, O., Munnich, A., … Melki, J. (1997). Correlation between severity and SMN protein level in spinal muscular atrophy. *Nature Genetics*, *16*(3), 265–269. https://doi.org/10.1038/ng0797-265

Lemercier, G., Bakalara, N., & Santarelli, X. (2003). On-column refolding of an insoluble histidine tag recombinant exopolyphosphatase from Trypanosoma brucei overexpressed in Escherichia coli. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *786*(1–2), 305–309. https://doi.org/10.1016/S1570-0232(02)00745-6

Lorson, C. L., Strasswimmer, J., Yao, J. M., Baleja, J. D., Hahnen, E., Wirth, B., … Androphy, E. J. (1998). SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nature Genetics*, *19*(1), 63–66. https://doi.org/10.1038/ng0598-63

Lunn, M. R., & Wang, C. H. (2008). Spinal muscular atrophy. *Lancet*, *371*(9630), 2120–2133. https://doi.org/10.1016/S0140-6736(08)60921-6

Monera, O. D., Kay, C. M., & Hodges, R. S. (1994). Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Science*, *3*(11), 1984–1991. https://doi.org/10.1002/pro.5560031110

Nagel, F., Dohm, C. P., Bähr, M., Wouters, F. S., & Dietz, G. P. H. (2008). Quantitative evaluation of chaperone activity and neuroprotection by different preparations of a cell-penetrating Hsp70. *Journal of Neuroscience Methods*, *171*(2), 226–232. https://doi.org/10.1016/j.jneumeth.2008.03.008

Nakase, I., Noguchi, K., Aoki, A., Takatani-Nakase, T., Fujii, I., & Futaki, S. (2017). Arginine-rich cell-penetrating peptide-modified extracellular vesicles for active macropinocytosis induction and efficient intracellular delivery. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-02014-6

Nishihara, K., Kanemori, M., & Yanagi, H. (2000). Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. *Applied and Environmental Microbiology*, *66*(3), 884–889. https://doi.org/10.1128/AEM.66.3.884-889.2000.Updated

Ottesen, E. W. (2017). ISS-N1 makes the first FDA-approved drug for spinal muscular atrophy. *Translational Neuroscience*, *8*(1), 1–6. https://doi.org/10.1515/tnsci-2017-0001

Pek, J. W., Anand, A., & Kai, T. (2012). Tudor domain proteins in development. *Development*, *139*(13), 2255–2266. https://doi.org/10.1242/dev.073304

Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J., & Weiss, M. S. (2013). Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, *69*(2), 209–214. https://doi.org/10.1107/S1744309112051858

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology*, *5*(APR), 1–

17. https://doi.org/10.3389/fmicb.2014.00172

Rossoll, W., Kröning, A.-K., Ohndorf, U.-M., Steegborn, C., Jablonka, S., & Sendtner, M. (2002). Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Human Molecular Genetics*, *11*(1), 93–105. https://doi.org/10.1093/hmg/11.1.93

Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., & Panda, A. K. (2015). Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microbial Cell Factories*, *14*(1), 1–10. https://doi.org/10.1186/s12934-015-0222-8

Singh, S. M., Sharma, A., Upadhyay, A. K., Singh, A., Garg, L. C., & Panda, A. K. (2012). Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form. *Protein Expression and Purification*, *81*(1), 75–82. https://doi.org/10.1016/j.pep.2011.09.004

Sørensen, H. P., & Mortensen, K. K. (2005). Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial Cell Factories*, *4*(Figure 1), 1–8. https://doi.org/10.1186/1475-2859-4-1

Uversky, V. N. (2015). Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Frontiers in Aging Neuroscience*, *7*(MAR), 1–6. https://doi.org/10.3389/fnagi.2015.00018

Werdnig, G. (1891). Zwei frühinfantile hereditäre Fälle von progressiver Muskelatrophie unter dem Bilde der Dystrophie, aber anf neurotischer Grundlage. *Archiv Für Psychiatrie Und Nervenkrankheiten*, *22*(2), 437–480. https://doi.org/10.1007/BF01776636

Willis, M. S., Hogan, J. K., Prabhakar, P., Liu, X., Tsai, K., Wei, Y., & Fox, T. (2005). Investigation of protein refolding using a fractional factorial screen: A study of reagent effects and interactions. *Protein Science*, *14*(7), 1818–1826. https://doi.org/10.1110/ps.051433205

Wingfield, T. P. (2016). Use of Protein Folding Reagents. *Current Protocols in Protein Science*, *May*(Appendix 3), A.3A.1-A.3A.8. https://doi.org/10.1002/0471140864.psa03as84

7. Appendix

List of Figures:

Figure 1: Overview of the effects of SMA 10

Figure 2: Different splicing of *SMN1* and *SMN2* 12

Figure 3: Localization of SMN1 in a neuronal cell 13

Figure 4: Overview of different therapeutic approaches 14

Figure 5: Transport mechanism of TAT peptide 14

Figure 6: Solubilization of inclusion bodies 16

Figure 7: The examination order of the transduction of different proteins 31

Figure 8: The examination order of the transduction of different proteins 32

Figure 9: Native lysis of SMN1 pellet 35

Figure 10: The native lysis of pellets produced in different media and coexpressed with chaperons 36

Figure 11: Native lysis of pellet expressed in *Origami*^{TM} DE3 37

Figure 12: Lysis of TAT-SMN1 pellet with guanidium-chloride 38

Figure 13: Testing of mild lysis methods 39

Figure 14: Lysis of TAT-SMN1 and SMN1 pellets with mild lysis methods without urea 40

Figure 15: Lysis of TAT-SMN1 pellet in 5 % DMSO and different concentrations of urea 41

Figure 16: Denaturing purification of TAT-SMN1 42

Figure 17: Denaturing purification of SMN1 42

Figure 18: Examination of TAT-SMN1 and SMN1 transduction by Western blot 43

Figure 19: Storage buffer test for TAT-SMN1 44

Figure 20: Additives influence on protein stability after ion exchange column 45

Figure 21: TAT-SMN1 purification with ion exchange column, TCEP and no concentration 46

Figure 22: Stability test of TAT-SMN1 and SMN1 purified in different ways 47

Figure 23: Examination of TAT-SMN1 and SMN1 transduction by Western blot 48

Figure24: Immunofluorescence pictures of TAT-SMN1 and SMN1 transduction 48

Figure 25-26: The effects of additives on protein stability after purification with 5 % n-propanol and 2 M urea 50, 51

Figure 27: The band submitted for mass spectrometry 52

Figure 28: DLS results of TAT-SMN1 53

List of tables:

Table 1: Types of spinal muscular atrophy and their defining characteristics 9

Table 2: Expression conditions of TAT-SMN1 23

Table 3: Combination of applied lysis methods, columns and storage buffers for purification of TAT-SMN1 29

Table 4: Concentrations achieved with additives after ion exchange column 45

Table 5: Result of mass spectrometry of the upper SDS-PAGE band of TAT-SMN1 52

Abstract

Spinal muscular atrophy (SMA) is the second most common autosomal recessive genetic disorder. Both the incidence rate (1 in 10.000 livebirths) and carrier frequency (1 in every 50 people) are high. SMA is the leading genetic cause of infant death characterized by degeneration of alpha motor neurons of the spinal cord, skeletal muscle atrophy and general weakness. The disease is caused by a homozygous disruption in the *SMN1* gene which results in low levels of the SMN1 protein. In order to devise a treatment for this disease we developed a fusion protein called TAT-SMN1. TAT (transactivator of transcription) is a cell penetrating peptide which is able to cross the cell membrane and the blood-brain barrier. We expressed and purified TAT-SMN1 from inclusion bodies in various ways and examined the transduction of these proteins into motor neuron-like cells. However, our results were not conclusive in terms of transduction because the purified protein was aggregating. We did not manage to express soluble protein even though we had attempted low expression temperature, different bacteria strains, co-expressing it with chaperons, supplementing additives to the expression media and the combination of these. We identified lysis methods which are able to disrupt the inclusion bodies. Denaturing lysis methods could solubilize the protein but surprisingly, so could mild lysis methods, meaning TAT-SMN1 is expressed into non-classical inclusion bodies. Even with the mild lysis method we needed to refold the protein and aggregation occurred at this point. We tried many different techniques for refolding and additives in the storage buffer. We found that the best storage buffer for supressing the aggregation in case of denaturing purification was TCEP and with the mild lysis methods the standard storage buffer without additives, however the protein was still aggregating to a significant extent. Being aware of the difficulties we faced during these experiments, the main focus for the future is to purify stable, soluble and non-aggregating TAT-SMN1 to high purity and reasonable concentration.

Zusammenfassung

Die spinale Muskelatrophie (SMA) ist die zweithäufigste autosomal-rezessive vererbte, genetisch bedingte Erkrankung. Sowohl die Inzidenzrate (1 von 10.000 Lebendgeburten) als auch die Trägerfrequenz (1 von 50 Menschen) sind sehr hoch. SMA ist die häufigste genetische Todesursache im Kindesalter, welche durch die Degeneration von Alpha-Motorneuronen im Rückenmark, Muskelatrophie des Skeletts und allgemeiner Schwäche gekennzeichnet ist. Diese Erkrankung wird verursacht durch eine homozygote Spaltung des *SMN1-Gens,* welches wiederum zu einer verminderten SMN1-Proteinproduktion führt.

Im Zuge der Erforschung nach Behandlungsmöglichkeiten für diese schwerwiegende Krankheit, versuchten wir ein sogenanntes Fusionsprotein (TAT-SMN1) dafür zu entwickeln. TAT (transactivator of transcription) ist ein zellpenetrierendes Peptid, welches die Fähigkeit besitzt die Bluthirnschranke zu durchdringen. Aus diesem Grund haben wir versucht, TAT-SMN1 erfolgreich zu exprimieren und in weiterer Folge aufzureinigen. Dafür haben wir diverse Ansatzweisen verfolgt, um in weiterer Folge die Transdtuktion dieses Proteins in verschiedene Zellmodelle auszutesten.

Unsere neuesten Untersuchungen am Protein zeigten jedoch, dass es aufgrund von Aggregation nicht in *in-vitro*-Versuchen transduziert werden konnte. Wir konnten keine Expression von gelöstem Protein bewerkstelligen, trotz umfassender Versuchansätze, wie der Verwendung von niedrigerer Temperaturen und verschiedene Bakterienstämme, sowie einer Coexpression mit Hilfe von Chaperonen als auch das Hinzufügen und die Kombination von verschiedenen Hilfsmitteln zum Kulturmedium.

Wir identifizierten des weiteren Lysierungsmethoden, mit Hilfe deren es möglich war die Einschlusskörperchen („inclusion bodies“) aufzubrechen. Wie zu erwarten konnten denaturierende Lysemethoden die Einschlusskörperchen aufbrechen, jedoch konnten

dies zu unserer Überraschung auch milde Lysemethoden bewerkstelligen was wiederum für uns bedeutete, dass TAT-SMN1 in sogenannten „non-classical“ Einschlusskörperchen exprimiert wird. Jedoch mussten wir bereits bei der Anwendung von milden Lysierungsmethoden das freigesetzte Protein wieder zurückfalten und dabei kam es zu der beschriebenen Aggregation von TAT-SMN1. Wir untersuchten diverse Methoden und Zusätze im Proteinlösungspuffer um eine erfolgreiche Wiederfaltung des Proteins zu gewährleisten und es kristallisierte sich dabei TCEP, als der am besten geeignete Aggregationsunterdrücker während einer denaturierenden Aufreinigung heraus. Des Weiteren entdeckten wir, dass innerhalb der milden Lysierung ein Proteinlösungspuffer ohne den Zusatz von Hilfsmitteln am besten geeignet war, wobei aber weiterhin eine Aggregation bis zu einem gewissen Grad bestehen blieb. Mit Rückblick auf die ganzen Schwierigkeiten, welche uns bei diesen Experimenten der Proteinentwicklung begleiteten, wird man den weiteren Fokus in Zukunft auf die Proteinaufreinigung von TAT-SMN1 mit hoher Reinheit und vernünftig anwendbaren Konzentrationen legen.