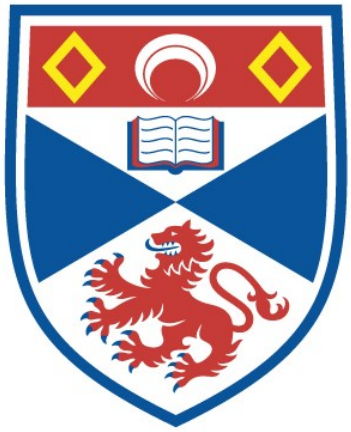
Quantitative interactomics to identify cellular pathways affected in spinal muscular atrophy

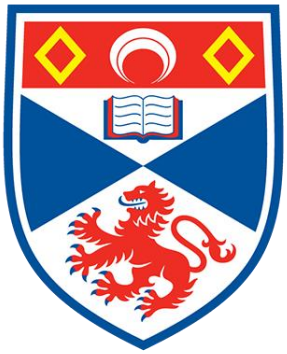
Bobby Beaumont

A thesis submitted for the degree of PhD
at the
University of St Andrews



2024

Quantitative Interactomics to Identify Cellular Pathways Affected in Spinal Muscular Atrophy



University of
St Andrews

Bobby Beaumont

Student ID: 190027627

Supervisor: Dr Judith E. Sleeman

March 2024

Declaration

Candidate's declaration

I, Bobby Beaumont, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 42,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2019.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 19/07/24 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 19/07/24 Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Bobby Beaumont, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date 19/07/24 Signature of candidate

Date 19/07/24 Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Bobby Beaumont, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date 19/07/24 Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date 19/07/24 Signature of candidate

Date 19/07/24 Signature of supervisor

Abstract

Spinal Muscular Atrophy (SMA) is an inherited neuromuscular disorder that results in muscle weakness, paralysis or - in most cases - death. Caused by insufficient production of the Survival of Motor Neuron (SMN) protein, this ubiquitously expressed protein is involved in numerous cellular processes from snRNP biogenesis to pre-mRNA splicing. Motor neurons are particularly susceptible to low levels of SMN, yet the precise pathomechanism behind their selective death as occurs in SMA remains to be elucidated.

To further understand the disease process, interactomic techniques offer useful methods to identify protein interactions which are disrupted in SMA. Using cell culture models of human cells expressing either SMN or SMNΔ7 (a canonically non-functional isoform of SMN), the proximity biotinylation capabilities of FLAG-tagged TurboID were utilised to identify transient protein interactions whilst FLAG-Immunoprecipitation allowed for simultaneous identification of stronger, covalent interactors. Concurrent SILAC (Stable Isotope Labelling by Amino Acids in Cell Culture) permitted downstream quantitative analysis of proteins identified via mass spectrometry (MS), allowing differential analysis of protein interaction levels.

Utilising bioinformatics tools such as STRING-DB, novel protein interactors – such as GAIT Complex components, Amino Acyl tRNA Synthetases, protein SON and SFXN1 - and affected pathways of potential relevance were obtained from MS datasets and further examined via techniques including colocalization imaging and quantitative western blot to reveal differential interactions between the SMN protein isoforms.

Acknowledgements

My first acknowledgement is to Dr Judith Sleeman, for consistently going above and beyond her role as my PhD advisor. Her knowledge, expertise and support over these last several years have made this thesis possible. Together we endured Covid, lockdown, moving labs and my frequent mental breakdowns; I expect we now share something of a trauma bond.

Secondly, I must thank Kim Vincent - my lab mate - for her continuous support and making the lab a happier place by being part of it.

I must also thank SPRINT-MND for funding this work and the University of St Andrews - particularly the BMS - for allowing me to be part of their community.

Finally, I must thank my mum for her continued support of the things I choose to do with my life. And of course, my cat Lucy, who loved to make her presence known during online Lab Meetings.

Bobby Beaumont

July 2024

Research data underpinning this thesis are available at https://doi.org/10.17630/909d53bc-65a2-4c5c-97ec-f3b9e5552b28

Contents

List of Figures and Tables

Presentations and Posters

SPRINT-MND Annual Conference (2023, 2022 and 2021)

St Andrews PGR Conference (2022 and 2021)

Abbreviations

AAV9 – Adeno-Associated Virus Serotype 9

AIO – All-In-One (in reference to the Tetracycline induction system used)

CB – Cajal Body

DAPI – 4',6-Diamidino-2-Phenylindole

DAVID – Database for Annotation, Visualisation, and Integrated Discovery

DMEM – Dulbecco's Modified Eagle Medium

DTT – Dithiothreitol

FBS – Foetal Bovine Serum

FDA – Food and Drug Administration (US)

FDR – False discovery rate

FITC – Fluorescein Isothiocyanate

GAPDH – Glyceraldehyde-3-Phosphate Dehydrogenase

GFP – Green Fluorescent Protein

GS – Goat Serum

HRP – Horseradish Peroxidase

IAA – Indole-3-acetic acid

ICC – Immunocytochemistry

IFN – Interferon

IP – Immunoprecipitation

IRE – Interferon Response Element

JIR – Jackson ImmunoResearch

LVX – Lentiviral

MCS – Multiple Cloning Site

MeCN – Methyl Cyanide (Acetonitrile)

MOI – Multiplicity of Infection

MS – Mass Spectrometry

NP-40 – Nonyl Phenoxypolyethoxylethanol

PAGE – Polyacrylamide Gel Electrophoresis

PBS – Phosphate Buffered Saline

PFA – Paraformaldehyde

PLS3 – Plastin-3

PPI – Protein-Protein Interactions

PTG – Proteintech Group

QC – Quality Control

RBP – RNA-Binding Protein

RCF – Relative Centrifugal Force

RIPA – Radioimmunoprecipitation Assay

RNA – Ribonucleic Acid

RNP – Ribonucleoproteins

SDS – Sodium Dodecyl-Sulphate

SHSY5Y – Human-Derived Neuroblastoma Cell Line

SILAC – Stable Isotope Labelling by Amino Acids in Cell Culture

SMA – Spinal Muscular Atrophy

SMN – Survival Motor Neuron Protein (Primary product of SMN1)

SMN1 – Survival of Motor Neuron 1 Gene, Telomeric

SMN2 – Survival of Motor Neuron 2 Gene, Centromeric

snRNAs – Small Nuclear Ribonucleic Acids

snRNPs – Small Nuclear Ribonucleoproteins

STAT1 – Signal Transducer and Activator of Transcription 1 Protein

STRING – Search Tool for Recurring Instances of Neighbouring Genes

SYNCRIP – Synaptotagmin-Binding, Cytoplasmic RNA-Interacting Protein (also known as hnRNP Q)

TBE – Tris-Borate-EDTA

TRITC – Tetramethylrhodamine

UV – Ultraviolet

WB – Western Blot

Chapter 1: Introduction

Spinal Muscular Atrophy (SMA) is an autosomal recessive neuromuscular disorder, generally recognized as the most common genetic cause of early infant mortality (Pearn, 1980). As a genetic disease, the carrier frequency of SMA is estimated to be around 1 in 47 (heterozygous) whilst the overall incidence of the disorder is between 1 in 6000 – 11000 live births (Sugarman et al., 2012).

1.1 Pathogenesis and Classification of SMA Subtypes

SMA is characterized by degradation and loss of alpha motor neurons in the anterior horn of the lower spinal cord. These neurons - which are primarily controlled by the descending motor tracts from higher brain centres - innervate the extrafusal muscle fibres of skeletal muscle and are responsible for initiating their contraction in response to both voluntary and involuntary (reflex) inputs (Powis & Gillingwater, 2016). Because of this, neurodegeneration here results in reduced motor input to muscle cells leading to progressive proximal muscular atrophy, loss of motor functionality and symmetrical paralysis. In severe cases, death is caused by respiratory failure due to insufficiency of the respiratory musculature (Edens et al., 2015; Kolb & Kissel, 2015). Since its initial description in the early 1900s by Werdnig and Hoffman, the broad range of disease severity observed between SMA patients prompted further classification of the disorder into 5 primary subtypes; these subtypes are dependent on factors such as age of symptom onset and motor milestones achieved by the patient, with Type 0 being the most severe and Type 4 the least (Tiziano et al., 2013).

Patients with Type 0 SMA display the most acute characteristics of the disease and are unlikely to live beyond 2 months post-partum due to breathing difficulties. Type 1 SMA, known as Werdnig-Hoffman disease, is similarly severe but has a life expectancy of around 2 years. Type 2 (Intermediate SMA) has an age of onset of 6 months and many patients survive into adolescence; 75% survive beyond 25 years of age with medical assistance (Rudnik-Schoneborn et al., 1996). This form of the disease is associated with muscle weakness and hyporeflexia, with most being unable to walk unaided. Type 3 SMA, known also Kugelberg-Welander disease, begins around 18 months and these patients have a normal lifespan but suffer some muscular abnormalities, such as hand tremors, whilst retaining the ability to walk. Type 4 (Adult Onset SMA) manifests in the 2nd or 3rd decade and these individuals continue to live relatively normal lives with little to no motor dysfunction. Types 1, 2 and 3 are the most commonly diagnosed forms of the disease (Talbot & Tizzano, 2017).

The typing system described above is based on observation and differential diagnosis and it is worth noting that this represents only one form of SMA disease classification; these subtypes can be further divided (type 3A, type 3B etc) and, given the current changing nature of the SMA landscape due to recent advances in treatment, may change even further still (Adami & Bottai, 2019).

1.2 The SMN1 & SMN2 Genes

While classifying disease based on symptom severity is useful in diagnosis, SMA is a genetic disorder and, until the 1990s, the gene responsible for the disease was unknown. In 1995, Lefebvre and colleagues pinpointed the SMA-causative gene to the telomeric region of Chromosome 5 (locus 5q13), identifying homozygous mutations in this gene as the primary cause of disease in 95% of SMA cases (Lefebvre et al., 1995). The gene was subsequently named SMN1 (Survival Motor Neuron 1). An almost identical gene, named SMN2, was discovered simultaneously. SMN2 is centromeric and paralogous to SMN1 except for 5 coding-region nucleotide changes which do not alter the amino acid sequence of the protein product: the survival motor neuron (SMN) protein (See Figure 1.1).

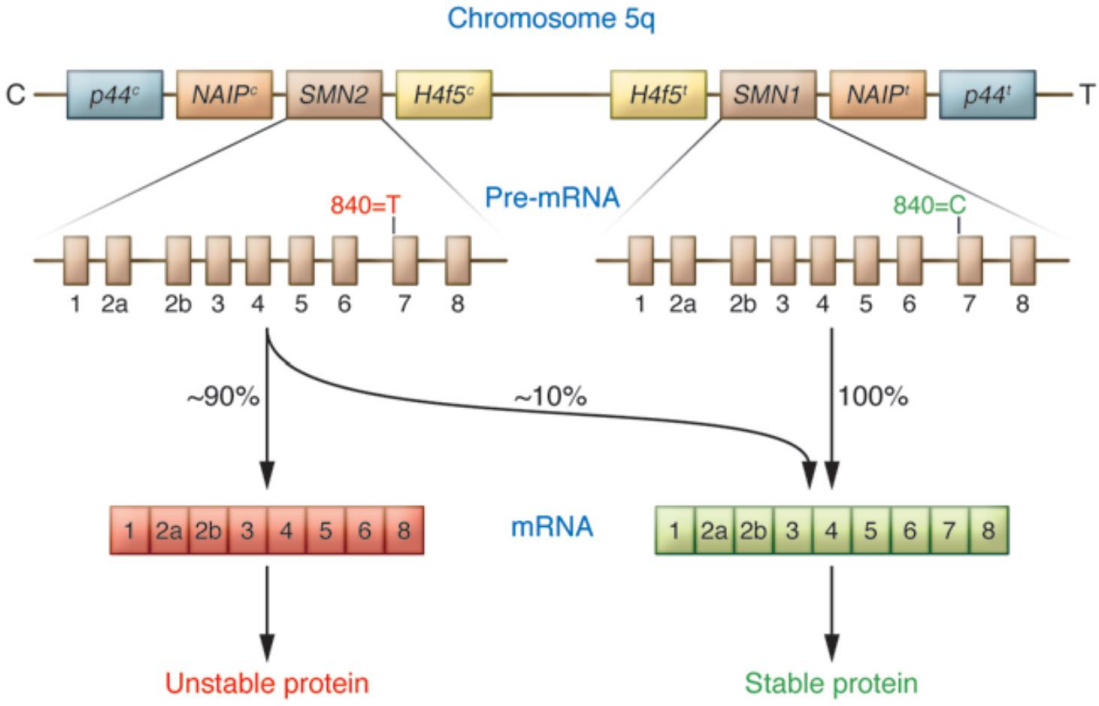


FIGURE 1.1: ORGANIZATION OF HUMAN CHROMOSOME 5Q WITH A FOCUS ON THE SMN1 & SMN2 GENES.
SMN is one of four genes within a duplicated, inverted region of 500 kilobase pairs on chromosome 5q. SMN1&2 are the only genes in the region which have been proven to contribute to the pathology

of SMA. The unstable protein produced by SMN2 alternative splicing is SMNΔ7. NAIP = NLR Family Apoptosis Inhibitory Protein. H4F5 = also known as SERF1A (small EDRK-rich factor 1A). P44 = also known as GTF2H2 (general transcription factor IIH subunit 2). Image courtesy of (Swoboda, 2011).

SMN protein is expressed ubiquitously and is essential for survival; it is involved in numerous cellular processes and complete knockout of SMN results in embryonic lethality in mice (Monani, 2000). For reasons that still remain unclear, depletion of SMN protein – as occurs in SMA – selectively effects the survival of alpha motor neurons of the spinal cord and brain stem. Reduced levels of SMN protein result in axonal degeneration, with the extent of the pathology inversely correlated to the levels of SMN produced (Lefebvre et al., 1997).

The SMN1 gene is responsible for production of the vast majority of total cellular SMN protein. As occurs in SMA, heterozygous dysfunction or deletion of this gene results in greatly reduced SMN protein production. Whilst SMN2 is functional and encodes for the same protein, a C to T transition in exon 7 of SMN2 locates to a region which functions as an Exonic Splice Enhancer site in the transcribed pre-mRNA. This results in alternative splicing of the SMN transcript which leads to exclusion of exon 7 and the formation of an unstable protein product known as SMNΔ7 (Lorson et al., 1999).

Classically considered non-functional, SMNΔ7 is described as being rapidly degraded and unable to compensate for the loss of full-length, native SMN. Around 90% of the protein produced from the SMN2 gene consists of the alternatively spliced SMNΔ7, whilst only 10%

is spliced correctly and forms native SMN (Lefebvre et al., 1997). This explains why complete knockout of SMN results in embryonic lethality in mice, whilst - in humans - those with SMA survive beyond this stage; deletions in SMN1 deplete the levels of SMN protein and lead to disease, but the products of SMN2 in humans provide enough of the full-length protein to avoid complete cell death. As mice do not have this backup SMN gene, early embryonic lethality occurs at the transition from morula to blastocyst (Schrank et al., 1997). This is of particular relevance in the study of SMA because it underpins the difficulty in creating effective mouse models of the disease which recapitulate the human SMA phenotype (Signoria et al., 2023). It also highlights the importance that cell-based models have with regards to understanding the underlying disease pathology in SMA research.

Because of this, the primary disease modifier of SMA is SMN2 copy number. An inverse correlation is generally observed between severity of the patient's SMA phenotype and number of SMN2 copies they have; as copy number increases, disease severity is reduced. Patients with SMA type 1, for example, may have only 1 or 2 copies of the SMN2 gene whilst Type 4 patients (Adult Onset) can have up to 6 (Harada et al., 2002). This dose-dependent effect of the number of SMN2 genes on the SMA phenotype has been the focus of many therapeutic strategies aimed at treatment of the disease (D'Ydewalle et al., 2017).

1.3 SMN Protein Structure & Function

The full length, native SMN protein is comprised of 294 amino acids (38kDa) and contains numerous functional motifs: these include a basic/lysine-rich domain, Tudor domain, YG box and a proline-rich region. Each of these regions enables the protein to carry out its numerous functions within the cell (Figure 1.2).

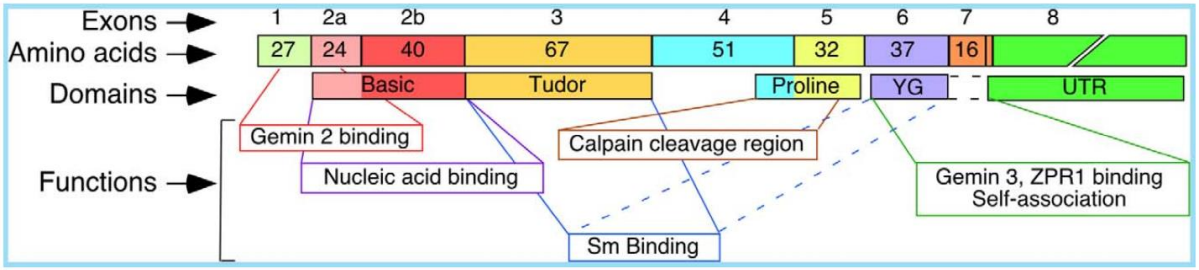


FIGURE 1.2: DIAGRAMMATIC REPRESENTATION OF SMN1 MRNA (TOP) AND FULL LENGTH SMN PROTEIN (BOTTOM).

The number of amino acids encoded by each exon is indicated. Exons as well as domains they encode are shown as boxes. Domain functions are indicated. UTR = Untranslated Region. Figure courtesy of *(Singh et al., 2012).*

For example, the lysine-rich region has been shown to interact with Gemin 2, a protein critical in the formation of small nuclear ribonucleoproteins (snRNPs). Simultaneously, the Tudor domain of SMN binds to Sm protein tails, proteins which play pivotal roles in the initiation of spliceosome formation. Of note, mutations in the Tudor domain of SMN are often found in SMA patients (Selenko et al., 2001).

These interactions highlight the (arguably) most important function of the SMN protein; its central role in the formation of snRNPs, the spliceosome and subsequent pre-mRNA splicing

(Figure 1.3). SMN - in association with the Unr-Interacting Protein (Unrip) and Gemins 2-8 - acts as a molecular chaperone in the formation of the snRNP core from their component Sm proteins and small nuclear ribonucleic acids (snRNAs). As part of a large macromolecular network known as the SMN complex, cytoplasmic SMN works in concert with other protein interactors to initiate the formation of the snRNP. In particular, the association between SMN and Gemin 5 enforces specificity to the snRNA through the direct binding of Gemin 5 to specific snRNAs in the cytoplasm (Yong et al., 2010). Upon assembly of the snRNP particle in the cytoplasm, phosphorylation dependant interactions release SMN from the SMN complex whereby it accompanies the nascent snRNP into the nucleus (Husedzinovic et al., 2014). Here, both SMN and the snRNPs localise to Cajal Bodies (CBs) and undergo maturation, with core snRNP proteins and snRNAs being added here. This localization is aided by the Tudor domain of SMN, which is known to interact with coilin, a protein marker specific to CBs (Lafarga et al., 2018).

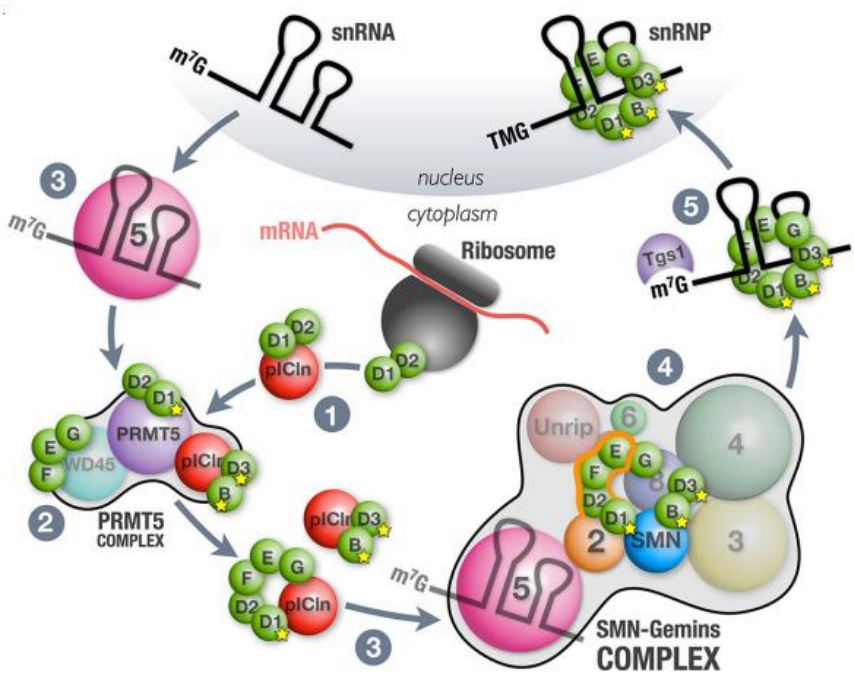


FIGURE 1.3: SIMPLIFIED REPRESENTATION OF THE ROLE OF THE SMN COMPLEX IN THE FORMATION OF SNRNPS.
Cooperation between the Protein Arginine N-Methyltransferase 5 (PRMT5) and SMN Complexes in the cytoplasm ensures accuracy of snRNP formation. 1) Once translated, the Sm D2/D1 heterodimer associates with the Methylosome Subunit pICln in the cytoplasm. 2) Several Sm proteins are then post-translationally modified by the PRMT5 complex. 3) The SMN Complex receives pre-organized subsets of Sm proteins from pICln and snRNAs from Gemin 5. 4) Within the SMN complex, most Sm proteins are bound by Gemin 2 until they are uploaded onto delivered snRNAs. 5) Trimethylguanosine Synthase 1 (TGS1) hypermethylates the caps of snRNPs prior to nuclear import. Figure adapted from *(Lanfranco et al., 2017).*

While this classical pathway of snRNP formation is well studied and characterized, a recent study has demonstrated an alternative pathway which does not rely on SMN-Gemin 5 initiation. In this pathway, the SMN complex is still proposed to play a central role but here it instead acts as a focal point for the congregation of RNAs and RNA-binding proteins (RBPs), prompting the exchange of ribonucleoproteins during snRNP assembly (So et al., 2016).

The vital function that SMN provides during splicing has also been demonstrated experimentally. A correlation exists between reduced levels of snRNPs in the cell - such as occurs when levels of SMN are reduced in SMA - and increased severity of the SMA phenotype, as analysed in mice spinal cords (Gabanella et al., 2007).

SMN reduction has been shown to have a particular impact on the functionality of the minor spliceosome (the U12 spliceosome) which functions to splice the rare class of Introns (ATAC introns) which are often associated with cellular alternative splicing events (Turunen et al.,

2013). In a severe SMA mouse model, RNA sequencing showed a significant increase in the number of retained U12-dependant introns across all tissues and demonstrated that this effect can be replicated in Hela cells via introduction of siRNA-mediated SMN knockdown (Doktor et al., 2017). Treatment with an anti-sense oligonucleotide (ASO) was found to ameliorate this effect in the mouse model, reversing many of the splicing deficits observed. More recently, virus-mediated snRNA gene delivery in both a severe SMA mouse model and mammalian cell culture knockdown models of SMN improved observed aberrant splicing and rescued the loss of proprioceptive sensory synapses onto motor neurons in mice. Importantly, this occurred without an increase in the level of SMN (Osman et al., 2020). For these reasons (as well as the previously noted splicing error that occurs during transcription of SMN2) SMA is sometimes classified as a spliceopathy.

1.4 SMN and the Cytoskeleton

Beyond its pivotal role in RNA splicing and the nuclear translocation of snRNPs, SMN has also been found to be an important component of other cellular trafficking dynamics. SMN has been shown to localize to the growth cones of extending motor axons and live-cell imaging has demonstrated bidirectional transport of SMN along the length of the axon (H. L. Zhang et al., 2003). SMN colocalises in extending neurites and dendrites with several RBPs critical for the transport of mRNA throughout the cell for local translation, such as HuD and IGF2 mRNA-binding Protein 1 (IMP1), further strengthening the role that SMN plays in RNA trafficking within the cell (Zhang et al., 2006). Similarly, in PC12 cells, the SMN protein has been shown to interact with hnRNP R to effectively transport localise beta-Actin mRNA to the extending neurites (Rossoll et al., 2003). Defects in the local translation of actin mRNA

would go towards explaining the decreased growth and neurite extension as observed in SMA, due to its requirements at the extending growth cone (Faravelli et al., 2023).

Given its presence within the growth cones of extending neurites and the previously discussed role in trafficking of beta-Actin mRNA, it is perhaps unsurprising that SMN has been shown to influence the dynamics of the cytoskeleton. The cytoskeleton is a key component in the function and architecture of the axon. SMN has been found to interact with the profilin proteins, key regulators of the actin cytoskeleton (Giesemann et al., 1999).

Profilins act to sequester actin monomers and dynamically regulate actin in response to stimuli. Profilin 2 has been found to preferentially interact with SMN, wherein SMN acts as a regulator of the profilin activity. Upon binding to actin monomers, profilin reduces the polymerization of f-actin. However, in the presence of SMN, this sequestering effect is reduced leading to increased f-actin polymerization and increased growth. In rat PC12 cells, a reduction in SMN levels leads to increased levels of actin-bound profilins, resulting in reduced f-actin and inhibited rates of growth (Sharma et al., 2005). Interestingly, Profilin 2 is also primarily expressed in neurons, linking altered actin dynamics to the neuronal-specific identity of SMA.

Although SMN2 copy number is considered the primary genetic modifier of SMA severity, other genes have also been found to modify the disease phenotype. Plastin-3 is a protein with roles in actin-bundling and growth cone stabilization. It is coded for by the PLS3 gene

located on the X chromosome and - via transcriptomic analysis of phenotypically discordant siblings with the same SMA genes - was found to be a protective modifier of SMA (Oprea et al., 2008). Overexpression studies of the protein have been shown to rescue the SMA neuronal phenotype and growth in intermediate Mouse models of the disease (Kaifer et al., 2017). There is, however, discussion surrounding the precise role of this protein; human patients with PLS3 mutations (leading to complete loss or reduction of Plastin-3 protein) do not display an SMA-like phenotype (Alrafiah et al., 2018). Instead, they show increased risks of diaphragmatic hernia and osteoporosis (Petit et al., 2023). Currently, the mechanism by which plastin-3 ameliorates the SMA phenotype is still unresolved.

1.5 The Pathobiology of SMA

While the precise pathomechanism underlying the neuronal cell death in SMA still remains undefined, motor neurons in animal models of the disease display distinct cellular phenotypes as the disease progresses.

In an RNA interference (RNAi) based model of SMA in *Caenorhabditis Elegans* (*C. Elegans*), early dysfunction of the neuromuscular junction (NMJ) is followed by visible defects in the axon, with processes including neuronal sprouting showing evident disruption (Gallotta et al., 2016).

Similar pathology has been observed in zebrafish models of the disease. Using morpholinos to reduce the level of endogenous SMN, affected fish display varying severities of

phenotype, including truncated motor axons with aberrant branching and incorrect innervation of neighbouring myotomes (Figure 1.4) (Carrel et al., 2006).

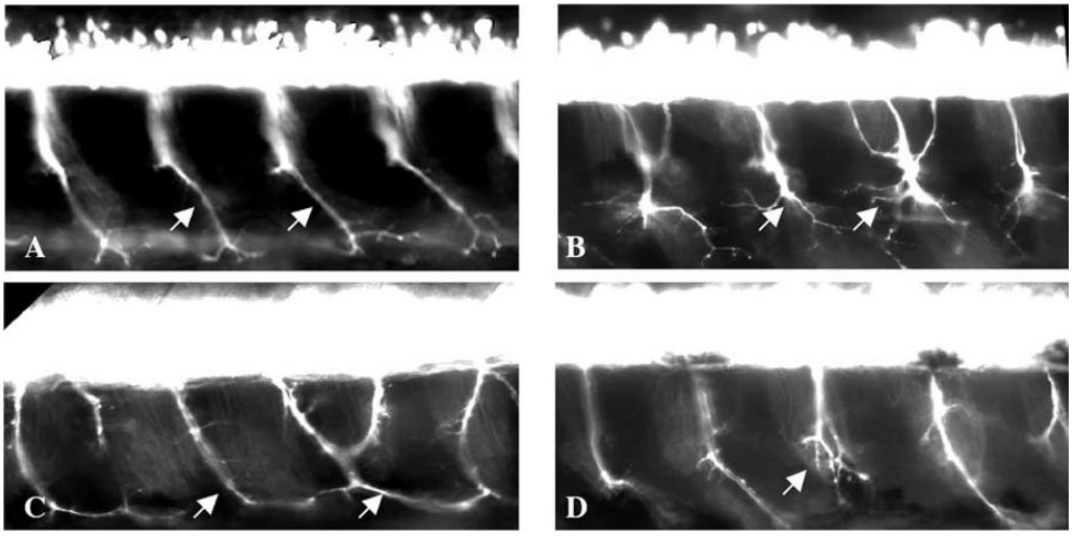


FIGURE 1.4 MOTOR AXON DEFECTS OBSERVED IN ZEBRAFISH LARVAE UPON SMN REDUCTION.
(A) In zebrafish larvae (2 days post-fertilisation), motor axon pathways follow a stereotypical pattern. (B) In larvae with reduced SMN, severely affected fish display multiple axonal truncations with or without branching. (C) Moderately affected fish display a more chimeric display of deficits, ranging from several neurons innervating a neighbouring myotome to (D) single truncations of the neuron. Image adapted from *(Carrel et al., 2006).*

In the $Smn^{2B/-}$ mouse model of SMA - which displays a moderate disease phenotype with increased lifespan compared to more severe mouse models - gross motor neuron defects are observed at later stages of the disease. Here, NMJ defects such as swelling of the presynaptic terminal have been observed to occur prior to the onset of axonal dysfunction, and loss of the motor neuron cell body is a late-stage event in this cascade (Courtney et al., 2019).

Together, these observations have given rise to the axonal theory of SMA pathogenesis, whereby presynaptic dysfunction at the NMJ caused by a lack of SMN triggers a “die back” mechanism that results in the eventual death of the motor neuron (Carrel et al., 2006).

1.6 A Role for SMNΔ7 in SMA

SMA is a primarily human disease and does not occur naturally in mammals. As described previously, complete knockout of SMN protein results in embryonic lethality in mice. Only humans and bonobos have been found to have copies of the SMN backup gene, SMN2, however chimpanzees can have multiple copies of the SMN1 gene (Swoboda, 2014).

As is evident from previous discussion, the intracellular role of SMN protein is both diverse and ubiquitous. A prime example of a “moonlighting” protein, it is an important component of numerous biochemical pathways and is integral to cellular function. As previously noted, the gene product of SMN2 – the “backup” SMN gene – is primarily SMNΔ7. Lacking exon 7 due to alternative splicing, the 16 amino acids of the C-terminus of the intact isoform are replaced with 4 amino acids from exon 8, EMLA (Le et al., 2005).

1.7 SMNΔ7 is an Unstable Protein Isoform

Described as unstable, SMNΔ7 has been shown to have a half-life of ~3hrs compared to SMN at ~8hrs. Investigating the reason for this instability, Cho and Dreyfuss (2010) used luciferase-based reporters fused to variations of SMN and SMNΔ7 to determine the amino

acid sequence responsible for SMNΔ7's instability. They concluded that the YG box domain and the EMLA C-terminal amino acids together act as a potent degron, or protein degradation signal, which when exposed at the C-terminus, initiates breakdown of the protein. This can help to explain the shorter half-life of the isoform (Cho & Dreyfuss, 2010).

Other factors have also been found to influence the stability of SMNΔ7. The isoform has been shown to oligomerize with native SMN in cell culture, as illustrated by several co-immunoprecipitation studies (Le et al., 2005). Native SMN, however, self-associates via motifs located in exons 2, 6 and 7 (Burnett et al., 2009). As SMNΔ7 lacks exon 7, it is not surprising that although it oligomerizes with native SMN, the efficiency of its oligomerization is reduced, contributing towards the increased instability observed.

Stability of SMNΔ7 has also been found to vary depending on cell type, although the precise cause for this is currently unknown. Expression of the isoform is much less stable in neurons than in other cell types – for example, fibroblasts - indicating that stability is also regulated in a tissue-specific manner, potentially contributing towards SMA's neuronal pathology (Vitte et al., 2007). Vitte and colleagues postulate that un-complexed, monomeric SMNΔ7 exposes the degron as described by Cho & Dreyfuss (Cho & Dreyfuss, 2010). Oligomerization with native SMN thus extends the half-life of the isoform by concealing the degron, increasing viability. In support of this, reducing endogenous levels of SMN within the cell is found to reduce the half-life of SMNΔ7 even further, to around only 65 minutes (Burnett et al., 2009). Targeting stability of SMNΔ7 and SMN complexes could provide a potential avenue of therapy in treatment of SMA. While oligomers of SMN have a half-life of hours,

when incorporated into larger, macromolecular structures like the SMN complex, this half-life extends to days (Burnett et al., 2009).

Oligomerization is not only essential for stability of SMN and SMNΔ7, it has also been shown to be necessary for proper localization of both proteins within the cell (Lorson et al., 1998). Native SMN is found concentrated in the nucleus, in either gems or Cajal bodies. It is also located diffusely throughout the cytoplasm where it oligomerises and forms the SMN complex in conjunction with several previously discussed cytoplasmic proteins (Young et al., 2001). SMNΔ7, on the other hand, is generally confined to the nucleus. This can be explained by the fact that exon 7 contains a cytoplasmic localization sequence; lacking exon 7, SMNΔ7 does not easily translocate beyond the nucleus (See Figure 1.5) (Honglai L. Zhang et al., 2003).

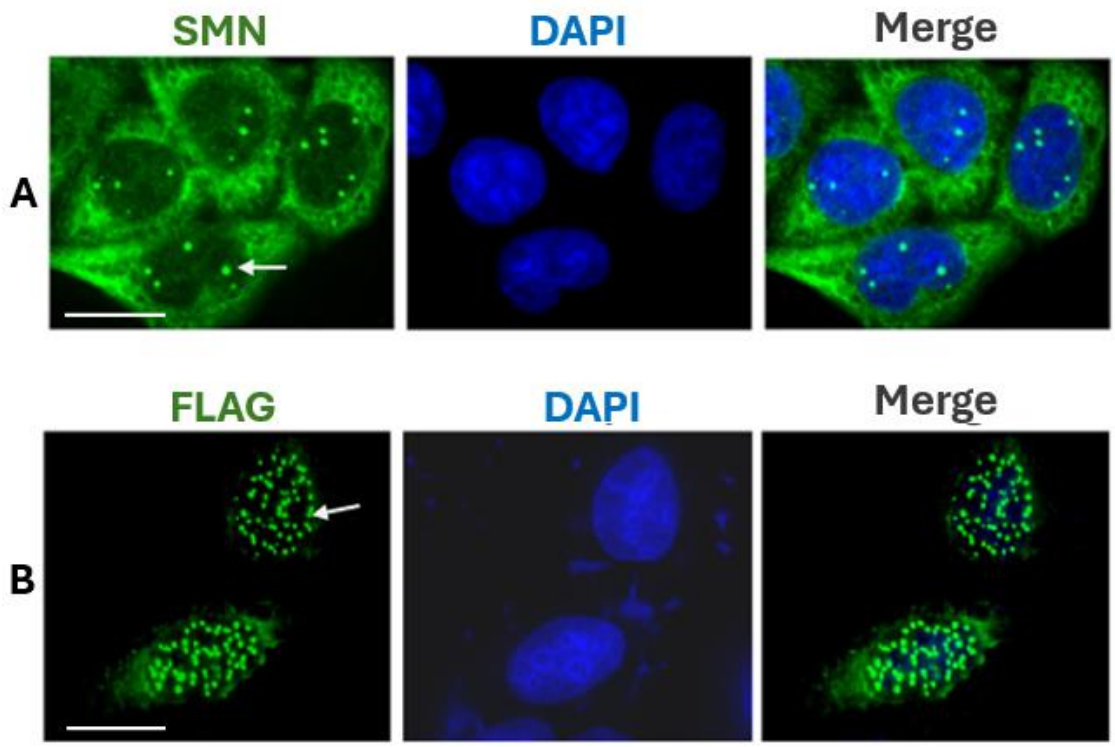


FIGURE 1.5: GFP-SMN AND FLAG-SMNΔ7 DEMONSTRATE DIFFERENT PATTERNS OF LOCALISATION IN HELA CELLS. A) GFP-Tagged SMN localises to both the nucleus in concentrated nuclear bodies (Cajal Bodies and gems) and diffuse within the cytoplasm. White arrow denotes nuclear bodies containing SMN. B) FLAG-tagged SMNΔ7 demonstrates a cellular localization almost exclusively nuclear, similarly concentrated in discrete specks as denoted by the white arrow, when stained with mouse Anti-FLAG. Scale = 20 microns. Image adapted from *(Han et al., 2012).*

Of particular note in the study of the SMNΔ7 isoform is research performed by Le and colleagues from 2005. Although SMNΔ7 is often described as being "non-functional", increasing levels of SMNΔ7 alone were here shown to improve the SMA phenotype in mice, increasing both weight and length of survival. This points towards SMNΔ7 retaining some functionality, particularly in a dose-dependent manner and when in complex with native SMN (Le et al., 2005).

Since Le *et al's* study, the SMNΔ7 mouse model has become popular and is used frequently in SMA research. Knockout of the mouse Smn gene and insertion of two copies of human SMN2 combined with an SMNΔ7 transgene results in mice which display a specific, semi-severe SMA phenotype; they produce very little native SMN protein. These mice survive up to two weeks after birth, further supporting the idea that SMNΔ7 retains some functionality *in vitro* (Eshraghi et al., 2016). It is worth noting, however, that the SMN2 genes would also produce a limited amount of full-length SMN.

1.8 Cellular Degradation of SMN and SMNΔ7

Studies of how the SMN protein – and SMNΔ7 – are degraded in the cell have shed light on the systems governing their disposal. Proteasome inhibitors such as lactacystin have been shown to increase the half-life of both SMN and SMNΔ7, confirming they are both targeted by the ubiquitin-proteasome system (Sanchez-Lanzas & Castano, 2017). Although both proteins are ubiquitinated prior to degradation by the proteasome, there are several deubiquitinating enzymes which have proven of interest to researchers targeting new avenues of SMA treatment, as SMN/SMNΔ7 deubiquitination leads to an increased half-life of the proteins (Abera et al., 2016). Of these deubiquitinating enzymes, Usp9x was found to be responsible for the deubiquitination and stabilization of native SMN, but not SMNΔ7. SMN is generally mono- or di-ubiquitinated whereas SMNΔ7 is polyubiquitinated, helping to explain the instability and rapid degradation times observed with the truncated protein compared to the native form (Han et al., 2012).

Interestingly, a non-SMN related variant of SMA known as X-linked Spinal Muscular Atrophy (XL-SMA) results from mutations in the Ubiquitin-Like Modifier Activating Enzyme 1 (UBA1) (Ramser et al., 2008). UBA1 functions to activate ubiquitin in the cell, initiating the ubiquitination cascade and activating the Ubiquitin-Proteosome System (UPS). Patients with mutations in the UBA1 gene display phenotypes similar to those observed in severe Type 1 SMA (Balak et al., 2017).

Notably, increasing levels of UBA1 via expression of an adeno-associated virus protein construct (AAV-UBA1) in SMA severe mice has been shown to ameliorate weight loss and increase survival, improving neuromuscular pathology; increased expression was also found to rescue motor axon pathology in zebrafish models of the disease (Powis et al., 2016). While this may seem counterintuitive (as reduction in UBA1 should lead to increased levels of SMN protein and vice versa) this demonstrates the complexity of the precise underlying pathology observed in SMA. UBA1 has also been implicated in neuroprotection and increased levels have been found to decrease Wallerian degeneration in mouse models of neuronal vulnerability (Wishart et al., 2008). From this, the widespread proteasomal dyshomeostasis that is seen in SMA may thus be a potential therapeutic target for future non-SMN related therapies.

1.9 SMA in the Era of Genetic Medicine

Recent advances in therapeutics aimed at the treatment of SMA now means that the disease – particularly SMA Type 1 – has a much better prognosis. The first drug to be

officially licensed for the treatment of SMA is Nusinersen, or Spinraza, an antisense oligonucleotide which works as a splicing corrector for the SMN2 gene product. Nusinersen binds to an intronic splice silencer (ISS) located in intron 7 of the SMN2 pre-mRNA (Bertini et al., 2017). This prevents the binding of associated splicing factors, increasing exon 7 retention and leading to a greater proportion of full-length SMN being produced (Singh et al., 2006). Delivered by intrathecal injection as it does not cross the blood-brain barrier, nusinersen has shown positive results in SMA Types 1, 2 and 3 patients, increasing both survivability and motor milestones achieved (Cho et al., 2023).

As nusinersen is delivered intrathecally, it is not without complications: back pain, lumbar-puncture syndrome and infection are just some of the potential side-effects from this method of delivery. Other promising treatments, however, have been developed which can be delivered via the systemic circulation. Risdiplam is a small molecule therapy which can be taken in pill form. It has been shown to increase the levels of full-length SMN in patients with SMA via splicing modification of the SMN2 gene (Schorling et al., 2019). Like Nusinersen, Risdiplam binds directly to SMN2 pre-mRNA and displaces protein factors which promote the removal of Exon 7. However, it does not bind at the same location on the unspliced pre-mRNA as Nusinersen (the Intron 7 ISS) and thus both therapies can potentially be used in combination. By binding to SMN2 pre-mRNA at two distinct sites (the Exonic Splicing Enhancer 2 in Exon 7 and the 5’ Splice Site in Intron 6) Risdiplam promotes the inclusion of Exon 7 into the SMN2 mRNA product (Baranello et al., 2021). Like Nusinersen, Risdiplam has been found to be effective at reducing the motor symptoms in SMA patients (Qiao et al., 2023).

Drug therapy, however, is only one avenue of treatment for the disease. Zolgensma is an Adeno Associated Viral serotype 9 (AAV9) vector encoding a copy of the wild-type SMN1 gene (Haggerty et al., 2019). One of the first FDA-approved gene therapies, Zolgensma replaces the non-functional SMN1 gene found in SMA patients with a working copy, allowing production of native SMN. While there is ongoing debate regarding efficacy of the vector's delivery route (systemic versus intrathecal), the drug has been shown to have promising results in Type 1 patients under 6 months of age (Schorling et al., 2019).

Treatment, however, does not come without challenges; the issues of whom to treat and when to treat being foremost amongst them. SMN is integral during gestation and in SMA type 1 motor neuron loss occurs rapidly in the first 6 months of life (Swoboda et al., 2005). Thus, the earlier a patient is identified and treatment is initiated the better the potential outcome. Newborn screening (NBS) for SMA is becoming more commonplace with up to 65% of European countries currently screening for SMA in their NBS panels (Gailite et al., 2022). Most drugs also have off-target effects, and these must also be taken into account.

1.10 SMA as a Systemic Disease

Potentially the most interesting aspect of the current revolution in SMA therapeutics is that treating motor neurons alone – as occurs during intrathecal treatment with Nusinersen or Zolgensma – and the changing prognosis of patients may further reveal the systemic nature of SMA as a disease. Other tissues - beyond alpha motor neurons and muscle fibres - have

been found to show distinct disease phenotypes in SMA, perhaps unsurprisingly considering the ubiquitous nature of the SMN protein.

Cardiac tissue abnormalities have been noted in human SMA patients, with congenital hypoplastic left heart syndrome being the most common described (Menke et al., 2008). These defects occur during development and support the idea that SMN is essential for organogenesis. Mouse models of the disease have also recapitulated these cardiac defects, which were observable in advance of any overt motor neuron loss. Beyond structural abnormalities, electrical defects have also been observed. SMA patients often show symptoms of arrythmia and electrical abnormalities have been demonstrated in the early stages of mouse models (Shababi et al., 2010). Similarly, microvasculature anomalies have been observed in both mouse models and human SMA patients, which were resolved with the use of an SMN-enhancing ASO (Zhou et al., 2022).

Pancreatic abnormalities are also found in those with SMA, with both mice and human pancreas tissue showing mis-localization of the pancreatic islet alpha cells and an increased number of beta cells. This may help to explain why hyperinsulinemia and an increased sensitivity to glucagon are both common in SMA patients (Simone et al., 2016).

Other metabolic defects have also been shown to occur in SMA, with alterations in fatty acid metabolism being foremost amongst these. Altered regulation of lipid metabolism has been noted in both human patients and mouse models of the disease. Human patients have

shown an increased susceptibility to dyslipidemia and abnormal liver steatosis, with similar observations present prior to denervation in mouse models of the disease (Deguise et al., 2019). Together, these findings suggest SMA is more than just a disease of the neuromuscular system. As treatment options for patients continue to progress and those with more severe forms of the disease live longer, these "off-target" symptoms will become better recognized and understood.

1.11 Inflammation and SMA

As in many other neurological disorders, inflammation has been shown to play a role in the pathology of SMA. In early stages of the disease, severe mouse models have been found to develop features consistent with systemic inflammation whilst changes in the size of the thymus and cytokine levels have also been observed (Wan et al., 2018).

In human patients, changes in both astrocytic and microglial activation - which are markers of inflammation - have been observed in several studies (McGivern et al., 2013) (Khayrullina et al., 2022). Specific immune signalling factors, including the cytokine Interferon Gamma (IFN-γ), have been demonstrated to upregulate levels of SMN protein *in vitro*, due to the presence of a putative interferon response element (IRE) in the SMN promoter region (Baron-Delage et al., 2000). This may be of potential clinical relevance, as it suggests that IFN-γ could be used in the treatment of SMA to increase levels of SMN.

While it is likely that the inflammation response contributes to SMA pathology and prognosis, the role that inflammation plays in the observed pathogenesis of SMA is still an area of active research.

1.12 Outstanding Questions in SMA Research

Although accumulating evidence points towards SMA being more systemic in nature than previously thought, one of the primary questions in SMA research is why the selective death of alpha motor neurons occurs in the disease. There are several potential reasons why this may be the case. Alpha motor neurons are large cells and reduced levels of SMN may be felt more strongly here; rescue of the SMA phenotype by SMN is dose-dependent, as can be seen from the ameliorating effects of SMN2 copy number (Harada et al., 2002). The roles of SMN in trafficking of mRNA along the axon may be integral to the neuron's synaptic stability; reduction in SMN levels will result in reduced intracellular trafficking and reduced local translation at the synapse, potentially explaining the neuronal die back observed at the neuromuscular junction (Fallini et al., 2012).

It has also been shown that the pre-mRNA splicing of SMN2 to full length SMN is less efficient in motor neurons compared to other neuronal populations (Ruggiu et al., 2012). While the cause for this is unknown, when factored together with other neuronal-specific interactions of the SMN protein – such as the previously discussed Profilin 2 association – it can be surmised that numerous negative factors act in concert to exert an effect which renders alpha motor neurons particularly susceptible to damage (Wang et al., 2019).

Untangling the contribution and weight of each of these factors to SMA pathology will likely be the focus of future research.

1.13 Interactomics and SMA

Proteomics describes the large-scale study of proteins with the goal of understanding and exploring the proteome. This proteome is comprised of the entire repertoire of proteins produced by an organism. Being dynamic, the proteome reflects changes in both cellular environment and physiological state (Al-Amrani et al., 2021). Mass Spectrometry (MS) is an analytical technique which measures the mass to charge ratio of ions; due to its sensitivity, accuracy and ability to analyse a broad range of protein types and post-translational modifications (PTMs), MS has emerged as a central tool in proteomics. By utilising MS, analysis of complex protein mixtures can be performed to elucidate specific properties including protein structure, function and interaction with high levels of precision (Nanni, 2022).

While studies have investigated the proteome of native SMN in the context of SMA, the functionality of SMNΔ7 has yet to be fully explored using currently available investigative techniques (Prescott, 2009). Targeted proteomic analysis has been performed for SMN, however production of an interactome – a map of all of the protein-protein interactions in which a target protein is involved – would allow a greater understanding of the role of the truncated protein (Fuller et al., 2016). Specific comparative interrogation of the SMNΔ7 isoform may shed new light on which functions - as defined by protein-protein interactions -

are retained. And, perhaps more importantly, which functions are lost due to the deletion of exon 7 and C-terminus amino acids. An understanding of which functions the isoform cannot perform may provide an important therapeutic insight for the design of future treatments of the disease, and allow for a greater understanding of SMN, SMNΔ7 and their contribution to SMA pathology.

1.14 Interactomic Techniques to Interrogate the Interactome of SMN and SMNΔ7

The FLAG tag is a short peptide sequence consisting of eight amino acids (DYKDDDDK) that can be fused to proteins for specific downstream applications in molecular biology. Adding a FLAG tag to a protein enables specific detection and study of the tagged proteins of interest. (Hopp et al., 1988). Antibodies that specifically recognize the FLAG tag can be used for various applications, such as FLAG-Immunoprecipitation (FLAG-IP). This is a variant of the common protein co-immunoprecipitation methodology: FLAG-IP utilises beads coated with FLAG-binding antibodies for efficient isolation of any FLAG-tagged proteins (in this case SMN and SMNΔ7) and their stable interacting partners. FLAG-tagged SMN and SMNΔ7 (along with their interacting protein partners) can thus be separated from complex protein mixtures and analysed by mass spectrometry (Fedorova & Dorogova, 2019).

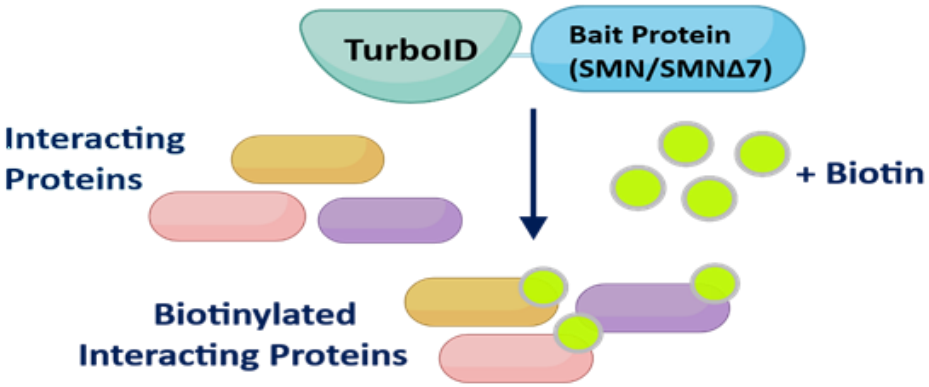


FIGURE 1.6: TURBOID CAPTURES PROTEIN INTERACTORS OF ATTACHED BAIT PROTEINS.

Upon the addition of biotin, interacting protein partners of a selected bait protein fused to TurboID become biotinylated, allowing for efficient pulldown of tagged proteins using Streptavidin.

TurboID is a proximity-dependant labelling technique which labels the interacting partners of a protein of interest (in this case, SMN/SMNΔ7) (Figure 1.6). The TurboID protein is a modified BirA biotin ligase which, when conjugated to the FLAG-SMN/SMNΔ7 protein construct, will biotinylate all proteins present within a 10-nanometer sphere upon the addition of biotin to culture media (Li et al., 2019). TurboID is a variation on the BioID (BirA) labelling technique, produced via targeted mutagenesis of the BirA ligase (Branon et al., 2018). This results in a variation of BirA which works more rapidly, reducing the time cells are exposed to biotin thus resulting in less false positives and "noise" that can occur in longer incubation times *in vitro*. Unlike FLAG-IP which reveals only stable PPIs, TurboID can capture transient, weaker protein interactions (Samavarchi-Tehrani et al., 2020).

As previously stated, the classic disease course of SMA is currently in flux due to changing treatment paradigms. It thus becomes increasingly important to elucidate the pathways involved in the pathology of the disease. As all SMA patients produce SMNΔ7, this makes it an attractive target for novel therapeutics. With the techniques described above - in tandem with downstream Mass Spectrometry analysis - it then becomes possible to decipher the differences in interactomes of SMN of SMNΔ7, further expanding the knowledgebase on a protein that has historically been both understudied and undervalued in the field of SMA research. Utilising these techniques makes it possible to design an experimental workflow that can produce datasets to elucidate which interactions are modified between full-length SMN and SMNΔ7. By culturing cells in SILAC (Stable Isotope Labelling by Amino Acids in Cell Culture) media, ratios can be compared between different SILAC conditions (Light/Medium/Heavy labels) to assign quantitative values to mass spectrometry data and precisely compare the differential interactivities of each protein isoform (Chen et al., 2015) (Figure 1.7).

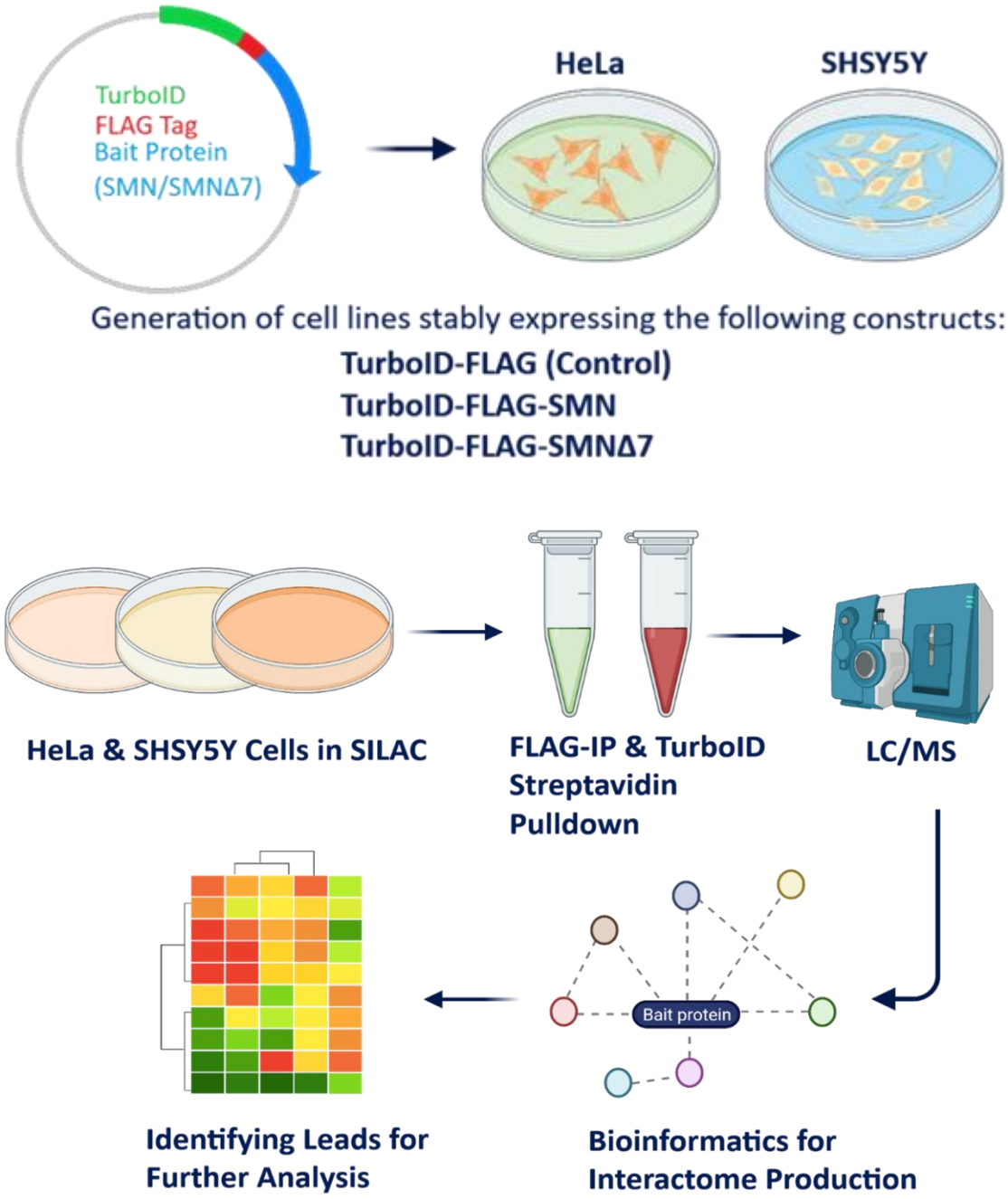


FIGURE 1.7: EXPERIMENTAL WORKFLOW TO IDENTIFY INTERACTING PARTNERS OF SMN AND SMNΔ7 USING TURBOID AND MASS SPECTROMETRY.

A graphical overview of the proposed experimental workflow, utilising human Hela and SHSY5Y stable cell lines expressing TurboID-SMN and TurboID-SMNΔ7 for FLAG-Immunoprecipitation/streptavidin pulldowns. Mass spectrometry (MS) of SILAC-labelled samples and subsequent data analysis of raw MS files allows for production of quantitative interactomes and identification of novel protein interactors.

Chapter 2: Materials and Methods

2.1 Plasmid Design and Generation

Plasmid generation was outsourced to Aruru Molecular and VectorBuilder. For in-depth discussion of design choices and plasmid maps, see Chapter 6.

2.1.1 All-In-One TurboID Plasmids

Upon extensive discussion with Aruru Molecular, plasmids expressing TurboID fused to both a FLAG tag and the proteins of interest (SMN and SMNΔ7) were designed and implemented into a modified lentiviral vector. It was agreed that Aruru would produce 3 deliverable plasmids as the final product: TurboID-FLAG (Control), TurboID-FLAG-SMN, TurboID-FLAG-SMNΔ7 (Table 2.1)

2.1.2 Lentiviral TurboID Plasmids

VectorBuilder produced a series of lentiviral plasmids which fit the prior mentioned specifications. Due to the “leakiness” inherent within an all-in-one lentiviral Tet-On system, these plasmids required a Helper plasmid containing both the Tet-On Transactivator (active in the presence of Doxycycline) and Tet Silencer (active in the absence of Doxycycline) to be transduced into cells prior to the TurboID viral particles, which express TurboID constructs under the control of a Tet Promoter upon addition of Doxycyline (1ug/ml). 4 different plasmids were delivered: pLVX Tet-On Helper (containing the Tet-responsive transactivator elements TetR and Tet Silencer), eGFP-mCherry, FLAG-TurboID-Control, FLAG-TurboID-SMN,

FLAG-TurboID-SMNΔ7 which were validated before shipping and passed VectorBuilder QC checks (Table 2.1). The three TurboID plasmids also contained a far-red fluorescent reporter (mKate2) to evaluate transduction efficiency.

Name	Type	Protein	Supplier	Antibiotic Resistance	Selection Marker	Fluorescent Marker	Tags
AIO-TurboID	Mammalian	TurboID Only	Aruru	Ampicillin	Puromycin	None	1xFLAG
AIO-TurboID-SMN	Mammalian	TurboID-SMN	Aruru	Ampicillin	Puromycin	None	1xFLAG
AIO-TurboID-SMNΔ7	Mammalian	TurboID-SMNΔ7	Aruru	Ampicillin	Puromycin	None	1xFLAG
pLV-FLAG-TurboID	Lentivirus	TurboID Only	VectorBuilder	Ampicillin	Puromycin	mKate2	5xFLAG
pLV-FLAG-TurboID-SMN	Lentivirus	TurboID-SMN	VectorBuilder	Ampicillin	Puromycin	mKate2	5xFLAG
pLV-FLAG-TurboID-SMNΔ7	Lentivirus	TurboID-SMNΔ7	VectorBuilder	Ampicillin	Puromycin	mKate2	5xFLAG
pLV-TetOn Helper	Lentivirus	Tet-On Transactivator & Silencer Elements	VectorBuilder	Ampicillin	Hygromycin	None	None
pLV-eGFP-mCherry	Lentivirus	eGFP and mCherry	VectorBuilder	Ampicillin	Puromycin	eGFP and mCherry	None

TABLE 2.1: LIST OF PLASMIDS USED IN EXPERIMENTS.

All plasmids are Dox inducible. AIO = All-in-One dox system (i.e., no 2-step transfection process with helper plasmids, as was necessary with the lentiviral constructs)

2.1.3 Validation of All-In-One TurboID Plasmids

2.1.3.1 Restriction Digest Reactions

Plasmid samples of AIO-TurboID-Control, AIO-TurboID-FLAG-SMN and AIO-TurboID-FLAG-SMNΔ7 were digested either with EcoRI alone (Promega) to linearise the plasmid (single digest) or EcoRI and HindIII (Promega) (double digest) for 1 hour at 37 degrees Celsius. Digestion reactions were quenched via heat inactivation of samples for 15 minutes at 65 degrees Celsius.

2.1.3.2 Gel Electrophoresis

Samples of digested constructs were loaded onto a 0.5% agarose gel stained with Safeview nucleic acid stain (Thermo). HyperLadder 1kb (Bioline) was used to measure band size. The gel was run in TBE buffer (500ml; Tris: 4.85 g, Glacial acetic acid: 1.4 mL, 0.05M EDTA (pH 8.0)) for 40 minutes at 90 volts and visualised beneath UV light using a transilluminator.

2.1.3.3 Plasmid DNA Mini-Preps and Purification:

Recombinant *E. coli* transformed with plasmids coding for each of the AIO TurboID plasmids were provided by Aruru Molecular as glycerol stocks. *E. coli* from each were streaked to give single colonies on agar plates supplemented with Neomycin. For each plasmid, a single colony was selected and inoculated in 10 mL of luria broth (LB) overnight at 37°C on a shaker. Plasmids were isolated using E.Z.N.A. Plasmid Mini Kit (Omega Bio-tek) following

manufacturer protocol. DNA purity and concentration was measured utilising a NanoDrop spectrophotometer.

2.2 Generation of Human Stable Cell Lines

HeLa cells (human cervical epithelial carcinoma) and SHSY5Y cells (human neuronal) were provided by Dr Judith Sleeman (University of St Andrews, UK) (Ross et al., 1983). A549 cells were provided as a gift from Dr Catherine Adamson (University of St Andrews, UK) (Giard et al., 1973). Unless otherwise stated, cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Corning) supplemented with 10% Foetal Bovine Serum (FBS) (Gibco) and 1% penicillin/streptomycin (Thermo). Cells were incubated at 37°C and 5% CO2 and passaged routinely upon reaching ~70% confluency. Cells were passaged at a 1:10 ratio for optimal health and proliferation.

For Interferon-Gamma Inhibitor of Translation (GAIT) Complex studies, 2000 units of recombinant human interferon-gamma (gifted by Dan Young, University of St Andrews) were added per 1mL of media for 24 hours.

2.2.1 TurboID Plasmid Transfection

HeLa and SHSY5Y cells were seeded at a concentration of $1x10^6$ cells in 6cm culture dishes and after 24 hours were transfected with 1µg of either AIO-TurboID-FLAG, AIO-TurboID-FLAG-SMN or AIO-TurboID-FLAG-SMNΔ7 using Effectene transfection reagent (QIAGEN) according to manufacturer's protocol. 48 hours post-transfection, cells were moved to T25

flasks for selection with either 200ng/ml (HeLa) or 500ng/ml (SHSY5Y) of puromycin. Selection medium was replaced every 2 to 3 days, for a total of 7 to 10 days. These cells were then either transferred to T75 flasks for expansion, or seeded, induced and stained as described below to verify plasmid incorporation and correct localisation.

2.2.2 Transient Transfections of GFP-SMNΔ7

For both GAIT Complex and Amino Acid tRNA Synthetase (AARS) Investigations, 0.5ug of GFP-SMNΔ7 plasmid (provided by Dr Judith Sleeman) was transfected (according to manufacturer's protocol) using Effectene transfection reagent into 10cm dishes which had been seeded 24 hrs prior with $8.5x10^5$ HeLa (AARS experiments) or SHSY5Y cells (GAIT Complex experiments). For cells transfected for downstream imaging, this initial seeding was performed in dishes lined with multiple glass coverslips. Cells were then processed after 48hrs.

2.2.3 TurboID Plasmid Transduction

2.2.3.1 Establishing Tet-On Helper Parental Cell Lines

$8.5x10^5$ HeLa or SHSY5Y cells were seeded into T25 flasks and after 24 hours, media (DMEM and 10% FBS) was replaced with serum-free equivalent. Cells were then transduced with a volume of either pLVX Tet-On Helper or pLVX eGFP-mCherry viral supernatant equivalent to a Multiplicity of Infection (MOI) of 10. After 6 hours, 20% FBS-Supplemented media was added to the flask. 72 hours post-transfection, cells were moved to T25 flasks for selection with 300ug/ml (HeLa) or 350ug/ml (SHSY5Y) of hygromycin. eGFP and mCherry expression

were visualised using the EVOS Imaging System to evaluate transduction method efficiency using the GFP and mCherry filters. Selection medium was replaced every 2 to 3 days for 7 to 10 days. Tet-On cells were then transferred to T75 flasks for expansion then either frozen down as cell stocks, expanded for harvesting for western blot analysis, or transduced with the pLVX TurboID plasmids.

2.2.3.2 Establishing TurboID Cell Lines Utilising Tet-On Helper Parent Cells

$8.5x10^5$ HeLa Tet-On and SHSY5Y Tet-On cells were seeded into T25 flasks and after 24 hours, media was replaced with serum-free equivalent. Cells were then transduced with a volume of either FLAG-TurboID, FLAG-TurboID-SMN or FLAG-TurboID-SMNΔ7 viral supernatant equivalent to an MOI of 10. After 6 hours, 20% FBS-Supplemented media was added to the flask. 72 hours post-transfection, cells were moved to T25 flasks for selection with either 200ng/ml (HeLa) or 500ng/ml (SHSY5Y) of puromycin. Selection medium was replaced every 2 to 3 days for 7 to 10 days. These cells were then transferred to T75 flasks for expansion then either frozen down as cell stocks, expanded for harvesting for western blot analysis, or seeded onto coverslips for immunofluorescent cell staining.

2.3 TurboID Induction of Protein Expression

Transduced cells post-selection with puromycin were seeded onto coverslips and induced with 1ug/ml doxycycline. After 48 hours induction, cells were incubated for 1 hour at 37 degrees Celsius with 150uM of Biotin (ThermoFisher).

For pLVX plasmids, visualisation of mKate2 fluorescent reporter was performed using the EVOS M5000 Imaging System (ThermoFisher). Cells were also fixed and stained with 488-Streptavidin (Licor) as described below.

2.4 Cell Staining

Coverslips were washed in PBS and 5ml of 3.7% w/v paraformaldehyde fixative buffer (1.85g PFA, 10ul KOH, 3.5ml H2O and bring to 50ml with PBS) was added to the 10cm dish before incubating for 10 minutes at room temperature. PFA was removed and the coverslips were washed twice with PBS and stored in PBS at 4°C. Cells were permeabilised using PBS/0.1% (v/v) triton-X100 (Sigma) on a shaker for 15 minutes and PBS/1% (v/v) goat serum (Thermo) was added for 20 minutes to block the coverslips. Coverslips were incubated in primary antibodies diluted with PBS and 1% Goat Serum (See Table 2.2) for 1 hr in a humidity chamber. Coverslips were then washed three times in PBS, for 10 minutes each time. The cells were incubated in secondary antibody (see Table 2.2) in PBS and 1% Goat Serum for 1 hour and washed once in PBS for 10 minutes. Coverslips were then counter-stained using DAPI (Sigma, 1.6ug/ml) for 10 minutes, then washed finally in dH2O for 10 minutes. Coverslips were mounted onto slides using 8ul of Prolong Gold Antifade Reagent (Life Technologies). For an experimental example of antibody titration for immunofluorescent staining, please see Supplementary Data S1.

Antibody & Supplier	Species	Primary or Secondary	ICC	ICC Dilution	WB	WB Dilution
GAPDH (PTG)	Mouse	Primary	✓	1:200	✓	1:1000
Alpha-Tubulin (PTG)	Rabbit	Primary	×	NA	✓	1:1000
Tubulin	Mouse	Primary	×	NA	✓	1:1000
STAT1 (Santa Cruz)	Mouse	Primary	×	NA	✓	1:1000
SMN (BD Biosciences)	Mouse	Primary	✓	1:250	✓	1:500
SMN (Mansma)	Mouse	Primary	✓	1:50	✓	1:300
SMN (PTG)	Rabbit	Primary	✓	1:200	×	NA
SMN (Santa Cruz)	Rabbit	Primary	✓	1:200	✓	1:700
SYNCRIP (PTG)	Rabbit	Primary	✓	1:250	✓	1:500
PML (Santa Cruz)	Rabbit	Primary	✓	1:200	×	NA
FLAG (Thermo)	Mouse	Primary	✓	1:250	✓	1:1000
FLAG (Sigma)	Mouse	Primary	✓	1:250	✓	1:500
FLAG (Thermo)	Rabbit	Primary	✓	1:200	✓	1:1000
TetR (Takara)	Mouse	Primary	✓	1:500	✓	1:1000
AF-680 Anti-Rabbit (JIR)	Goat	Secondary	×	NA	✓	1:20000
AF-800 Anti-Mouse IgG (Licor)	Goat	Secondary	×	NA	✓	1:20000
AF-790 Anti-Mouse IgG (JIR)	Goat	Secondary	×	NA	✓	1:20000
AF-488 Streptavidin (Licor)	NA	Secondary	✓	1:350	×	NA
AF-680 Streptavidin (JIR)	NA	Secondary	×	NA	✓	1:1000
RARS (PTG)	Mouse	Primary	✓	1:500	✓	1:1000
IARS (PTG)	Rabbit	Primary	✓	1:100	✓	1:1000

TABLE 2.2: LIST OF ANTIBODIES USED FOR IMMUNOCYTOCHEMISTRY (ICC) AND WESTERN BLOT (WB).

Dilutions, species and supplier included. NA = not applicable as not used in this capacity in any experiments.

For GAIT complex experiments, induced (2000u/ml human recombinant interferon-gamma) and uninduced control SHSY5Y cell lines were seeded onto coverslips and fixed in 3.7% PFA prior to staining which was carried out as described above. Antibodies used in these experiments were: rabbit anti-SYNCRIP (PTG), rabbit anti-SMN (PTG), rabbit anti-GAPDH (PTG), mouse anti-SMN (Mansma), rabbit anti-Tubulin (PTG) and mouse anti-Tubulin (Santa Cruz).

For AARS studies, Hela cell lines (including transiently transfected Hela GFP-SMNΔ7) were seeded onto coverslips and fixed in 3.7% PFA prior to staining with relevant antibodies as described above. Antibodies used in these experiments were: mouse anti-RARS (PTG), rabbit anti-IARS (PTG), mouse anti-SMN (BD Biosciences) and rabbit anti-SMN (PTG)

For TurboID cell lines, coverslips containing induced cells (Doxycycline 1ug/ml for 48hrs, Biotin 150mM for 1hr) were fixed as described above. Antibodies used in these experiments were: streptavidin-488 (Licor), rabbit anti-SMN (Santa Cruz), mouse anti-SMN (Mansma), rabbit anti-FLAG (Sigma), mouse anti-FLAG (Invitrogen).

2.3 Fixed Cell Microscopy

Images were obtained utilising a DeltaVision RT microscope (Applied Precision) and the Olympus 100x NA 1.4 UPLS Apo objective. Varying exposure times were used depending on the cells imaged and antibodies used. A peak intensity of between 2600 and 3600 was set to reduce potential overexposure and binning of 2x2 selected. A z-stack of images were taken at a spacing of 0.2 um. The FITC (531 nm), TRITC (624 nm), DAPI (447 nm) and Cy5 (675 nm) filters were used for imaging.

Alternatively, images were obtained using the EVOS M5000 imaging system. Utilising the x40 objective, images were obtained using the DAPI (447 nm), Cy5 (692 nm), GFP (524 nm) and Texas Red (628 nm) filters.

2.4 Image Analysis

Images were processed using Volocity software version 6.3 (PerkinElmer). Images obtained from the DeltaVision microscope were deconvolved using the “Iterative Deconvolution” function. Colocalisation analysis was performed between deconvolved image channels of interest to generate the Pearson Correlation Coefficient in triplicate deconvolved images, with thresholds determined utilising the in-built automatic threshold function as designed by Costes *et al* (Costes et al., 2004).

Any adjustment of image brightness or switching of channel colour for the generation of figures was performed in Fiji software (https://fiji.sc).

2.5 Gel Electrophoresis & Western Blotting

2.5.1 Preparation of Whole Cell Lysates

Cells were pelleted at 250 RCF and pellets were resuspended and washed 3 times in ice-cold PBS. Pellets were then resuspended in 250-400µl cold Lysis Buffer and incubated on ice for 10 minutes. Lysates were then passed through QiaShredder columns (Qiagen) and the eluate was spun at 4°, 16000 RCF for 30 minutes to clarify. Total protein was measured via Bradford Assay (Expedeon). Lysates were diluted in water (1:20) to a volume final of 100 µL. 1.5 ml of Bradford Reagent (Sigma) was added, and samples were incubated in the dark for ten minutes. Absorbance was then measured at 595 nm in a Biophotometer (Eppendorf). Lysates were flash frozen in liquid nitrogen and stored at -80°C.

2.5.2 SDS-PAGE

Either 10% SDS-PAGE gels were hand-poured (See Table 2.3 for recipes) or NuPAGE 4-12% pre-cast gels (ThermoFisher) were utilised. 4X loading dye containing either 10% B-Mercaptoethanol (200mM Tris-HCl, pH 6.8; 8% Sodium Dodecyl Sulphate (SDS), 0.4% bromophenol blue, 40% glycerol) or 50mM DTT was added to lysates and heated to either 95°C for 10 minutes (samples containing B-Mercaptoethanol) or 65°C for 25 minutes (samples containing DTT). 40µg of protein was loaded alongside a 1kD molecular weight

marker Ladder (ThermoFisher) on the gel. Hand-poured gels were then run for 80 minutes at 150v, whereas NuPage pre-cast gels were run for 45 minutes at 200v.

Component	Resolving Gel	Stacking Gel
Acrylamide	6.66 ml	1200 ul
dH2O	8.3 ml	4.7 ml
Rx Buffer pH 8.8	5 ml	-
Rx Buffer pH 6.8	-	2 ml
TEMED	40 ul	40 ul
10% APS	80 ul	40 ul
Total Volume	**19.96 ml**	**6.7 ml**

TABLE 2.3: RECIPES FOR HAND-POURED ACRYLAMIDE GELS.

2.5.3 Western Blotting

Proteins were transferred to an Amersham Protran Nitrocellulose membrane (pore size 0.45 µm) (GE Healthcare) using a Turbo-Blot semi-dry blotter (BIORAD) utilising the appropriate pre-configured programme and runtime for the blot setup (e.g., for 1 blot, the 30-minute, Midi gel function was selected). Transfers were performed in 1X blotting buffer (48mM Tris base, 34mM Glycine, 0.04% SDS) supplemented with 20% methanol. To confirm efficient transfer, the membrane was stained with ponceau solution (SIGMA) for 5 minutes, washed with dH2O, scanned and subsequently de-stained with PBS. The blot was blocked in 5% milk and PBS for 1 hour at 4°C and incubated in primary antibodies (See Table 2.2) overnight at 4°C in 5% milk and PBS-T (0.05% V/V; 500ml PBS, 25ul Tween-20). The blot was washed for 3 times at 5 minutes in PBS-T. Secondary antibodies were added (see Table 2.2) in 5% milk

and PBS-T. incubating the blot for an hour at room temp. Finally, blots were washed three times for 5 minutes in PBS-T and imaged using the Odyssey® CLx imaging system (Licor). For Quantitation, bands of interest were normalised to levels of tubulin utilising Empiria Studio software (Licor).

2.6 SILAC Cell Culture

Light, Medium and Heavy SILAC media were prepared as previously described, supplemented with dialyzed FBS to avoid amino acid carry-over (See Table 2.4 for exact recipes) (Deng et al., 2019).

Condition	Media	Components
Light	DMEM (Corning)	- 10% dialyzed FBS (Gibco)
(R0K0)		- 1% Penicillin/Streptomycin (P/S) (Sigma)
		- 1mL Leucine (52mg/ml) (Sigma)
		- 0.5mL R0 arginine (84mg/ml) (Thermo)
		- 0.5mL K0 lysine (146mg/ml) (Thermo)
Medium	DMEM	- 10% dialyzed FBS
(R6K4)		- 1% Penicillin/Streptomycin (P/S)
		- 1mL Leucine (52mg/ml)
		- 0.5mL R6 arginine (84mg/ml) (Thermo)
		- 0.5mL K4 lysine (146mg/ml) (Thermo)
Heavy	DMEM	- 10% dialyzed FBS
(R10K8)		- 1% Penicillin/Streptomycin (P/S)
		- 1mL Leucine (52mg/ml)
		- 0.5mL R10 arginine (84mg/ml) (Thermo)
		- 0.5mL K8 lysine (146mg/ml) (Thermo)

TABLE 2.4: SILAC MEDIA RECIPES.

Components required for the creation of 500mls of SILAC Light (R0K0) (yellow cells in table), Medium (R6K4) (blue cells in table) and Heavy (R10K8) (green cells in table) media.

Cells were cultured for more than 3 passages in their corresponding media. Biological replicates were obtained from 2 different passages for each cell line, with the labels

switched between conditions (Table 2.5). Once labelled, cells were dissociated using Enzyme-Free Cell Dissociation Buffer (Gibco), spun at 200 RCF to pellet, washed and resuspended 3 times in ice-cold PBS, then flash frozen in liquid nitrogen before further processing. For the specific SILAC Labelling Scheme, please see Chapter 6, Table 6.3.

2.7 SILAC Sample Preparation for Mass Spectrometry Analysis

2.7.1 Cell Lysis

Pellets were resuspended in up to 1ml of fresh Low Salt RIPA buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate and one protease inhibitor tablet) (Roche) then sonicated on ice at 5 microns for 3 x 15 seconds, with 10 seconds of rest on ice between. Total protein was measured via Bradford Assay (Expedeon) using a Biophotometer (Eppendorf). Lysates were stored at -80°C.

2.7.2 FLAG Immunoprecipitation

FLAG-IP was performed according to the general immunoprecipitation protocol designed for Protein A Beads (Thermo) described elsewhere (Gerace & Moazed, 2015). In short, 2mg of input protein (derived from doxycycline-induced TurboID Hela Cell lines) was pre-cleared with 20µl of Sepharose beads via gentle rotation (15RPM) in an end-over-end shaker for 30 min at 4°C. Beads were then pelleted and 40µl of the supernatant was saved as the "Precleared" samples.

The remaining supernatant was added to either 30ul Protein A Beads (Thermo) and 2ug/ml of Invitrogen Mouse Anti-FLAG antibody (the “Antibody” samples) or 30ul of Protein A Beads only (the “No Antibody” samples – i.e. negative controls). All samples were incubated at 4˚C overnight with gentle rotation (15 RPM) on an end-over-end shaker. Beads were spun down at 2000 RCF and 40ul of the supernatant was kept as the “Bead Unbound” sample. The beads were then washed in 500µl of low salt RIPA buffer for a total of 4 times.

For western blotting, 4X Lithium Dodecyl Sulfate (LDS) Sample Buffer (NuPage) was supplemented with DTT (VWR) to a final concentration of 50 mM and beads were heated for 25 mins to reduce proteins at 65^{o}. For Mass Spectrometry, proteins were not eluted from the beads: after the final low salt RIPA buffer wash, 30ul of each bead sample (from Heavy/Medium/Light cell lysate immunoprecipitation reactions) were combined in a single Eppendorf and washed 5 times with 50mM ammonium bicarbonate (Thermo) solution.

2.7.3 Streptavidin Pulldown

Streptavidin pulldown of TurboID biotinylated proteins has been described previously (Branon et al., 2018). 2 mg of input protein per sample (derived from 48hr doxycycline-induced TurboID Hela Cell lines) was pre-cleared with 20µl of Sepharose beads (ThermoFisher) with gentle rotation (15RPM) in an end-over-end shaker for 30 min at 4^{o}C. Beads were then pelleted via gentle centrifugation (2000 RCF for 1 minute) and 40ul of this supernatant was saved as the “Precleared” samples. Magnetic Streptavidin-Agarose Beads (Pierce) were washed 3 times in low salt RIPA buffer prior to 30ul being added to the remaining precleared cell lysates. Once added, beads and lysates were mixed end-over-end

at 4°C overnight. Samples were then pelleted via magnetic rack, and 40µl was saved as an "Unbound" sample.

For western blotting, beads were washed 4 times with low salt RIPA buffer, and proteins were eluted using 30ul of elution buffer (30mM Biotin in 2% SDS) and 20ul of 4xLDS Buffer (NuPage) with 50mM DTT (VWR). Beads were heated for 25 min to reduce proteins at 65°C. The bound and unbound samples were checked via western blot using AF800-Streptavidin (JIR) to determine the depletion efficiency of the biotinylated proteins from lysates.

For mass spectrometry, proteins were not eluted from the beads. Instead, after the final RIPA buffer wash, 30ul of each bead sample (the Heavy/Medium/Light sample) were combined in a single Eppendorf and washed 5 times with 50mM ammonium bicarbonate (Thermo) solution.

2.8 Mass Spectrometry Analysis

Trypsin (Sigma) digestion of beads was performed overnight at 37°C. The final sample volume was reduced to 20µl by concentrating the samples in a Speedvac (Thermo Scientific).

Peptides were subjected to LC-MS/MS using an Ultimate 3000 RSLC (Thermo Fisher Scientific) coupled to an Orbitrap Fusion Lumos mass spectrometer with a FAIMS interface

(Thermo Fisher Scientific). Peptides were injected onto a reverse-phase trap (Pepmap100 C18 5µm 0.3 × 5 mm) for pre-concentration and desalted with Loading Buffer (100% water, 0.05% Trifluoroacetic Acid (TFA)) at 10µL/min for 3 minutes.

The peptide trap was then switched in line with the analytical column (Easy-spray Pepmap RSLC C18 2um, 50cm x75um ID). Peptides were eluted from the column using a linear solvent gradient using the following gradient: linear 2–40% of Buffer B (80% acetonitrile, 20% water and 0.1% formic acid) over 80 min, linear 40–60% of Buffer B for 15 min, sharp increase to 95% Buffer B within 0.1min, isocratic 95% of Buffer B for 10 min, sharp decrease to 2% Buffer B within 0.1min and isocratic 2% Buffer B for 20 min. The FAIMS interface was set to -45V and -65V at standard resolution. The mass spectrometer was operated in Data Dependant Acquisition (DDA) positive ion mode with a cycle time of 1.5 seconds. The Orbitrap was selected as the MS1 detector at a resolution of 120000 with a scan range of from M/Z 375 to 1500. Peptides with charge states 2 to 5 were selected for fragmentation in the orbitrap using HCD as collision energy.

2.9 Data Analysis and Bioinformatics

2.9.1 MaxQuant

Raw MS data files were searched against the UniProt Human Proteome database (using the Human Proteome Fasta (Swissprot Only) retrieved from Uniprot dated 23.06.23) using MaxQuant software v2.5.0.0 (http:/www.maxquant.org). The following criteria were selected in MaxQuant prior to running the programme: peptide tolerance = 10 ppm, trypsin

as the enzyme (2 missed cleavages allowed) and carboxyamidomethylation of cysteine as a fixed modification. Variable modifications consisted of oxidation of methionine and N-terminal acetylation. SILAC labels of Arg6 (R6), Arg10 (R10), Lys4 (K4) and Lys8 (K8) were selected. A minimum of 1 unique peptide, minimum ratio count of 2 and quantitation based on razor and unique peptides were selected. Peptide and Protein FDR chosen was 0.01.

2.9.2 Data Processing

Data was processed using Microsoft Excel to produce ratios, proportions and percentages. Bar charts were then generated using either Prism 6 (GraphPad, La Jolla) or Microsoft Excel from the processed data. Statistical analysis was also performed in Microsoft Excel (paired and unpaired t-testing) and Perseus software. Proteins displaying significant changes in interactivity (P Value <0.05 and Fold Change <0.5 and >1.5) were selected for upload to String-DB.

Volcano plots were created using VolcanoseR (https://huygens.science.uva.nl/VolcaNoseR) by uploading relevant protein interactor lists along with their corresponding -log10 P-Values and log2 Fold Change values. Fold change (FC) cutoffs were set to show proteins with only FCs above 0.5 and below -1 (the log2 equivalent of an FC <0.5 and >1.5) (Goedhart & Luijsterburg, 2020). The significance cutoff was set to 1.3 (the log10 equivalent of a 0.05 cutoff). The top 15 hits - ranked by the largest FC in either direction - were labelled with their gene name.

Heat maps were created using Heatmapper (https://heatmapper.ca) by uploading the significantly different protein interactors (P-Value <0.05) and their corresponding P-Values (-log10 transformed) and log2 Fold Change values (Babicki et al., 2016). Clustering Method was set to Average Linkage and Distance Measurement Method was set to Euclidean. Clustering was applied to Rows only.

2.9.3 STRING-DB & Gene Ontology

Gene names were entered into String (version 12.0) to elucidate existing interactions between protein sets from the literature, selecting medium to high confidence interactors (0.4–0.8) and filtering interactions so that only experimentally-derived interactions were retained in the output. String was also utilised to produce Gene Ontology (GO) data from the discrete protein lists provided. These were sorted by network strength and the experimental PPI prediction score was used as a measure of biological significance. GO Component, GO Molecular Function and GO Process were exported and refined into final GO lists in Excel.

Chapter 3: Data Analysis of Existing GFP-Trap Datasets

3.1 Introduction

GFP-Trap is a technique for immunoprecipitating GFP and any interacting proteins from whole cell lysates for various downstream analysis techniques including Western Blot and Mass Spectrometry (MS) analysis (Yilmazer et al., 2023). Plasmids containing a GFP molecule fused to both SMN and SMNΔ7 had previously been transfected into Hela cell lines in the Sleeman lab (Figure 3.1). The fused GFP-SMN protein is assumed to retain the functionality of native SMN, due to its correct localisation within the cell as well as its ability to interact and immunoprecipitate with classic SMN interacting proteins (Prescott et al., 2014). Analysis of both western blot and imaging data (not shown) of GFP cell lines was performed prior to MS, to ensure cells were expressing similar levels of the GFP-Protein constructs. Datasets resulting from Mass Spectrometry analysis of these Hela GFP-SMN and GFP-SMNΔ7 pulldown samples were then produced.

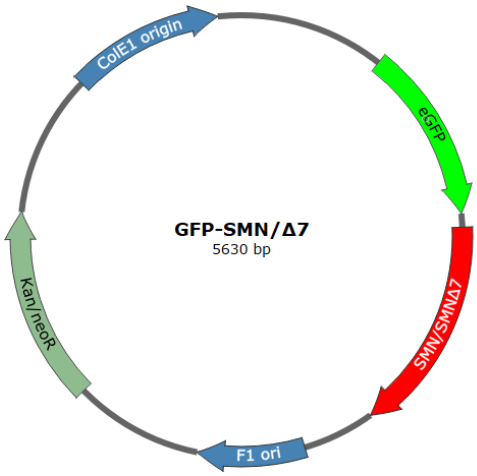


FIGURE 3.1: SIMPLIFIED PLASMID MAP OF THE EGFP-SMN/SMNΔ7 CONSTRUCT.

Expresses eGFP fused to either SMN or SMNΔ7. The eGFP-Control plasmid lacks the SMN/SMNΔ7 segment.

These data were not quantitative in nature due to the lack of metabolic labelling and the lack of biological replicates (N=1). This meant that statistical analysis was not possible to determine the veracity of the findings. However, the data did contain intensity values for all proteins and from these a potential fold change could be calculated.

Using Microsoft Excel, a fold change value was calculated between each condition (GFP Only Control, GFP-SMN and GFP-SMNΔ7) and this was then used to indicate directionality of changes between conditions i.e., a fold change of <1 was considered a decrease in interactivity whilst >1 denoted increased protein interactivity. Protein presence between each condition was also analyzed, allowing for the identification of unique interactions in each condition. While not conclusive, the numerical data generated still provides a useful - if broad - starting point for an examination of the differences between the interaction landscapes of SMN and SMNΔ7.

3.2 The Interactomic Landscape Differs Between GFP-SMN and GFP-SMNΔ7

The analysis identified 389 different proteins across all conditions, however this was reduced to 369 after removal of contaminants via cross-referencing the data with the MaxQuant contaminants fasta file. The GFP Control condition identified 34 interacting partners; these interactors were removed from GFP-SMN and GFP-SMNΔ7 protein shortlists as they represent GFP interacting proteins. GFP-SMN and GFP-SMNΔ7 associated with 248 and 301 proteins respectively (Figure 3.2). The substantial variation in the number of identified proteins suggests protein variant-specific interactions.

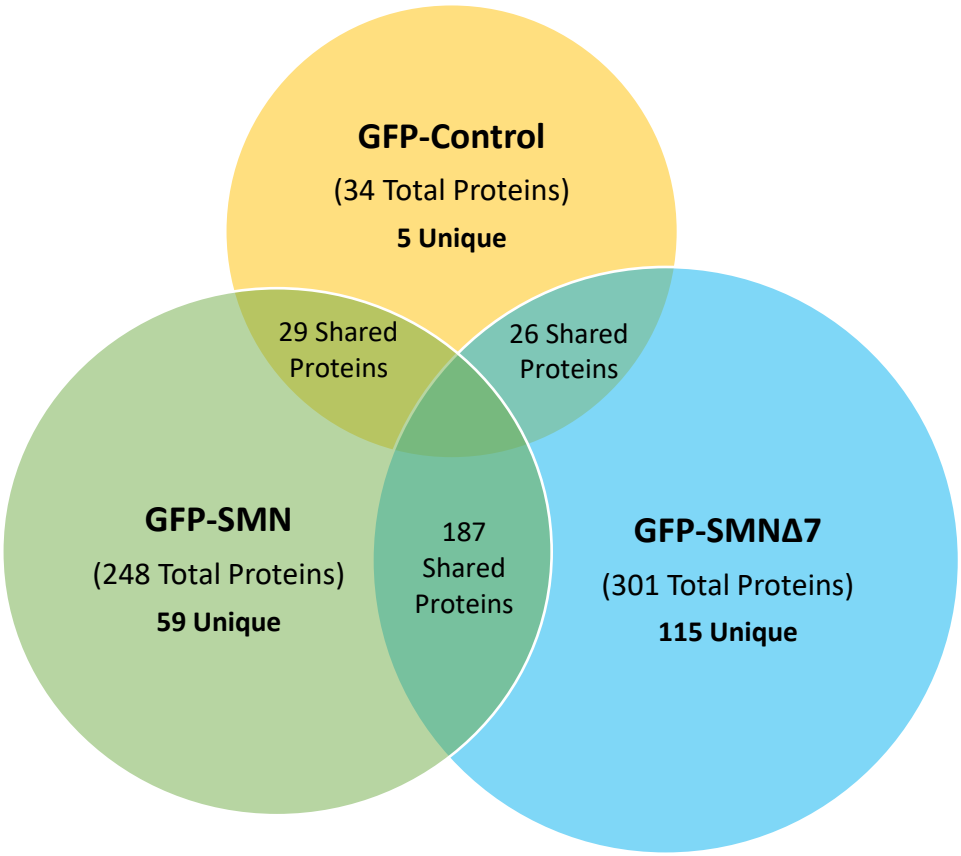


FIGURE 3.2: VENN DIAGRAM OVERVIEW OF PROTEIN COMPOSITION IN THE GFP-TRAP DATASETS.

Also noted are the numbers of protein in common between conditions, total protein number per condition and the number of unique proteins.

Several proteins were shared between conditions, however differential interaction analysis between GFP-SMNΔ7 and GFP-SMN revealed notable changes in protein interactivity. GFP-SMNΔ7 exhibited 23 proteins with decreased interactivity compared to GFP-SMN, suggesting a loss of interacting partners specific to the Δ7 isoform. Conversely, GFP-SMNΔ7 displayed 41 proteins with increased interactivity compared to GFP-SMN, suggesting interactions gained in the Δ7 isoform. 122 proteins displayed fold changes (FC) between 0.5 and 1.5, indicative of a consistent interaction pattern (Figure 3.3).

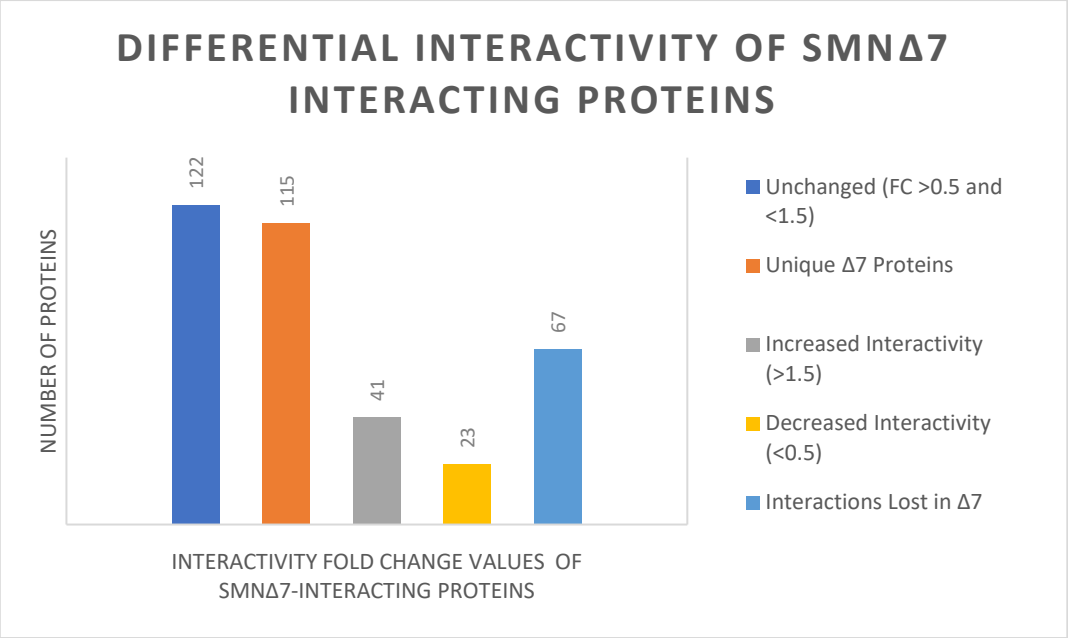


FIGURE 3.3: BAR CHART DEMONSTRATING THE DIFFERENCES IN PROTEIN ABUNDANCE OF GFP-SMNΔ7 INTERACTORS.

Fold-Change data derived from comparison to the GFP-SMN protein data. Illustrates the numbers of proteins differentially interacting in comparison to the GFP-SMN datasets. Also includes unique proteins and lost interactions (i.e. proteins identified in GFP-SMN pulldown data but missing in the GFP-SMNΔ7 dataset).

To elucidate functional enrichments between conditions, a comprehensive Gene Ontology (GO) analysis of the proteins expressed within the GFP-SMN and GFP-SMNΔ7 conditions was performed using String-DB. String-DB is a powerful bioinformatics tool that facilitates the exploration of protein-protein interactions and functional associations, integrating experimental and predicted protein interaction data (Szklarczyk et al., 2019). This allows for the construction of qualitative interaction networks among proteins identified during proteomic analysis.

All proteins associated with GFP-SMN and GFP-SMNΔ7 were uploaded to String-DB as gene lists, allowing for GO analysis of the overall interactome pulled down in each condition. GFP-SMN and GFP-SMNΔ7 showed statistically significant differences in the biological processes, molecular functions and cellular components highlighted during analysis (Table 3.1).

#Term ID	Term Description	Network Strength	False Discovery Rate
GFP-SMN			
GO Component			
GO:0005754	Mitochondrial proton-transporting ATP synthase	1.91	0.0158
GO:0035632	Mitochondrial prohibitin complex	1.91	0.0158
GO:0097452	GAIT complex	1.79	0.0013
GO:0005832	Chaperonin-containing T-complex	1.76	2.41e-08
GO:0097149	Centralspindlin complex	1.73	0.0234
GO Process			
GO:0019240	Citrulline biosynthetic process	1.79	0.004
GO:1904851	Positive regulation of establishment of protein localization to telomere	1.76	1.27e-07
GO:1904871	Positive regulation of protein localization to Cajal body	1.71	1.98e-07
GO:0000915	Actomyosin contractile ring assembly	1.69	0.0062
GO:0090666	scaRNA localization to Cajal body	1.69	0.0062
GO Function			
GO:0002135	CTP binding	1.91	0.0415
GO:0004088	Carbamoyl-phosphate synthase (glutamine-hydrolyzing) activity	1.91	0.0415
GO:0030623	U5 snRNA binding	1.91	0.0415
GO:0032564	dATP binding	1.91	0.0415
GO:0051920	Peroxiredoxin activity	1.71	2.85e-05
GFP-SMNΔ7			
GO Component			
GO:0005754	Mitochondrial proton-transporting ATP synthase	1.83	0.0196
GO:0031595	Nuclear proteasome complex	1.83	0.0196
GO:0035632	Mitochondrial prohibitin complex	1.83	0.0196
GO:0005832	Chaperonin-containing T-complex	1.73	2.91e-09
GO:0097452	GAIT complex	1.7	0.0023
GO Process			
GO:0046084	Adenine biosynthetic process	1.83	0.0439
GO:0075506	Entry of viral genome into host nucleus through nuclear pore complex via importin	1.83	0.0439
GO:1905215	Negative regulation of RNA binding	1.83	0.0439
GO:1904851	Positive regulation of establishment of protein localization to telomere	1.78	4.67e-10
GO:1904871	Positive regulation of protein localization to Cajal body	1.74	7.75e-10

GO Function			
GO:0036402	Proteasome-activating activity	1.75	2.48e-05
GO:0008379	Thioredoxin peroxidase activity	1.59	0.001
GO:0048027	mRNA 5-UTR binding	1.49	2.17e-11
GO:1990948	Ubiquitin ligase inhibitor activity	1.48	0.002
GO:0001094	TFIID-class transcription factor complex binding	1.43	0.0027

TABLE 3.1: GO ANALYSIS RESULTS OF THE FULL GFP-SMN AND GFP-SMNΔ7 PROTEOMES.

Sorted by the network strength in 3 sub-categories: Cellular Component, Biological Process and Molecular Function.

Sorted by Network Strength, this measure describes how large the enrichment effect is in a particular dataset. It constitutes the ratio between the number of proteins in a network that are annotated with a term and the number of proteins that would be expected to be annotated with this term in a random protein network of the same size. A value of 1 to 1.5 is strong network confidence, whereas values closer to 2 are considered excellent indicators. When paired with low FDRs, this serves to ensure a level of stringency within the results of the analysis.

Several GO terms, including those related to mRNA splicing and protein localization were common to both datasets; this was expected given the previously discussed canonical roles that SMN plays in the splicing process (Chaytow et al., 2018). Similarly, SMN is known to interact intimately with proteins of the Cajal Bodies and this functional enrichment in the GO analysis results can be taken as positive confirmation that the GFP-Trap pulldown technique had worked efficiently (Renvoise, 2006).

The presence of novel GO terms related to infectious processes in both datasets, particularly the reoccurrence of the GAIT Complex entry, is intriguing; previous published data connecting SMN to this functional complex has not been found. Functional enrichment in terms related to mitochondrial complexes were also found in both datasets, which is relevant given the role that mitochondrial dysfunction is known to play in SMA (Zilio et al., 2022). However, as many of the proteins within both the GFP-SMN and GFP-SMNΔ7 pulldowns are shared between protein variants (See Figure 3.2), it is unsurprising that many of the functional enrichments between the two are similar.

To further probe the differences between the two protein variants, String-DB was used to build interaction networks from the same GFP-SMN and GFP-SMNΔ7 interactomes noted previously. This resulted in the production of large, interconnected protein networks which showed statistically significant Protein-Protein Interaction (PPI) scores. For example, the GFP-SMN network (PPI Enrichment P-Value < 1.0e-16) contained more than double the number of edges (i.e. known interactions) as would be expected for the same number of random proteins, suggesting high probability of functional connectedness within the dataset (Figure 3.4).

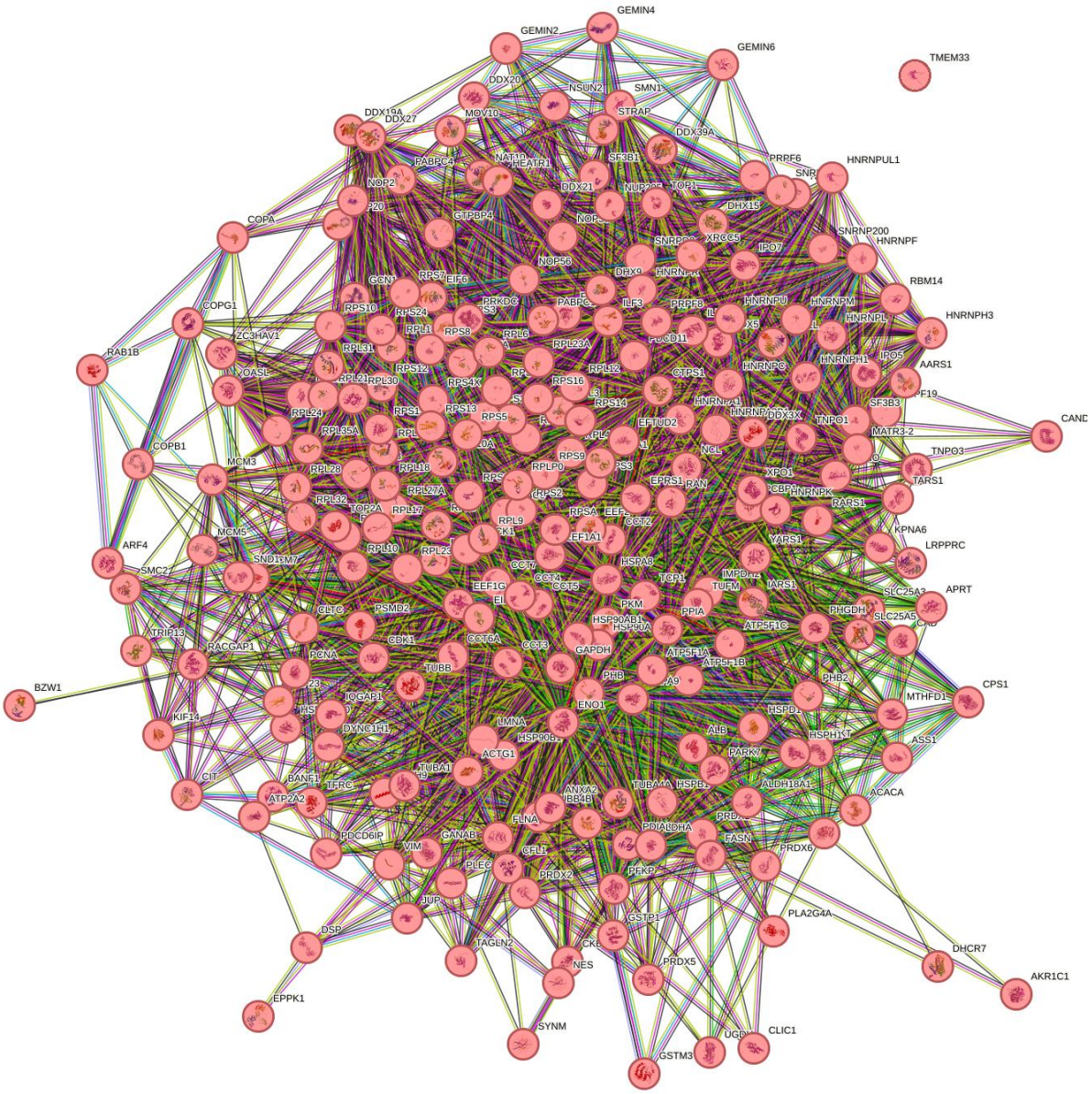


FIGURE 3.4: GFP-SMN PULLDOWN INTERACTION NETWORK.

Each node (circle) represents proteins identified in the GFP-SMN pulldown, whereas each line represents an existing identified interaction between the two connected proteins.

Given the high level of interconnectedness displayed within all datasets, K-Means Clustering – an algorithm which allows the proteins in a network to be segregated into distinct functional clusters – was performed to allow for up to 3 primary clusters per protein variant,

highlighting both the similarity and variability of the interactomes of these proteins (Szklarczyk et al., 2021) (Figure 3.5).

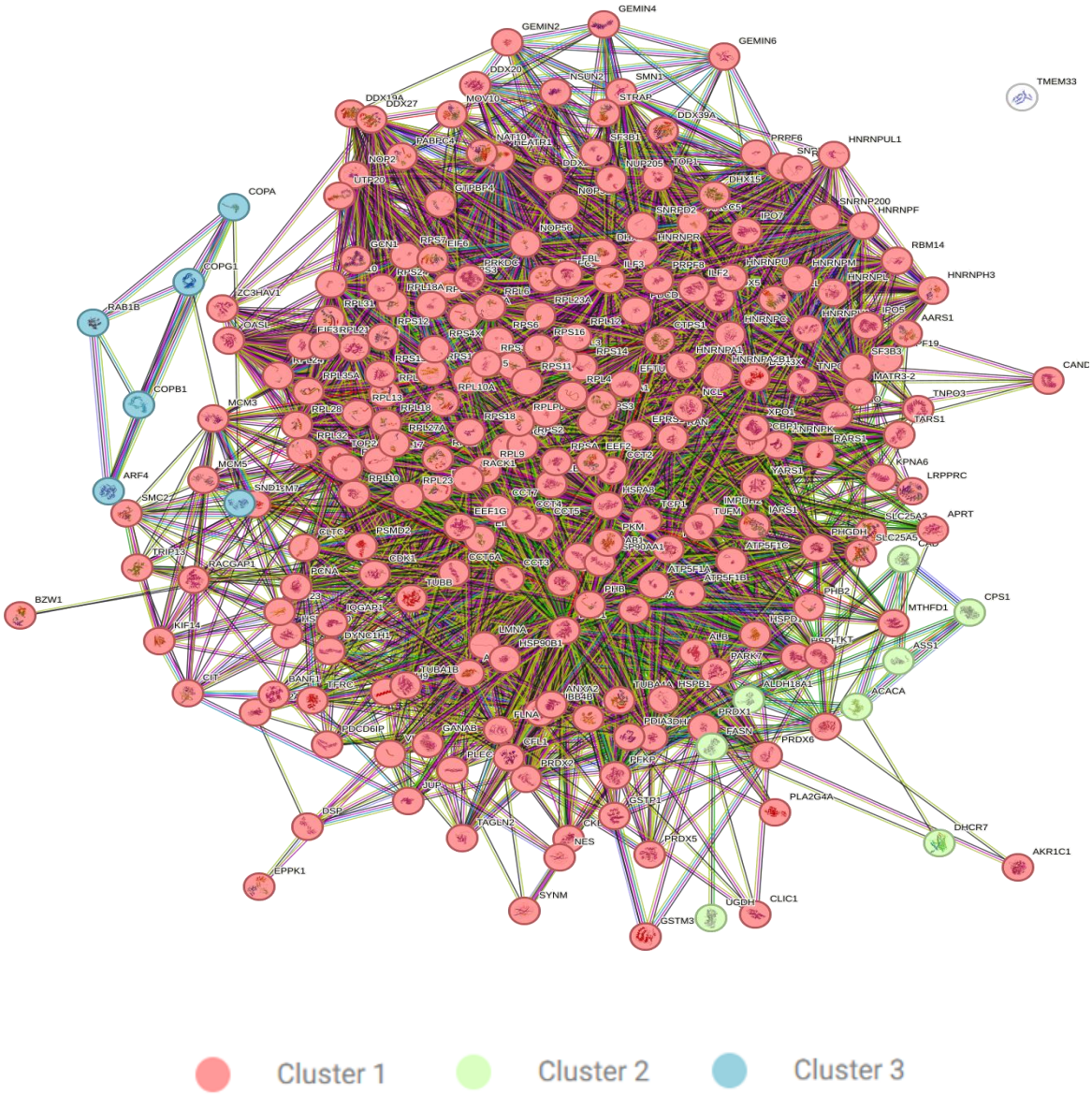


FIGURE 3.5: SMN PROTEIN PULLDOWN INTERACTION NETWORK WITH K-MEANS CLUSTERING APPLIED.

Each coloured node represents proteins belonging to a particular functional cluster whereas each line represents an existing identified interaction between the two connected proteins.

The primary clusters revealed by such processing in the GFP-SMN pulldown were related to splicing and protein localisation (Cluster 1 – 227 proteins). In particular, terms such as regulation of protein transport to Cajal Bodies, U5 snRNA binding and large ribosomal subunit localisation were highlighted. Cluster 2 focused on amino acid binding and ligase activity (8 Proteins) whilst Cluster 3 associated with vesicular transport and the COPI vesicle system (6 Proteins).

Similarly, the GFP-SMNΔ7 Pulldown's first cluster (260 proteins) was also primarily associated with protein localization, with highlighted terms including localisation to the telomere and - like the GFP-SMN data - localisation to the Cajal Bodies. Unlike the GFP-SMN dataset, however, it also highlighted an involvement with the negative regulation of RNA binding and various nuclear pore-related functions. Cluster 2 (23 Proteins) had a similar focus on nuclear pore dynamics, alongside – interestingly - viral entry into the genome via nuclear pore entry. Cluster 3 (with 8 proteins) showed an overrepresentation of proteins involved in fatty acid metabolism.

From the gene ontology analysis of the complete GFP-SMN and GFP-SMNΔ7 interactomes, it is clear that – whilst many similarities in the interaction partners of the two proteins exist at the functional level – there are discrete differences between the pair.

3.3 Both GFP-SMN and GFP-SMNΔ7 Interact with Distinct Sets of Proteins

The identification of unique proteins across experimental conditions holds significant importance in elucidating the distinct molecular signatures associated with each protein. These unique protein interactors not only serve as potential biomarkers but also offer insights into the specific biological processes, pathways, and cellular functions that interact differentially with each protein variant.

Upon examination of the data, it was clear that specific interactions were unique to both the GFP-SMN and GFP-SMNΔ7 variants. While the 115 unique interactors of SMNΔ7 may be considered interactions *gained* by the loss of exon 7, so might the 59 unique interactions of SMN be considered interacting proteins *lost* to the SMNΔ7 isoform because of its different amino acid composition (Table 3.2).

Protein Variant	Number of Unique Proteins
SMN	59
SMNΔ7	115

TABLE 3.2: THE NUMBER OF UNIQUE PROTEINS IDENTIFIED IN EITHER GFP-SMN OR GFP-SMNΔ7 PULLDOWN DATA.

Derived from proteins with intensity values only in either the GFP-SMN or GFP-SMNΔ7 experimental condition.

Gene Ontology analysis of these unique proteins revealed differential biological pathway enrichment, indicating that both GFP-SMN and GFP-SMNΔ7 occupy overlapping, yet distinct, interactomic domains (Table 3.3).

#Term ID	Term Description	Network Strength	False Discovery Rate
GFP-SMN Unique Proteins			
GO Component			
GO:0005642	Annulate lamellae	2.1	0.0088
GO:0017101	aminoacyl-tRNA synthetase multienzyme complex	1.94	0.00055
GO:0005916	Fascia adherens	1.85	0.0198
GO:0005682	U5 snRNP	1.83	5.25e-05
GO:0030126	COPI vesicle coat	1.69	0.0356
GO Process			
GO:0006450	Regulation of translational fidelity	1.6	0.0246
GO:1903241	U2-type prespliceosome assembly	1.6	0.0246
GO:0000387	Spliceosomal snRNP assembly	1.51	0.0037
GO:0098901	Regulation of cardiac muscle cell action potential	1.5	0.0404
GO:0006418	tRNA aminoacylation for protein translation	1.47	0.005
GO Function			
GO:0030623	U5 snRNA binding	2.5	0.0092
GO:0086083	Cell adhesive protein binding involved in bundle of His cell-Purkinje myocyte communication	2.1	0.0236
GO:0030620	U2 snRNA binding	2.02	0.03
GO:0003916	DNA topoisomerase activity	1.96	0.0377
GO:0140142	Nucleocytoplasmic carrier activity	1.52	0.0169
GFP-SMNΔ7 Unique Proteins			
GO Component			
GO:0031595	Nuclear proteasome complex	2.24	0.0083
GO:0031597	Cytosolic proteasome complex	2.02	0.00063
GO:0008540	Proteasome regulatory particle, base subcomplex	1.98	9.08e-07
GO:0005838	Proteasome regulatory particle	1.84	4.51e-10
GO:0005663	DNA replication factor C complex	1.84	0.0257
GO Process			
GO:0031595	Nuclear proteasome complex	2.24	0.0083
GO:0031597	Cytosolic proteasome complex	2.02	0.00063
GO:0008540	Proteasome regulatory particle, base subcomplex	1.98	9.08E-07
GO:0005838	Proteasome regulatory particle	1.84	4.51E-10
GO:0005663	DNA replication factor C complex	1.84	0.0257
GO Function			
GO:0036402	Proteasome-activating activity	2.16	6.73e-07
GO:0016742	Hydroxymethyl-, formyl- and related transferase activity	1.71	0.0072
GO:0048027	mRNA 5-UTR binding	1.52	0.00017

GO:0019843	rRNA binding	1.19	0.00072
GO:0051082	Unfolded protein binding	1.02	0.0012

TABLE 3.3: GO ANALYSIS RESULTS OF THE GFP-SMN AND GFP-SMNΔ7 UNIQUE PROTEINS.

Shortlist derived from proteins with intensity values only in either the GFP-SMN or GFP-SMNΔ7 experimental conditions. Results are sorted by the network strength in 3 sub-categories: Biological Process, Molecular Function and Cellular Component/Compartment. FDR<0.05.

The GFP-SMN Unique Interactor GO analysis reinforces the idea that many of the proteins lost to GFP-SMNΔ7 are splicing related. The inclusion of processes such as snRNP assembly and translational fidelity highlight many of the basic functions of SMN that are known to be lost in disease and thus involved in the etiology of SMA (Ji et al., 2021). On the other hand, enriched processes such as tRNA aminoacylation are not classically associated with SMA.

The Unique Interactor analysis of GFP-SMNΔ7 highlights several processes relating to the proteosome and protein degradation. Considering the short half-life of the Δ7 isoform, this is not surprising (Burnett et al., 2009). However, novel functions such as unfolded protein binding and mRNA 5'-UTR binding are not generally associated with SMNΔ7.

3.4 The tRNA Synthetases and Multisynthetase Complex (MSC)

A functional role for SMN in cellular processes such as translation and transcription has been well established (Faravelli et al., 2023). Given the results of the above GFP-Trap data analysis - within which several aspects of protein-specific processes were significantly represented in both GFP-SMN and GFP-SMNΔ7 datasets - particular attention was paid to the protein family comprised of the tRNA Synthetases. These proteins function as amino

acid-specific adaptors and are essential in the translation of mRNA to protein (for more detail, see Chapter 5). Several of these proteins form part of a large, multiprotein hub known as the Multisynthetase Complex (MSC) and have been previously linked to several neuromuscular or neurodegenerative conditions (Ognjenović & Simonović, 2018). Their interaction with SMN and a role in SMA pathology, however, has not been directly examined.

Examining the dataset for these proteins showed a differential regulation pattern or different patterns of interaction between the GFP-SMN and GFP-SMNΔ7 datasets (Table 3.4). Some, like the Alanine tRNA Synthetase were found only in the GFP-SMN pulldown, whilst others such as the Serine tRNA Synthetase were only present in the GFP-SMNΔ7 condition. Both Tyrosine and Threonine tRNA Synthetases were found to be present in each condition but upregulated in the SMNΔ7 pulldown dataset (Table 3.4). Given this analysis - along with the enrichment of terms related to aminoacylation in the GO analysis - it becomes clear that investigating the relevance of this family of proteins and their interactions with the SMN protein are worth pursuing in the future.

tRNA Synthetase/Ligase	**Gene Name**	**Accession**	**SMN**	**SMNΔ7**
Tyrosine--tRNA Synthetase	YARS	SYYC_HUMAN	Decreased Interaction	Increased Interaction
Phenylalanine--tRNA Synthetase	FARSA	SYFA_HUMAN	Not Present	Present in SMNΔ7 Only
Serine--tRNA Synthetase	SARS	SYSC_HUMAN	Not Present	Present in SMNΔ7 Only
tRNA-splicing Synthetase RtcB	RTCB	RTCB_HUMAN	Not Present	Present in SMNΔ7 Only
Alanine--tRNA Synthetase	AARS	SYAC_HUMAN	Present in SMN Only	Not Present
Arginine--tRNA Synthetase	RARS	SYRC_HUMAN	Present in SMN Only	Not Present

Bifunctional glutamate/proline--tRNA Synthetase	EPRS	SYEP_HUMAN	Present in SMN Only	Not Present
Isoleucine--tRNA Synthetase	IARS	SYIC_HUMAN	Present in SMN Only	Not Present
Threonine--tRNA Synthetase	TARS	SYTC_HUMAN	Decreased Interaction	Increased Interaction

TABLE 3.4: IDENTIFIED TRNA SYNTHETASES/LIGASES AND THEIR INTERACTION PATTERNS IN THE GFP-TRAP SMN AND SMNΔ7 PULLDOWNS.

Noted are the Protein Name, Gene Name, Accession and whether the specific tRNA Synthetase was present and/or differentially regulated in either the GFP-TRAP SMN and SMNΔ7 pulldown datasets.

3.5 The GAIT Complex

A role for the SMN protein in inflammation has been previously discussed (See Section 1.11). In light of this, the appearance of enriched cellular processes involved in both virus replication and viral resistance in the Gene Ontology data are perhaps not entirely inconsistent with previously described SMN functions. However, the inclusion of the Gamma-Interferon Inhibitor of Translation (GAIT Complex) in these results is novel, and upon further research, the literature has not previously linked the SMN protein or SMA to this cellular complex.

The GAIT Complex is comprised of four proteins and plays a role in inhibiting the translation of specific mRNAs in response to gamma interferon (IFN-γ) signalling, a cytokine produced in response to viral infection. It works to fine-tune the cellular response to infection by regulating the translation of specific genes (Mukhopadhyay et al., 2009) (for more detail, see Chapter 4).

Interestingly, the previously discussed tRNA Synthetases have recently also been implicated in the process of viral infection (Feng & Zhang, 2022). As noted, analysis of tRNA Synthetases within the dataset revealed differential expression patterns of these proteins between conditions. The EPRS tRNA Synthetase – which is also a constituent protein of the GAIT complex - was identified only in the GFP-SMN pulldown. This suggested a loss of interaction unique to the SMNΔ7 isoform. Utilising the String Interaction Network and expanding the network to uncover node proteins connecting SMN to EPRS revealed SYNCRIP (also known as HNRNPQ) as a linking interactor (Figure 3.6) Interestingly, this protein was identified only in the GFP-SMNΔ7 pulldown.

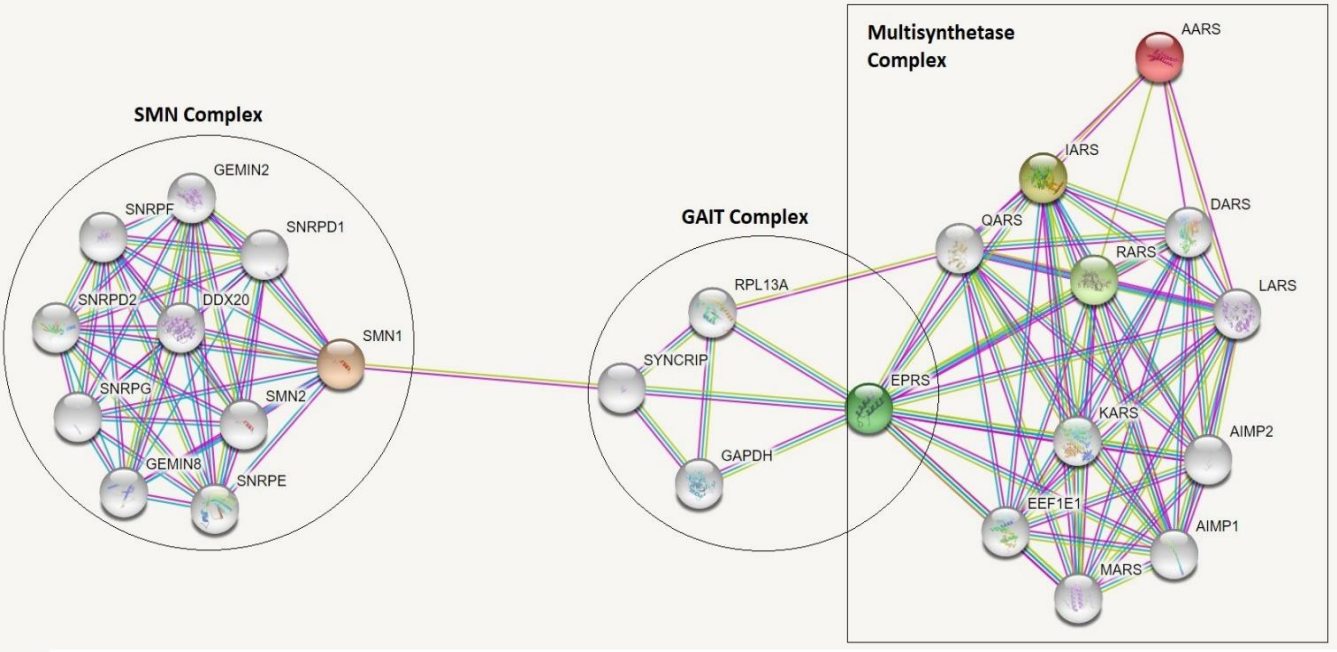


FIGURE 3.6: INTERACTOMIC LANDSCAPE OF SMN, THE GAIT COMPLEX AND THE MSC.

The SMN Complex interacts with the GAIT complex via a known interaction with SYNCRIP, which in turn interacts with the Multisynthetase complex via interaction with EPRS. Highlighted nodes include SMN and notable amino acid synthetases.

Both EPRS and SYNCRIP are components of the GAIT complex, along with RPL13A (found to be slightly reduced in the Δ7 condition) and GAPDH (increased by an almost double 1.8-fold change in Δ7 compared to native SMN). Given that each member of the GAIT complex differentially interacted in some capacity between the two overexpressed conditions, and that the GAIT complex was identified in the GO analysis, this would suggest an interesting focus for upcoming investigation.

3.6 Conclusion

Although primarily qualitative, this GFP-Trap dataset analysis has allowed for the formation of several avenues of continued research towards delineating new pathways and cellular processes relevant to SMA. From further exploration of the tRNA Synthetases and the GAIT Complex, it may be possible to find novel – and disease-relevant – biological interactors of the SMN protein. This may help to unravel some of the as-yet unknown mechanisms of SMA aetiology.

Chapter 4: Investigating the GAIT Complex and its Relevance to SMA

Disclaimer: This work was undertaken with the help of former student Joanna Troka, as part of her Undergraduate Research Project.

4.1 Overview of the Interferon-Stimulated GAIT Complex

Interferon signalling pathways are a critical component of the human immune response. After viral infection, one of several defensive strategies employed by the immune system is the release of pleiotropic cytokines known as Interferons (IFN). Interferon works by interacting with receptors on the cell surface, initiating intracellular cascades of signalling proteins to result in eventual induction of genes which improve the cell's ability to resist viral entry (De Andrea et al., 2002).

There are two main sub-types of IFN: Type 1 IFNs signal via the IFNAR receptor, and Type 2 IFNs which signal via the IFN-γ receptor. While type 1 IFNs contain over 20 different variants of IFN cytokines, type 2 IFNs are comprised solely of IFN-γ (Interferon-Gamma) (Platanias, 2005).

The interferon-γ-activated inhibitor of translation (GAIT) system is a cellular mechanism whereby translational control of several functionally related genes occurs in response to IFN-γ signalling. In myeloid cells, the presence of IFN-γ induces the formation of the

multiprotein GAIT complex (Mukhopadhyay et al., 2009).This heterotetrametric protein complex binds to GAIT responsive elements in 3'-untranslated regions (UTRs) of several mRNAs related to inflammation - such as ceruloplasmin and Vascular Endothelial Growth Factor A (VEGF-A) - thus repressing their translation (Ray et al., 2009). In humans, the GAIT complex consists of 4 proteins, all of which have other necessary functions within the cell beyond being GAIT complex components (Figure 4.1).

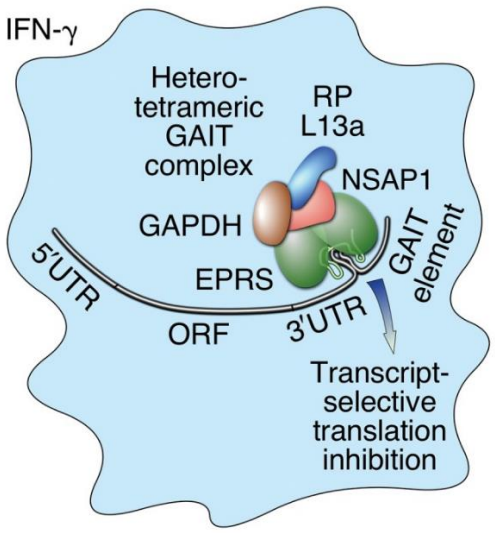


FIGURE 4.1: ILLUSTRATED OVERVIEW OF THE GAIT COMPLEX AND ASSOCIATED INHIBITION OF MRNA TRANSLATION. When the GAIT complex binds to the 3'UTR GAIT element of select inflammation-related mRNAs, translation is reduced. NSAP1 is an alternative name for the protein referred to as SYNCRIP below. Figure adapted from (Mukhopadhyay et al., 2009)

The four proteins composing the GAIT complex are the Bifunctional glutamate/proline-tRNA Synthetase (EPRS), synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and ribosomal protein L13a (L13a). EPRS is a tRNA synthetase that attaches the amino acids glutamate and proline to their

respective transfer RNAs (tRNAs), ensuring accurate translation during protein synthesis (Yao & Fox, 2020). SYNCRIP is a ribonucleoprotein known to be involved primarily in RNA metabolism, with roles in processes such as mRNA trafficking, mRNA stability and translation (Kim et al., 2013). GAPDH is a key enzyme in the glycolysis cascade, converting cellular glucose into energy. In addition to this, GAPDH has also been implicated in processes such as regulation of gene expression and apoptosis (Sirover, 1997). RPL13a is an essential protein component of the ribosome, thus integral to protein synthesis. It forms part of the large ribosomal subunit and contributes to the formation of the peptidyl transferase centre, where peptide bond formation occurs (Chaudhuri et al., 2007).

While each of these proteins have their canonical roles within the cell, upon exposure to IFN-γ they undergo changes in protein interactivity. A network of IFN-γ-stimulated kinases regulates recruitment and assembly of these GAIT complex constituents. Activation of cyclin-dependent kinase 5 (Cdk5), mammalian target of rapamycin complex 1 (mTORC1), and S6K1 kinases induce EPRS release from the multisynthetase complex (MSC) to join SYNCRIP in formation of a 'pre-GAIT' complex (Arif et al., 2012). Simultaneously, the DAPK-ZIPK kinase axis phosphorylates RPL13a, inducing its release from the 60S ribosomal subunit and subsequent binding to GAPDH. Both subcomplexes join to form the functional GAIT complex. Each constituent has a specific role in the GAIT system: EPRS binds the GAIT element in target mRNAs, SYNCRIP negatively regulates mRNA binding, RPL13a binds to eIF4G thus blocking ribosome recruitment and GAPDH shields L13a from degradation via the proteosome (Mazumder et al., 2003). All of these processes occur in response to IFN-γ signalling.

4.2 SHSY5Y Neuroblastoma Cells Respond to Interferon-Gamma (IFN-γ)

SHSY5Y neuroblastoma cells were chosen as a model system due to their relevance to SMA; as a neuronal cell line, these cells represent a more disease-relevant system than classic Hela cells. Although previous research suggests that SHSY5Y cells demonstrate an IFN-γ response *in vitro*, it was necessary to establish that the in-house SHSY5Y cells undergoing IFN-γ induction also respond to the cytokine as predicted (Shelke et al., 2017). Using western blot, STAT1 was selected as a positive marker of Interferon response due to its use in various studies of IFN-γ signalling (Hu & Ivashkiv, 2009). A549 lung carcinoma epithelial cells - which are known to respond to IFN-γ - were chosen as a positive control. Both SHSY5Y and A549 cells were dosed with 2000u/ml of IFN-γ for 24 hours as described in previous studies of IFN-γ signalling; these cells were harvested, and lysates were subsequently run on a western blot and probed with an anti-STAT1 antibody (Herbst et al., 2011).

The interferon-induced conditions demonstrated clear banding at 91 kDa as would be expected in STAT1 induction (Figure 4.2.1A). Although these bands were clear, this result was statistically insignificant when quantifying the bands (normalized to Tubulin) in the SHSY5Y cell line (P=0.055, N=3) (Figure 4.2.2B). Because of this, another modality to test IFN-γ induction was chosen to support these results.

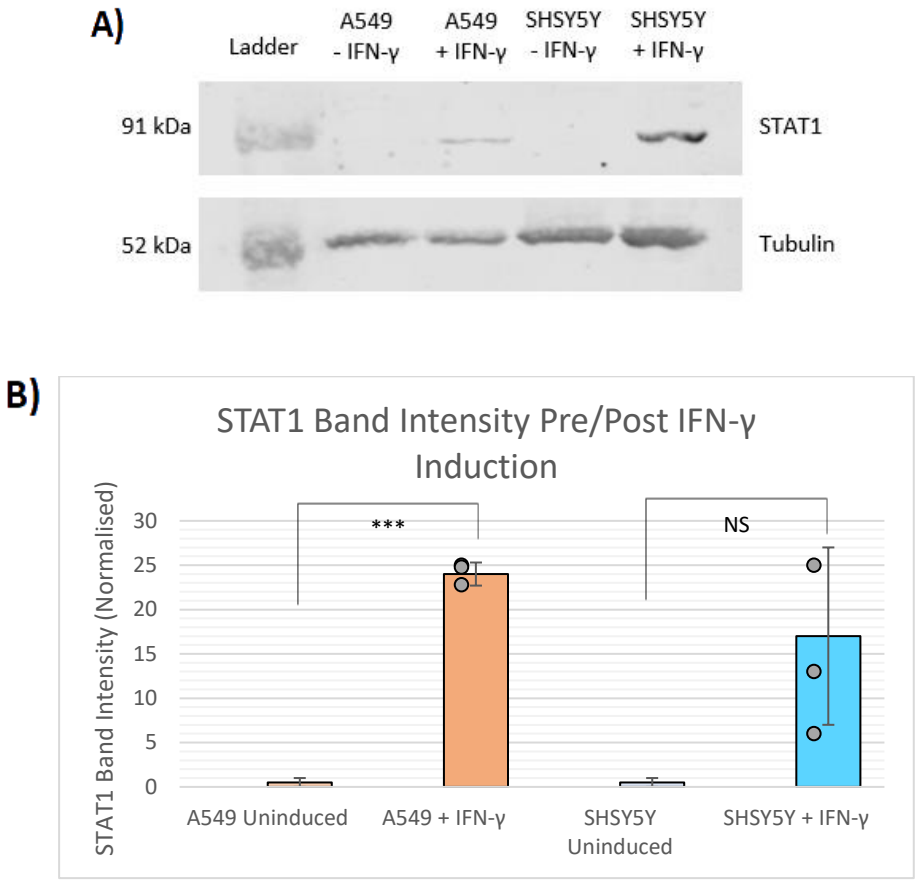


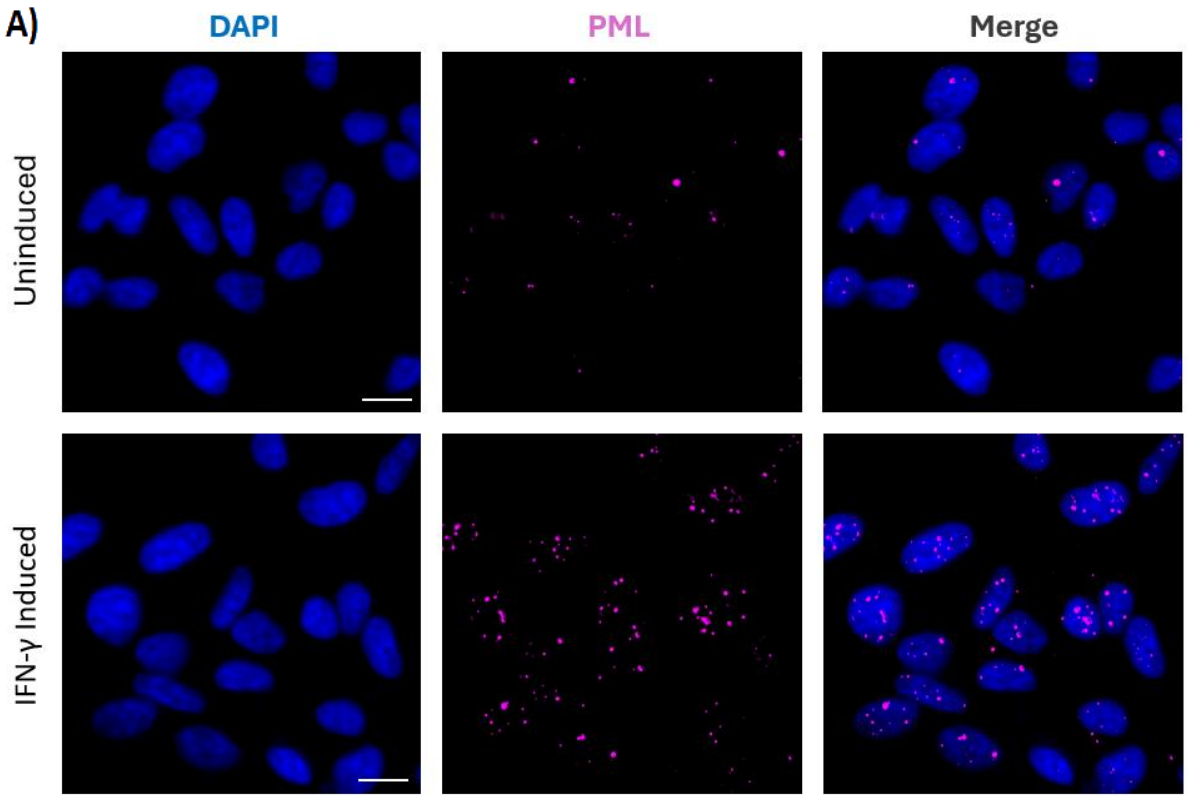
FIGURE 4.2.1: IFN-GAMMA INDUCES THE EXPRESSION OF STAT1 IN A549 AND SHSY5Y CELLS.

A) Representative Western Blot stained against STAT1 protein and Tubulin in lysates derived from uninduced A549 and SHSY5Y cells alongside SHSY5Y and A549 cells induced with 2000u/ml human recombinant IFN-γ for 24 hrs. B) Bar graph displaying STAT1 band intensity (normalized to tubulin) in uninduced and induced A549 and SHSY5Y cells. A549 cells are represented by the orange bars whereas SHSY5Y cells are represented by the blue bars. (N=3, P=0.055, NS = non-significant and P=0.0003, **** = denotes level of statistical significance). Error bars represent standard deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

A common indicator of positive interferon response is the increased formation of distinct proteinaceous bodies composed of promyelocytic leukaemia protein (PML) within the nuclei of IFN-γ induced cells (Fabunmi et al., 2001). PML-specific antibody staining of coverslips

with both induced and uninduced SHSY5Y cells allowed for examination of the associated changes in the numbers of PML bodies formed (Figure 4.2.2A).

Manual counting of PML nuclear bodies in 100 representative cells (obtained from 3 separate coverslips) revealed a significant increase in the number of nuclear bodies formed after IFN-γ induction, in comparison to the uninduced SHSY5Y cells (1 PML body per uninduced cell vs 5 PML bodies per cell in the induced line; $P<0.0001$) (Figure 4.2.2B)



B)

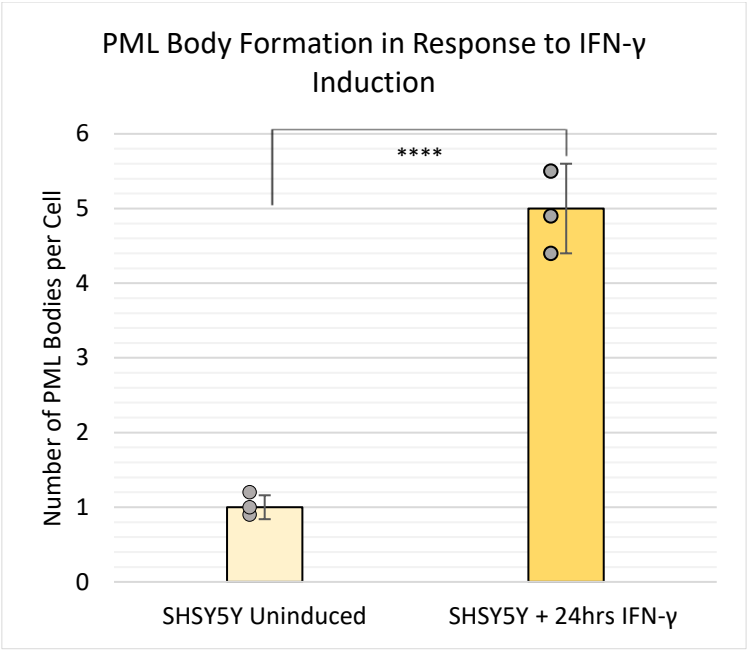


FIGURE 4.2.2: THE NUMBER OF PML BODIES IN THE NUCLEUS INCREASES AFTER INDUCTION WITH IFN-GAMMA.

A) Representative fluorescent imaging of uninduced SHSY5Y cells (top row) and SHSY5Y cells induced with IFN-γ (2000u/ml human recombinant IFN-γ for 24 hrs) (bottom row) stained against PML (magenta). The nucleus is counterstained with DAPI (blue). Scale = 10 microns. Non-deconvolved images were obtained on a Deltavision microscope at 100x magnification and represent a single slice from the centre of a Z-stack. B) Bar graph displaying the number of PML Bodies manually counted in uninduced (light yellow bar) and induced (dark yellow bar) SHSY5Y cells (2000u/ml human recombinant IFN-γ for 24 hrs) (N=100 cells from 3 separate coverslips, P=0.0001, **** = denotes level of statistical significance) Error bars represent standard error of the mean (SEM). Individual data points are represented by grey circles located to the left of error bars.

From these results, it can be surmised that SHSY5Y cells do respond as predicted to IFN-γ induction.

4.3 Selected GAIT Complex Components and SMN Do Not Show Marked Localisation Changes in Response to IFN-γ Induction.

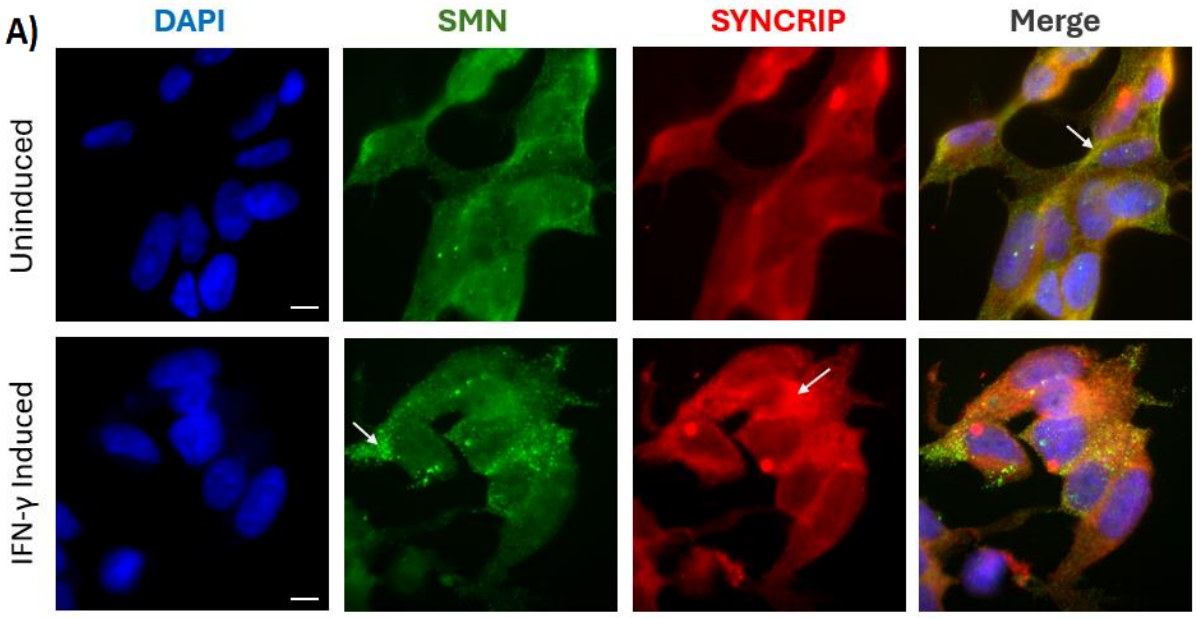
After confirmation that SHSH5Y cells demonstrate the hallmarks indicative of an inflammatory response to IFN-γ, two component proteins from the GAIT complex were chosen for further investigation: SYNCRIP and GAPDH.

SYNCRIP, also known as hnRNP Q, is an RNA-binding protein that has been previously linked to SMA and has known interactions with the SMN protein. Increasing levels of SYNCRIP in mice was found to promote Exon 7 inclusion in SMN2 pre-mRNA, suggesting a potential role as a splicing modulator of SMN transcripts (Chen et al., 2008). Rescue of SMA motor neurons was accomplished by overexpression of SYNCRIP in both *c. elegans* and mouse models of SMA, where it was found to bind to SMN and increase levels of full length SMN protein (Rizzo et al., 2019).

GAPDH is intimately involved in the metabolism of glucose in the cell. Human and mouse studies have demonstrated aberrant glucose metabolism in SMA patients, resulting in conditions such as hyperglycemia. Similarly, splenic defects have been observed in both mouse models of SMA and SMA patients (Deguise et al., 2019). In humans, SMA Type I patients have been found to have increased levels of GAPDH transcripts compared to SMA Type II and Type III patients (Crawford et al., 2012).

These proteins were selected because of their potential relevance to processes known to be disrupted in SMA. Neither GAPDH nor SYNCRIP of these proteins had been studied before with respect to their roles in the GAIT complex from an SMA-centric viewpoint. Furthermore, both proteins had been identified in the GFP-Trap dataset.

In order to investigate whether induction with IFN-γ induces localisation shifts in these component proteins of the GAIT complex, SHSY5Y cells were seeded onto coverslips and induced with IFN-γ-or left as negative, uninduced controls. Immunocytochemistry was carried out using antibodies against GAPDH, SYNCRIP and SMN then imaged using the DeltaVision microscope at 100x magnification. Overt changes in protein localization – as well as SMN co-localisation – were then determined.



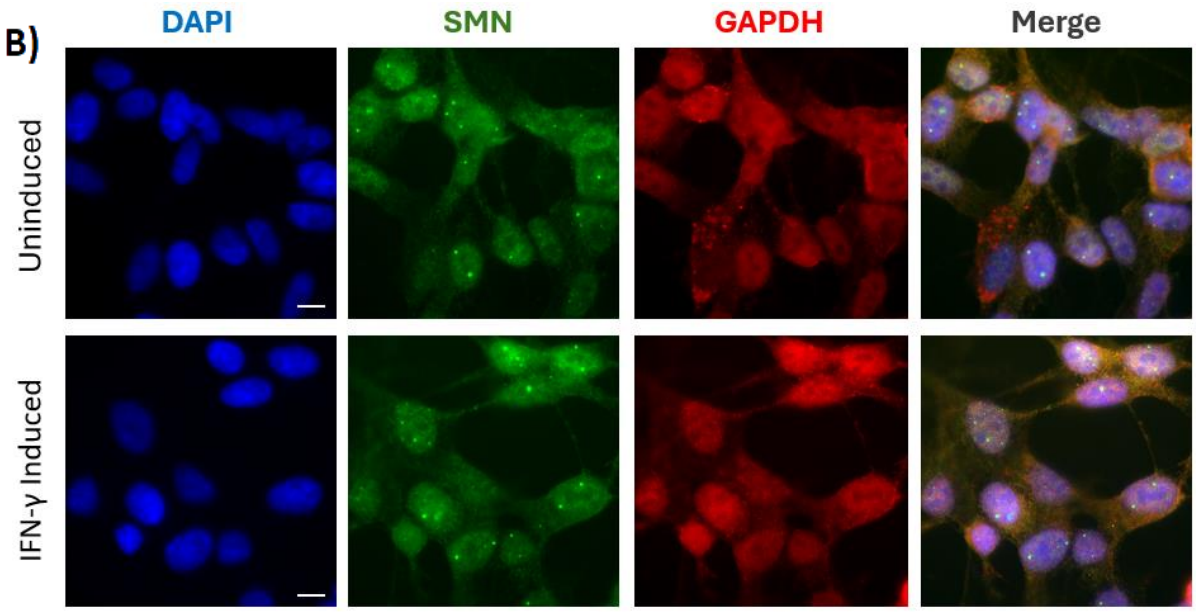


FIGURE 4.3.1: IMMUNOFLUORESCENT IMAGING OF GAIT COMPLEX COMPONENTS AND SMN DOES NOT REVEAL DISTINCT CHANGES BETWEEN IFN-Y INDUCED AND UNINDUCED SHSY5Y CELLS.

A) Representative fluorescent imaging of uninduced SHSY5Y cells (top row) and SHSY5Y cells induced with IFN-γ (2000u/ml human recombinant IFN-γ for 24 hrs) (bottom row) stained against SMN (green) and SYNCRIP (red). Nuclei are counterstained with DAPI (blue). The white arrow in the SMN Induced panel (bottom row, green) highlights extranuclear SMN speckles. The white arrow in the SYNCRIP Induced panel (bottom row, red) highlights an area of high concentration. The white arrow in the uninduced Merge panel (top row, merge) highlights yellow areas of colocalisation. B) Representative fluorescent imaging of uninduced SHSY5Y cells (top row) and SHSY5Y cells induced with IFN-γ (2000u/ml human recombinant IFN-γ for 24 hrs) (bottom row) stained against SMN (green) and GAPDH (red). Nuclei are counterstained with DAPI (blue). Nucleus counterstained with DAPI (blue). Scale = 10 microns. Non-deconvolved images were obtained on a Deltavision microscope at 100x magnification and represent a single slice from the centre of a Z-stack.

SYNCRIP naturally localizes within the cytosol (Chen et al., 2008). As can be seen in Figure 4.3.1A, this localisation does not drastically change upon IFN-γ induction. The SYNCRIP signal does appear to be stronger in the induced cells, where it appears more prone to form discrete puncta as opposed to the diffuse cytoplasmic localisation demonstrated in the

uninduced cell line (see white arrow in the SYNCRIP induced image). SYNCRIP also seems to localize to an extent with diffuse cytoplasmic SMN (see white arrow in SYNCRIP/SMN merge panel) however no localisation can be observed in the SMN-concentrated nuclear Gem structures. As discussed in Chapter 1, nuclear SMN often localizes to the regions of nuclear Gems, distinct membraneless organelles associated with the Cajal Bodies of the nucleus and illustrated as bright green specks in the nuclei in the above images (Clelland et al., 2009). There did appear to be puncta of cytoplasmic SMN in the induced condition, which can occur after cellular stress (see white arrow in the SMN induced panel) (Hua & Zhou, 2004).

GAPDH is a ubiquitous protein and is present across the cell, including the nucleus and cytoplasm (Sirover, 1997). Similarly to SYNCRIP, Figure 4.3.1B shows that GAPDH demonstrates some cytoplasmic overlap with SMN (as seen from the yellow in the merged image) however little to no GAPDH signal colocalizes to the nuclear Gems, implying a lack of nuclear SMN colocalisation. As with SYNCRIP, IFN-γ induction does not seem to necessitate a compartmental shift and the distribution of GAPDH is similar in both induced and non-induced SHSY5Y cells. However, the induced cells do tend to demonstrate a higher nuclear concentration of the protein compared to the uninduced. However, as there were no distinct trends in either SYNCRIP or GAPDH localisation changes across all three biological replicates, it is not possible to conclude that IFN-γ induces localisation changes in these proteins in SHSY5Y cells from these results.

Utilising Volocity Image Analysis software, the Deltavision images were deconvolved and colocalisation analysis was performed between the deconvolved image channels

representing either SMN and SYNCRIP, or SMN and GAPDH, pre and post IFN-γ induction. The colocalisation function measures the overlapping pixel intensities between channels and produces a numerical value – the Pearson Correlation Coefficient (PCC) – between 0 and 1. A PCC of 1 is indicative of total colocalisation, whereas a 0 value indicates no colocalisation between the examined proteins (Åberg & Robinson, 2021). The PCC in triplicate deconvolved images was calculated using automatic thresholding and an unpaired T-Test performed to identify significant changes between the induced and uninduced conditions (Figure 4.3.2) (Costes et al., 2004).

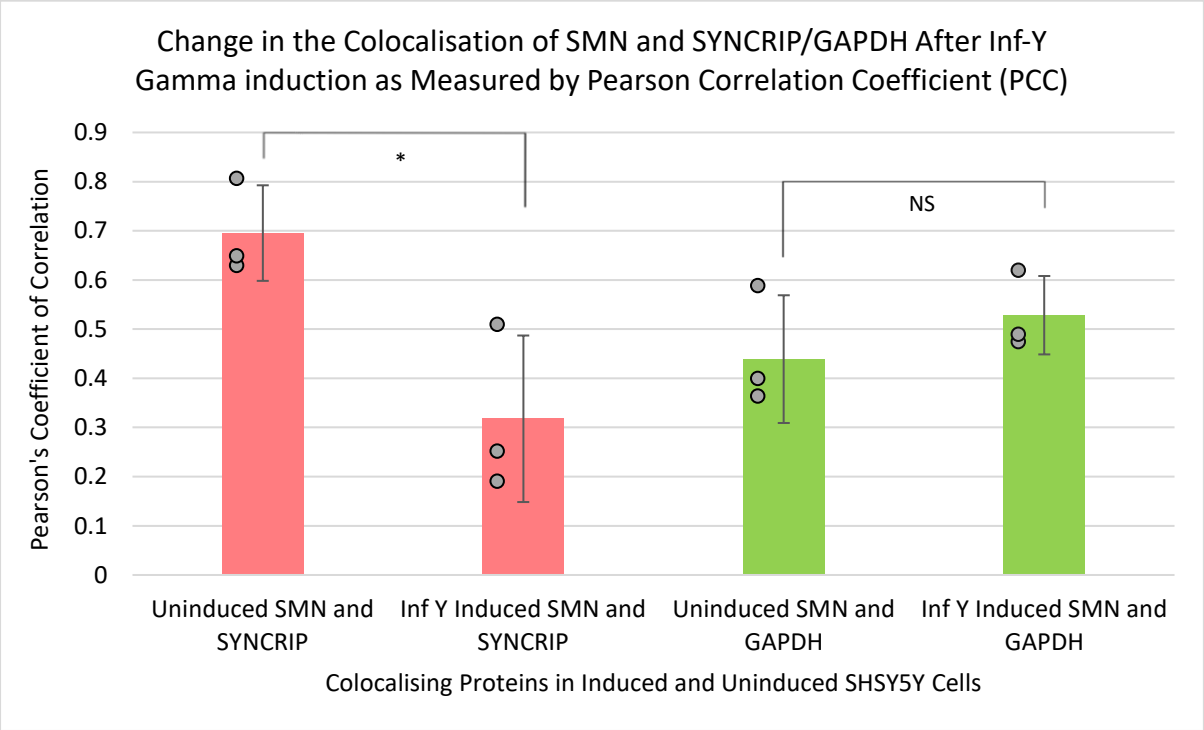


FIGURE 4.3.2: COLOCALISATION ANALYSIS OF SMN WITH GAPDH AND SYNCRIP IN UNINDUCED AND IFN-Y INDUCED SHSY5Y CELLS IDENTIFIES SOME DIFFERENCES IN COLOCALISATION.

Bar graph displaying the average Pearson Correlation Coefficient calculated between uninduced and Induced SHSY5Y Cells, specifically targeting localisation changes between SMN and SYNCRIP or SMN and GAPDH. SMN and SYNCRIP are represented by the pink bars (N=3, P=0.04, * = significant) while SMN and GAPDH are represented by the green bars (N=3, P=0.3, NS = non-significant). Input images were obtained on a Deltavision microscope at 100x magnification and deconvolved using Volocity

Image Analysis software. Error bars represent standard Deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

While a slight increase in GAPDH and SMN colocalisation was observed post IFN-γ induction, this result was not statistically significant (P=0.3). Of note, SYNCRIP's localisation with SMN was significantly decreased post IFN-Y induction (P=0.04).

Further utilisation of Volocity software allowed for a comparison of the PCC values generated between the Nuclear image channel (DAPI) and the GAIT Complex protein image channels, as well as the SMN image channel. This permitted examination of any nuclear-specific localisation changes of these proteins in response to IFN-Y induction in SHSY5Y cells (Figure 4.3.3).

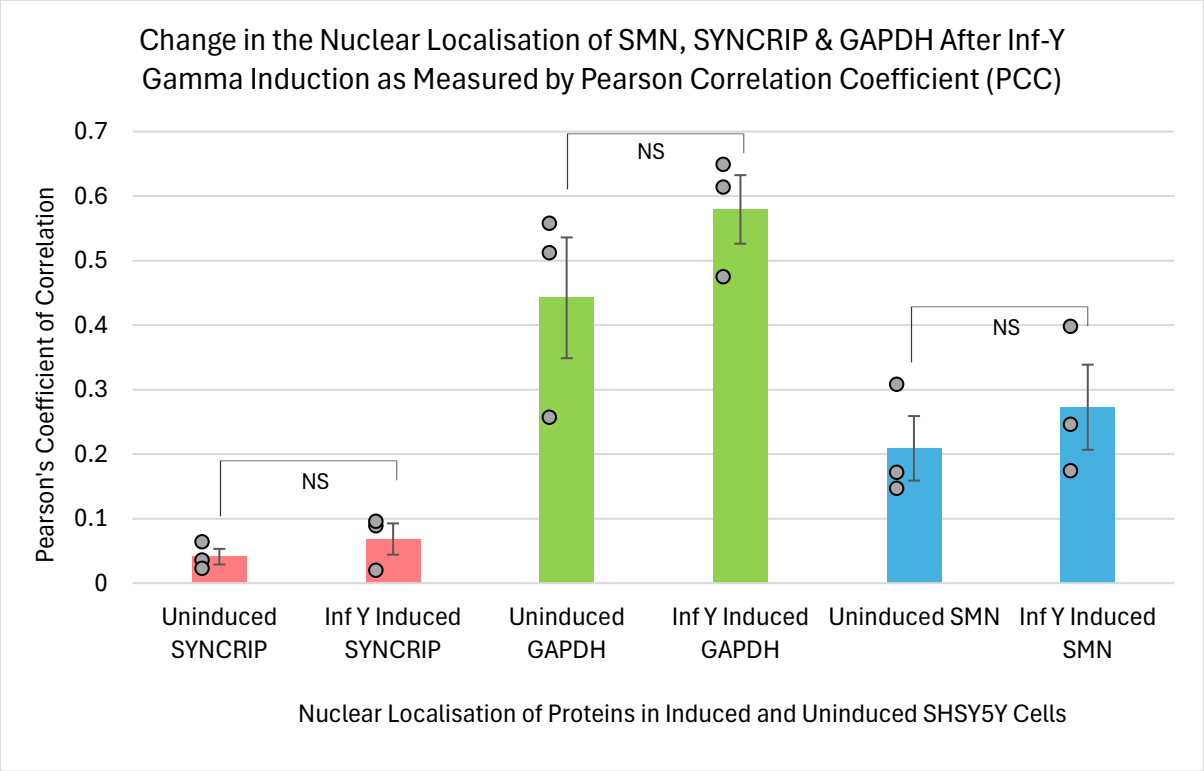
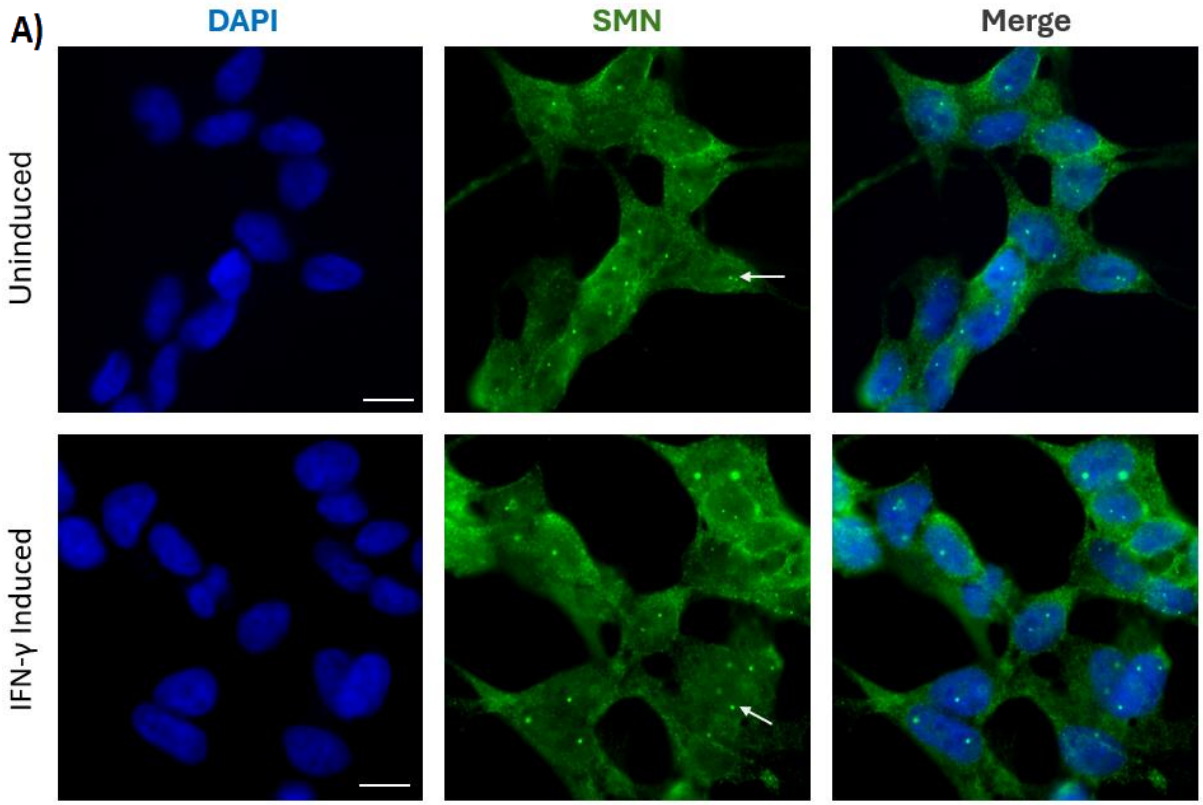


FIGURE 4.3.3: NUCLEAR LOCALISATION ANALYSIS OF SYNCRIP, GAPDH AND SMN IN UNINDUCED AND IFN-Y INDUCED SHSY5Y CELLS IDENTIFIES NO MAJOR DIFFERENCES IN NUCLEAR LOCALISATION.

Bar graph displaying the average Pearson Correlation Coefficient (PCC) calculating the nuclear localisation changes of SYNCRIP, GAPDH and SMN in both uninduced and Inf-Y Induced SHSY5Y Cells. The SYNCRIP-specific PCCs are represented by the pink bars (N=3, P=0.38, NS = non-significant) while the GAPDH-specific PCCs are represented by the green bars (N=3, P=0.40, NS). The SMN-specific PCCs in are represented by the blue bars (N=3, P=0.17, NS). Input images were obtained on a Deltavision microscope at 100x magnification and deconvolved using Volocity Image Analysis software. Error bars represent standard deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

Whilst SYNCRIP, GAPDH and SMN all demonstrated a relative increase in levels of nuclear localisation post-induction, none of these results were statistically significant (P=0.38, 0.40 and 0.17, respectively).

To further investigate whether IFN-γ induces a shift of cytoplasmic SMN to the nuclear compartment, the number of SMN-positive nuclear gems (white arrows in Figure 4.3.3) were counted in both uninduced and IFN-γ-induced SHSY5Y cells (N=100) from three separate coverslips (Figure 4.3.4A).



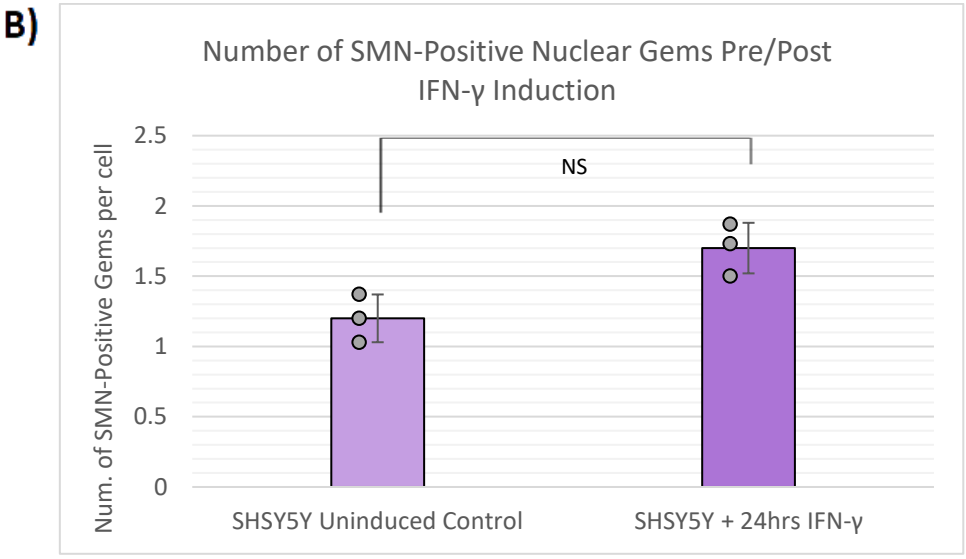


FIGURE 4.3.4: THE NUMBER OF SMN-POSITIVE NUCLEAR GEMS DOES NOT SIGNIFICANTLY INCREASE IN RESPONSE TO IFN-GAMMA INDUCTION.

A) Representative fluorescent imaging of uninduced SHSY5Y cells (top row) and SHSY5Y cells induced with IFN-γ (2000u/ml human recombinant IFN-γ for 24 hrs) (bottom row) stained against SMN (green). The white arrow in the SMN panels (green) highlights the nuclear gems containing SMN in

the nucleus, counterstained with DAPI (blue). Scale = 10 microns. Non-deconvolved images were obtained on a Deltavision microscope at 100x magnification and represent a single slice from the centre of a Z-stack. B) Bar graph displaying the number of SMN-positive nuclear gems manually counted in uninduced (light purple bar) and induced (dark purple bar) SHSY5Y cells (N=100 cells from 3 seperate coverslips, P=0.09, NS = non-significant) Error bars represent standard error of the mean (SEM). Individual data points are represented by grey circles located to the left of error bars.

Although IFN-γ induction did appear to moderately increase the number of SMN positive nuclear Gems (1.2 gems per cell in the uninduced cells versus 1.7 gems per cell after IFN-γ induction) this result was not statistically significant (T-Test P=0.09) (Figure 4.3.3B).

4.4 Protein Levels of Selected GAIT Complex Components and SMN Do Not Clearly Change After IFN-γ Induction in SHSY5Y Cells.

To determine if the protein levels of either SMN or the selected GAIT complex proteins were changed upon IFN-γ induction, cell lysates were harvested after 24hrs of treatment and immunoblot analysis using antibodies against SMN, SYNCRIP and GAPDH proteins were performed. Quantitative analysis of protein levels was determined via normalization to Tubulin isoforms.

SYNCRIP did appear to show an increase in protein concentration in the induced condition across several of the blots (N=3) (See Figure 4.4.1A) and this was supported by the quantification calculations, which showed an almost 6 times higher level of SYNCRIP in the

induced condition (2.3 Normalized Intensity (NI) in induced vs the 0.4 NI of uninduced controls) However, this value was not statistically significant (P=0.28) (Figure 4.4.1.B).

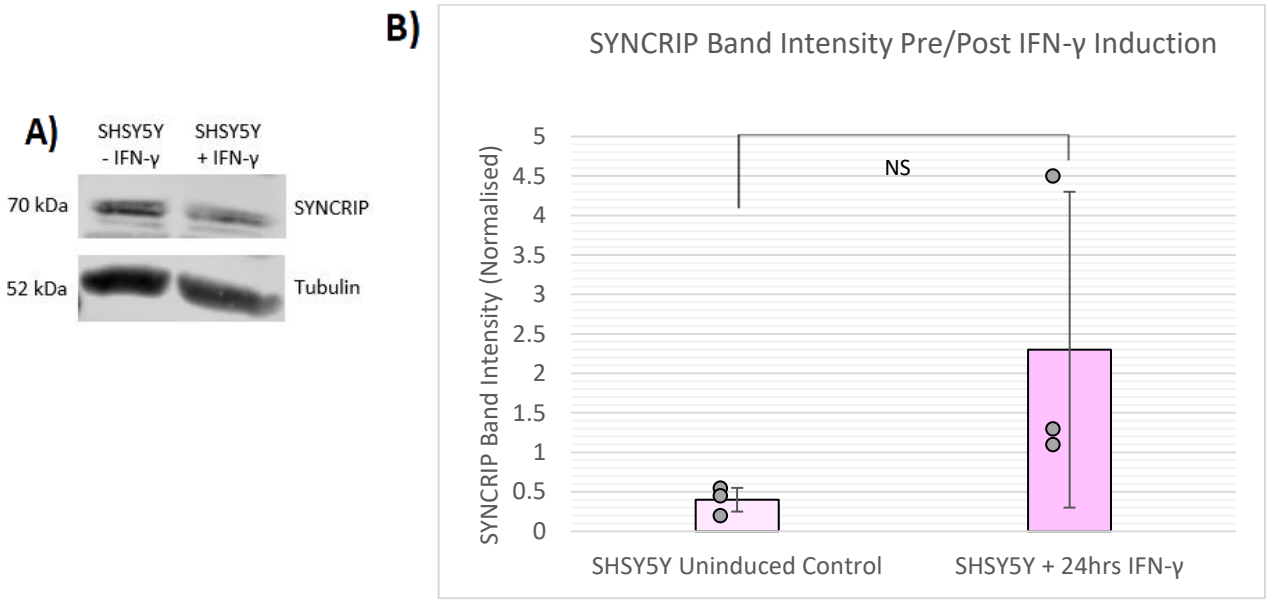


FIGURE 4.4.1: LEVELS OF SYNCRIP PROTEIN DO NOT CHANGE SIGNIFICANTLY AFTER IFN-GAMMA INDUCTION IN SHSY5Y CELLS.

A) Representative Western Blot stained against SYNCRIP protein and Tubulin in lysates derived from uninduced SHSY5Y cells and SHSY5Y cells induced with 2000u/ml human recombinant IFN-γ for 24 hrs. B) Bar graph displaying SYNCRIP band intensity (normalized to tubulin) of SYNCRIP in uninduced (light pink bar) and induced (dark pink bar) SHSY5Y cells. (N=3, P=0.28, NS = non-significant). Error bars represent standard deviation (SD). Individual data points are represented by grey circles to the left of Error Bars.

Similarly, GAPDH also appeared to show an increase in protein concentration in the induced condition across the blots (N=3) (Figure 4.4.2A). Again, this was supported by the quantification calculations, which showed a modest increase (1.3 NI in the induced vs 0.2 NI in the uninduced control) Again, this value was not statistically significant (P=0.12) (Figure 4.4.2B).

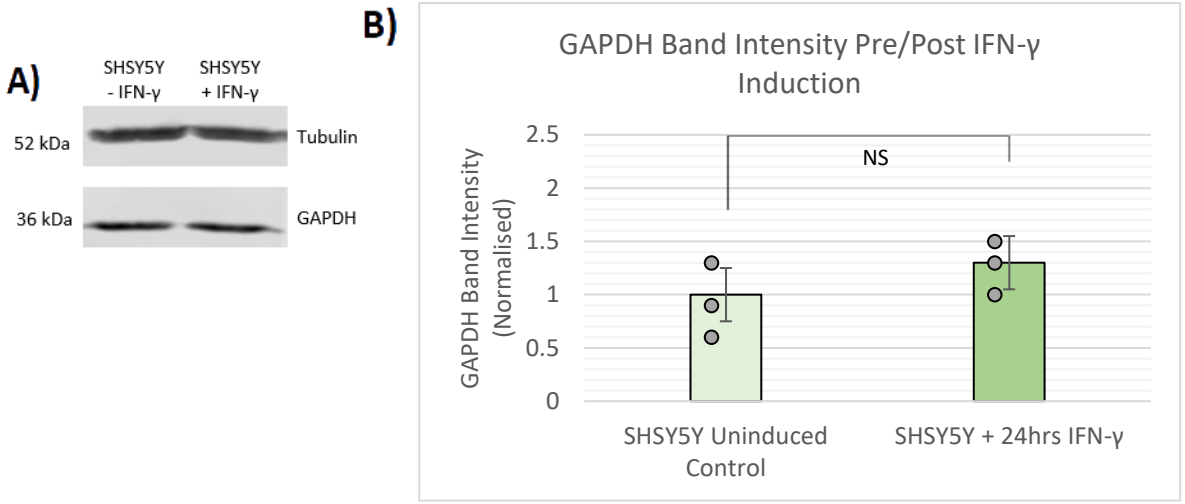


FIGURE 4.4.2: LEVELS OF THE GAPDH PROTEIN DO NOT CHANGE SIGNIFICANTLY AFTER IFN-GAMMA INDUCTION IN SHSY5Y CELLS.

A) Representative Western Blot stained against GAPDH protein and Tubulin in lysates derived from uninduced SHSY5Y cells and SHSY5Y cells induced with 2000u/ml human recombinant IFN-γ for 24 hrs. B) Bar graph displaying GAPDH band intensity (normalized to tubulin) of GAPDH in uninduced (light green bar) and induced (dark green bar) SHSY5Y cells. (N=3, P=0.012, NS = non-significant) Error bars represent standard deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

SMN also appeared to show a consistent increase in protein concentration in the induced condition across the blots (N=3) (Figure 4.4.3A). Again, this was supported by the quantification calculations, which showed an increase of over double in the induced lysates (2.5 NI in the induced vs 1.0 NI in the uninduced control) Again, this value was not statistically significant (P=0.08) (Figure 4.4.3B).

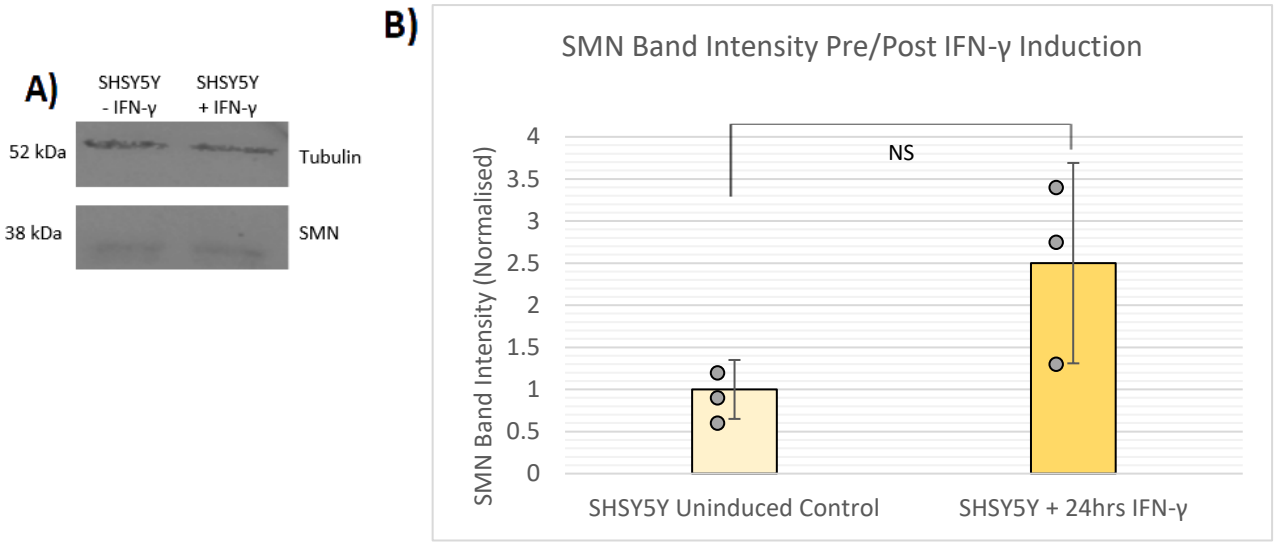


FIGURE 4.4.3: LEVELS OF THE SMN PROTEIN DO NOT CHANGE SIGNIFICANTLY AFTER IFN-GAMMA INDUCTION IN SHSY5Y CELLS.

A) Representative Western Blot stained against SMN protein and Tubulin in lysates derived from uninduced SHSY5Y cells and SHSY5Y cells induced with 2000u/ml human recombinant IFN-γ for 24 hrs. B) Bar graph displaying SMN band intensity (normalized to tubulin) of SMN in uninduced (light yellow bar) and induced (dark yellow bar) SHSY5Y cells. (N=3, P=0.08, NS = non-significant) Error bars represent standard deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

4.5 The Interaction Between GAIT Complex Proteins and GFP-SMN is Inconclusive

To clarify whether IFN-γ induction influences the interaction between SMN and GAPDH or SYNCRIP, the GFP-SMN plasmid discussed in Chapter 3 was transiently transfected into SHSY5Y cells. Whole cell lysates were harvested after 48 hours transfection and 24 hours of IFN-γ induction. Lysates were then processed and underwent GFP-Trap Immunoprecipitation (Figure 4.5).

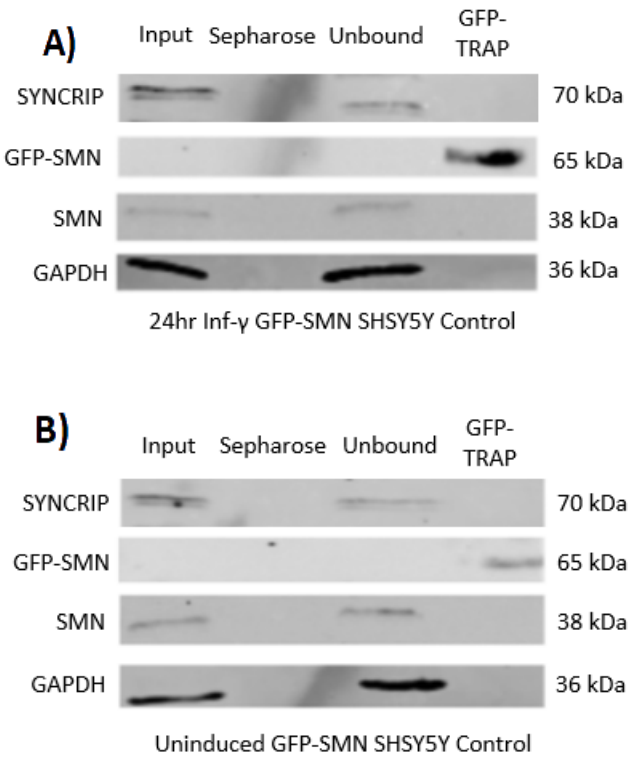


FIGURE 4.5: GFP-TRAP IMMUNOPRECIPITATION OF GFP-SMN IN A) UNINDUCED AND B) 24HRS IFN-Γ INDUCED SHSY5Y TRANSFECTED WITH GFP-SMN REVEALS NO INTERACTION BETWEEN GFP-SMN AND GAPDH OR SYNCRIP IN EITHER CONDITION.

Western blot of lysates derived from SHY5Y cells transiently transfected with GFP-SMN and upon which GFP-Trap Immunoprecipitation was performed. Signals for both GAPDH and SYNCRIP were missing in the GFP-Trap Lane, suggesting no interaction with GFP-SMN. Blots were probed with Rabbit Anti-SMN (SantaCruz) Mouse Anti-GAPH (PTG) and Rabbit Anti-SYNCRIP (PTG) antibodies (N=3).

From the above, it can be concluded that both SYNCRIP and GAPDH do not appear to pull-down with GFP-SMN in either IFN- γ induced or control conditions. However, as endogenous SMN did not pull down with GFP-SMN as would be expected (as shown by a lack of the SMN signal in the GFP-Trap lane) this calls into question the validity of these results. An attempt at repeating the experiment with GFP-SMNΔ7 failed due to low transfection efficiency of the GFP-SMNΔ7 plasmid in SHSY5Y cells. Due to time constraints, it was not possible to resolve these issues thus the data presented can be considered inconclusive.

4.6 Discussion

The existence of the formation of the GAIT complex in response to IFN- γ in neuronal cell lines is understudied within the literature. Previous studies have demonstrated that GAIT complex formation occurs in myeloid cells, however a similar response in neuronal cells has not been investigated; these results represent an initial investigation of its kind in the field (Mukhopadhyay et al., 2009). This investigation begins to shed light on the impact of IFN-γ on SMN and two key proteins within the GAIT complex - SYNCRIP and GAPDH – and further establishes that SHSY5Y cells are indeed IFN-γ competent. SHSY5Y neuroblastoma cells could thus be useful as a model for greater understanding of the neuron-specific cytokine response and, in turn, enhance our understanding of the role neuroinflammation plays in SMA aetiology.

While there were no overt trends observed in the qualitative image analysis of colocalisation changes after IFN-γ induction, further analysis utilising the quantitative generation of Pearson Correlation Coefficients did uncover a significant reduction in the colocalisation of SYNCRIP and SMN in IFN-γ induced SHSY5Y.

Furthermore, it was observed that SMN, SYNCRIP, and GAPDH experienced slight increases in protein level - although not statistically significant - following IFN-γ treatment. Given the ambiguous interpretation of results regarding SYNCRIP and GAPDH expression, these findings provide a foundation for future investigations into the GAIT complex in neuron-like

cells. Consequently, it is crucial to ascertain whether the GAIT complex functionality exists in neuronal cells and contributes to the reduction of pro-inflammatory transcripts.

Given results which show a lack of increased protein expression but a decrease in protein localisation between SMN and SYNCRIP, it is worth noting that - concerning the two remaining GAIT complex protein components (RPL13A and EPRS) - a 'depot' hypothesis has been proposed. In this hypothesis, these proteins are assumed to be released from their other canonical macromolecular complexes in order to function within the GAIT complex, rather than being generated via *de novo* synthesis after IFN-γ induction (Arif et al., 2018). Likewise, the incorporation of SYNCRIP and GAPDH into the GAIT complex most likely does not require new protein synthesis and would thus explain the above results. This would similarly explain the significant reduction in induced SYNCRIP and SMN colocalisation as identified by the Pearson Correlation Coefficient.

IFN-γ has been utilised in the clinic as an approved treatment since the 1980s, for a variety of conditions ranging from leukaemia to scleroderma (Miller et al., 2009). Its use in the context of SMA has, however, not been studied. It is possible that IFN-γ may have potential beneficial effect in attenuating the SMA phenotype, beyond a role in GAIT complex induction. IFN-γ upregulates expression of the guanylate-binding protein 1(GBP1), a GTPase with viral inhibitory properties. In a study by Xu *et al*., cells overexpressing GBP1 were found to contain higher concentrations of the SMNΔ7 protein, with no change observed at the level of SMN2 mRNA. This effect was also found in peripheral-blood derived patient samples; an increase in the level of GBP1 also correlated with the number of SMN2 copies in

the cell (Xu et al., 2023). From these data, they concluded that the IFN-γ stimulated protein stabilised the SMNΔ7 protein isoform, protecting it from degradation and thereby increasing the active cellular concentration. Whether this may prove beneficial in the context of SMA treatment remains to be seen, however Le *et al* demonstrated that increasing expression of SMNΔ7 alone in SMA mice ameliorates symptoms of SMA (Le et al., 2005).

On the other hand, IFN-γ is a natural cytokine and intimately involved in the processes of inflammation and host viral response. Cerebrospinal fluid (CSF) analysis of Type 1 SMA patients found significantly increased levels of IFN-γ (and other pro-inflammatory cytokines) compared to both Type 2 and Type 3 patients (Nuzzo et al., 2023). SMA Type 1 is the most severe clinical subtype of SMA, so baseline levels of inflammation – and thus neuroinflammatory markers - are likely higher in this cohort. Treatment with nusinersen was found to reduce the circulating levels of these markers, including IFN-γ, 6 months after the onset of therapy (Bonanno et al., 2022).

While the data generated in this study of the interplay between GAIT complex components and SMN are inconclusive, the observed outcomes suggest that IFN-γ may not be a viable therapeutic option in the treatment of SMA since it fails to substantially elevate SMN levels. This result stands in contrast to the findings of Baron-Delage and colleagues who - using human astrocytoma (A172) and human oligodendroglioma (HOG) cells- reported a significant induction of SMN protein expression as early as 8 hours post-IFN-γ treatment (Baron-Delage et al., 2000). Interestingly, this effect began to fade by the 24-hour time

point, therefore future investigations using shorter induction times would be beneficial. While the results presented earlier in this chapter are suggestive of an increase in the level of SMN in response of IFN-γ, the lack of significance demonstrates that further research is required.

Future investigations may consider the use of fluorescently tagged GAIT component proteins and live neuronal cell imaging; a measure of the colocalization changes in real-time post IFN-γ treatment may help to shed light on complex formation dynamics in these cells. Furthermore, the use of targeted qPCR to analyse the levels of SMN transcripts in induced and uninduced cells would be of interest. Putative mention of interferon response elements in the SMN promoter were reported in the literature (Echaniz-Laguna et al., 1999). An examination of this data, utilising modern technology and techniques that have improved since the time of the initial publication, would be of benefit to the broader SMA research community.

Chapter 5: Investigating Select tRNA Synthetases and their Relevance to SMA

Disclaimer: This work was undertaken with the help of student Susannah Leese, as part of her Undergraduate Research Project.

5.1 Overview of the tRNA Synthetases and the Multisynthase Complex

During protein translation, tRNA molecules pair to their cognate amino acid via an esterification reaction in a process catalysed by a family of enzymes known as the Amino Acyl tRNA Synthetases (AARSs). This process is initiated by activation of the amino acid, whereby the amino acid reacts with adenosine triphosphate (ATP) to form an aminoacyl adenylate intermediate (aminoacyl-AMP), releasing pyrophosphate. Subsequently, the aminoacyl-AMP is transferred to the 3' end of its cognate tRNA, establishing an ester bond at the 3'-OH group of the tRNA's adenosine residue. Now "charged," the tRNA proceeds to bind to elongation factors and is then presented to the ribosome for accurate amino acid deposition on the growing peptide chain (Banik & Nandi, 2012).

Further to their function in charging tRNAs, several of these tRNA synthetases form a large, multimolecular complex known as the Multisynthase Complex (MSC). In addition to a number of non-enzymatic proteins, such as AIMP1 and AIMP2, 9 of the AARS enzymes complete this complex (Figure 5.1). The MSC is theorized to increase tRNA proximity to

elongation factors and the ribosome, thereby increasing efficiency of translation (Kim & Kang, 2022).

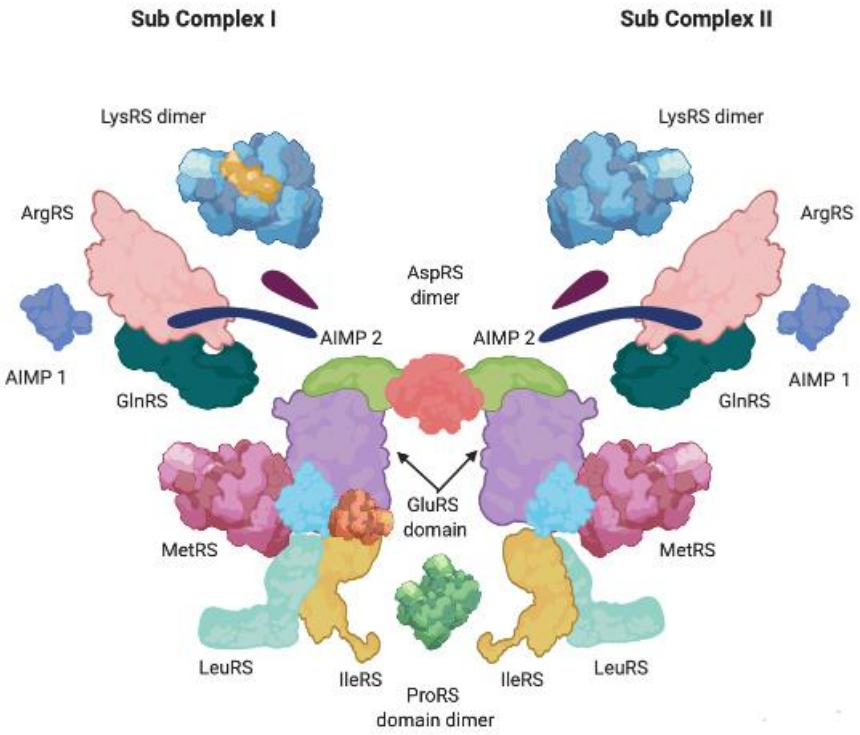


FIGURE 5.1: PROPOSED ORGANISATION OF THE MULTISYNTHASE COMPLEX (MSC).

All identified components of the MSC, including the tRNA Synthetases and interacting proteins AIMP1 and AIMP2. This model of the MSC proposes two monomeric Sub Complexes, each with a single copy of the relevant tRNA Synthetase. In this model, AspRS, LysRS and ProRS exist as dimers (Khan et al., 2020) (Image adapted with permission from Wikimedia Commons).

There are currently 39 identified AARS enzymes in humans, each responsible for the loading of a specific amino acid onto the relevant tRNA. 20 of these are cytoplasmic whilst 19 are mitochondrial. Cytoplasmic tRNA synthetases are involved in the synthesis of the bulk of cellular protein content, whereas mitochondrial variants synthesise ~13 proteins that are key to oxidative phosphorylation (Zheng et al., 2022)

Beyond this traditional role in translation, non-canonical functions have been described for several of the tRNA synthetases, with involvement in processes as diverse as cell division to inflammation (Turvey et al., 2022). Indeed, EPRS' role as an integral component of the GAIT complex and IFN-γ response has been previously discussed (See Chapter 4). Perhaps unsurprisingly, mutations in genes encoding AARSs are implicated in a broad spectrum of human disease, from tumorigenesis to neurodegeneration (Ognjenović & Simonović, 2018).

Of note, mutations within the gene encoding the glycyl-tRNA synthetase (GARS) have been linked to both Charcot-Marie-Tooth (CMT) disease and distal Spinal Muscular Atrophy Type 5 (dSMA-V). (Antonellis et al., 2003). Both diseases share similarities with SMA; loss of function mutations in a ubiquitously expressed protein leads to a phenotype reflected particularly in neurons and resulting in neuromuscular deficits (Vinogradova et al., 2021). This makes the tRNA synthetases prime candidates for further investigation, with regards to potential links to SMA.

5.2 Selection of tRNA Synthetases for Further Investigation

The list of differentially interacting tRNA synthetases discussed in Chapter 3 (See Table 3.4) describes the 9 tRNA synthetases that pulled down across the GFP-SMN and GFP-SMNΔ7 GFP-Trap experiments. Of interest, two proteins – Arginine-tRNA Synthetase and Isoleucine-tRNA Synthetase, herein referred to as RARS and IARS respectively - were both present in the GFP-SMN pulldown but missing from the GFP-SMNΔ7 dataset. This total loss of interactivity between the SMN and SMNΔ7 variant is intriguing, as the underlying pathology

of SMA is most likely reliant on the loss of function/interaction between SMN/SMNΔ7 variants and other proteins. Both RARS and IARS fit this previously established pathogenic pattern: interactivity appears to be present with native SMN but lost to the SMNΔ7 isoform.

IARS and RARS have no known interactions with the SMN protein, as verified by String-DB and a search of the literature, thus they represent novel proteins of interest when examined from an SMA-centric perspective. They do, however, form established interactions with other tRNA synthetases – such as the aforementioned GARS – linking them indirectly to the previously discussed neuropathies CMT and dSMA-V (Figure 5.2).

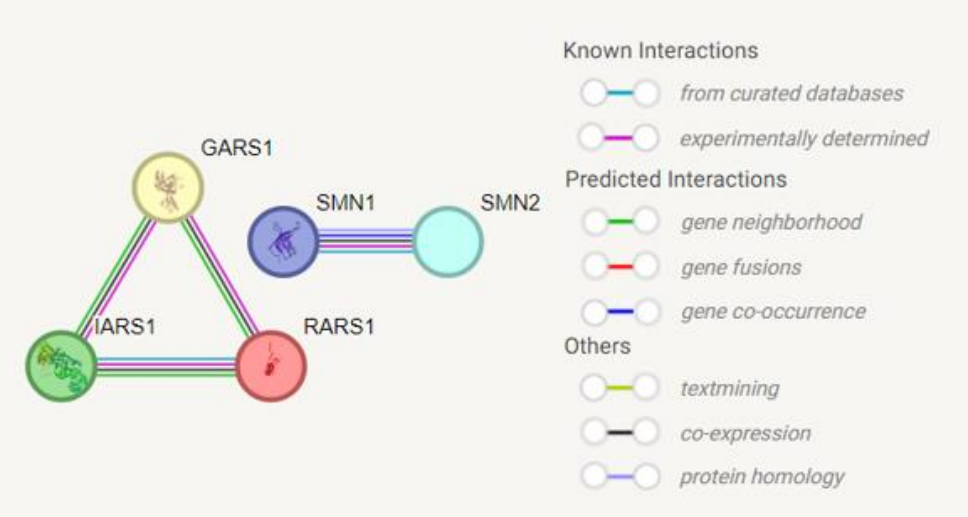


FIGURE 5.2: STRING-DB RESULTS SHOW NO KNOWN INTERACTIONS BETWEEN THE SMN PROTEIN ISOFORMS AND IARS, RARS AND GARS.

Edges (shown here as coloured lines) between node proteins (shown as circles) represent evidence supporting protein interaction; the Key defines these types of evidence. Edges between the SMN proteins and IARS, RARS and GARS are non-existent. Figure derived from String-DB.

Beyond their interactions with the GARS protein, mutations in both IARS and RARS have been found to cause defined neuronal effects. Mutations in RARS are known to cause hypomyelination in human patients (Wolf et al., 2014). Mutations in IARS, on the other hand, have been linked to growth retardation, impaired intellectual development, hypotonia, and hepatopathy (GRIDHH), implicating this protein in both neuronal and muscular dysfunction (Kopajtich et al., 2016). For these reasons, both IARS and RARS represent proteins of interest and relevance and were selected to undergo further investigation.

5.3 Colocalization Analysis of RARS and IARS with SMN in Both Hela and GFP-SMNΔ7-Transfected Hela Cells Reveals Differences in Interactivity Between Protein Variants

In order to investigate whether RARS or IARS colocalize with SMN, Hela cells were seeded onto coverslips before fixation. Antibodies specific to IARS, RARS and SMN were used in cell staining and, as before, imaged using the DeltaVision Microscope at 100x magnification. Localisation patterns – as well as any SMN co-localisation – were then determined.

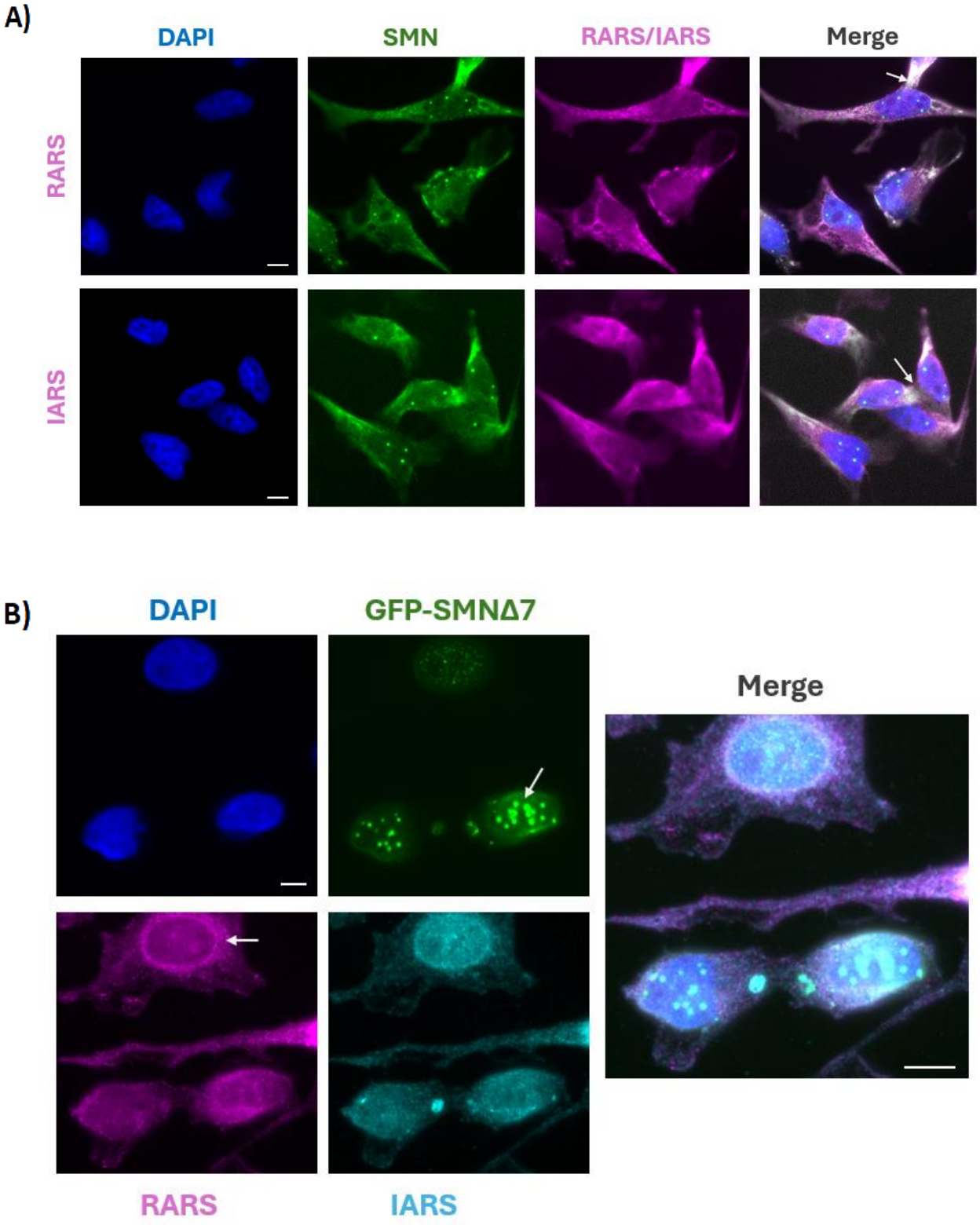


FIGURE 5.3.1: IMMUNOFLUORESCENT IMAGING OF RARS, IARS AND SMN/GFP-SMNΔ7 REVEALS AREAS OF COLOCALIZATION.

A) Fluorescent imaging of Hela cells antibody-stained with either mouse Anti-SMN (BD Biosciences) or rabbit Anti-SMN (Santa Cruz) (green) alongside mouse Anti-RARS (PTG) (magenta – top row) and

IARS (PTG) (magenta – bottom row). White arrows in merge images represent areas of observed colocalisation. B) Fluorescent imaging of antibody-stained mouse anti-RARS (PTG) (magenta) and rabbit anti-IARS (PTG) (cyan) in conjunction with transiently transfected GFP-SMNΔ7 (green) expressed in HeLa cells. The white arrow in the GFP-SMNΔ7 (green) panel highlights the punctate, nuclear localisation of SMNΔ7. The white arrow in the RARS panel (magenta) highlights the nuclear corona formation of RARS in this image. Scale = 10 microns. Non-deconvolved images were obtained on a Deltavision microscope at 100x magnification and represent a single slice from the centre of a Z-stack. Nucleus counterstained with DAPI (blue).

Both RARS and IARS have a primarily cytoplasmic localisation, given their canonical roles in translation of mRNA into protein. Examination of both proteins in conjunction with SMN revealed little to no specific colocalisation in nuclear Gems, however, there did appear to be some overlap with diffuse SMN in the cytosol in both cases (Figure 5.3.1A) (See white arrows in merged images).

Of note, in the GFP-SMNΔ7 Hela cells – where foci containing GFP-SMNΔ7 are abundant in the nucleus (see white arrow in the GFP-SMNΔ7 (green) panel) – IARS appears more localised in the nucleus than its cytoplasmic distribution in the untransfected cells (Figure 5.3.1B). Similarly, RARS appears much more prominently clustered within the perinuclear space, forming a nuclear corona (see white arrow in the RARS (magenta) panel). This suggests that there is a difference in localisation in the GFP-SMNΔ7 condition, whereby the tRNA synthetases appear more directed towards the nucleus. This is noteworthy because GFP-SMNΔ7, when transiently transfected, is primarily nuclear in distribution.

Utilising Volocity Image Analysis software, Deltavision images (including those shown above) were deconvolved and colocalisation analysis was performed between the deconvolved image channels representing either SMN and RARS/IARS, or GFP-SMNΔ7 and RARS/IARS. The colocalisation function measuring the overlapping pixel intensities between channels and produced a numerical value – the Pearson Correlation Coefficient (PCC) – between 0 and 1. The PCC in triplicate deconvolved images was calculated and a T-Test performed to identify significant changes between the colocalisation of the tRNA synthetase proteins and the two SMN isoforms (Figure 5.3.2).

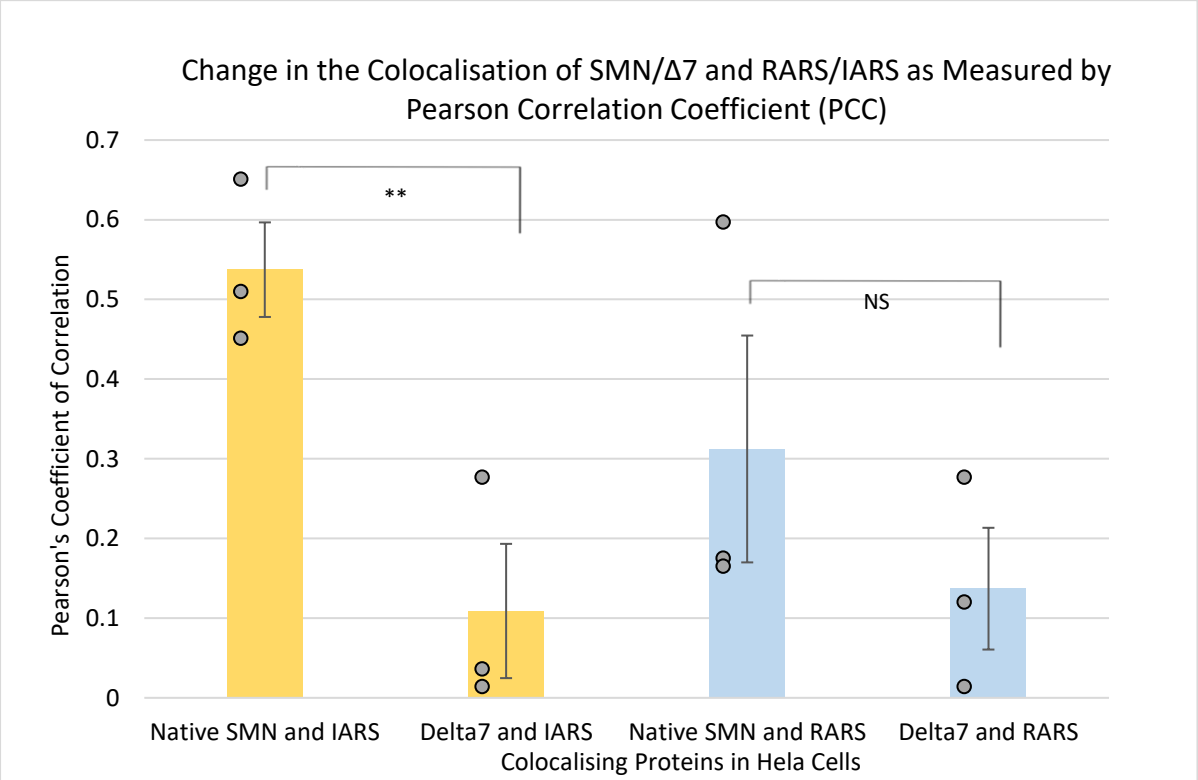


FIGURE 5.3.2: COLOCALISATION ANALYSIS OF SMN AND GFP-SMNΔ7 WITH RARS AND IARS SHOWS A SIGNIFICANT DIFFERENCE IN COLOCALISATION.

Bar graph displaying the average Pearson Correlation Coefficient (PCC) calculated between either SMN and RARS/IARS, or GFP-SMNΔ7 and RARS/IARS. The IARS-specific PCCs in are represented by the yellow bars (N=3, P=0.008, ** = significant) while the RARS-specific PCCs are represented by the

blue bars (N=3, P=0.053, NS = non-significant). Input images were obtained on a Deltavision microscope at 100x magnification and deconvolved using Volocity Image Analysis software. Error bars represent standard deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

Notably, a statistically significant difference in the interaction of the IARS protein with SMN and the SMNΔ7 isoform was found during image analysis (P=0.008). IARS was found to colocalise to a lesser extent with the SMNΔ7 variant, implying a loss of interaction specific to SMNΔ7. A similar loss of interaction with the SMNΔ7 variant was also observed for the RARS protein, however this data was just above the cutoff for statistical significance (P=0.053). Together with the qualitative image analysis, these quantitative data demonstrate that there are differences in the interaction landscape between SMN and SMNΔ7 with respect to the RARS and IARS proteins. However, as the SMNΔ7 protein primarily localises to the nucleus, this is perhaps unsurprising.

Further utilisation of Volocity software allowed for comparison of the PCC generated between the Nuclear image channel (DAPI) and the specific tRNA synthetase image channels. This data permitted examination of any nuclear-specific localisation changes of the AARSs in the transiently transfected GFP-SMNΔ7 cell line (compared to their nuclear localisation in the native, untransfected Hela cells) (Figure 5.3.3).

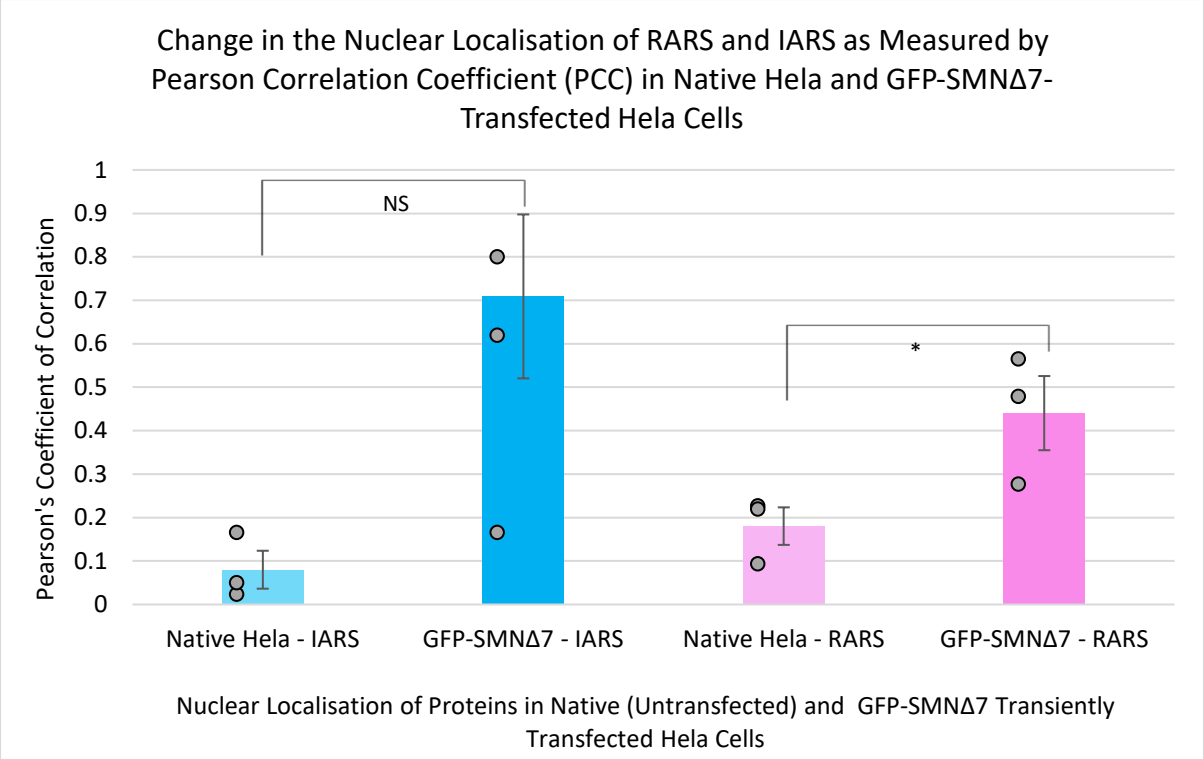


FIGURE 5.3.3: NUCLEAR LOCALISATION ANALYSIS OF RARS AND IARS IN NATIVE HELA AND TRANSIENTLY TRANSFECTED GFP-SMNΔ7 HELA CELLS REVEALS A SIGNIFICANT DIFFERENCE IN LOCALISATION.

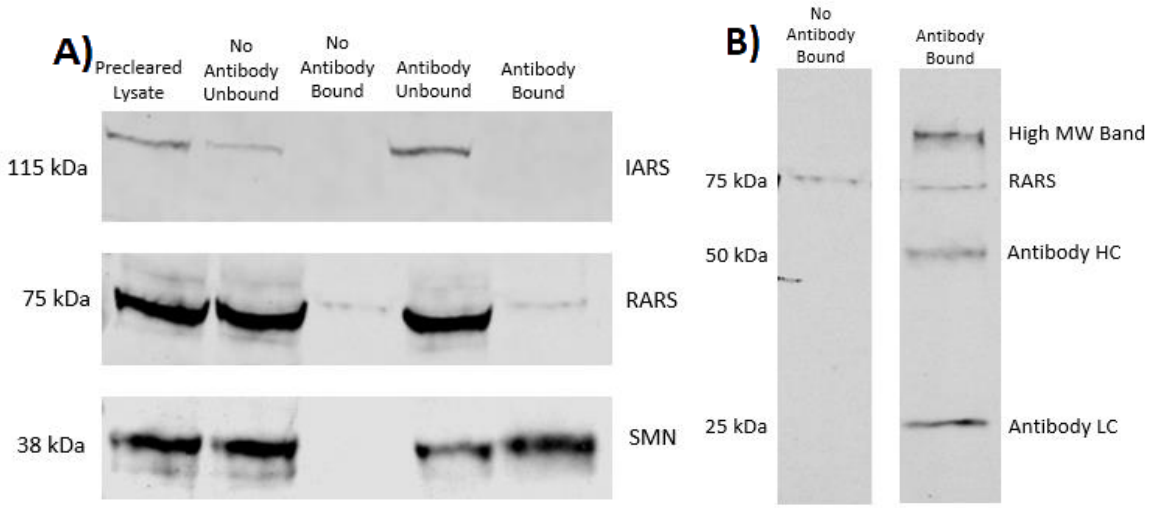
Bar graph displaying the average Pearson Correlation Coefficient (PCC) calculating either the nuclear localisation of RARS/IARS in native Hela cells or in equivalent GFP-SMNΔ7 transfected Hela cells. The IARS-specific PCCs are represented by the blue bars (N=3, P=0.067, NS = non-significant) while the RARS-specific PCCs are represented by the purple bars (N=3, P=0.036, * = significant). Input images were obtained on a Deltavision microscope at 100x magnification and deconvolved using Volocity Image Analysis software. Error bars represent standard deviation (SD). Individual data points are represented by grey circles located to the left of error bars.

Whilst IARS appeared to exhibit an increased level of nuclear localisation in the GFP-SMNΔ7 Hela cell line, this was not statistically significant (P=0.067), a result most likely attributable to the wide spread of the data points. RARS, however, did demonstrate a statistically significant increase in nuclear localisation in the GFP-SMNΔ7 cell line, when compared to native Hela controls (P=0.036). Together with the previous whole cell localisation analysis, these data indicate that there does appear to be differences

in the interactivity profiles between the AARSs – particularly IARS – and the SMN/SMNΔ7 proteins. Given the relatively increased abundance of nuclear-localised IARS and RARS in the GFP-SMNΔ7 cell line, this effect does not seem to be entirely due to the nuclear-specific localisation of GFP-SMNΔ7.

5.4 Immunoprecipitation and GFP-Trap of Hela and GFP-SMNΔ7 Hela Cells Reveal Differential Interactions with RARS and IARS

To probe the proposed interactions between SMN, IARS and RARS at the protein level, Hela cell lysates were immunoprecipitated utilising an antibody against SMN. The resulting fractions underwent western blotting (Figure 5.4.1). This represented the native SMN variant pulldown. In order to compare this with an SMNΔ7 variant pulldown, the GFP-SMNΔ7 plasmid discussed in Chapter 3 was transiently transfected into Hela cells; lysates were harvested after 48 hours of transfection. Harvested lysates were processed and underwent GFP-Trap Immunoprecipitation (Figure 5.4.2).



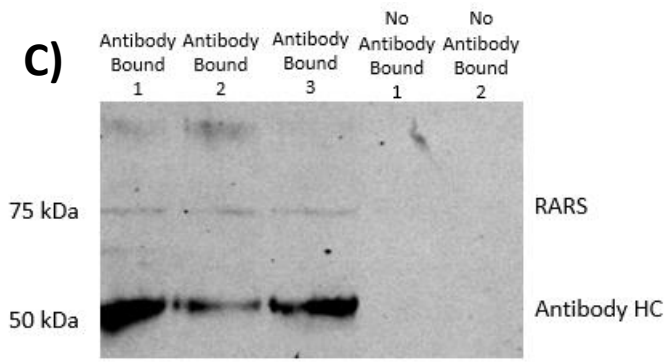


FIGURE 5.4.1: IMMUNOPRECIPITATION PULLDOWN OF SMN FROM HELA CELLS DOES NOT REVEAL ANY CONCLUSIVE INTERACTION WITH RARS AND IARS PROTEINS.

A) Western blot of Hela SMN IP demonstrating the lack of an IARS signal in the No-Antibody Bound control lane and the Antibody Bound Lane. A signal present in these lanes for RARS is identified in both lanes, suggesting a RARS-specific interaction with the protein-G Beads. SMN is also confirmed to have immunoprecipitated, as shown by its presence in the Antibody Bound lane. B) Full lane blot of Hela SMN IP antibody bound lane, revealing the high Molecular Weight (MW) band above the 75kda RARS band and the antibody HC and LC. The No Antibody Bound lane is also shown to demonstrate that this band does not appear in this fraction. C) Repeat of the Hela SMN IP antibody bound lane using harsher denaturing conditions, demonstrating the high molecular weight band has diminished and the non-specific band present in the no-antibody bound control lane is no longer present. Blots were probed with Rabbit Anti-IARS (PTG) and Mouse Anti-RARS (PTG) antibodies, then re-probed with Mouse Anti-SMN (MANSMA). HC= Heavy Chain; LC = Light Chain (N=3).

IARS does not appear to pull down with native SMN, as can be inferred by the lack of appropriate signal in the “Antibody Bound” lane for this protein (Figure 5.4.1A). RARS, on the other hand, does appear to have been pulled down. However, as a similar RARS signal has also appeared in the “No Antibody -Bound” control lane, this suggests that RARS may be interacting with the Protein G beads and not the SMN antibody. Yet, as can be seen in Figure 5.4.1B, just above the expected 75kDa weight of RARS in the Antibody-Bound fraction is an intense signal - in the RARS channel - that does not exist in the No-Antibody control fraction. After accounting for the heavy and light chains of the Mouse SMN antibody used in

the pulldown, it is likely that this heavier signal was either RARS or complexed antibody. It seemed probable that a portion of these proteins were either not fully denatured or remained bound to some interactor, most likely attributable to the milder conditions of the denaturing process (65^{o} for 25 minutes versus the 95^{o} for 10 minutes used in the other, non-bead bound fractions).

After re-boiling the Antibody Bound and No-Antibody Bound control fractions for a further 5 mins at 95^{o} it became clear that this band consisted primarily of heavy chain antibody and not RARS protein (Figure 5.4.1C). In the repeated western, the 75kDa signal of RARS persists in the Antibody Bound Lane, however, the same weight signal in the No-Antibody Bound lane has disappeared. This may suggest that RARS does in fact pull down with native SMN, however given the confounding factor of a potential bead-based interaction, further testing is required.

If RARS does indeed pull down with full-length SMN and they partake in mutual interactions within the cell, this agrees with the previous GFP-Trap dataset analysed in Chapter 3. The lack of an IARS signal in this pulldown is at odds with said data, whereby IARS was detected as an interacting partner of GFP-SMN.

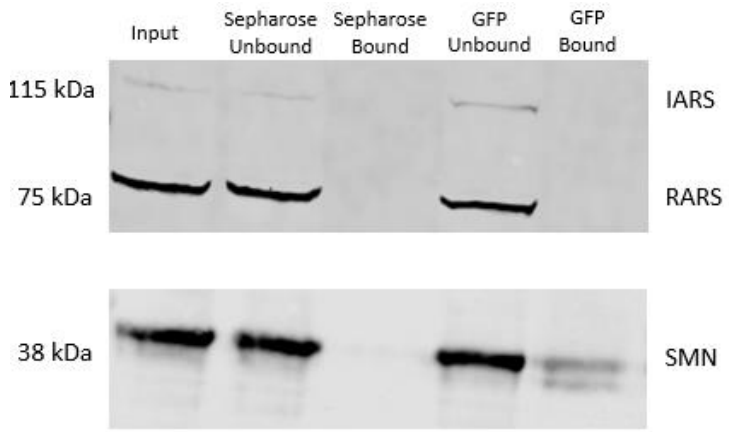


FIGURE 5.4.2: THE GFP-TRAP PULLDOWN OF GFP-SMNΔ7 FROM TRANSIENTLY TRANSFECTED HELA CELLS DOES NOT REVEAL ANY INTERACTION BETWEEN RARS, IARS AND GFP-SMNΔ7.

Western blot of lysates derived from Hela cells transiently transfected with GFP-SMNΔ7 and upon which GFP-Trap Immunoprecipitation was performed. Signals for both RARS and IARS were missing in the GFP-Bound Lane, suggesting no interaction with GFP-SMNΔ7. SMN was probed as a control to ensure pulldowns were effective. Blots were probed with Rabbit Anti-IARS (PTG) and Mouse Anti-RARS (PTG) antibodies, then re-probed with Mouse Anti-SMN (BD Biosciences) (N=3).

The GFP-Trap pulldown of GFP-SMNΔ7 revealed a clearer lack of interaction between RARS, IARS and the SMNΔ7 protein isoform, as there was no interaction between the proteins and the GFP-Trap (or plain Sepharose) beads (Figure 5.4.2). From this data it can be concluded that IARS and RARS do not interact with GFP-SMNΔ7.

5.5 Discussion

Initial mass spectrometry analysis of GFP-Trap pulldown datasets showed that interactions between RARS, IARS and SMN were unique to the native SMN isoform and lost in the SMNΔ7 experimental condition (as described in Chapter 3, Table 3.4).

The targeted immunoprecipitations performed above agree with a lack of interaction of these proteins with SMNΔ7, however they do not confirm the presence of an interaction with native SMN. RARS in particular - due to a potential interaction with the Protein G Beads used in the native pulldowns - requires further examination. While native SMN was chosen for examination in the above AARS analysis, it is worth noting that the initial pulldown Mass Spectrometry datasets were generated via pulldown of GFP-SMN; future investigations may want to consider the use of both GFP-SMN and native SMN to probe this potential interactor.

IARS was not immunoprecipitated with native SMN in Hela cell lysates, contrary to its appearance in the GFP-SMN dataset. It is possible that the level of IARS protein pulled down was merely beyond the range of detection of western blotting. Mass spectrometry is highly sensitive and can identify interactors that do not appear on general western blot (Aebersold et al., 2013).

In agreement with this idea, the change in colocalisation values (as defined by the Pearson Correlation Coefficient) between IARS and the two isoforms of SMN demonstrates a statistically significant reduction in localisation between proteins; this may imply a level of interaction lost in the SMNΔ7 condition that exists between the native SMN protein and IARS. The RARS protein showed a similar pattern of decreased localisation, however this data was not statistically significant.

Because RARS and IARS are novel potential SMN interactors, they have not been thoroughly investigated through an SMA-centric lens; their contribution - if any - to the pathology observed in SMA is unclear. However, the specific involvement of other Amino Acid tRNA Synthetases in the pathogenesis of neuromuscular disease provides a prospective framework for consideration.

Charcot-Marie-Tooth Disease, Type 2D (CMT2D) is an inherited neurological disorder caused by mutations in the GARS gene which encodes for glycyl-tRNA Synthetase. Patients display muscular denervation, motor and sensory deficits along with muscular weakness and atrophy. Similarly, Distal Spinal Muscular Atrophy Type 5 (dSMA-V) is a neurological disease closely linked to CMT2D, and indeed the two conditions show significant overlap. It is characterized by progressive weakness of the distal muscles and atrophy predominantly affecting the upper limbs. Like CMT2D, dSMA-V is also caused by mutations in the GARS gene.

Precisely how mutations in the genes encoding AARSs result in specific neuropathies like those described above are not well understood. However, disruptions in protein synthesis, tRNA charging, or other cellular functions mediated by RARS and IARS could potentially compromise neuronal integrity and function via effects on these essential processes.

Mutations in the RARS gene have been associated with hypomyelination disorders, whereas mutations in the IARS gene have been implicated in forms of hereditary spastic paraplegia (HSP), a group of inherited neurological disorders characterized by progressive weakness and spasticity of the lower limbs (Nafisinia et al., 2017). IARS mutations have been found to cause both intellectual disability and muscular hypotonia in affected patients (Kopajtich et al., 2016).

At which functional level the AARSs and SMN interact has thus far not been elucidated. SMN has been previously shown to interact with the translation machinery through interactions with both the ribosome and elongation factors (Sharma et al., 2024). SMN-primed ribosomes have also been demonstrated to directly affect the translation of specific transcripts, while knockdown of SMN has been shown to significantly reduce the level of protein synthesis in primary motor neurons *in vitro* (Lauria et al., 2020). Given its evident relevance to translational processes, it is possible that interactions native to SMN and the AARSs at this level are affected in SMA, contributing to the decreased translation and ribosome stalling as seen in the disease (Bernabò et al., 2017).

Yet several AARS enzymes have been shown to perform non-canonical, non-tRNA charging-related functions. As previously discussed in Chapter 4, EPRS – the bifunctional glutamyl-prolyl-tRNA synthetase - moonlights as a constituent of the GAIT complex in response to induction of interferon-gamma (Arif et al., 2018). Similarly, IARS knockout mouse models display phenotypes with distinct mitochondrial deficits, unexpected given the context of the protein's canonical function (Watanabe et al., 2023).

Of interest would be further examination of the aforementioned EPRS protein – combining both the tRNA synthetase and GAIT complex angles. As is evident from any interactomic investigation, a protein's function - and dysfunction – is reliant on its response to the proteins surrounding them. Combining the two lanes of investigation may lead to potentially interesting results.

Chapter 6: Generation and Analysis of TurboID Stable Cell Lines

6.1 Utilising an All-In-One TetOn Plasmid System to Generate Stable Cell Lines Expressing TurboID Fused to FLAG-SMN/SMNΔ7

6.1.1 Overview of TurboID

As discussed in Chapter 1, TurboID is a proximity-dependant labelling technique which labels the interacting partners of a protein of interest (in this case, SMN/SMNΔ7). The TurboID enzyme is a variant of BirA, a biotin ligase which - when conjugated to a bait protein such as SMN or SMNΔ7 - will biotinylate all proteins within a 10-nanometer sphere upon the addition of biotin to culture media (Li et al., 2019). Unlike immunoprecipitation reactions such as FLAG-IP - which reveal only stable protein interactors - TurboID can capture transient, weaker protein interactions (Samavarchi-Tehrani et al., 2020).

6.1.2 Design Specifics of an All-In-One TurboID Plasmid Line

In order to proceed with TurboID experiments, it was decided that the plasmids for generating stable cell lines for downstream proteomic analysis should contain several key elements:

1) Express the TurboID construct, linked to a FLAG tag and the protein of interest with a sequence corresponding to SMN or SMNΔ7 proteins.

2) Contain a puromycin selection cassette in order to reduce the time required for selection (compared to geneticin).

3) Contain a doxycycline inducible promoter to offset any potential toxicity i.e. a Tet-On system. A recent study has shown that use of a Tet-On system in TurboID can reduce off-target, background biotinylation (May et al., 2020).

4) Be backwards compatible with a range of in-house plasmids i.e. contain appropriate restriction sites surrounding (and within) the TurboID-FLAG-Protein construct so that future proteins of interest can be inserted into the vector easily. For example, the eGFP-SMN construct can be inserted in place of the TurboID-FLAG construct to produce a dox-inducible SMN system.

5) Contain stop codons in all three reading frames downstream of the TurboID-FLAG-Protein construct as a transcriptional failsafe measure.

These elements were combined into a set of TurboID plasmids delivered by Aruru Molecular, herein referred to as the All-in-One (AIO) TurboID Plasmid Line (Figure 6.1.2). These plasmids are considered All-in-One because in using the TetOn 3G system, no extra TetOn Helper plasmid was required for stable cell line generation: all components necessary for the Tetracycline inducible system are present within each plasmid. This reduces the necessity of stepwise transfections as occurs in 2-Plasmid Systems, however the increased size of the genetic material delivered in AIO plasmids can result in reduced transfection efficiencies (Randolph et al., 2017).

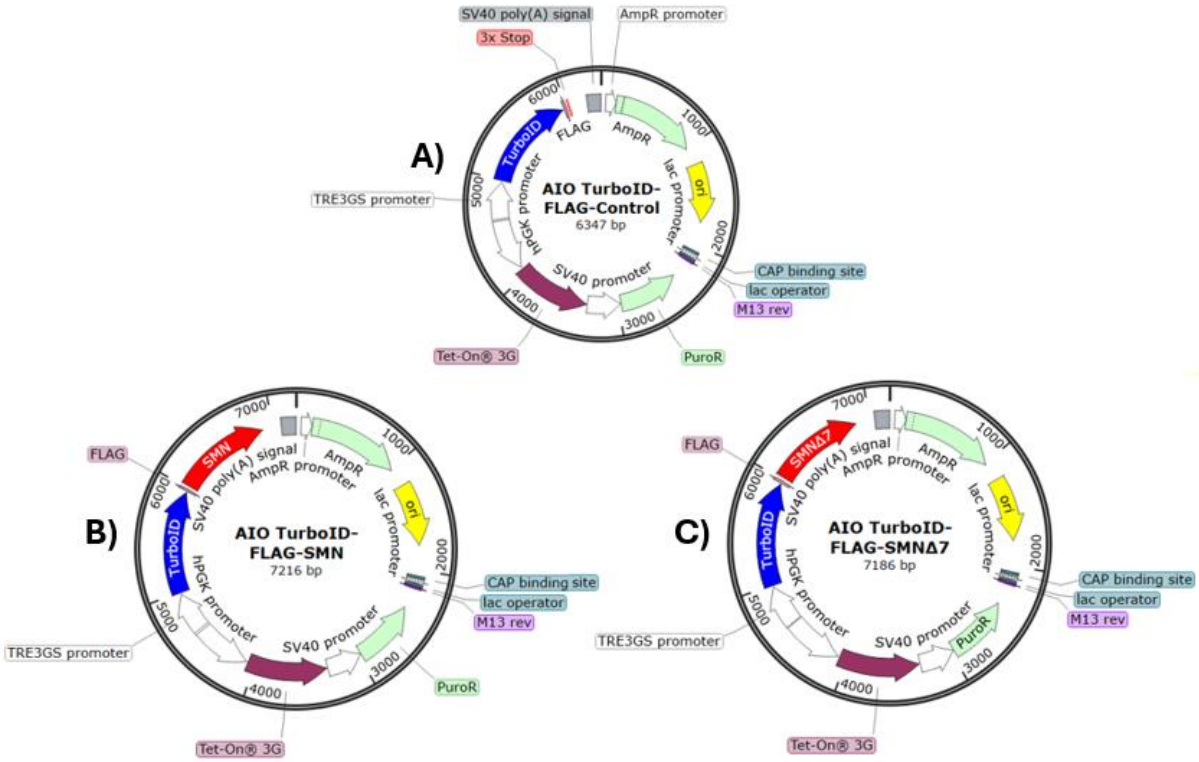


FIGURE 6.1.2: PLASMID MAPS OF THE ALL-IN-ONE (AIO) TETON TURBOID PLASMIDS.

A) Map of AIO Plasmid designed to express FLAG-tagged TurboID (control plasmid) B) Map of AIO Plasmid designed to express FLAG-tagged TurboID-SMN. C) Map of AIO Plasmid designed to express FLAG-tagged TurboID-SMNΔ7

6.1.3 Restriction Enzyme Digests Validate the All-in-One Plasmids

In order to validate the identity of the plasmids, the AIO TurboID Plasmids were digested using restriction enzymes: either with EcoRI alone (linearization/single digest) or EcoRI and HindIII (double digest). These digests were run on a 0.5% agarose gel for 40 minutes in TEB buffer, stained with Safeview and visualized under UV light (Figure 6.1.3).

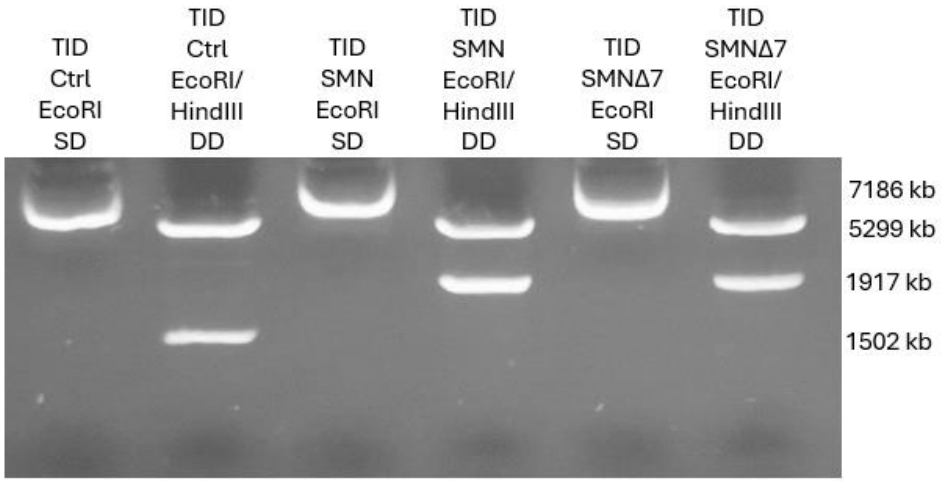


FIGURE 6.1.3: RESTRICTION DIGESTIONS OF THE AIO TURBOID PLASMID LINE.

0.5% agarose gel stained with Safeview, run for 40 minutes in TBE Buffer. Plasmids were digested using either EcoRI alone (SD = Single Digest i.e. linearization of plasmid) or EcoRI and HindIII (DD = Double Digest). TID = TurboID.

Bands at the predicted locations were present post-digestion, confirming that the AIO plasmids were the correct sizes and had digested as their plasmid maps had predicted. Plasmid stocks were then expanded via miniprep, and glycerol stocks were frozen at -80.

6.1.4 Generation of Stable Cell Lines Using the All-in-One TurboID Plasmids

1ug of each AIO TurboID Plasmid was transfected into both Hela and SHSY5Y cells using Effectene Transfection Reagent. Low transfection efficiencies were problematic in the generation of stable cell lines in both Hela and SHSY5Y lines (<10% efficiency in Hela and <5% in SHSY5Y). Post-selection with puromycin (400ng/ml), the majority of surviving cells did not express the constructs in any appreciable fashion as tested by both immunostaining and western blot. However, increasing doses of selection antibiotic did not induce cell death

in these lines, implying that they had been transfected with the antibiotic resistance gene. Comparative transfections with other, validated plasmids (GFP-C1 and GFP-SMN) demonstrated that the transfection technique was successful, however numerous variations of this technique – changing the DNA concentrations, the volume of transfection reagent, seeding density of cells prior to transfection – had no discernible effect on increasing expression of the AIO transfected TurboID plasmids. PFA fixation of induced and biotinylated cells on coverslips allowed cell lines to be immunostained and imaged on the DeltaVision microscope (Figure 6.1.4).

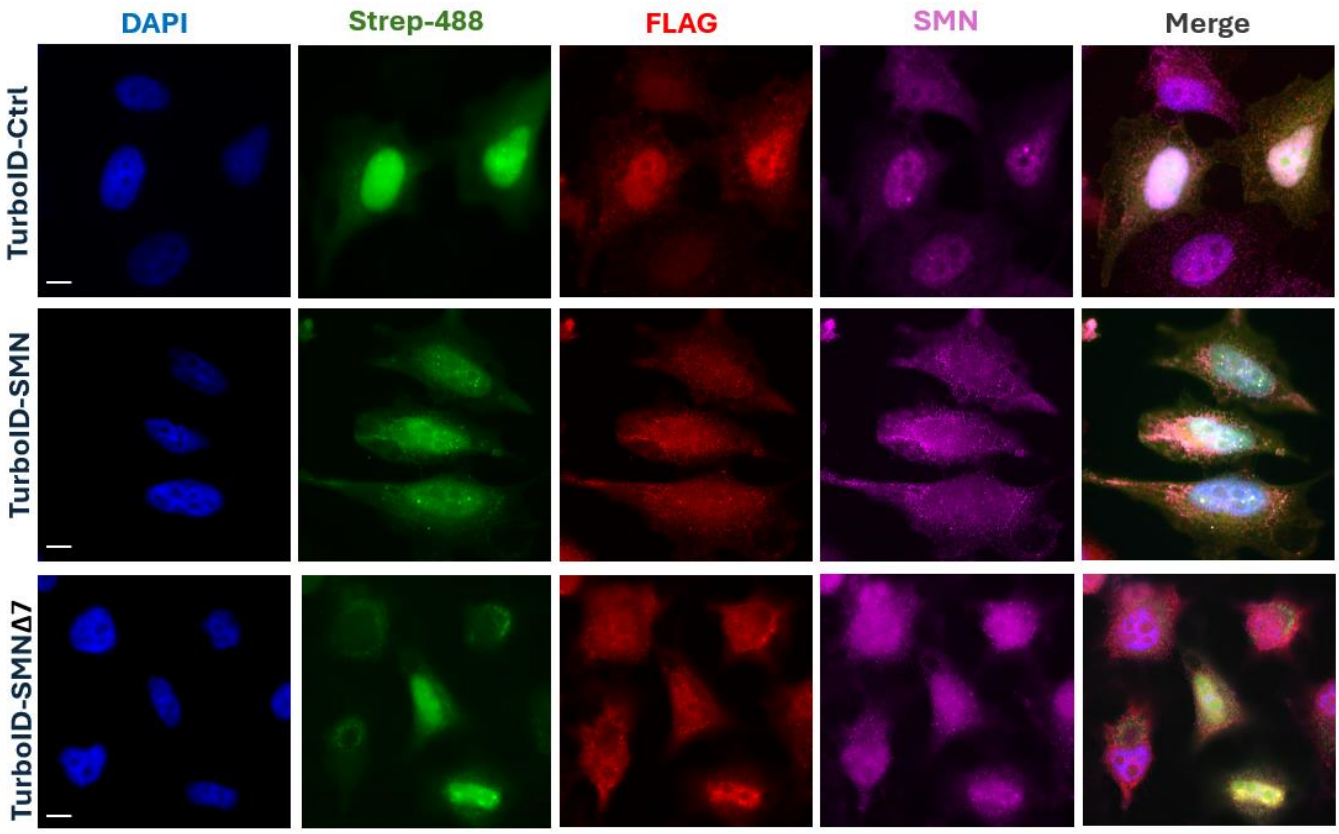


FIGURE 6.1.4: IMMUNOFLUORESCENT IMAGING OF INDUCED AND BIOTINYLATED HELA TURBOID CELL LINES SHOWS EVIDENCE OF BIOTINYLATION AND PROTEIN EXPRESSION.

Representative fluorescent imaging of 48hr induced (1ug/ml doxycycline) and 1 hr Biotin incubation (150uM at 37 degrees Celsius) TurboID Hela cell lines. TurboID-Control (top), TurboID-SMN (centre)

and TurboID-SMNΔ7 (bottom) cells were stained with Streptavidin-488 (green) (Licor) to indicate biotinylation, mouse Anti-FLAG (Invitrogen) (red) and rabbit anti-SMN (Santa Cruz) (magenta). Nuclei were counterstained with DAPI (blue). Scale = 10 microns. Deconvolved images were obtained on the DeltaVision microscope at a magnification of 100x and represent the centre slice of a Z-Stack.

Although there was evidence of biotinylation and protein expression, these cells were not well distributed and represented less than 1% of the total cells on each coverslip.

This small number of cells which did appear to express were sparse and resistant to clonal isolation; cell death occurred within days upon colony isolation and attempted expansion, and for those that did propagate, any initial level of expression diminished after subculture. SHSY5Y cell lines did not appear to show any expression at all when examined via immunostaining.

6.1.5 The All-in-One TurboID Plasmids Do Not Maintain Sufficient Levels of Expression in Hela or SHSY5Y Cells.

Further to imaging and attempted subcloning/expansion, western blots to probe the expression levels of TurboID-Control, TurboID-SMN and TurboID-SMNΔ7 proved negative in both Hela and SHSY5Y transfected cell lines (Figure 6.1.5). This seemed to be a result of the low initial transfection efficiency paired with low expression levels of the proteins of interest.

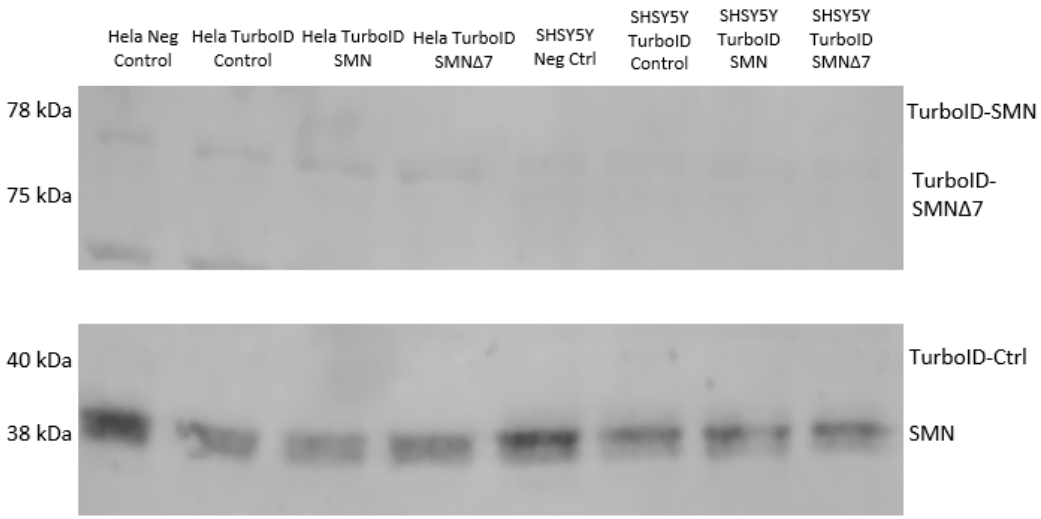


FIGURE 6.1.5: THE AIO TURBOID PLASMIDS DO NOT MAINTAIN SUFFICIENT LEVELS OF EXPRESSION IN HELA OR SHSY5Y CELLS.

Western blot of the induced AIO TurboID Plasmid Lines. All samples were induced with doxycycline (1ug/ml) for 48 hours. Stained with mouse Anti-SMN (BD Biosciences) and rabbit Anti-FLAG (Sigma). Negative controls represent untransfected Hela/SHSY5Y cells. Input = 40ug.

From these results, it became clear that the proteins of interest were not being expressed in sufficient concentration - (or by a sufficient number of cells) - to enable further downstream analysis of TurboID-SMN/SMNΔ7 interactomes.

6.1.6 The All-in-One TurboID Plasmids Could Not Successfully Generate Stable Hela or SHSY5Y Cell Lines

It became evident that over time these cells lost or repressed their limited expression of the TurboID constructs, resulting in cell lines which did not produce enough protein for targeted pulldowns via either FLAG IP or Streptavidin. Whilst the antibiotic resistance gene within the

plasmid appeared to function well and integrate into transfected cells, the TurboID constructs either lacked consistent levels of expression or were silenced in ensuing generations.

Due to the lockdowns of 2020 - when these plasmids were generated - they were not subject to the stringent level of Quality Control as would be expected from an externally supplied product. Although they were sequenced prior to shipping and validated in the lab via restriction digest, their ability to be efficiently transfected and express in living cells was not validated prior to delivery.

Due to this failure, as confirmed via both repeated immunostaining and western blot analysis, it was decided to pursue another avenue of TurboID expression, this time utilising Lentiviral vectors.

6.2 Utilising a Lentiviral TetOn Plasmid System to Generate Stable Cell Lines Expressing FLAG-TurboID Fused to SMN/SMNΔ7

6.2.1 Design Specifics of a Lentiviral TurboID Plasmid Line

Further to the failure of the previous plasmid line to transfect and express efficiently in Hela and SHSY5Y cells, it was decided that a new line of plasmids would be generated with a few key changes. These plasmids would:

1) Utilise a lentiviral transduction system.

2) Include a second generation TetOn system – like the previous plasmids, the TurboID constructs were inducible in the presence of Doxycycline. Unlike the previous plasmids, all components of the TetOn System were not present within the one plasmid. The Tet Repressor protein, which cannot bind to the TetOn transcription elements when dox is absent, was coded into one plasmid (known herein as the TetOn Helper plasmid) which must first be transduced into cells prior to TurboID plasmid transduction. This plasmid also encodes a silencer which binds to upstream promoter elements in the absence of doxycycline, preventing expression and reducing any leakage inherent in the system. (Agha-Mohammadi et al., 2004). Utilising a two-plasmid system resulted in a reduction in size of the plasmid encoding the TurboID constructs, allowing for space available for the inclusion of the following features.
3) Include a fluorescent marker for deduction of transduction efficiency. This protein would be expressed to allow for easier assessment of transduction. mKate2 was chosen because it is a red fluorescent protein which is non-toxic and has shown positive results in previous live cell imaging (Shcherbo et al., 2009).
4) Contain an expanded FLAG tag (from a single FLAG motif to 5) which was moved from the centre of the TurboID-Bait Protein construct (wherein it had been situated in the previous plasmid line) to the start of the construct; this was in case the previous centralised tag location (between the TurboID and the bait proteins) had interfered with the ability of the previous constructs to fold properly.

These elements were combined into the second set of TurboID plasmids, delivered by Vector Builder, herein referred to as the LV (Lentiviral) TurboID Plasmids (Figure 6.2.1).

Vector Builder provided several certificates of authenticity, and their in-house QC meant that we could be reasonably sure that each plasmid was delivered as described.

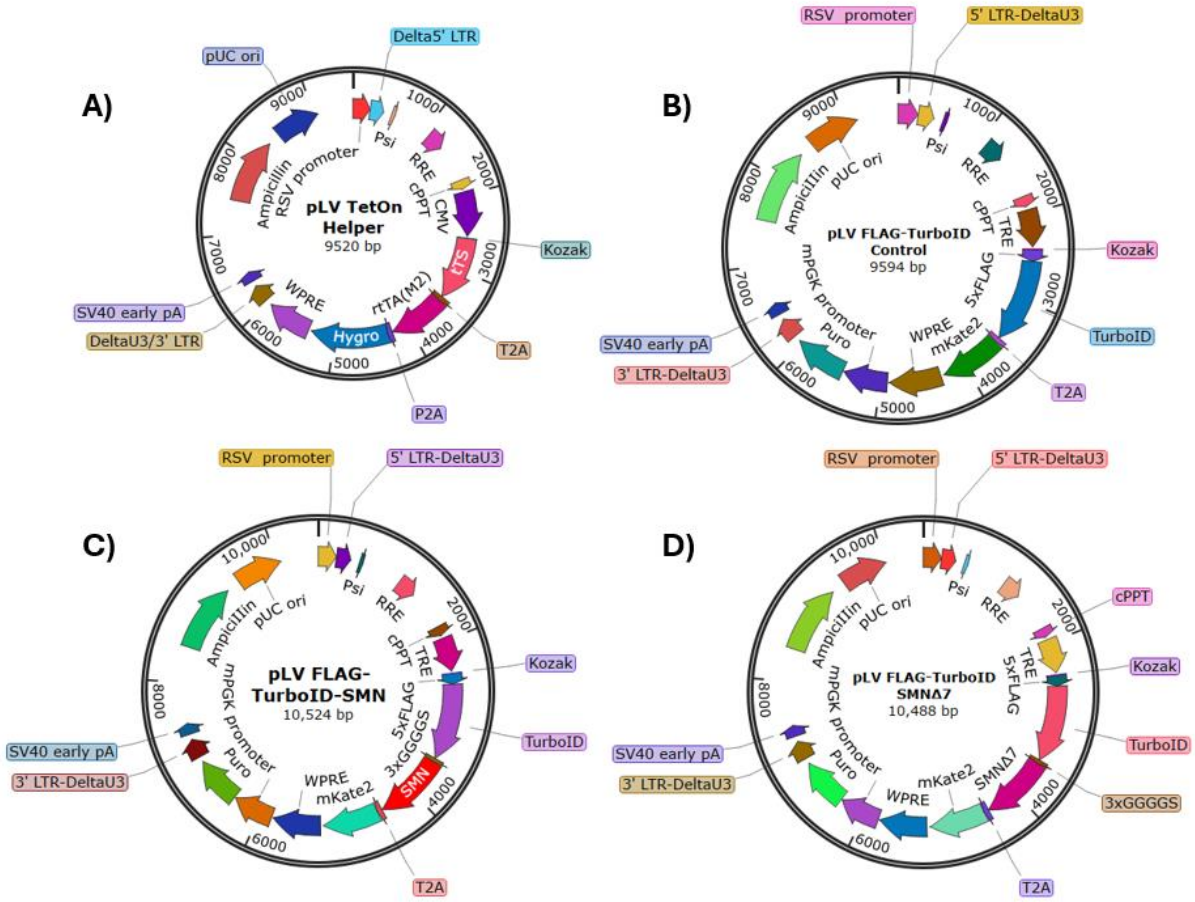


FIGURE 6.2.1: PLASMID MAPS OF THE 4 LENTIVIRAL (LV) TETON TURBOID PLASMIDS.

A) Map of the TetOn Helper plasmid. B) Map of LV Plasmid designed to express FLAG-tagged TurboID (control plasmid) C) Map of LV Plasmid designed to express FLAG-tagged TurboID-SMN. C) Map of LV Plasmid designed to express FLAG-tagged TurboID-SMNΔ7.

6.2.2 Generation of Stable Cell Lines Using the Lentiviral TurboID Plasmids

Initial Hela and SHSY5Y cells were transduced with the TetOn Helper Plasmid (Multiplicity of Infection (MOI) = 10), containing the Tet Repressor protein construct; these cells underwent

selection with hygromycin (300ug/ml (HeLa) or 350ug/ml (SHSY5Y)) and, once sufficient levels of this protein were identified on western blot, these parental cell lines were then used to generate the TurboID cell lines (Fig 6.2.2).

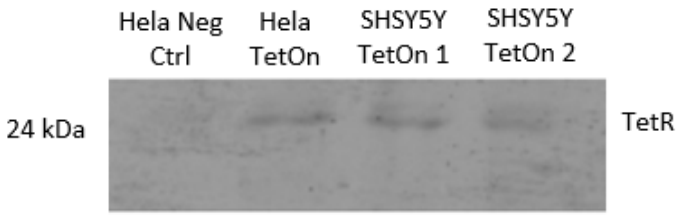


FIGURE 6.2.2: HELA AND SHSY5Y CELLS TRANSDUCED WITH THE TETON HELPER LV PLASMID EXPRESS THE TET REPRESSOR PROTEIN.

Western blot of lysates derived from an untransduced Hela cell line (acting as a negative control), a Hela TetOn cell line and two separate SHSY5Y TetOn cell lines. The blot was probed with mouse Anti-Tet antibody (Takara) to identify expression of the Tet Repressor protein in cells transduced with the LV TetOn Helper plasmid. Input = 40ug.

While expression levels were not particularly high, high concentrations of the repressor protein are not required for functionality of the TetOn system (Agha-Mohammadi et al., 2004). As they were of sufficient expression to be identified via western blot, sufficient levels of expression were assumed. The Hela TetOn and SHSY5Y TetOn 1 cell lines were then expanded, seeded into three 6cm plates and transduced (MOI = 10) with the FLAG-TurboID-Ctrl, FLAG-TurboID-SMN and FLAG-TurboID-SMNΔ7 LV plasmids respectively.

6.2.3 The Lentiviral TurboID Plasmids Express Correctly and Maintain Sufficient Expression Levels in Hela Cells

Post-Selection with Puromycin (400ng/ml), cells from each TurboID cell line were seeded onto coverslips and induced with 1ug/ml of Doxycycline for 48hrs prior to biotinylation. Optimisation of biotinylation was determined via dosing coverslips with varying

concentrations of biotin (50-250uM) and incubating for various timepoints (10 minutes, 20 minutes, 30 minutes and 1 hour). The dose/time which gave the clearest signal when staining with Streptavidin-489 (i.e. visible expression and localisation with minimal off-target background) was chosen for future experiments. From these data (unshown), 150uM of Biotin was added to media for 1 hour at 37 degrees Celsius, following which cells were fixed with 3.7% paraformaldehyde.

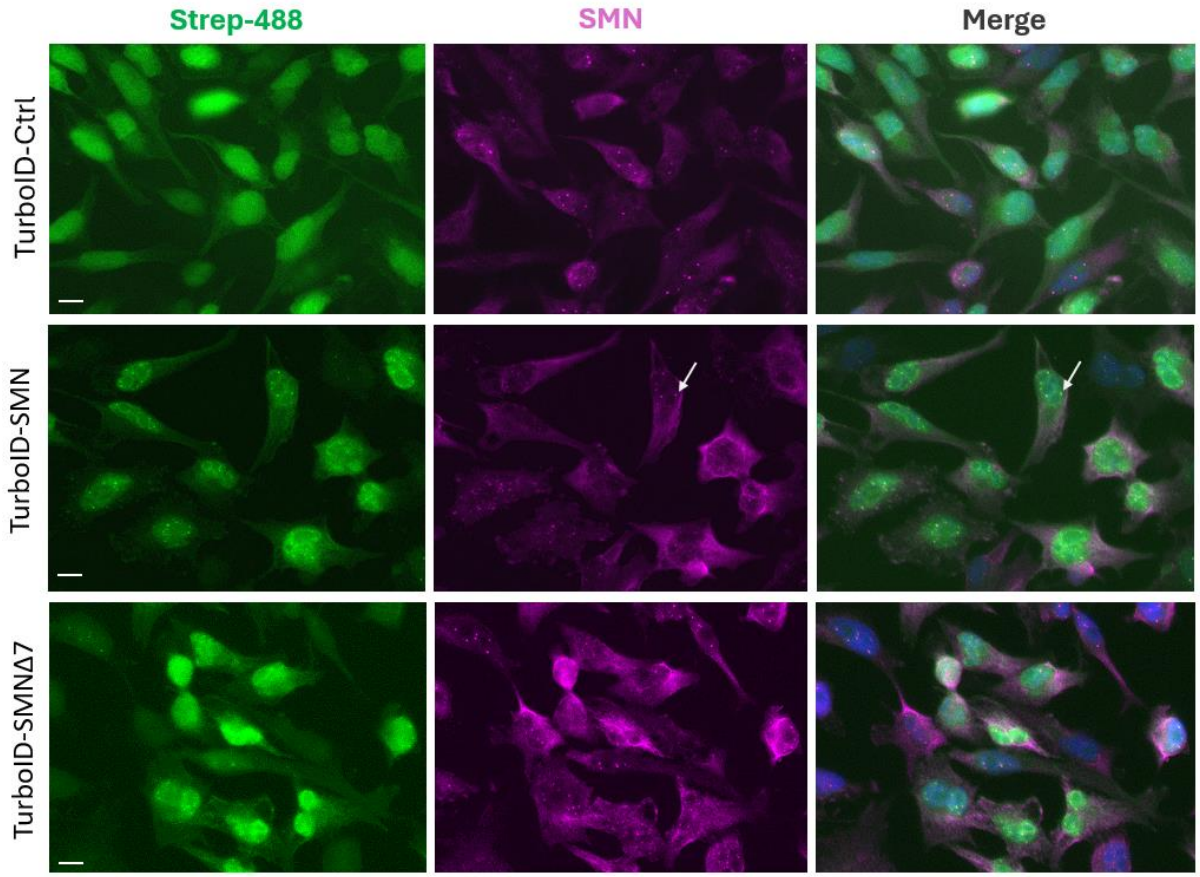


FIGURE 6.2.3: IMMUNOFLUORESCENT IMAGING OF INDUCED HELA TURBOID CELL LINES SHOWS BIOTINYLATION AND LOCALISATION PATTERNS CONSISTENT WITH SMN AND SMNΔ7.

Representative fluorescent imaging of 48hr induced (1ug/ml doxycycline) and 1 hr biotin incubation (150uM) TurboID Hela cell lines. TurboID-Control (top), TurboID-SMN (centre) and TurboID-SMNΔ7 (bottom) cells were stained with Streptavidin-488 (Licor) (Green) to indicate biotinylation and mouse anti-SMN (BD Biosciences) (magenta). Nuclei are counterstained with DAPI (blue). Scale = 10 microns. Images were obtained on the EVOS M5000 Imaging System at a magnification of 40x. White arrows indicate nuclear SMN-containing gems of Cajal Bodies.

Each of the Hela TurboID cell lines demonstrated sufficient levels of biotinylation in all three experimental conditions (TurboID-Control, TurboID-SMN and TurboID-SMNΔ7) when induced for 48 hours with 1ug/ml of doxycycline. Biotinylation with 150uM of biotin for 1 hour at 37 degrees Celsius gave a strong, specific biotinylation signal consistent the with signal expected for each cell lines' transduced construct. The biotinylation signal observed in the TurboID-Control cell line was consistent with a diffuse, non-specific cellular localisation. The TurboID-SMN line demonstrated biotinylation consistent with nuclear SMN gem staining (see white arrows in Figure 6.2.2, middle row) as well as a level of diffuse, cytoplasmic staining whereas the biotinylation present in the TurboID-SMNΔ7 line was primarily concentrated in the nucleus, consistent with expected SMNΔ7 expression.

Unfortunately, transduced SHSY5Y cell lines did not adequately express the constructs; while there was a minimal level of expression as denoted by mKate2 fluorescence, further subculture and expansion would be required. Due to time constraints, these cell lines were frozen at -80 and transferred to liquid nitrogen for future refinement.

6.3 Optimising Pulldown Conditions for Mass Spectrometry Analysis of TurboID SILAC Lysates

In order to compare the differential interactomes of proteins captured by FLAG Immunoprecipitation and Streptavidin Pulldown, each technique would require optimisation prior to their use in downstream quantitative proteomic analysis.

6.3.1 FLAG Immunoprecipitation

Various FLAG antibodies were utilised to immunoprecipitate the FLAG-TurboID proteins – including Sigma Mouse Anti-FLAG and Thermo Mouse Anti-FLAG - which demonstrated either inefficient pulldown or no pulldown capability at all. Immunoprecipitation reactions using 2ug/ml of Invitrogen Mouse Anti-FLAG antibody were successful (Figure 6.3.3.1). Utilisation of a low salt (150mM) RIPA buffer as both cell lysis buffer and wash buffer, and elution conditions optimal for the Protein A beads (boil at 65 degrees for 20 minutes in 4x LDS Sample Buffer) resulted in efficient capture of the TurboID proteins.

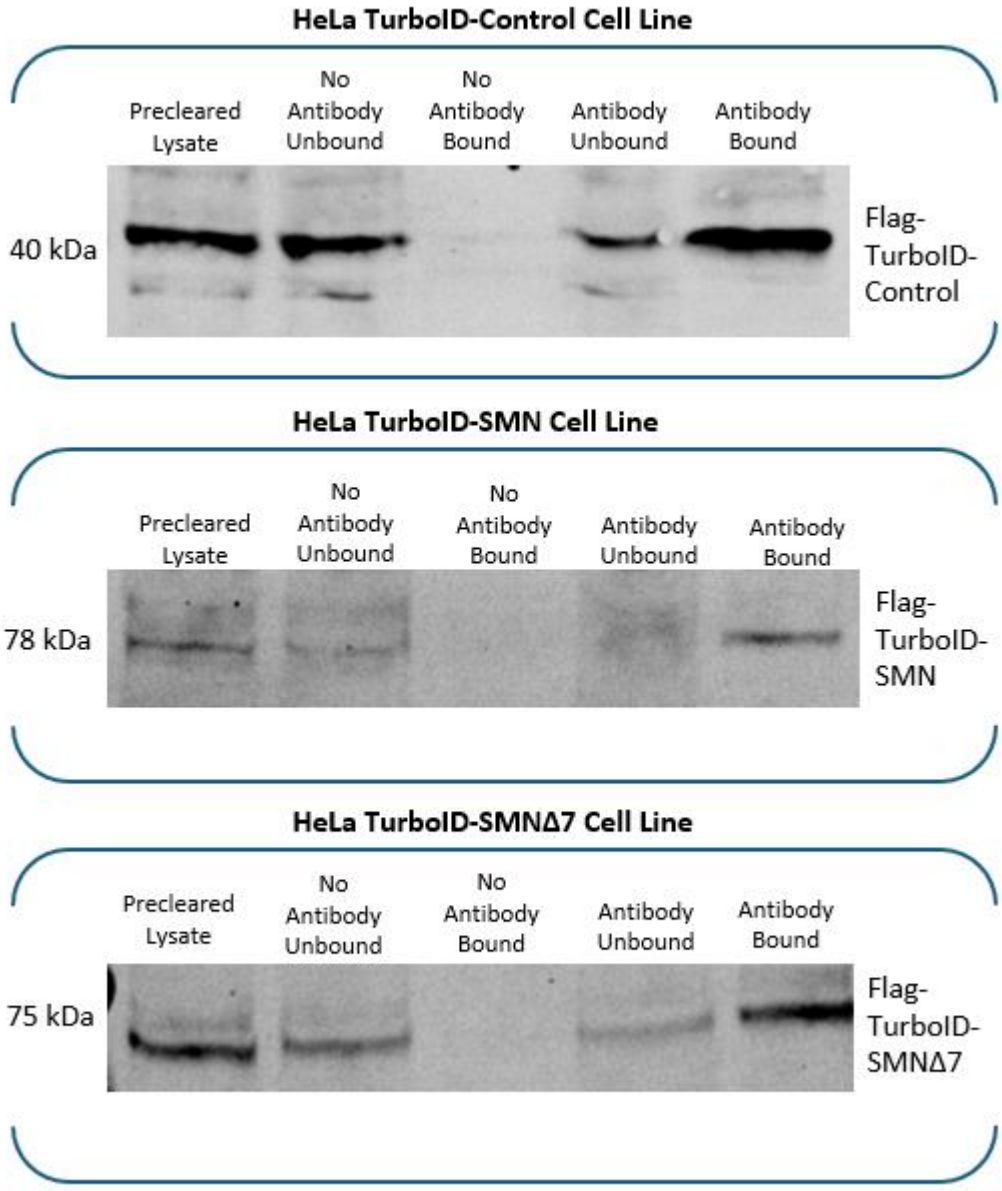


FIGURE 6.3.1: FLAG IMMUNOPRECIPITATION OF HELA FLAG-TURBOID-CONTROL (TOP), HELA FLAG-TURBOID-SMN (CENTRE) AND HELA FLAG-TURBOID-SMNΔ7 CELL LINES.

Immunoprecipitation was performed using overnight incubation (at 4 degrees Celsius) with Protein A beads and a mouse Anti-FLAG antibody (Invitrogen) and gentle rotation. Precleared lysates were incubated with Protein A beads for 30 minutes. The "No Antibody" lanes represent the negative controls i.e lysates were incubated with beads only. The "Antibody" lanes indicate that lysates were incubated with both beads and antibody. Starting input = 1mg of cell lysate. Blots were probed with a rabbit Anti-FLAG (Sigma) antibody.

6.3.2 Streptavidin Pulldown

Cell lysates were incubated overnight with 30ul of Pierce Streptavidin Magnetic Beads. After washing in Low Salt RIPA buffer, lysates were eluted from the beads utilising Elution buffer (30mM Biotin, 2% SDS) and 4x LDS sample buffer. This elution method, which had worked well in the FLAG IP, gave rise to few biotinylated proteins of molecular weight above 50kDa identifiable on a western blot probed with fluorescent streptavidin (Figure 6.3.2) Non-specific Streptavidin staining can be observed in both the "Input" and "Bead Unbound" lanes; these proteins, however, did not appear to bind to the Streptavidin Magnetic Beads.

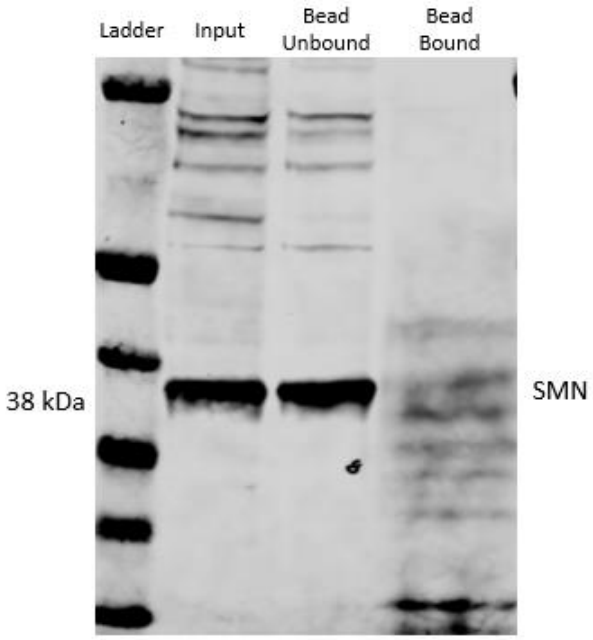


FIGURE 6.3.2: INDUCED AND BIOTINYLATED TURBOID CELL LYSATES EXHIBITED AN UNUSUAL PULLDOWN PATTERN AFTER STAINING WITH STREPTAVIDIN-680.

Cell lysates derived from 48hr dox-induced (1ug/ml) TurboID-Control Hela cell line. Incubated with 150uM of Biotin for 1hr at 37 degrees Celsius. Blots were probed with Streptavidin-680 (JIR) and mouse Anti-SMN (BD Sciences).

Several adaptations of the elution method were subsequently tested, in order to achieve a more “normalised” distribution of eluted biotinylated proteins (Figure 6.3.3). While this was achieved, there did appear to be some amount of protein degradation. The use of new buffers with new protease inhibitor tablets did not resolve this issue. Because of this, it was decided that prior to Mass Spectrometry, samples should not be eluted from the beads; instead, they would be digested on the beads using trypsin before being injected into the mass spectrometer (i.e. on-bead digestion).

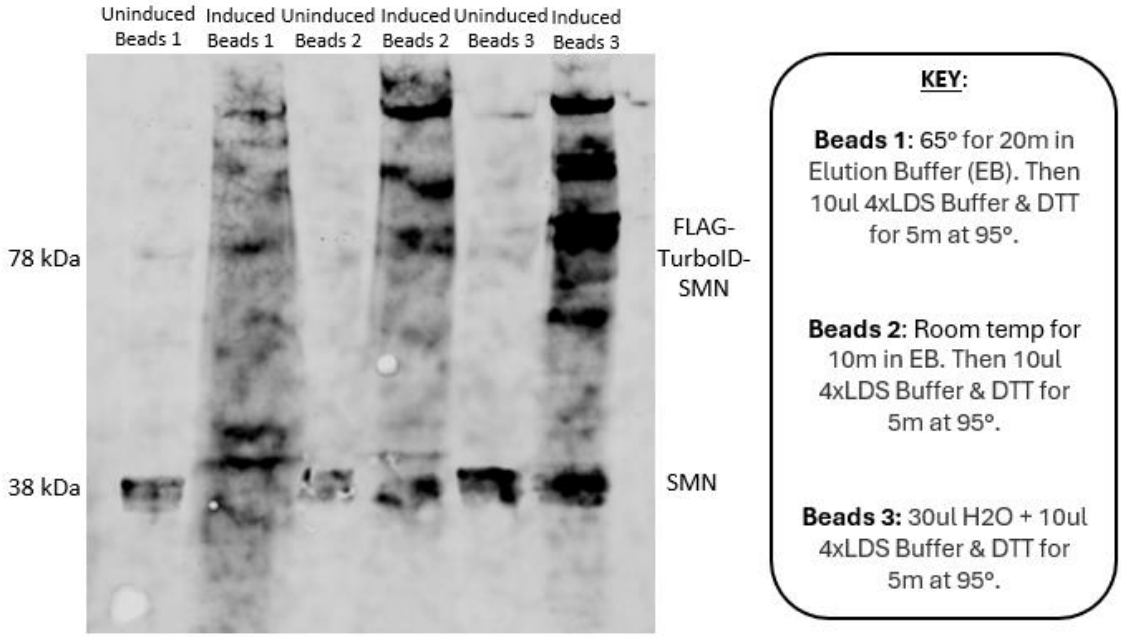


FIGURE 6.3.3: MODIFYING THE ELUTION CONDITIONS OF PIERCE MAGNETIC STREPTAVIDIN BEADS RESULTS IN A MORE NORMALISED SPREAD OF ELUTED BIOTINYLATED PROTEINS.

Induced (1ug/ml of Dox for 48hrs) and Uninduced (Dox negative) FLAG-TurboID-SMN Hela cells were biotinylated for 1 hr (150uM Biotin) at 37 degrees Celsius. Whole cell lysates underwent streptavidin pulldown using 30ul Pierce Magnetic Streptavidin beads as described and eluted with various adaptations to the elution protocol (see figure key for specifics). Blot was probed with Streptavidin-680 (JIR) to detect biotinylated proteins and mouse Anti-SMN (BD Sciences).

In preparation for the Pulldowns and subsequent On-Bead Digestion, the Hela FLAG-TurboID-Ctrl, FLAG-TurboID-SMN and FLAG-TurboID-SMNΔ7 cell lines were cultured in SILAC media supplemented with 10% dialysed FBS according to the labelling scheme in Table 6.3, for a total of 4 passages.

Cell Type	Experimental Condition	Replicate Number	SILAC Label
Hela	FLAG-TurboID-Ctrl	1	Light
	FLAG-TurboID-SMN	1	Medium (Arg6 & Lys4)
	FLAG-TurboID-SMNΔ7	1	Heavy (Arg10 & Lys8)
	FLAG-TurboID-Ctrl	2	Heavy (Arg10 & Lys8)
	FLAG-TurboID-SMN	2	Light
	FLAG-TurboID-SMNΔ7	2	Medium (Arg6 & Lys4)

TABLE 6.3: SILAC LABELLING SCHEME.

Each cell line was labelled with a different SILAC label and these labels were switched between sets of biological replicates (N=2).

Cells were induced with 1ug/ml of Doxycyline for 48hrs then biotinylated with addition of 150mM Biotin to cell culture media for 1 hr at 37 degrees Celsius. Pellets were subsequently washed, resuspended and lysed in Low Salt RIPA buffer then clarified via high speed centrifugation (16000 RCF), with half of each lysate used for the FLAG IP with the other half used for the Streptavidin Pulldown, after calculation of protein concentration via Bradford assay.

FLAG-IP and Streptavidin Pulldown were performed concurrently as described previously. 30ul of Magnetic Streptavidin beads/Protein A beads from each pulldown in each SILAC condition (Heavy/Medium/Light) were combined after the final wash step to produce the

final sample for mass spectrometry analysis. Prior to delivery to Mass Spectrometry, each combined sample was washed 5 times with 10% ammonium bicarbonate. On-Bead digestion was performed using an overnight trypsin digestion prior to injection on the mass spectrometer.

Chapter 7: Data Analysis of FLAG Immunoprecipitation and Streptavidin Pulldown TurboID SILAC Datasets

7.1 Streptavidin Pulldown is a More Effective Method for Interactome Production than FLAG Immunoprecipitation

Following MS/MS analysis, raw data files were processed using Maxquant and analysis was performed on the ProteinGroups.txt output files using Microsoft Excel.

In short, proteins marked as contaminants and those identified by reverse ID only (the Maxquant equivalent of Decoy proteins) were removed resulting in the initial protein shortlist. Next, proteins for which there were no intensity values in all three SILAC fields (Light/Medium/Heavy) were also removed. Finally, SILAC ratios were combined between replicates (N=2) and the average value of these were used as the final SILAC ratios. Statistical analysis was performed as described using an unpaired student t-test on the final ratio arrays to find proteins with statistically significant differences in interactivity between each experimental condition (Aguilan et al., 2020) (Figure 7.1.1).

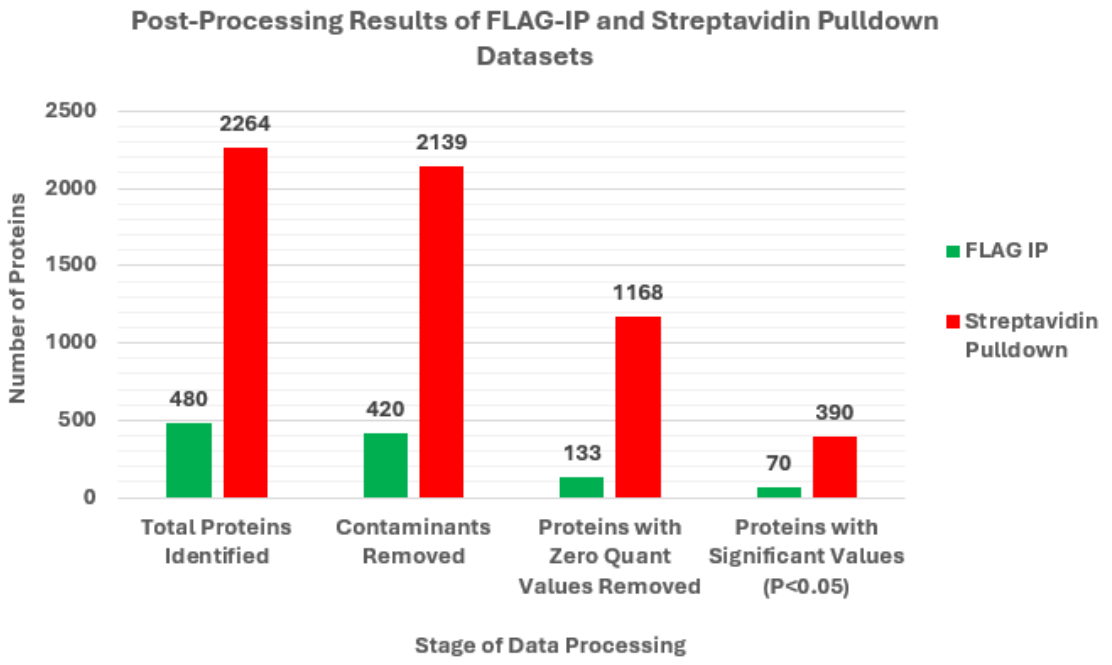


FIGURE 7.1.1: BOTH FLAG-IP AND STREPTAVIDIN PULLDOWN DATASETS CONTAIN PROTEINS WITH STATISTICALLY SIGNIFICANT CHANGES IN INTERACTIVITY BETWEEN CONDITIONS.

FLAG IP (green bars) and Streptavidin Pulldown (red bars) are displayed with the number of proteins remaining after each stage of data processing above the bar.

76.6% of the proteins identified in the FLAG IP were also present in the Streptavidin Pulldown. The Streptavidin Pulldown, however, resulted in over 5 times more proteins identified after contaminant removal compared to the FLAG IP (420 proteins in FLAG IP vs 2139 in Streptavidin Pulldown).

Due to the removal of any identified proteins with zero quantitative values (i.e a zero SILAC intensity or a NaN value) the remaining proteins in both the FLAG IP and Streptavidin Pulldown displayed a similar distribution, with most being shared between all three experimental conditions (Figure 7.1.2). This allowed for the calculation of fold changes in interactivity between all remaining proteins in each TurboID cell line.

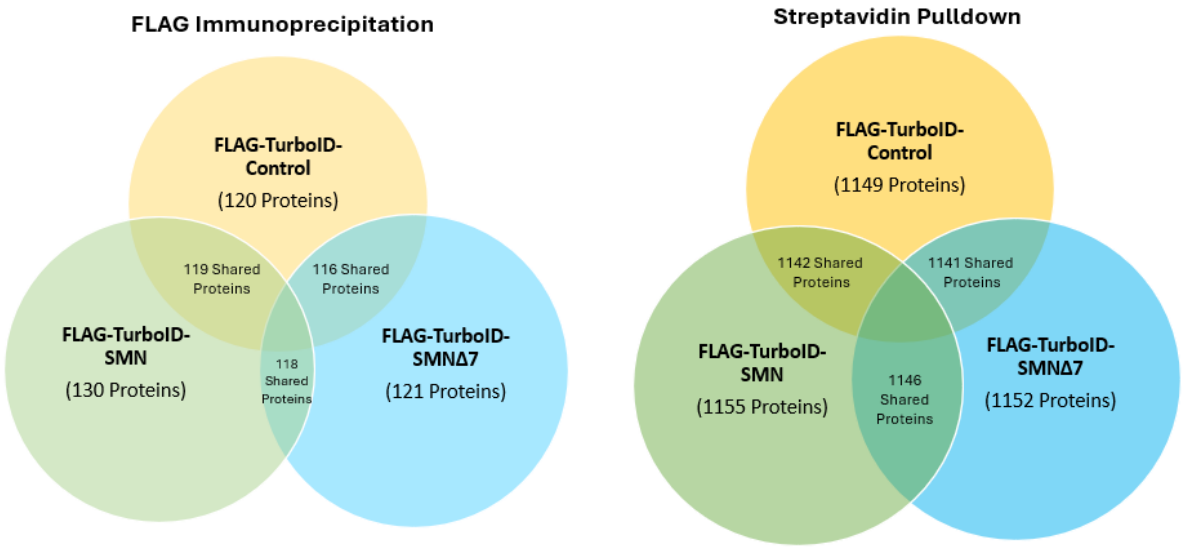


FIGURE 7.1.2: FLAG IP AND STREPTAVIDIN PULLDOWN DEMONSTRATE A SIMILAR PROTEIN DISTRIBUTION BETWEEN EXPERIMENTAL CONDITIONS.

Venn diagram overview of the final quantitative protein content of both the FLAG Immunoprecipitation dataset (left) and the Streptavidin Pulldown dataset (right). The number of proteins in common between experimental conditions are noted in the areas of relevant overlap. Diagrams are not drawn to scale.

Given the difference in number of initially identified proteins (480 in FLAG IP vs 2264 in Streptavidin Pulldown) and the fact that three quarters of the proteins identified in the smaller FLAG IP dataset were also present in the Streptavidin Pulldown dataset, it was clear that the Streptavidin Pulldown is a more effective technique to probe interactomic variations between experimental conditions than FLAG IP alone.

7.2 The TurboID-SMNΔ7 Interactome Displays Highly Decreased Interactivity with the Proteins of the TurboID-SMN Interactome

Heatmaps were produced to gain a visual overview of the interaction changes in proteins with significantly altered interactivity ($P<0.05$) (Figure 7.2.1). Both the FLAG IP and

Streptavidin Pulldown interactors show very similar patterns of differential interactivity across the three Fold Change ratios. In both datasets, TurboID-SMNΔ7 shows a greatly decreased level of interaction with other protein interactors of TurboID-SMN.

Both TurboID-SMNΔ7 and TurboID-SMN show a generalised increase in the level of interactivity with the identified proteins across both FLAG IP and Streptavidin Pulldown datasets, in comparison to TurboID-Control. As the TurboID-Control protein is expressed at baseline throughout the cell with no specific localisation, promiscuous biotinylation of background proteins is to be expected. The increased level of interactivity with these proteins in both TurboID-SMN and TurboID-SMNΔ7 would be expected in comparison to the baseline TurboID-Control interactors.

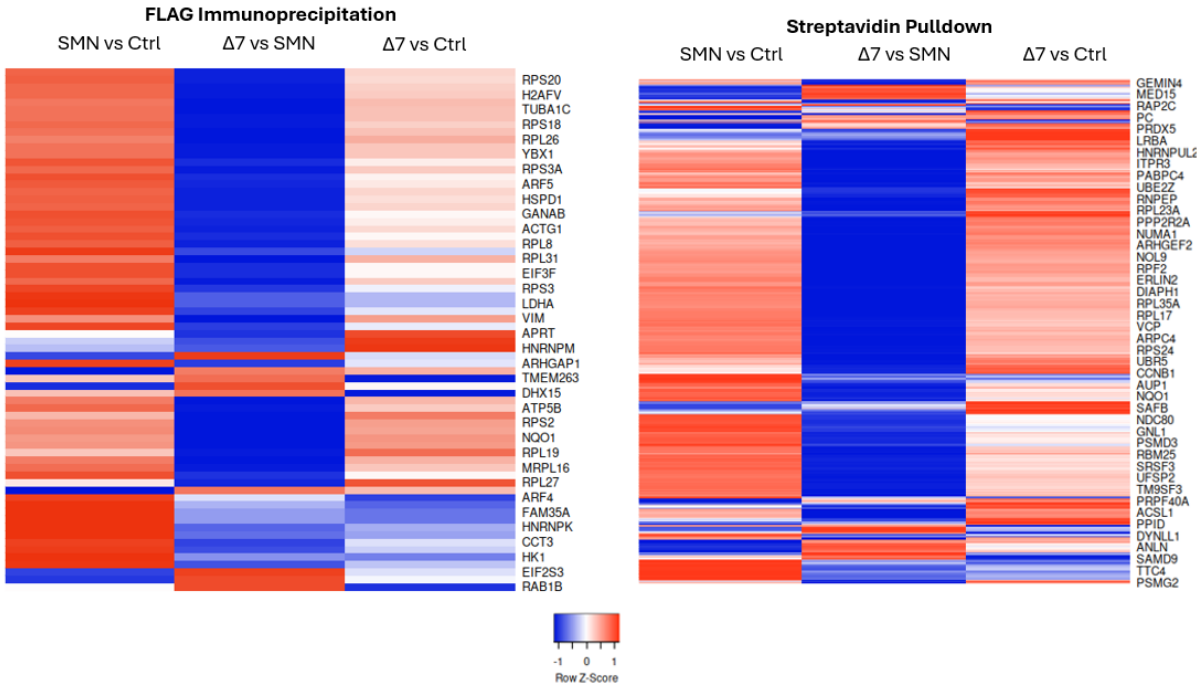


FIGURE 7.2.1: FLAG IP AND STREPTAVIDIN PULLDOWN DATASETS SHOW A SIMILAR OVERALL PATTERN OF INTERACTIVITY CHANGES IN EACH EXPERIMENTAL CONDITION.

Heatmaps comparing the fold change interactivity ratios between SMN v Ctrl, Δ7 v SMN and Δ7 v Ctrl in both the FLAG Immunoprecipitation dataset (left) and the Streptavidin Pulldown dataset (right). Each row represents a protein with a statistically significant change in interactivity profile. Heatmaps were scaled to the row values in order to show relationships between protein profiles across conditions more distinctly, resulting in Z-Scores as defined by the key displayed at the bottom of the figure. Blue shades indicate a decreased interaction between proteins, whereas red shades indicate increased levels of interaction. Each column represents a SILAC ratio, equivalent to a fold change for each protein in each row. Heatmaps were clustered according to the row (i.e. protein fold change) using the Euclidean algorithm.

While the heatmaps contained only significantly altered proteins from each experimental condition (P Value <0.05), the general trends remain consistent with regards to differential interactivity patterns across the FLAG IP and Streptavidin Pulldown datasets.

To more closely examine the significant interaction patterns of the proteins differentially interacting with TurboID-SMNΔ7 and TurboID-SMN, volcano plots were created, utilising the fold change and statistical significance data obtained from the FLAG IP and Streptavidin Pulldown datasets (Figure 7.2.2).

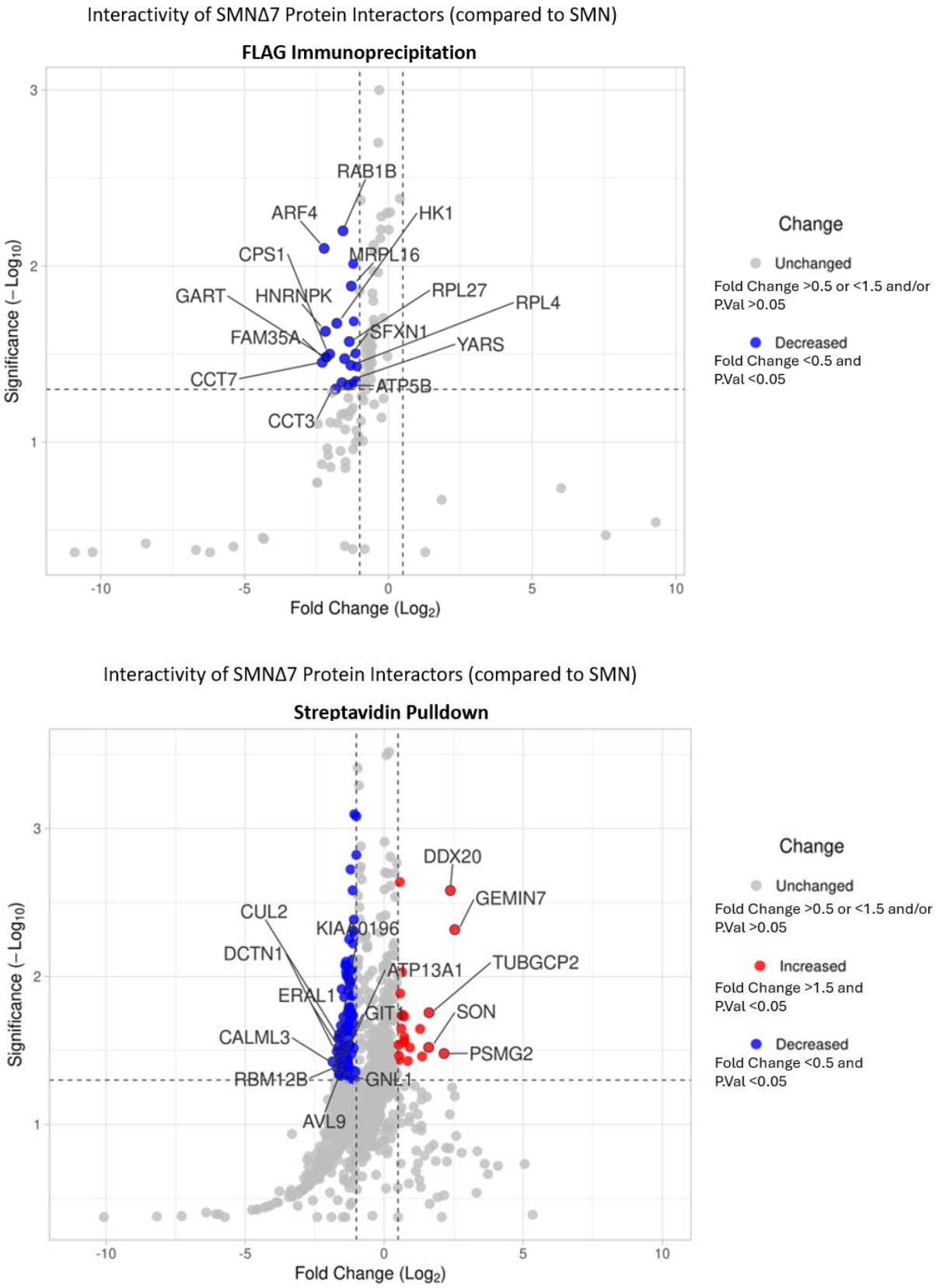


FIGURE 7.2.2: THE STREPTAVIDIN PULLDOWN DATA DEMONSTRATES BETTER INTERACTOME COVERAGE THAN FLAG IP.

Volcano plot showing differentially regulated proteins of interest in the Δ7 v SMN condition. Both the FLAG Immunoprecipitation data (top) and the Streptavidin Pulldown data (bottom) are displayed. Fold change (FC) cutoffs were set to show proteins with only FCs above 0.5 and below -1 (the log2 equivalent of an FC <0.5 and >1.5) The significance cutoff was set to 1.3 (the log10 equivalent of a 0.05 cutoff). The top 15 hits - ranked by the largest FC in either direction - were labelled with their gene name. Directionality of FC indicated by the colours defined in the key (found to the right of plots).

Both datasets display a similar profile insofar as that the majority of significantly interacting proteins do not reach the cut off value (<0.5 or >1.5 Fold Change) necessary for highlighting. Interestingly, the FLAG IP data does not contain any proteins with increased interactivity which fit the parameters for inclusion in the graph; only decreased interactors attain the level of significance and meet the fold change cutoff.

The Streptavidin Pulldown dataset, on the other hand, reveals a much wider scope of interactors – both increased and decreased in the TurboID-SMNΔ7 condition compared to TurboID-SMN.

7.3 The Identified TurboID-SMNΔ7 Interactomes are Enriched in Identified Protein-Protein Interactions

In order to probe the interactomes of the differentially interacting proteins between TurboID-SMNΔ7 and TurboID-SMN, relevant lists of proteins from both the FLAG IP and Streptavidin Pulldown datasets were uploaded to String-DB. Utilising the SILAC ratio comparing TurboID-SMNΔ7 to TurboID-SMN (Δ7/SMN), the selected proteins comprised

those which met the stringency cutoffs (<0.5 FC >1.5 and P.Value <0.05). These expanded protein lists can be found in the Supplementary Data section (S2).

Results of the interactome analysis of the SMNΔ7 Significantly Decreased FLAG IP interactors (21 proteins, FC <0.5 and P Value <0.05) revealed a high level of existing interactions identified between the input proteins (Figure 7.3.1). This suggests biological relevance and, as these proteins were pulled down using FLAG IP, can be considered stable interactors of the SMN protein, displaying decreased interactivity with SMNΔ7. PPI Enrichment scores indicate whether a set of input proteins are biologically connected, calculated by comparing the number of edges present in the input protein list to what would be expected given a set of random proteins the same size as the input protein list length. Smaller PPI P-Values indicate biological connection. In this case, A PPI enrichment P-Value assigned by STRING-DB to the protein list (P-Val=2.82e-13), further confirmed significant biological connection as a group of proteins.

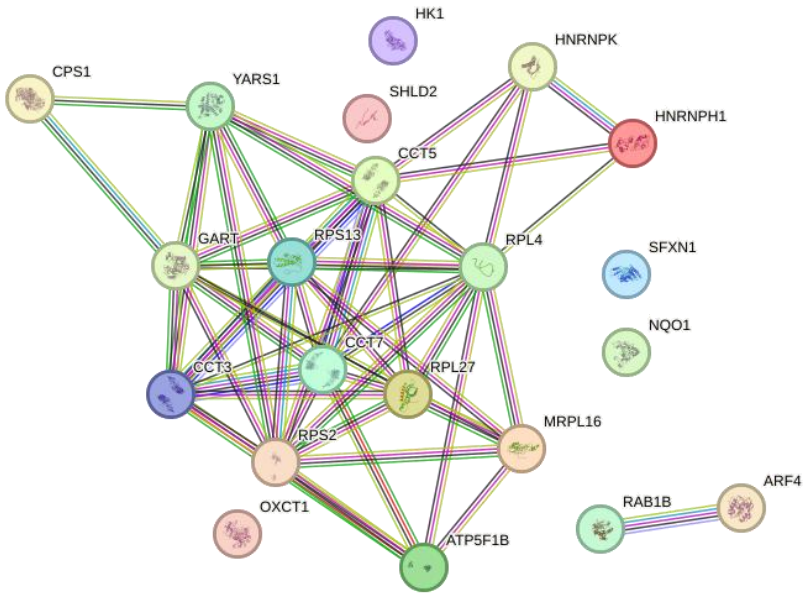


FIGURE 7.3.1: THE TURBOID-SMNΔ7 DECREASED INTERACTIVITY PROTEIN INTERACTION NETWORK FROM THE FLAG IP DATASET DEMONSTRATES HIGH LEVELS OF INTERCONNECTIVITY.

Edges (shown here as coloured lines) between the 21 node proteins (shown as coloured circles) represent existing lines of evidence of protein-protein interaction.

The Increased SMNΔ7 Interaction Protein Data from the Streptavidin Pulldown contained 17 proteins with an FC > 1.5 and a P Value < 0.05 (Figure 7.3.2). Fewer existing interactions between proteins were identified than in the downregulated datasets. As the Streptavidin Pulldown will pull down more transient interactors of SMN and SMNΔ7, these proteins are of particular interest because they represent novel, previously unidentified interactors. The PPI enrichment P-Value assigned by STRING-DB to the protein list (P-Val= 0.283) did not indicate significant biological enrichment. However, given the smaller size of the protein list as well as the understudied nature of SMNΔ7 interactions, these proteins may purely represent novel, unidentified interactions unknown to String-DB. There is still identifiable and significant biological relevance however, as covered later in Section 7.4.

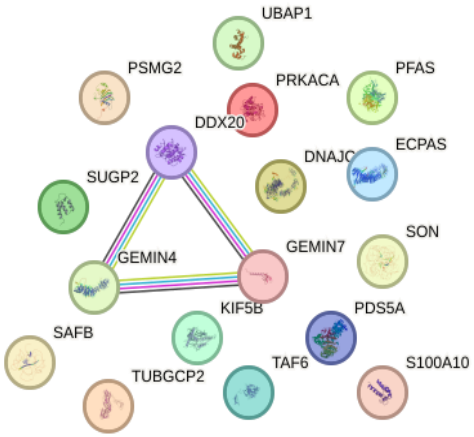


FIGURE 7.3.2: THE TURBOID-SMNΔ7 INCREASED INTERACTIVITY PROTEIN INTERACTION NETWORK FROM THE STREPTAVIDIN PULLDOWN DATASET DEMONSTRATES LOW LEVELS OF INTERCONNECTIVITY.

Edges (shown here as coloured lines) between node 17 node proteins (shown as coloured circles) represent existing evidence of protein-protein interaction.

The Decreased SMNΔ7 Interaction Protein Data from the Streptavidin Pulldown contained 91 proteins with an FC < 0.5 and a P Value < 0.05 (Figure 7.3.3). This dataset - like the FLAG IP Decreased Interaction dataset - displays high levels of existing protein interaction between input proteins. However, like the Increased Interaction Dataset from the same Streptavidin Pulldown, several isolated node proteins represent novel interactors with no existing known links to the other proteins. The PPI enrichment P-Value (P-Val= 1.57e-09) further implied these proteins were biologically connected beyond chance.

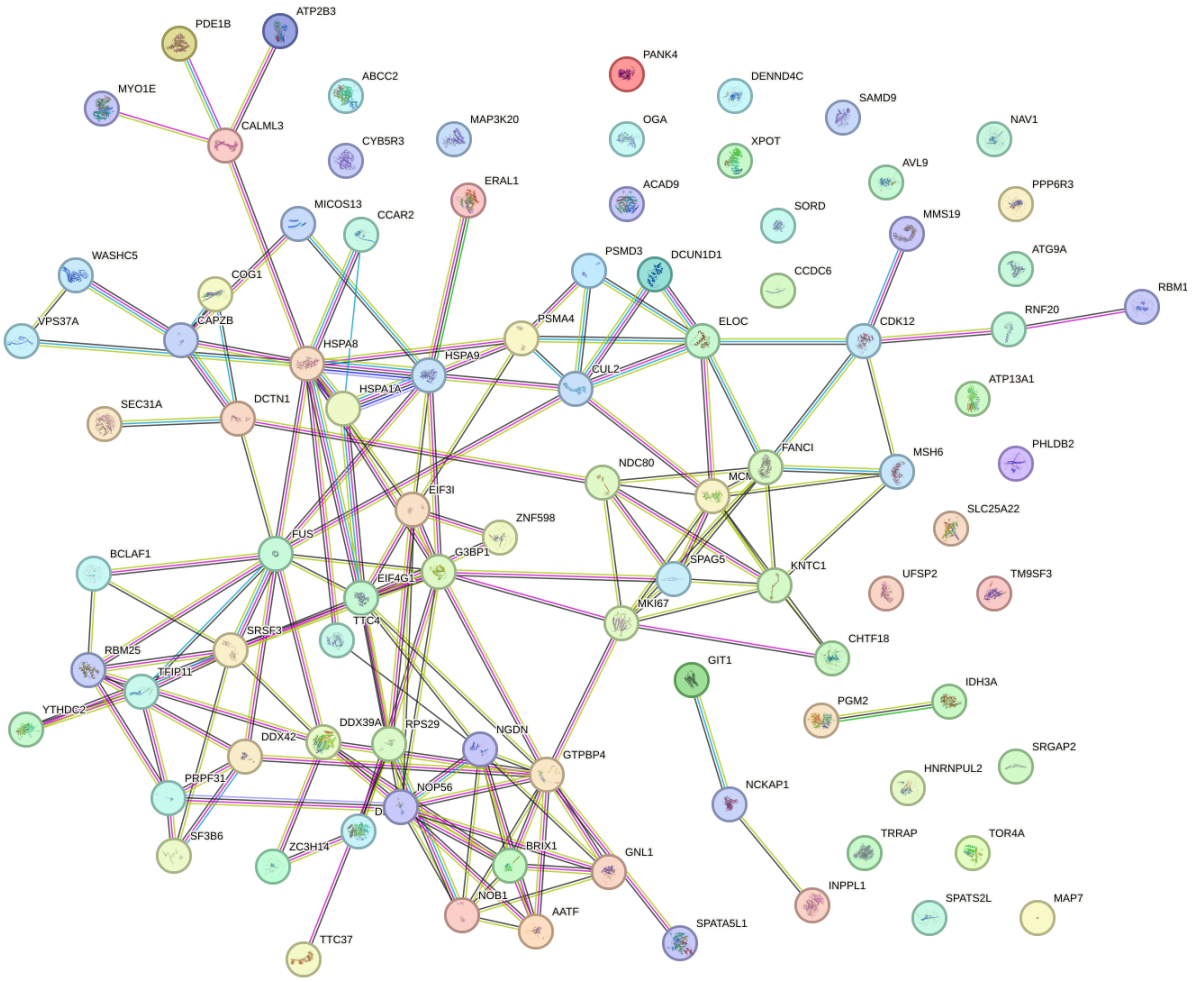


FIGURE 7.3.3: THE TURBOID-SMNΔ7 DECREASED INTERACTIVITY PROTEIN INTERACTION NETWORK FROM THE STREPTAVIDIN PULLDOWN DATASET IS HIGHLY INTERCONNECTED.

Edges (shown here as coloured lines) between the 91 node proteins (shown as coloured circles) represent existing evidence of protein-protein interaction.

There was little protein overlap between the FLAG IP SMNΔ7 Significantly Decreased Proteins and the Streptavidin Pulldown SMNΔ7 Significantly Decreased Proteins; 4 of the FLAG IP Significantly Decreased proteins (HNRNPK, HNRNPH1, NQO1 and RPL4) were also present in the Streptavidin Decreased dataset, however these proteins did not meet the stringency Fold Change cutoffs. This difference between datasets can be attributed to the

much larger initial number of proteins pulled down in the Streptavidin Data as opposed to the smaller number of proteins identified in the FLAG IP dataset.

However, the much larger number of significantly downregulated proteins across both FLAG IP and Streptavidin Pulldown-derived SMNΔ7 Interactomes supports the idea that the loss of the SMNΔ7 protein isoform's functionality is due to a loss of, or decrease in, protein interactivity.

7.4 The TurboID-SMNΔ7 Interactome Is Enriched in Proteins Involved in Protein and RNA Metabolism

In order to assign a biological context to the above significant protein shortlists, Gene Ontology analysis was performed utilising String-DB (Table 7.4).

#Term ID	Term Description	Network Strength	False Discovery Rate
GO Component	**FLAG IP – Proteins with DECREASED Interactivity in TurboID-SMNΔ7**		
GO:0005832	Chaperonin-containing T-complex	2.41	0.0008
GO:0022626	Cytosolic ribosome	1.53	0.0018
GO:0044391	Ribosomal subunit	1.36	0.00096
GO:0015934	Large ribosomal subunit	1.36	0.0412
GO:0005925	Focal adhesion	1.01	0.0181
GO Process			
GO:1904851	Positive regulation of establishment of protein localization to telomere	2.41	0.0062
GO:1904871	Positive regulation of protein localization to Cajal body	2.37	0.0062
GO:1904874	Positive regulation of telomerase RNA localization to Cajal body	2.23	0.0062
GO:0032212	Positive regulation of telomere maintenance via telomerase	1.88	0.011
GO:0007339	Binding of sperm to zona pellucida	1.82	0.0122
GO Function			
GO:0140662	ATP-dependent protein folding chaperone	1.92	0.008
GO:0003735	Structural constituent of ribosome	1.4	0.0018
GO:0016874	Ligase activity	1.31	0.0195
GO:0003723	RNA binding	0.82	3.08e-05

GO:0035639	Purine ribonucleoside triphosphate binding	0.67	0.015
GO Component	**Streptavidin Pulldown – Proteins with INCREASED Interactivity in TurboID-SMNΔ7**		
GO:0032797	SMN complex	2.54	0.00031
GO:0097504	Gemini of coiled bodies	2.46	0.00031
GO:0016604	Nuclear body	0.92	0.0211
GO:0031981	Nuclear lumen	0.49	0.0211
GO:0005654	Nucleoplasm	0.49	0.0385
GO Process			
GO:0000387	Spliceosomal snRNP assembly	1.95	0.0138
GO:0071826	Ribonucleoprotein complex subunit organization	1.34	0.035
GO:0008380	RNA splicing	1.19	0.0216
GO:0016071	mRNA metabolic process	1.18	0.00027
GO:0006397	mRNA processing	1.18	0.0049
GO Compartment			
GOCC:0032797	SMN complex	2.54	0.00035
GOCC:0097504	Gemini of coiled bodies	2.46	0.00035
GOCC:0016604	Nuclear body	1.18	0.0082
GOCC:0043227	Membrane-bounded organelle	0.31	0.0186
GOCC:0005829	Cytosol	0.58	0.021
GO Component	**Streptavidin Pulldown – Proteins with DECREASED Interactivity in TurboID-SMNΔ7**		
GO:0015030	Cajal body	1.15	0.0239
GO:0035770	Ribonucleoprotein granule	0.75	0.0296
GO:0016607	Nuclear speck	0.74	0.00064
GO:1990904	Ribonucleoprotein complex	0.64	0.00058
GO:0016604	Nuclear body	0.58	0.00099
GO Process			
GO:0090063	Positive regulation of microtubule nucleation	1.9	0.0133
GO:0043484	Regulation of RNA splicing	0.97	0.0047
GO:0071826	Ribonucleoprotein complex subunit organization	0.9	0.0083
GO:0008380	RNA splicing	0.87	0.00044
GO:0022618	Ribonucleoprotein complex assembly	0.86	0.0362
GO Function			
GO:0003724	RNA helicase activity	1.14	0.0085
GO:0016887	ATP hydrolysis activity	0.98	1.96e-07
GO:0004386	Helicase activity	0.93	0.018
GO:0140657	ATP-dependent activity	0.82	1.45e-06
GO:0017111	Nucleoside-triphosphatase activity	0.74	1.49e-05

TABLE 7.4: GO ANALYSIS RESULTS OF THE TURBOID-SMNΔ7 SIGNIFICANTLY ENRICHED INTERACTORS IN BOTH FLAG IP AND STREPTAVIDIN PULLDOWN DATASETS.

Shortlist derived from SILAC quantitative values (the Δ7/SMN Fold Change Ratio) with an <0.5 FC >1.5 and P<0.05. Increased interactivity (FC>1.5) protein analysis is in orange, whilst decreased interactivity (FC<0.5) protein analyses are in blue. Results are sorted by the network strength in 3

sub-categories: Biological Process, Molecular Function and Cellular Component/Compartment. FDR<0.05.

In the FLAG IP SMNΔ7 Decreased Interactor GO analysis, there is an over-representation of terms related to protein synthesis and folding, as well as protein localisation and regulation. This decreased interaction with essential protein metabolic machinery may highlight the inability of SMNΔ7 to recapitulate SMN's recognised role in protein trafficking and ribosome maintenance.

In the Streptavidin Pulldown SMNΔ7 Increased Interactor GO analysis, RNA metabolism and processing are highlighted, including processes such as splicing which are known to be affected in SMA. There is an over-representation of terms related to the nucleus, consistent with SMNΔ7's primarily nuclear localisation. The formation of complexes – including the SMN complex and RNP complexes – are canonically associated with the SMN protein, and the overabundance of several of these protein complex components in association with SMNΔ7 may suggest a compensatory mechanism for the isoform's inability to fully integrate into these complexes. SMNΔ7 still contains many of the functional domains inherent to SMN and the canonical binding of interactors - such as the Gemin proteins - may lead to the truncated isoform exhibiting a sequestering effect; it can bind to classical interactors, however being unable to efficiently exit the nucleus, it cannot partake in the same canonical pathways as the full-length protein.

In the Streptavidin Pulldown SMNΔ7 Decreased Interactor GO analysis, RNA processing and organization are highlighted along with an overrepresentation of terms related to RNA splicing which, whilst SMN plays an indirect role in the process, is commonly considered SMN's canonical function. Furthermore, several terms relating to enzymatic activities - such as "RNA helicase activity" - highlight the enzymatic functional interactors decreased in the SMNΔ7 condition, several of which, including the aforementioned RNA helicase, are also crucial for RNA metabolism and processing.

These results align with the qualitative data analysis performed on the GFP-Trap datasets (see Chapter 3). One difference between that dataset and the current TurboID datasets is the lack of unique proteins present between experimental conditions in the above TurboID data. This is most likely attributable to the difference in Mass Spectrometry equipment utilised in the GFP-Trap data acquisition, as well as the software used in raw data analysis and the non-quantitative nature of the GFP-Trap input samples. The removal of several identified proteins in the current datasets due to lack of data values (NaN values – Not a Number) most likely limited the identification of proteins unique to each cell line. However, the qualitative values remaining for all experimental conditions allowed for a precise comparative analysis of the remaining protein interactors.

7.5 Gait Complex Proteins and Amino Acid tRNA Synthetases Are Represented in the TurboID Datasets

All four protein components of the GAIT complex (GAPDH, SYNCRIP, RPL13 and EPRS) were identified in the Streptavidin Pulldown as decreased interactions in the SMNΔ7 cell line. None of these were significant, however.

The FLAG IP, on the other hand, only identified GAPDH and RPL13. While RPL13 did not have quantitative values, GAPDH was significantly decreased in this dataset (P Value = 0.02) however the effect was above the stringency cutoff (0.7 FC versus 0.5 FC cutoff)

The Streptavidin Pulldown also identified 16 Amino Acid tRNA Synthetase proteins – including RARS and IARS - all of which were downregulated in the SMNΔ7 cell line, except IARS which was notably unchanged in comparison to the SMN line. None of these were statistically significant apart from the Asparagine-tRNA synthase (NARS) (P Value = 0.03) however the effect was minor (FC=0.8)

Eight of these AARSs were also identified in the FLAG IP dataset, however six had no quantitative values and one - Phenylalanine-tRNA Synthetase (FARSA) - did not meet the FC stringency cutoffs whilst still being significantly decreased in the SMNΔ7 line (P value = 0.02). The Tyrosine-tRNA Synthetase (YARS) was both significantly decreased (P Value = 0.04) and met the stringency cutoff (FC of 0.34). It is worth noting that mutations in the YARS gene are associated with Dominant Intermediate Charcot-Marie-Tooth (DI-CMT)

disease, a neuropathy associated with both demyelination, reduced nerve impulse conduction and axonal degeneration (Hyun et al., 2014).

7.6 Discussion and Avenues of Future Research

An interesting aspect of both the FLAG IP and Streptavidin Pulldown datasets are the significantly enriched SMNΔ7 interactors which *do not* meet the Fold Change Cutoff as required by stringency i.e. those which exhibit a fold change value above 0.5 or below 1.5. In particular, the Streptavidin Pulldown data shows numerous interactors which fall into this category (See Figure 7.2.2, the centre grey "Unchanged" data points between the increased (red) and decreased (blue) data). This significant enrichment of "minor" fluctuations is of interest, because SMN and SMNΔ7 are closely related isoforms, sharing over 90% structural homology.

While a loss of binding to canonical SMN interactors by SMNΔ7 is assumed to be the primary causative factor for its inability to recapitulate the function of SMN in SMA patients, these data suggest that most of the interactions in the SMNΔ7 isoform are not completely lost but simply decreased. There is a significant decrease observed in the interactivity between SMNΔ7 and several proteins (as illustrated most clearly in the Streptavidin Pulldown dataset), suggesting that it is not a total loss of interactivity that hinders the functionality of SMNΔ7. Although it must be reiterated that, due to the lack of unique interactors identified in the datasets, this conclusion is not definitive. However, loss of

functionality in SMNΔ7 may be due to a shift in the isoform's axis of interactivity, with lower baseline interactions with several SMN specific interactors and, interestingly, increased interaction with others such as Gemins 4 and 7, two proteins which are found in both the nucleus and the cytoplasm (Cauchi, 2010). These - along with DDX20 - are protein components of the SMN complex, integral to the formation of the core of snRNPs in the cytoplasm. The role of these proteins in the nucleus is less clear, however their increased interaction with SMNΔ7 may suggest that SMNΔ7 is still incorporated into the SMN complex, most likely due to its ability to dimerise – albeit less effectively - with native, full-length SMN.

Given the mechanism of action of TurboID as a biotin ligase, however, it is important to note that there must be a measure of discretion when interpreting results; the biotinylation and enrichment of proteins via pulldown comprise interactors present in a 10-nanometre radius of the bait protein. While the proteins may be in the right place at the right time, it does not necessitate that there is a direct interaction between the two. Hence it is important to apply biological context and previous knowledge when interpreting TurboID results datasets.

In this case, the Gene Ontology results clearly show that this shift in interactivity has significant effects on the isoform's ability to take part in the canonical functions of SMN; in particular, associations with both RNA and protein metabolic functions were significantly reduced in the SMNΔ7 cell line.

Comparing the effectiveness of the pulldowns demonstrates a more thorough proteome coverage obtained in the Streptavidin Pulldown in comparison to the FLAG IP. FLAG IP identified an initial 480 protein interactors across the TurboID cell lines, as opposed to the 2264 identified in the Streptavidin Pulldown data. This almost five times larger interactome allowed for a more detailed analysis of differential interactivity between TurboID conditions, whilst demonstrating the potential that proximity biotinylation has in deciphering the complexities of protein interaction networks, particularly in combination with quantitative labelling techniques such as SILAC. To our knowledge, this study represents the first of its kind comparing both FLAG IP and Streptavidin Pulldown-generated SILAC data generated from the same starting samples in the context of an interactomic investigation.

A concurrent TurboID-specific Immunoprecipitation would be interesting to add as a comparison, particularly with respect to the FLAG IP; however, due to the limited number of TurboID-specific antibodies on the market - and even fewer which are validated for immunoprecipitation reactions - this may prove difficult and cost-prohibitive.

Beyond this, there are still areas for improvement. A repeat of the mass spectrometry data acquisition to address the NaN/missing quantitative values would have been of benefit, particularly in the identification of the unique interactomes specific to each TurboID cell line. Similarly, a comparison between an equivalent neuronal cell line – such as SHSY5Y neuroblastoma cells – as was initially planned would have been of interest to identify any neuron-specific pathways or protein interactors of relevance to SMA. Unfortunately, due to time constraints, these analyses were not performed.

7.6.1 Characterisation of Novel Interactors with Potential Relevance to SMA

As is evident from the analysis of the differentially interactive datasets, several proteins with no established previous interaction with the SMN protein - or existing relevance to the study of SMA - were identified. Using both String-DB and literature search to confirm no existing connections, these proteins are notable in that they may represent viable new avenues for future research. While SMA is now treatable due to the advent of splicing modifying therapeutics and gene replacement therapy, these drugs are expensive and focus solely on increasing levels of SMN (Chaytow et al., 2021). There is an unmet need for therapeutics which aim to restore cellular function in SMA independently of SMN. Novel interactors, like those described below, may offer innovative new avenues to guide the focus of future research of clinical relevance.

7.6.1.1 Protein SON

Protein SON is an RNA and DNA-binding protein which acts as an mRNA splicing cofactor and promotes the efficient splicing of transcripts with weak splice sites. Functionally, SON is essential for the correct splicing of several pre-mRNAs critical for normal development of the nervous system and patients with disease-causing mutations in the SON gene are diagnosed with Zhu-Tokita-Takenouchi-Kim (ZTTK) syndrome (OMIM: 617140) (Lu et al., 2014).

ZTTK syndrome is a severe multisystem developmental disorder characterized by, amongst other symptoms, intellectual disability, musculoskeletal abnormalities and nervous system problems, such as epilepsy. It is thought to arise from the inability of mutant SON to efficiently promote the splicing of the over 500 transcripts normally processed with the aid of the SON protein (Kim et al., 2016). Knockdown of SON in mice results in deficits in neuronal migration and dendritic spine formation, processes also known to be affected in SMA (Ueda et al., 2020). These symptoms are caused by a lack of the SON protein, however any analysis of a role for the protein in relation to the pathology of SMA has thus far not been performed.

The SON protein shows significantly increased interactivity with the SMNΔ7 isoform as identified by Streptavidin Pulldown. As SMNΔ7 is primarily nuclear in localisation, it is situated in close proximity to the nuclear-speckle localised SON in the cell (Sharma et al., 2010).

As discussed in Chapter 1, SMN's incorporation into the cytoplasmic SMN complex is a critical event in the formation of the snRNPs of the spliceosome. The minor spliceosome in particular is affected by the reduction of SMN in SMA, with U12 introns being preferentially retained in several disease relevant transcripts and widespread alternative splicing deficits observed in SMA mouse models (Doktor et al., 2017). Similarly, mutations in the SON gene have been shown – via genome-wide exome array experiments – to induce defects in alternative splicing (Sharma et al., 2011). While there is currently little to no information regarding the potential effect of SON on the splicing of SMN2 pre-mRNA, it is possible that

the increased interaction observed between the SON protein and SMNΔ7 plays a role in SMNΔ7's lack of functionality.

Precisely how this could occur is an open question. Could SMNΔ7's inability to effectively recapitulate the functions of native SMN lead to the localised concentration of splicing factors such as SON and Gemins 4 and 7? And does increased interactivity with these nuclear splicing factors further limit SMNΔ7's ability to exit the nucleus, interfering with functionality required in the cytoplasm?

Analysis of separate nuclear and cytoplasmic fractions of TurboID-SMNΔ7 cell lysates would be of interest to address such questions, to analyse differential interactivity between different cellular compartments. How the SON protein interacts with both native SMN and SMNΔ7 in the context of an SMA-relevant disease model would also be worthy of further investigation.

It is worth noting that SON has also been implicated in the host anti-viral immune response. SON has been shown to repress the Hepatitis B promoter and prevent transcription *in vitro* (Sun et al., 2001). In macrophages, Type 1 Interferon signalling (specifically, Interferon-Beta signalling) has also been found to depend on SON activity, and Interferon-induced gene expression is disrupted by knockdown of the protein (Gregory et al., 2020). While previous discussion on the GAIT Complex proteins (see Chapter 4) focussed on the role of Interferon-Gamma induction and the potential significance of those novel interactors to SMA, SON's

antiviral role may also be of relevance given the neuroinflammation seen in the disease (Nuzzo et al., 2023).

7.6.1.2 Tyrosyl-tRNA Synthetase and Hexokinase 1

Further to the discussion in Chapter 5 of RARS and IARS - two aminoacyl tRNA synthetases (AARS) - the Tyrosyl-tRNA Synthetase (YARS) protein was identified as a novel SMNΔ7 interactor with significantly decreased interactivity in the FLAG IP screen.

Mutations in the YARS1 gene are the cause of Charcot Marie Tooth Dominant-Intermediate C (CMTDIC), a subtype of CMT with patients displaying symptoms including nerve conduction impairments, axonal degeneration and demyelination (Hyun et al., 2014). How this protein interacts with SMN, or a potential role in the context of SMA, has thus far not been examined.

Like the other novel AARS interactors discussed previously, the precise pathological mechanisms by which mutations in this protein family results in distinct neuronal disease phenotypes are still unclear and an area of active research. However, given the similarities between CMT and SMA, the suggestion that there may be shared, unidentified pathomechanisms between these conditions is not without merit. Of note, some variants of the GARS protein - the disease-causative protein, mutations in which are responsible for CMT type 2D (CMT2D) - have been shown to gain novel, toxic functionality in mouse models of the disease (He et al., 2015). In this case, these mutant variants aberrantly bind to the

Neuropilin 1 (Nrp1) receptor; this novel interaction consequently interferes with the binding of the cognate vascular endothelial growth factor (VEGF) at this receptor, resulting in neuronal impairments as observed in the disease.

While mutations such as these are not explicitly relevant to SMA pathogenesis, they demonstrate that the precise contribution of the AARS proteins to neuronal pathology need not rely solely on an inability to fulfil their canonical, tRNA-charging functions. Previous discussion of the GAIT complex novel interactors in Chapter 4 also relied upon the non-canonical functions of AARSs, specifically Glutamyl-Prolyl-tRNA Synthetase (EPRS). Interestingly, one of the transcripts downregulated by the Interferon-Gamma activated GAIT complex is the previously mentioned VEGF (Ray et al., 2009). Functionality of these proteins beyond their canonical domains widens the scope of how these proteins – and their observed interactions with SMN and SMNΔ7 – may influence the SMA disease process, if at all.

Another protein that demonstrated significantly decreased interactivity with SMNΔ7 in this dataset is Hexokinase-1 (HK1). This enzyme catalyses the formation of glucose-6-phosphate, the initial step in most pathways of glucose metabolism. Like YARS, mutations in this protein are associated with a subtype of CMT: specifically, CMT Type 4G (CMT4G).

CMT4G is an autosomal recessive condition, characterized by lower limb weakness and muscle atrophy which results in walking difficulties. Considered a myelinopathy, patients

display reduced nerve conduction velocities as well as length-dependent axon degeneration (Sevilla et al., 2013).

While a link between HK1 and SMA has not been established thus far, in a BioID2-derivered interactomic study of murine NSC34 cells (neuroblastoma and spinal cord neurons), Hexokinase-2 (HK2) was identified as a novel interacting partner of SMN (Detering, 2021). While HK1 is expressed ubiquitously, HK2 is restricted to expression in specific tissues, including muscle and heart (Wilson, 2003).

As discussed in Chapter 1, metabolic defects have been observed in SMA and patients with the disease display various symptoms of glucose dysfunction (Deguise et al., 2021). Hypoglycaemia has been reported in several SMA patients, and mouse models of the disease have been shown to develop a similar metabolic phenotype (Bowerman et al., 2012). The decreased interaction between SMNΔ7 and HK1 may thus implicate this protein as a significant target in deciphering the as-yet unexplained role that SMN plays in glucose metabolism.

Of note, both HK1 and 2 are unique amongst the four hexokinase isoforms in that they bind to the outer mitochondrial membrane (OMM). HK1 has been implicated in the cellular defence against oxidative stress, a process which is disease relevant as stress markers have been shown to be increased in the neurons of Type 1 SMA patients (Hayashi et al., 2002)

7.6.1.3 Sideroflexin1 and Mitochondrial Dysfunction

Like HK1 and YARS, Sideroflexin1 (SFXN1) demonstrated significantly decreased interactivity with SMNΔ7 in the FLAG-IP dataset. SFXN1 is a nuclear-encoded mitochondrial serine transporter which localises to the inner mitochondrial membrane (IMM) and is integral to 1-Carbon metabolism. This biochemical process is a requirement for the downstream synthesis of many cellular metabolites, including amino acids and nucleotides (Petrova et al., 2023).

CRISPR-based knockout of SFXN1 has been shown to lead to a reduced glycine content in human Jurkat cells (Kory et al., 2018). At first glance this may seem unrelated to the current discussion, however this discovery gains more relevance upon consideration that mutations in the previously mentioned glycine tRNA Synthetase (GARS) are responsible for CMT Type 2D. Could reduced glycine availability in the cell lead to a similar phenotype as that seen in patients with GARS mutations? While GARS and SFXN1 localise to different parts of the cell, mutations in each protein results in the lack of glycine being incorporated during translation, whether via reduced cellular glycine content (SFXN1 KO) or via reduced glycine-tRNA charging (GARS mutants). Although this is a novel link between these two proteins, how this dysfunction interacts with the pathology as observed in SMA is still unknown.

No existing research has directly investigated the interaction between SMN and SFXN1, however - using genome-wide RNA-Seq in human induced pluripotent stem cells (hIPSCs) derived from Type 1 SMA patients - levels of SFXN1 mRNA were found to be reduced in SMA patients (Ng et al., 2015). While there is no current evidence indicating glycine depletion in

SMA patients, this could be an interesting area of future investigation considering the notable metabolic perturbations observed in the disease.

As a mitochondrial protein, SFXN1 localises to an organelle of emerging importance in the pathology of SMA. Mitochondrial dysfunction, and dysfunction in bioenergetic processes previously discussed, have been noted in both SMA patients and mouse models of the disease (Zilio et al., 2022). Fragmented, swollen mitochondria of a reduced size and number are a common feature of SMA neurons in mouse models of the disease and are associated with cellular apoptosis, the process assumed to be the primary cause of cell death in SMA (Miller et al., 2016).

Altered mitochondrial structure is associated with an increase in oxidative stress within the cell, a symptom also observed in the spinal motor neurons of SMA patients post-mortem (Hayashi et al., 2002). The precise role of SMN in maintaining healthy mitochondria has yet to be elucidated, however, there is evidence of reduced mRNA levels of mitochondrion-related proteins after SMN knockdown in mice motor neurons (Saal et al., 2014). As SMN has identified roles in splicing, transcription and mRNA transport, it is possible that a reduction in SMN could lead to the decreased transcription of nuclear-encoded proteins integral to the function of mitochondria, resulting in the mitochondrial phenotype as observed in SMA motor neurons.

As previously mentioned, Hexokinase-1 localises to the mitochondria. Furthermore, IARS - an AARS with differential SMN/SMNΔ7 interactivity investigated in Chapter 5 - has been

demonstrated to cause increased levels of cellular reactive oxygen species and mitochondrial defects upon knockdown in mice (Watanabe et al., 2023). Several mitochondria-related proteins, including SFXN1, were found to show significantly decreased interactivity with SMNΔ7 in both the FLAG IP and Streptavidin Pulldown datasets. Whilst this and previous research implicate the mitochondria as a locus of particular dysfunction in SMA, precisely where this dysfunction occurs in the pathogenesis of the disease has thus far not been defined and requires further elucidation.

7.7 Conclusion

SMN and SMNΔ7 are protein isoforms of particular clinical relevance to the pathology of SMA. Classically considered non-functional, using a variety of techniques including colocalisation analysis and quantitative interactomics, we have demonstrated that this functional loss specific to SMNΔ7 may be a result of overall reduced interactivity with proteins relevant to the cellular functions of SMN. In particular, those functions involving RNA metabolism, spliceosome formation and protein translation are significantly affected.

Novel protein interactors – such as GAIT Complex components, select Amino Acid tRNA Synthetases, Protein SON and Sideroflexin1 - with potential relevance to SMA pathology have been identified via targeted bioinformatic analysis, including the use of String-DB and Gene Ontology interrogation.

There is still an unmet need for SMN-independent adjunct therapies for SMA; while disease modifying therapies have shown great success in the past few years, the disease itself is by no means “cured.” These novel interacting proteins – and their implicated cellular pathways – may be worthy of future investigation to unravel the roles they might play in the pathogenesis of the disease.

References

Abera, M. B., Xiao, J., Nofziger, J., Titus, S., Southall, N., Zheng, W., Moritz, K. E., Ferrer, M., Cherry, J. J., Androphy, E. J., Wang, A., Xu, X., Austin, C., Fischbeck, K. H., Marugan, J. J., & Burnett, B. G. (2016). ML372 blocks SMN ubiquitination and improves spinal muscular atrophy pathology in mice. *JCI Insight*, *1*(19). https://doi.org/10.1172/jci.insight.88427

Åberg, C., & Robinson, A. (2021). Single-molecule localisation microscopy: accounting for chance co-localisation between foci in bacterial cells. *Eur Biophys J*, *50*(7), 941-950. https://doi.org/10.1007/s00249-021-01555-z

Adami, R., & Bottai, D. (2019). Spinal Muscular Atrophy Modeling and Treatment Advances by Induced Pluripotent Stem Cells Studies. *Stem Cell Rev Rep*, *15*(6), 795-813. https://doi.org/10.1007/s12015-019-09910-6

Aebersold, R., Burlingame, A. L., & Bradshaw, R. A. (2013). Western blots versus selected reaction monitoring assays: time to turn the tables? *Mol Cell Proteomics*, *12*(9), 2381-2382. https://doi.org/10.1074/mcp.E113.031658

Agha-Mohammadi, S., O'Malley, M., Etemad, A., Wang, Z., Xiao, X., & Lotze, M. T. (2004). Second-generation tetracycline-regulatable promoter: repositioned tet operator elements optimize transactivator synergy while shorter minimal promoter offers tight basal leakiness. *J Gene Med*, *6*(7), 817-828. https://doi.org/10.1002/jgm.566

Aguilan, J. T., Kulej, K., & Sidoli, S. (2020). Guide for protein fold change and p-value calculation for non-experts in proteomics [10.1039/D0MO00087F]. *Molecular Omics*, *16*(6), 573-582. https://doi.org/10.1039/D0MO00087F

Al-Amrani, S., Al-Jabri, Z., Al-Zaabi, A., Alshekaili, J., & Al-Khabori, M. (2021). Proteomics: Concepts and applications in human medicine. *World J Biol Chem*, *12*(5), 57-69. https://doi.org/10.4331/wjbc.v12.i5.57

Alrafiah, A., Karyka, E., Coldicott, I., Iremonger, K., Lewis, K. E., Ning, K., & Azzouz, M. (2018). Plastin 3 Promotes Motor Neuron Axonal Growth and Extends Survival in a Mouse Model of Spinal Muscular Atrophy. *Mol Ther Methods Clin Dev*, *9*, 81-89. https://doi.org/10.1016/j.omtm.2018.01.007

Antonellis, A., Ellsworth, R. E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S. Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L. T., Sivakumar, K., Ionasescu, V., Funalot, B., Vance, J. M., Goldfarb, L. G., Fischbeck, K. H., & Green, E. D. (2003). Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. *American journal of human genetics*, *72*(5), 1293-1299. https://doi.org/10.1086/375039

Arif, A., Chatterjee, P., Moodt, R. A., & Fox, P. L. (2012). Heterotrimeric GAIT complex drives transcript-selective translation inhibition in murine macrophages. *Mol Cell Biol*, *32*(24), 5046-5055. https://doi.org/10.1128/mcb.01168-12

Arif, A., Yao, P., Terenzi, F., Jia, J., Ray, P. S., & Fox, P. L. (2018). The <scp>GAIT</scp> translational control system. *WIREs RNA*, *9*(2), e1441. https://doi.org/10.1002/wrna.1441

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., & Wishart, D. S. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res*, *44*(W1), W147-153. https://doi.org/10.1093/nar/gkw419

Balak, C. D., Hunter, J. M., Ahearn, M. E., Wiley, D., D'Urso, G., & Baumbach-Reardon, L. (2017). Functional characterizations of rare UBA1 variants in X-linked Spinal Muscular Atrophy. *F1000Research*, *6*, 1636. https://doi.org/10.12688/f1000research.11878.1

Banik, S. D., & Nandi, N. (2012). Mechanism of the activation step of the aminoacylation reaction: a significant difference between class I and class II synthetases. *J Biomol Struct Dyn*, *30*(6), 701-715. https://doi.org/10.1080/07391102.2012.689701

Baranello, G., Darras, B. T., Day, J. W., Deconinck, N., Klein, A., Masson, R., Mercuri, E., Rose, K., El-Khairi, M., Gerber, M., Gorni, K., Khwaja, O., Kletzl, H., Scalco, R. S., Seabrook, T., Fontoura,

P., & Servais, L. (2021). Risdiplam in Type 1 Spinal Muscular Atrophy. *New England Journal of Medicine*, *384*(10), 915-923. https://doi.org/10.1056/nejmoa2009965

Baron-Delage, S., Abadie, A., Echaniz-Laguna, A., Melki, J., & Beretta, L. (2000). Interferons and IRF-1 induce expression of the survival motor neuron (SMN) genes. *Mol Med*, *6*(11), 957-968.

Bernabò, P., Tebaldi, T., Groen, E. J. N., Lane, F. M., Perenthaler, E., Mattedi, F., Newbery, H. J., Zhou, H., Zuccotti, P., Potrich, V., Shorrock, H. K., Muntoni, F., Quattrone, A., Gillingwater, T. H., & Viero, G. (2017). In Vivo Translatome Profiling in Spinal Muscular Atrophy Reveals a Role for SMN Protein in Ribosome Biology. *Cell Rep*, *21*(4), 953-965. https://doi.org/10.1016/j.celrep.2017.10.010

Bertini, E., Hwu, W. L., Reyna, S. P., Farwell, W., Gheuens, S., Sun, P., Zhong, Z. J., & De Vivo, D. C. (2017). Efficacy and safety of nusinersen in infants with presymptomatic spinal muscular atrophy (SMA): Interim results from the NURTURE study. *European Journal of Paediatric Neurology*, *21*, e14. https://doi.org/10.1016/j.ejpn.2017.04.1218

Bonanno, S., Cavalcante, P., Salvi, E., Giagnorio, E., Malacarne, C., Cattaneo, M., Andreetta, F., Venerando, A., Pensato, V., Gellera, C., Zanin, R., Arnoldi, M. T., Dosi, C., Mantegazza, R., Masson, R., Maggi, L., & Marcuzzo, S. (2022). Identification of a cytokine profile in serum and cerebrospinal fluid of pediatric and adult spinal muscular atrophy patients and its modulation upon nusinersen treatment. *Frontiers in cellular neuroscience*, *16*, 982760. https://doi.org/10.3389/fncel.2022.982760

Bowerman, M., Swoboda, K. J., Michalski, J. P., Wang, G. S., Reeks, C., Beauvais, A., Murphy, K., Woulfe, J., Screaton, R. A., Scott, F. W., & Kothary, R. (2012). Glucose metabolism and pancreatic defects in spinal muscular atrophy. *Ann Neurol*, *72*(2), 256-268. https://doi.org/10.1002/ana.23582

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N., & Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology*, *36*(9), 880-887. https://doi.org/10.1038/nbt.4201

Burnett, B. G., Munoz, E., Tandon, A., Kwon, D. Y., Sumner, C. J., & Fischbeck, K. H. (2009). Regulation of SMN Protein Stability. *Molecular and Cellular Biology*, *29*(5), 1107-1115. https://doi.org/10.1128/mcb.01262-08

Carrel, T. L., McWhorter, M. L., Workman, E., Zhang, H., Wolstencroft, E. C., Lorson, C., Bassell, G. J., Burghes, A. H. M., & Beattie, C. E. (2006). Survival Motor Neuron Function in Motor Axons Is Independent of Functions Required for Small Nuclear Ribonucleoprotein Biogenesis. *The Journal of Neuroscience*, *26*(43), 11014-11022. https://doi.org/10.1523/jneurosci.1637-06.2006

Cauchi, R. J. (2010). SMN and Gemins: 'we are family' … or are we?: insights into the partnership between Gemins and the spinal muscular atrophy disease protein SMN. *BioEssays*, *32*(12), 1077-1089. https://doi.org/10.1002/bies.201000088

Chaudhuri, S., Vyas, K., Kapasi, P., Komar, A. A., Dinman, J. D., Barik, S., & Mazumder, B. (2007). Human ribosomal protein L13a is dispensable for canonical ribosome function but indispensable for efficient rRNA methylation. *Rna*, *13*(12), 2224-2237. https://doi.org/10.1261/rna.694007

Chaytow, H., Faller, K. M. E., Huang, Y. T., & Gillingwater, T. H. (2021). Spinal muscular atrophy: From approved therapies to future therapeutic targets for personalized medicine. *Cell Rep Med*, *2*(7), 100346. https://doi.org/10.1016/j.xcrm.2021.100346

Chaytow, H., Huang, Y.-T., Gillingwater, T. H., & Faller, K. M. E. (2018). The role of survival motor neuron protein (SMN) in protein homeostasis. *Cellular and molecular life sciences : CMLS*, *75*(21), 3877-3894. https://doi.org/10.1007/s00018-018-2849-1

Chen, H. H., Chang, J. G., Lu, R. M., Peng, T. Y., & Tarn, W. Y. (2008). The RNA binding protein hnRNP Q modulates the utilization of exon 7 in the survival motor neuron 2 (SMN2) gene. *Mol Cell Biol*, *28*(22), 6929-6938. https://doi.org/10.1128/mcb.01332-08

Chen, X., Wei, S., Ji, Y., Guo, X., & Yang, F. (2015). Quantitative proteomics using SILAC: Principles, applications, and developments. *PROTEOMICS*, *15*(18), 3175-3192. https://doi.org/10.1002/pmic.201500108

Cho, J., Lee, J., Kim, J., Lee, H., Kim, M. J., Lee, Y. J., Yum, M. S., Byun, J. H., Lee, C. G., Lee, Y. M., Lee, J., & Chae, J. H. (2023). Nusinersen demonstrates effectiveness in treating spinal muscular atrophy: findings from a three-year nationwide study in Korea. *Front Neurol*, *14*, 1294028. https://doi.org/10.3389/fneur.2023.1294028

Cho, S., & Dreyfuss, G. (2010). A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes & Development*, *24*(5), 438-442. https://doi.org/10.1101/gad.1884910

Clelland, A. K., Kinnear, N. P., Oram, L., Burza, J., & Sleeman, J. E. (2009). The SMN protein is a key regulator of nuclear architecture in differentiating neuroblastoma cells. *Traffic*, *10*(11), 1585-1598. https://doi.org/10.1111/j.1600-0854.2009.00972.x

Costes, S. V., Daelemans, D., Cho, E. H., Dobbin, Z., Pavlakis, G., & Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys J*, *86*(6), 3993-4003. https://doi.org/10.1529/biophysj.103.038422

Courtney, N. L., Mole, A. J., Thomson, A. K., & Murray, L. M. (2019). Reduced P53 levels ameliorate neuromuscular junction loss without affecting motor neuron pathology in a mouse model of spinal muscular atrophy. *Cell Death Dis*, *10*(7), 515. https://doi.org/10.1038/s41419-019-1727-6

Crawford, T. O., Paushkin, S. V., Kobayashi, D. T., Forrest, S. J., Joyce, C. L., Finkel, R. S., Kaufmann, P., Swoboda, K. J., Tiziano, D., Lomastro, R., Li, R. H., Trachtenberg, F. L., Plasterer, T., Chen, K. S., & on behalf of the Pilot Study of Biomarkers for Spinal Muscular Atrophy Trial, G. (2012). Evaluation of SMN Protein, Transcript, and Copy Number in the Biomarkers for Spinal Muscular Atrophy (BforSMA) Clinical Study. *PLoS One*, *7*(4), e33572. https://doi.org/10.1371/journal.pone.0033572

D'Ydewalle, C., Ramos, D. M., Pyles, N. J., Ng, S.-Y., Gorz, M., Pilato, C. M., Ling, K., Kong, L., Ward, A. J., Rubin, L. L., Rigo, F., Bennett, C. F., & Sumner, C. J. (2017). The Antisense Transcript SMN-AS1 Regulates SMN Expression and Is a Novel Therapeutic Target for Spinal Muscular Atrophy. *93*(1), 66-79. https://doi.org/10.1016/j.neuron.2016.11.033

De Andrea, M., Ravera, R., Gioia, D., Gariglio, M., & Landolfo, S. (2002). The interferon system: an overview. *Eur J Paediatr Neurol*, *6 Suppl A*, A41-46; discussion A55-48. https://doi.org/10.1053/ejpn.2002.0573

Deguise, M.-O., Baranello, G., Mastella, C., Beauvais, A., Michaud, J., Leone, A., De Amicis, R., Battezzati, A., Dunham, C., Selby, K., Warman Chardon, J., McMillan, H. J., Huang, Y.-T., Courtney, N. L., Mole, A. J., Kubinski, S., Claus, P., Murray, L. M., Bowerman, M., . . . Kothary, R. (2019). Abnormal fatty acid metabolism is a core component of spinal muscular atrophy. *Annals of clinical and translational neurology*, *6*(8), 1519-1532. https://doi.org/10.1002/acn3.50855

Deguise, M.-O., Chehade, L., & Kothary, R. (2021). Metabolic Dysfunction in Spinal Muscular Atrophy. *International Journal of Molecular Sciences*, *22*(11), 5913. https://doi.org/10.3390/ijms22115913

Deng, J., Erdjument-Bromage, H., & Neubert, T. A. (2019). Quantitative Comparison of Proteomes Using SILAC. *Current protocols in protein science*, *95*(1), e74-e74. https://doi.org/10.1002/cpps.74

Detering, N. T. (2021). Posttranslational modifications of the Spinal Muscular Atrophy gene product SMN: Implications for therapy and pathophysiology [PhD Thesis]. *Institute of Neuroanatomy and Cell Biology, Hannover Medical School*.

Doktor, T. K., Hua, Y., Andersen, H. S., Brøner, S., Liu, Y. H., Wieckowska, A., Dembic, M., Bruun, G. H., Krainer, A. R., & Andresen, B. S. (2017). RNA-sequencing of a mouse-model of spinal

muscular atrophy reveals tissue-wide changes in splicing of U12-dependent introns. *Nucleic Acids Res*, *45*(1), 395-416. https://doi.org/10.1093/nar/gkw731

Echaniz-Laguna, A., Miniou, P., Bartholdi, D., & Melki, J. (1999). The Promoters of the Survival Motor Neuron Gene (SMN) and Its Copy (SMNc) Share Common Regulatory Elements. *The American Journal of Human Genetics*, *64*(5), 1365-1370. https://doi.org/https://doi.org/10.1086/302372

Edens, B. M., Ajroud-Driss, S., Ma, L., & Ma, Y.-C. (2015). Molecular mechanisms and animal models of spinal muscular atrophy. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1852*(4), 685-692. https://doi.org/10.1016/j.bbadis.2014.07.024

Eshraghi, M., McFall, E., Gibeault, S., & Kothary, R. (2016). Effect of genetic background on the phenotype of theSmn2B/-mouse model of spinal muscular atrophy. *Human Molecular Genetics*, ddw278. https://doi.org/10.1093/hmg/ddw278

Fabunmi, R. P., Wigley, W. C., Thomas, P. J., & DeMartino, G. N. (2001). Interferon gamma regulates accumulation of the proteasome activator PA28 and immunoproteasomes at nuclear PML bodies. *J Cell Sci*, *114*(Pt 1), 29-36. https://doi.org/10.1242/jcs.114.1.29

Fallini, C., Bassell, G. J., & Rossoll, W. (2012). Spinal muscular atrophy: The role of SMN in axonal mRNA regulation. *1462*, 81-92. https://doi.org/10.1016/j.brainres.2012.01.044

Faravelli, I., Riboldi, G. M., Rinchetti, P., & Lotti, F. (2023). The SMN Complex at the Crossroad between RNA Metabolism and Neurodegeneration. *International Journal of Molecular Sciences*, *24*(3), 2247. https://doi.org/10.3390/ijms24032247

Fedorova, S. A., & Dorogova, N. V. (2019). Protein trap: a new Swiss army knife for geneticists? [journal article]. *Molecular Biology Reports*. https://doi.org/10.1007/s11033-019-05181-z

Feng, M., & Zhang, H. (2022). Aminoacyl-tRNA Synthetase: A Non-Negligible Molecule in RNA Viral Infection. *Viruses*, *14*(3), 613. https://doi.org/10.3390/v14030613

Fuller, H. R., Gillingwater, T. H., & Wishart, T. M. (2016). Commonality amid diversity: Multi-study proteomic identification of conserved disease mechanisms in spinal muscular atrophy. *Neuromuscul Disord*, *26*(9), 560-569. https://doi.org/10.1016/j.nmd.2016.06.004

Gabanella, F., Butchbach, M. E. R., Saieva, L., Carissimi, C., Burghes, A. H. M., & Pellizzoni, L. (2007). Ribonucleoprotein Assembly Defects Correlate with Spinal Muscular Atrophy Severity and Preferentially Affect a Subset of Spliceosomal snRNPs. *2*(9), e921. https://doi.org/10.1371/journal.pone.0000921

Gailite, L., Sterna, O., Konika, M., Isakovs, A., Isakova, J., Micule, I., Setlere, S., Diriks, M., & Auzenbaha, M. (2022). New-Born Screening for Spinal Muscular Atrophy: Results of a Latvian Pilot Study. *Int J Neonatal Screen*, *8*(1). https://doi.org/10.3390/ijns8010015

Gallotta, I., Mazzarella, N., Donato, A., Esposito, A., Chaplin, J. C., Castro, S., Zampi, G., Battaglia, G. S., Hilliard, M. A., Bazzicalupo, P., & Di Schiavi, E. (2016). Neuron-specific knock-down of SMN1 causes neuron degeneration and death through an apoptotic mechanism. *Human Molecular Genetics*, ddw119. https://doi.org/10.1093/hmg/ddw119

Gerace, E., & Moazed, D. (2015). Affinity Pull-Down of Proteins Using Anti-FLAG M2 Agarose Beads. *Methods in enzymology*, *559*, 99-110. https://doi.org/10.1016/bs.mie.2014.11.010

Giard, D. J., Aaronson, S. A., Todaro, G. J., Arnstein, P., Kersey, J. H., Dosik, H., & Parks, W. P. (1973). In Vitro Cultivation of Human Tumors: Establishment of Cell Lines Derived From a Series of Solid Tumors2. *JNCI: Journal of the National Cancer Institute*, *51*(5), 1417-1423. https://doi.org/10.1093/jnci/51.5.1417

Giesemann, T., Rathke-Hartlieb, S., Rothkegel, M., Bartsch, J. W., Buchmeier, S., Jockusch, B. M., & Jockusch, H. (1999). A Role for Polyproline Motifs in the Spinal Muscular Atrophy Protein SMN. *Journal of Biological Chemistry*, *274*(53), 37908-37914. https://doi.org/10.1074/jbc.274.53.37908

Goedhart, J., & Luijsterburg, M. S. (2020). VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Scientific reports*, *10*(1). https://doi.org/10.1038/s41598-020-76603-3

Gregory, D. J., DeLoid, G. M., Salmon, S. L., Metzger, D. W., Kramnik, I., & Kobzik, L. (2020). SON DNA-binding protein mediates macrophage autophagy and responses to intracellular infection. *FEBS Lett*, *594*(17), 2782-2799. https://doi.org/10.1002/1873-3468.13851

Haggerty, D. L., Grecco, G. G., Reeves, K. C., & Atwood, B. (2019). Adeno-associated viral vectors in neuroscience research. *Molecular Therapy - Methods & Clinical Development*. https://doi.org/10.1016/j.omtm.2019.11.012

Han, K.-J., Foster, D. G., Zhang, N.-Y., Kanisha, K., Dzieciatkowska, M., Sclafani, R. A., Hansen, K. C., Peng, J., & Liu, C.-W. (2012). Ubiquitin-specific Protease 9x Deubiquitinates and Stabilizes the Spinal Muscular Atrophy Protein-Survival Motor Neuron. *Journal of Biological Chemistry*, *287*(52), 43741-43752. https://doi.org/10.1074/jbc.M112.372318

Harada, Y., Sutomo, R., Sadewa, A. H., Akutsu, T., Takeshima, Y., Wada, H., Matsuo, M., & Nishio, H. (2002). Correlation between SMN2 copy number and clinical phenotype of spinal muscular atrophy: three SMN2 copies fail to rescue some patients from the disease severity. *J Neurol*, *249*(9), 1211-1219. https://doi.org/10.1007/s00415-002-0811-4

Hayashi, M., Araki, S., Arai, N., Kumada, S., Itoh, M., Tamagawa, K., Oda, M., & Morimatsu, Y. (2002). Oxidative stress and disturbed glutamate transport in spinal muscular atrophy. *Brain Dev*, *24*(8), 770-775. https://doi.org/10.1016/s0387-7604(02)00103-1

He, W., Bai, G., Zhou, H., Wei, N., White, N. M., Lauer, J., Liu, H., Shi, Y., Dumitru, C. D., Lettieri, K., Shubayev, V., Jordanova, A., Guergueltcheva, V., Griffin, P. R., Burgess, R. W., Pfaff, S. L., & Yang, X. L. (2015). CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. *Nature*, *526*(7575), 710-714. https://doi.org/10.1038/nature15510

Herbst, S., Schaible, U. E., & Schneider, B. E. (2011). Interferon Gamma Activated Macrophages Kill Mycobacteria by Nitric Oxide Induced Apoptosis. *PLoS One*, *6*(5), e19105. https://doi.org/10.1371/journal.pone.0019105

Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Pat Cerretti, D., Urdal, D. L., & Conlon, P. J. (1988). A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. *Bio/Technology*, *6*(10), 1204-1210. https://doi.org/10.1038/nbt1088-1204

Hu, X., & Ivashkiv, L. B. (2009). Cross-regulation of signaling pathways by interferon-gamma: implications for immune responses and autoimmune diseases. *Immunity*, *31*(4), 539-550. https://doi.org/10.1016/j.immuni.2009.09.002

Hua, Y., & Zhou, J. (2004). Survival motor neuron protein facilitates assembly of stress granules. *FEBS Lett*, *572*(1-3), 69-74. https://doi.org/10.1016/j.febslet.2004.07.010

Husedzinovic, A., Oppermann, F., Draeger-Meurer, S., Chari, A., Fischer, U., Daub, H., & Gruss, O. J. (2014). Phosphoregulation of the human SMN complex. *European Journal of Cell Biology*, *93*(3), 106-117. https://doi.org/https://doi.org/10.1016/j.ejcb.2014.01.006

Hyun, Y. S., Park, H. J., Heo, S. H., Yoon, B. R., Nam, S. H., Kim, S. B., Park, C. I., Choi, B. O., & Chung, K. W. (2014). Rare variants in methionyl- and tyrosyl-tRNA synthetase genes in late-onset autosomal dominant Charcot-Marie-Tooth neuropathy. *Clin Genet*, *86*(6), 592-594. https://doi.org/10.1111/cge.12327

Ji, C., Bader, J., Ramanathan, P., Hennlein, L., Meissner, F., Jablonka, S., Mann, M., Fischer, U., Sendtner, M., & Briese, M. (2021). Interaction of 7SK with the Smn complex modulates snRNP production. *Nature communications*, *12*(1). https://doi.org/10.1038/s41467-021-21529-1

Kaifer, K. A., Villalón, E., Osman, E. Y., Glascock, J. J., Arnold, L. L., Cornelison, D. D. W., & Lorson, C. L. (2017). Plastin-3 extends survival and reduces severity in mouse models of spinal muscular atrophy. *2*(5). https://doi.org/10.1172/jci.insight.89970

Khan, K., Baleanu-Gogonea, C., Willard, B., Gogonea, V., & Fox, P. L. (2020). 3-Dimensional architecture of the human multi-tRNA synthetase complex. *Nucleic Acids Research*, *48*(15), 8740-8754. https://doi.org/10.1093/nar/gkaa569

Khayrullina, G., Alipio-Gloria, Z. A., Deguise, M. O., Gagnon, S., Chehade, L., Stinson, M., Belous, N., Bergman, E. M., Lischka, F. W., Rotty, J., Dalgard, C. L., Kothary, R., Johnson, K. A., & Burnett, B. G. (2022). Survival motor neuron protein deficiency alters microglia reactivity. *Glia*. https://doi.org/10.1002/glia.24177

Kim, D. Y., Kim, W., Lee, K. H., Kim, S. H., Lee, H. R., Kim, H. J., Jung, Y., Choi, J. H., & Kim, K. T. (2013). hnRNP Q regulates translation of p53 in normal and stress conditions. *Cell Death Differ*, *20*(2), 226-234. https://doi.org/10.1038/cdd.2012.109

Kim, J. H., Shinde, D. N., Reijnders, M. R. F., Hauser, N. S., Belmonte, R. L., Wilson, G. R., Bosch, D. G. M., Bubulya, P. A., Shashi, V., Petrovski, S., Stone, J. K., Park, E. Y., Veltman, J. A., Sinnema, M., Stumpel, C., Draaisma, J. M., Nicolai, J., Yntema, H. G., Lindstrom, K., . . . Ahn, E. Y. E. (2016). De Novo Mutations in SON Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectual-Disability Syndrome. *American journal of human genetics*, *99*(3), 711-719. https://doi.org/10.1016/j.ajhg.2016.06.029

Kim, M. H., & Kang, B. S. (2022). Structure and Dynamics of the Human Multi-tRNA Synthetase Complex. In J. R. Harris & J. Marles-Wright (Eds.), *Macromolecular Protein Complexes IV: Structure and Function* (pp. 199-233). Springer International Publishing. https://doi.org/10.1007/978-3-031-00793-4_6

Kolb, S. J., & Kissel, J. T. (2015). Spinal Muscular Atrophy. *Neurologic Clinics*, *33*(4), 831-846. https://doi.org/10.1016/j.ncl.2015.07.004

Kopajtich, R., Murayama, K., Janecke, A. R., Haack, T. B., Breuer, M., Knisely, A. S., Harting, I., Ohashi, T., Okazaki, Y., Watanabe, D., Tokuzawa, Y., Kotzaeridou, U., Kölker, S., Sauer, S., Carl, M., Straub, S., Entenmann, A., Gizewski, E., Feichtinger, R. G., . . . Staufner, C. (2016). Biallelic IARS Mutations Cause Growth Retardation with Prenatal Onset, Intellectual Disability, Muscular Hypotonia, and Infantile Hepatopathy. *American journal of human genetics*, *99*(2), 414-422. https://doi.org/10.1016/j.ajhg.2016.05.027

Kory, N., Wyant, G. A., Prakash, G., Uit de Bos, J., Bottanelli, F., Pacold, M. E., Chan, S. H., Lewis, C. A., Wang, T., Keys, H. R., Guo, Y. E., & Sabatini, D. M. (2018). SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism. *Science*, *362*(6416). https://doi.org/10.1126/science.aat9528

Lafarga, V., Tapia, O., Sharma, S., Bengoechea, R., Stoecklin, G., Lafarga, M., & Berciano, M. T. (2018). CBP-mediated SMN acetylation modulates Cajal body biogenesis and the cytoplasmic targeting of SMN. *Cellular and molecular life sciences : CMLS*, *75*(3), 527-546. https://doi.org/10.1007/s00018-017-2638-2

Lanfranco, M., Vassallo, N., & Cauchi, R. J. (2017). Spinal Muscular Atrophy: From Defective Chaperoning of snRNP Assembly to Neuromuscular Dysfunction. *Front Mol Biosci*, *4*, 41. https://doi.org/10.3389/fmolb.2017.00041

Lauria, F., Bernabò, P., Tebaldi, T., Groen, E. J. N., Perenthaler, E., Maniscalco, F., Rossi, A., Donzel, D., Clamer, M., Marchioretto, M., Omersa, N., Orri, J., Dalla Serra, M., Anderluh, G., Quattrone, A., Inga, A., Gillingwater, T. H., & Viero, G. (2020). SMN-primed ribosomes modulate the translation of transcripts related to spinal muscular atrophy. *Nature Cell Biology*. https://doi.org/10.1038/s41556-020-00577-7

Le, T. T., Pham, L. T., Butchbach, M. E., Zhang, H. L., Monani, U. R., Coovert, D. D., Gavrilina, T. O., Xing, L., Bassell, G. J., & Burghes, A. H. (2005). SMNDelta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Hum Mol Genet*, *14*(6), 845-857. https://doi.org/10.1093/hmg/ddi078

Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., Le Paslier, D., Frézal, J., Cohen, D., Weissenbach, J., Munnich, A., & 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., Dreyfuss, G., & 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

Li, P., Meng, Y., Wang, L., & Di, L. J. (2019). BioID: A Proximity-Dependent Labeling Approach in Proteomics Study. *Methods Mol Biol*, *1871*, 143-151. https://doi.org/10.1007/978-1-4939-8814-3_10

Lorson, C. L., Hahnen, E., Androphy, E. J., & Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proceedings of the National Academy of Sciences*, *96*(11), 6307-6311. https://doi.org/10.1073/pnas.96.11.6307

Lorson, C. L., Strasswimmer, J., Yao, J. M., Baleja, J. D., Hahnen, E., Wirth, B., Le, T., Burghes, A. H., & Androphy, E. J. (1998). SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nat Genet*, *19*(1), 63-66. https://doi.org/10.1038/ng0598-63

Lu, X., Ng, H. H., & Bubulya, P. A. (2014). The role of <scp>SON</scp> in splicing, development, and disease. *WIREs RNA*, *5*(5), 637-646. https://doi.org/10.1002/wrna.1235

May, D. G., Scott, K. L., Campos, A. R., & Roux, K. J. (2020). Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells*, *9*(5), 1070. https://doi.org/10.3390/cells9051070

Mazumder, B., Sampath, P., Seshadri, V., Maitra, R. K., DiCorleto, P. E., & Fox, P. L. (2003). Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. *Cell*, *115*(2), 187-198. https://doi.org/10.1016/s0092-8674(03)00773-6

McGivern, J. V., Patitucci, T. N., Nord, J. A., Barabas, M.-E. A., Stucky, C. L., & Ebert, A. D. (2013). Spinal muscular atrophy astrocytes exhibit abnormal calcium regulation and reduced growth factor production. *Glia*, *61*(9), 1418-1428. https://doi.org/10.1002/glia.22522

Menke, L. A., Poll-The, B. T., Clur, S. A., Bilardo, C. M., van der Wal, A. C., Lemmink, H. H., & Cobben, J. M. (2008). Congenital heart defects in spinal muscular atrophy type I: a clinical report of two siblings and a review of the literature. *Am J Med Genet A*, *146a*(6), 740-744. https://doi.org/10.1002/ajmg.a.32233

Miller, C. H. T., Maher, S. G., & Young, H. A. (2009). Clinical Use of Interferon-γ. *Annals of the New York Academy of Sciences*, *1182*(1), 69-79. https://doi.org/10.1111/j.1749-6632.2009.05069.x

Miller, N., Shi, H., Zelikovich, A. S., & Ma, Y.-C. (2016). Motor neuron mitochondrial dysfunction in spinal muscular atrophy. *Human Molecular Genetics*, *25*(16), 3395-3406. https://doi.org/10.1093/hmg/ddw262

Monani, U. R. (2000). The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in Smn-/- mice and results in a mouse with spinal muscular atrophy. *9*(3), 333-339. https://doi.org/10.1093/hmg/9.3.333

Mukhopadhyay, R., Jia, J., Arif, A., Ray, P. S., & Fox, P. L. (2009). The GAIT system: a gatekeeper of inflammatory gene expression. *Trends Biochem Sci*, *34*(7), 324-331. https://doi.org/10.1016/j.tibs.2009.03.004

Nafisinia, M., Sobreira, N., Riley, L., Gold, W., Uhlenberg, B., Weiß, C., Boehm, C., Prelog, K., Ouvrier, R., & Christodoulou, J. (2017). Mutations in RARS cause a hypomyelination disorder akin to Pelizaeus–Merzbacher disease. *European Journal of Human Genetics*, *25*(10), 1134-1141. https://doi.org/10.1038/ejhg.2017.119

Nanni, P. (2022). Mass Spectrometry in Proteomics: Technologies, Methods, and Research Applications for the Life Sciences. *CHIMIA*, *76*(1-2), 73. https://doi.org/10.2533/chimia.2022.73

Ng, S. Y., Soh, B. S., Rodriguez-Muela, N., Hendrickson, D. G., Price, F., Rinn, J. L., & Rubin, L. L. (2015). Genome-wide RNA-Seq of Human Motor Neurons Implicates Selective ER Stress Activation in Spinal Muscular Atrophy. *Cell Stem Cell*, *17*(5), 569-584. https://doi.org/10.1016/j.stem.2015.08.003

Nuzzo, T., Russo, R., Errico, F., D'Amico, A., Tewelde, A. G., Valletta, M., Hassan, A., Tosi, M., Panicucci, C., Bruno, C., Bertini, E., Chambery, A., Pellizzoni, L., & Usiello, A. (2023). Nusinersen mitigates neuroinflammation in severe spinal muscular atrophy patients. *Communications Medicine*, *3*(1), 28. https://doi.org/10.1038/s43856-023-00256-2

Ognjenović, J., & Simonović, M. (2018). Human aminoacyl-tRNA synthetases in diseases of the nervous system. *RNA Biology*, *15*(4-5), 623-634. https://doi.org/10.1080/15476286.2017.1330245

Oprea, G. E., Krober, S., McWhorter, M. L., Rossoll, W., Muller, S., Krawczak, M., Bassell, G. J., Beattie, C. E., & Wirth, B. (2008). Plastin 3 Is a Protective Modifier of Autosomal Recessive Spinal Muscular Atrophy. *320*(5875), 524-527. https://doi.org/10.1126/science.1155085

Osman, E. Y., Van Alstyne, M., Yen, P. F., Lotti, F., Feng, Z., Ling, K. K., Ko, C. P., Pellizzoni, L., & Lorson, C. L. (2020). Minor snRNA gene delivery improves the loss of proprioceptive synapses on SMA motor neurons. *JCI Insight*. https://doi.org/10.1172/jci.insight.130574

Pearn, J. (1980). Classification of spinal muscular atrophies. *Lancet*, *1*(8174), 919-922. https://doi.org/10.1016/s0140-6736(80)90847-8

Petit, F., Longoni, M., Wells, J., Maser, R. S., Bogenschutz, E. L., Dysart, M. J., Contreras, H. T. M., Frénois, F., Pober, B. R., Clark, R. D., Giampietro, P. F., Ropers, H. H., Hu, H., Loscertales, M., Wagner, R., Ai, X., Brand, H., Jourdain, A. S., Delrue, M. A., . . . High, F. A. (2023). PLS3 missense variants affecting the actin-binding domains cause X-linked congenital diaphragmatic hernia and body-wall defects. *American journal of human genetics*, *110*(10), 1787-1803. https://doi.org/10.1016/j.ajhg.2023.09.002

Petrova, B., Maynard, A. G., Wang, P., & Kanarek, N. (2023). Regulatory mechanisms of one-carbon metabolism enzymes. *Journal of Biological Chemistry*, *299*(12), 105457. https://doi.org/10.1016/j.jbc.2023.105457

Platanias, L. C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nature Reviews Immunology*, *5*(5), 375-386. https://doi.org/10.1038/nri1604

Powis, R. A., & Gillingwater, T. H. (2016). Selective loss of alpha motor neurons with sparing of gamma motor neurons and spinal cord cholinergic neurons in a mouse model of spinal muscular atrophy. *Journal of Anatomy*, *228*(3), 443-451. https://doi.org/10.1111/joa.12419

Powis, R. A., Karyka, E., Boyd, P., Côme, J., Jones, R. A., Zheng, Y., Szunyogova, E., Groen, E. J. N., Hunter, G., Thomson, D., Wishart, T. M., Becker, C. G., Parson, S. H., Martinat, C., Azzouz, M., & Gillingwater, T. H. (2016). Systemic restoration of UBA1 ameliorates disease in spinal muscular atrophy. *JCI Insight*, *1*(11). https://doi.org/10.1172/jci.insight.87908

Prescott, A. R., Bales, A., James, J., Trinkle-Mulcahy, L., & Sleeman, J. E. (2014). Time-resolved quantitative proteomics implicates the core snRNP protein SmB together with SMN in neural trafficking. *J Cell Sci*, *127*(Pt 4), 812-827. https://doi.org/10.1242/jcs.137703

Qiao, Y., Chi, Y., Gu, J., & Ma, Y. (2023). Safety and Efficacy of Nusinersen and Risdiplam for Spinal Muscular Atrophy: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Brain Sci*, *13*(10). https://doi.org/10.3390/brainsci13101419

Ramser, J., Ahearn, M. E., Lenski, C., Yariz, K. O., Hellebrand, H., von Rhein, M., Clark, R. D., Schmutzler, R. K., Lichtner, P., Hoffman, E. P., Meindl, A., & Baumbach-Reardon, L. (2008). Rare missense and synonymous variants in UBE1 are associated with X-linked infantile spinal muscular atrophy. *American journal of human genetics*, *82*(1), 188-193. https://doi.org/10.1016/j.ajhg.2007.09.009

Randolph, L. N., Bao, X., Zhou, C., & Lian, X. (2017). An all-in-one, Tet-On 3G inducible PiggyBac system for human pluripotent stem cells and derivatives. *Scientific reports*, *7*(1), 1549. https://doi.org/10.1038/s41598-017-01684-6

Ray, P. S., Jia, J., Yao, P., Majumder, M., Hatzoglou, M., & Fox, P. L. (2009). A stress-responsive RNA switch regulates VEGFA expression. *Nature*, *457*(7231), 915-919. https://doi.org/10.1038/nature07598

Renvoise, B. (2006). Distinct domains of the spinal muscular atrophy protein SMN are required for targeting to Cajal bodies in mammalian cells. *Journal of Cell Science*, *119*(4), 680-692. https://doi.org/10.1242/jcs.02782

Rizzo, F., Nizzardo, M., Vashisht, S., Molteni, E., Melzi, V., Taiana, M., Salani, S., Santonicola, P., Di Schiavi, E., Bucchia, M., Bordoni, A., Faravelli, I., Bresolin, N., Comi, G. P., Pozzoli, U., & Corti, S. (2019). Key role of SMN/SYNCRIP and RNA-Motif 7 in spinal muscular atrophy: RNA-Seq and motif analysis of human motor neurons. *Brain*, *142*(2), 276-294. https://doi.org/10.1093/brain/awy330

Ross, R. A., Spengler, B. A., & Biedler, J. L. (1983). Coordinate morphological and biochemical interconversion of human neuroblastoma cells. *J Natl Cancer Inst*, *71*(4), 741-747.

Rossoll, W., Jablonka, S., Andreassi, C., Kröning, A. K., Karle, K., Monani, U. R., & Sendtner, M. (2003). Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. *J Cell Biol*, *163*(4), 801-812. https://doi.org/10.1083/jcb.200304128

Rudnik-Schoneborn, S., Forkert, R., Hahnen, E., Wirth, B., & Zerres, K. (1996). Clinical spectrum and diagnostic criteria of infantile spinal muscular atrophy: further delineation on the basis of SMN gene deletion findings. *Neuropediatrics*, *27*(1), 8-15. https://doi.org/10.1055/s-2007-973741

Ruggiu, M., McGovern, V. L., Lotti, F., Saieva, L., Li, D. K., Kariya, S., Monani, U. R., Burghes, A. H. M., & Pellizzoni, L. (2012). A Role for SMN Exon 7 Splicing in the Selective Vulnerability of Motor Neurons in Spinal Muscular Atrophy. *32*(1), 126-138. https://doi.org/10.1128/mcb.06077-11

Saal, L., Briese, M., Kneitz, S., Glinka, M., & Sendtner, M. (2014). Subcellular transcriptome alterations in a cell culture model of spinal muscular atrophy point to widespread defects in axonal growth and presynaptic differentiation. *Rna*, *20*(11), 1789-1802. https://doi.org/10.1261/rna.047373.114

Samavarchi-Tehrani, P., Samson, R., & Gingras, A.-C. (2020). Proximity dependent biotinylation: key enzymes and adaptation to proteomics approaches. *Molecular & Cellular Proteomics*, mcp.R120.001941. https://doi.org/10.1074/mcp.r120.001941

Sanchez-Lanzas, R., & Castano, J. G. (2017). Lysine-Less Variants of Spinal Muscular Atrophy SMN and SMNDelta7 Proteins Are Degraded by the Proteasome Pathway. *Int J Mol Sci*, *18*(12). https://doi.org/10.3390/ijms18122667

Schorling, D. C., Pechmann, A., & Kirschner, J. (2019). Advances in Treatment of Spinal Muscular Atrophy - New Phenotypes, New Challenges, New Implications for Care. *J Neuromuscul Dis*. https://doi.org/10.3233/jnd-190424

Schrank, B., Götz, R., Gunnersen, J. M., Ure, J. M., Toyka, K. V., Smith, A. G., & Sendtner, M. (1997). Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc Natl Acad Sci U S A*, *94*(18), 9920-9925. https://doi.org/10.1073/pnas.94.18.9920

Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U., & Sattler, M. (2001). SMN tudor domain structure and its interaction with the Sm proteins. *Nat Struct Biol*, *8*(1), 27-31. https://doi.org/10.1038/83014

Sevilla, T., Martínez-Rubio, D., Márquez, C., Paradas, C., Colomer, J., Jaijo, T., Millán, J. M., Palau, F., & Espinós, C. (2013). Genetics of the Charcot-Marie-Tooth disease in the Spanish Gypsy population: the hereditary motor and sensory neuropathy-Russe in depth. *Clin Genet*, *83*(6), 565-570. https://doi.org/10.1111/cge.12015

Shababi, M., Habibi, J., Yang, H. T., Vale, S. M., Sewell, W. A., & Lorson, C. L. (2010). Cardiac defects contribute to the pathology of spinal muscular atrophy models. *Human Molecular Genetics*, *19*(20), 4059-4071. https://doi.org/10.1093/hmg/ddq329

Sharma, A., Lambrechts, A., Hao le, T., Le, T. T., Sewry, C. A., Ampe, C., Burghes, A. H., & Morris, G. E. (2005). A role for complexes of survival of motor neurons (SMN) protein with gemins and

profilin in neurite-like cytoplasmic extensions of cultured nerve cells. *Exp Cell Res*, *309*(1), 185-197. https://doi.org/10.1016/j.yexcr.2005.05.014

Sharma, A., Markey, M., Torres-Muñoz, K., Varia, S., Kadakia, M., Bubulya, A., & Bubulya, P. A. (2011). Son maintains accurate splicing for a subset of human pre-mRNAs. *J Cell Sci*, *124*(Pt 24), 4286-4298. https://doi.org/10.1242/jcs.092239

Sharma, A., Takata, H., Shibahara, K., Bubulya, A., & Bubulya, P. A. (2010). Son is essential for nuclear speckle organization and cell cycle progression. *Mol Biol Cell*, *21*(4), 650-663. https://doi.org/10.1091/mbc.e09-02-0126

Sharma, G., Paganin, M., Lauria, F., Perenthaler, E., & Viero, G. (2024). The SMN-ribosome interplay: a new opportunity for Spinal Muscular Atrophy therapies. *Biochemical Society Transactions*, *52*(1), 465-479. https://doi.org/10.1042/bst20231116

Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M., Lukyanov, S., Zaraisky, A. G., Davidson, M. W., & Chudakov, D. M. (2009). Far-red fluorescent tags for protein imaging in living tissues. *Biochem J*, *418*(3), 567-574. https://doi.org/10.1042/bj20081949

Shelke, G. V., Jagtap, J. C., Kim, D. K., Shah, R. D., Das, G., Shivayogi, M., Pujari, R., & Shastry, P. (2017). TNF-α and IFN-γ Together Up-Regulates Par-4 Expression and Induce Apoptosis in Human Neuroblastomas. *Biomedicines*, *6*(1). https://doi.org/10.3390/biomedicines6010004

Signoria, I., van der Pol, W. L., & Groen, E. J. N. (2023). Innovating spinal muscular atrophy models in the therapeutic era. *Dis Model Mech*, *16*(9). https://doi.org/10.1242/dmm.050352

Simone, C., Ramirez, A., Bucchia, M., Rinchetti, P., Rideout, H., Papadimitriou, D., Re, D. B., & Corti, S. (2016). Is spinal muscular atrophy a disease of the motor neurons only: pathogenesis and therapeutic implications? *Cellular and molecular life sciences : CMLS*, *73*(5), 1003-1020. https://doi.org/10.1007/s00018-015-2106-9

Singh, N. K., Singh, N. N., Androphy, E. J., & Singh, R. N. (2006). Splicing of a Critical Exon of Human Survival Motor Neuron Is Regulated by a Unique Silencer Element Located in the Last Intron. *Molecular and Cellular Biology*, *26*(4), 1333-1346. https://doi.org/10.1128/mcb.26.4.1333-1346.2006

Singh, N. N., Seo, J., Rahn, S. J., & Singh, R. N. (2012). A Multi-Exon-Skipping Detection Assay Reveals Surprising Diversity of Splice Isoforms of Spinal Muscular Atrophy Genes. *PLoS One*, *7*(11), e49595. https://doi.org/10.1371/journal.pone.0049595

Sirover, M. A. (1997). Role of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. *J Cell Biochem*, *66*(2), 133-140.

So, B. R., Wan, L., Zhang, Z., Li, P., Babiash, E., Duan, J., Younis, I., & Dreyfuss, G. (2016). A U1 snRNP–specific assembly pathway reveals the SMN complex as a versatile hub for RNP exchange. *Nature Structural & Molecular Biology*, *23*(3), 225-230. https://doi.org/10.1038/nsmb.3167

Sugarman, E. A., Nagan, N., Zhu, H., Akmaev, V. R., Zhou, Z., Rohlfs, E. M., Flynn, K., Hendrickson, B. C., Scholl, T., Sirko-Osadsa, D. A., & Allitto, B. A. (2012). Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: clinical laboratory analysis of >72 400 specimens. *European Journal of Human Genetics*, *20*(1), 27-32. https://doi.org/10.1038/ejhg.2011.134

Sun, C. T., Lo, W. Y., Wang, I. H., Lo, Y. H., Shiou, S. R., Lai, C. K., & Ting, L. P. (2001). Transcription repression of human hepatitis B virus genes by negative regulatory element-binding protein/SON. *J Biol Chem*, *276*(26), 24059-24067. https://doi.org/10.1074/jbc.M101330200

Swoboda, K. J. (2011). Of SMN in mice and men: a therapeutic opportunity. *121*(8), 2978-2981. https://doi.org/10.1172/jci58752

Swoboda, K. J. (2014). SMN-targeted therapeutics for spinal muscular atrophy: are we SMArt enough yet? , *124*(2), 487-490. https://doi.org/10.1172/jci74142

Swoboda, K. J., Prior, T. W., Scott, C. B., McNaught, T. P., Wride, M. C., Reyna, S. P., & Bromberg, M. B. (2005). Natural history of denervation in SMA: Relation to age,SMN2 copy number, and function. *Annals of Neurology*, *57*(5), 704-712. https://doi.org/10.1002/ana.20473

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Christian. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607-D613. https://doi.org/10.1093/nar/gky1131

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & Christian. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, *49*(D1), D605-D612. https://doi.org/10.1093/nar/gkaa1074

Talbot, K., & Tizzano, E. F. (2017). The clinical landscape for SMA in a new therapeutic era. *Gene Therapy*, *24*(9), 529-533. https://doi.org/10.1038/gt.2017.52

Tiziano, F. D., Melki, J., & Simard, L. R. (2013). Solving the puzzle of spinal muscular atrophy: What are the missing pieces? *American Journal of Medical Genetics Part A*, *161*(11), 2836-2845. https://doi.org/10.1002/ajmg.a.36251

Turunen, J. J., Niemelä, E. H., Verma, B., & Frilander, M. J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA*, *4*(1), 61-76. https://doi.org/10.1002/wrna.1141

Turvey, A. K., Horvath, G. A., & Cavalcanti, A. R. O. (2022). Aminoacyl-tRNA synthetases in human health and disease. *Front Physiol*, *13*, 1029218. https://doi.org/10.3389/fphys.2022.1029218

Ueda, M., Matsuki, T., Fukada, M., Eda, S., Toya, A., Iio, A., Tabata, H., & Nakayama, A. (2020). Knockdown of Son, a mouse homologue of the ZTTK syndrome gene, causes neuronal migration defects and dendritic spine abnormalities. *Mol Brain*, *13*(1), 80. https://doi.org/10.1186/s13041-020-00622-4

Vinogradova, E. S., Nikonov, O. S., & Nikonova, E. Y. (2021). Associations between Neurological Diseases and Mutations in the Human Glycyl-tRNA Synthetase. *Biochemistry (Moscow)*, *86*(S1), S12-S23. https://doi.org/10.1134/s0006297921140029

Vitte, J., Fassier, C., Tiziano, F. D., Dalard, C., Soave, S., Roblot, N., Brahe, C., Saugier-Veber, P., Bonnefont, J. P., & Melki, J. (2007). Refined characterization of the expression and stability of the SMN gene products. *Am J Pathol*, *171*(4), 1269-1280. https://doi.org/10.2353/ajpath.2007.070399

Wan, B., Feng, P., Guan, Z., Sheng, L., Liu, Z., & Hua, Y. (2018). A severe mouse model of spinal muscular atrophy develops early systemic inflammation. *Hum Mol Genet*, *27*(23), 4061-4076. https://doi.org/10.1093/hmg/ddy300

Wang, Y., Xu, C., Ma, L., Mou, Y., Zhang, B., Zhou, S., Tian, Y., Trinh, J., Zhang, X., & Li, X.-J. (2019). Drug screening with human SMN2 reporter identifies SMN protein stabilizers to correct SMA pathology. *Life Science Alliance*, *2*(2), e201800268. https://doi.org/10.26508/lsa.201800268

Watanabe, M., Shishido, K., Kanehira, N., Hiura, K., Nakano, K., Okamura, T., Ando, R., Sasaki, H., & Sasaki, N. (2023). Molecular and Pathological Analyses of IARS1-Deficient Mice: An IARS Disorder Model. *International Journal of Molecular Sciences*, *24*(8), 6955. https://www.mdpi.com/1422-0067/24/8/6955

Wilson, J. E. (2003). Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. *J Exp Biol*, *206*(Pt 12), 2049-2057. https://doi.org/10.1242/jeb.00241

Wishart, T. M., Pemberton, H. N., James, S. R., McCabe, C. J., & Gillingwater, T. H. (2008). Modified cell cycle status in a mouse model of altered neuronal vulnerability (slow Wallerian degeneration; Wlds). *Genome Biology*, *9*(6), R101. https://doi.org/10.1186/gb-2008-9-6-r101

Wolf, N. I., Salomons, G. S., Rodenburg, R. J., Pouwels, P. J., Schieving, J. H., Derks, T. G., Fock, J. M., Rump, P., van Beek, D. M., van der Knaap, M. S., & Waisfisz, Q. (2014). Mutations in RARS cause hypomyelination. *Ann Neurol*, *76*(1), 134-139. https://doi.org/10.1002/ana.24167

Xu, M., Lin, M.-C., Liu, J., Zuo, J., & Li, Z.-H. (2023). *Interferon-Stimulated Gene Gbp1 Protects Smn2 from Degradation*. Elsevier BV. https://dx.doi.org/10.2139/ssrn.4605481

Yao, P., & Fox, P. L. (2020). Aminoacyl-tRNA synthetases in cell signaling. In (pp. 243-275). Elsevier. https://doi.org/10.1016/bs.enz.2020.04.002

Yilmazer, I., Abt, M. R., Liang, Y., Seung, D., Zeeman, S. C., & Sharma, M. (2023). Determining Protein-Protein Interaction with GFP-Trap Beads. *Methods Mol Biol*, *2564*, 317-323. https://doi.org/10.1007/978-1-0716-2667-2_17

Yong, J., Kasim, M., Bachorik, J. L., Wan, L., & Dreyfuss, G. (2010). Gemin5 Delivers snRNA Precursors to the SMN Complex for snRNP Biogenesis. *Molecular Cell*, *38*(4), 551-562. https://doi.org/10.1016/j.molcel.2010.03.014

Young, P. J., Le, T. T., Dunckley, M., Nguyen, T. M., Burghes, A. H., & Morris, G. E. (2001). Nuclear gems and Cajal (coiled) bodies in fetal tissues: nucleolar distribution of the spinal muscular atrophy protein, SMN. *Exp Cell Res*, *265*(2), 252-261. https://doi.org/10.1006/excr.2001.5186

Zhang, H., Xing, L., Rossoll, W., Wichterle, H., Singer, R. H., & Bassell, G. J. (2006). Multiprotein complexes of the survival of motor neuron protein SMN with Gemins traffic to neuronal processes and growth cones of motor neurons. *J Neurosci*, *26*(33), 8622-8632. https://doi.org/10.1523/jneurosci.3967-05.2006

Zhang, H. L., Pan, F., Hong, D., Shenoy, S. M., Singer, R. H., & Bassell, G. J. (2003). Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization. *J Neurosci*, *23*(16), 6627-6637.

Zhang, H. L., Pan, F., Hong, D., Shenoy, S. M., Singer, R. H., & Bassell, G. J. (2003). Active Transport of the Survival Motor Neuron Protein and the Role of Exon-7 in Cytoplasmic Localization. *The Journal of Neuroscience*, *23*(16), 6627-6637. https://doi.org/10.1523/jneurosci.23-16-06627.2003

Zheng, T., Luo, Q., Han, C., Zhou, J., Gong, J., Chun, L., Xu, X. Z. S., & Liu, J. (2022). Cytoplasmic and mitochondrial aminoacyl-tRNA synthetases differentially regulate lifespan in Caenorhabditis elegans. *iScience*, *25*(11), 105266. https://doi.org/10.1016/j.isci.2022.105266

Zhou, H., Hong, Y., Scoto, M., Thomson, A., Pead, E., MacGillivray, T., Hernandez-Gerez, E., Catapano, F., Meng, J., Zhang, Q., Hunter, G., Shorrock, H. K., Ng, T. K., Hamida, A., Sanson, M., Baranello, G., Howell, K., Gillingwater, T. H., Brogan, P., . . . Muntoni, F. (2022). Microvasculopathy in spinal muscular atrophy is driven by a reversible autonomous endothelial cell defect. *The Journal of Clinical Investigation*, *132*(21). https://doi.org/10.1172/JCI153430

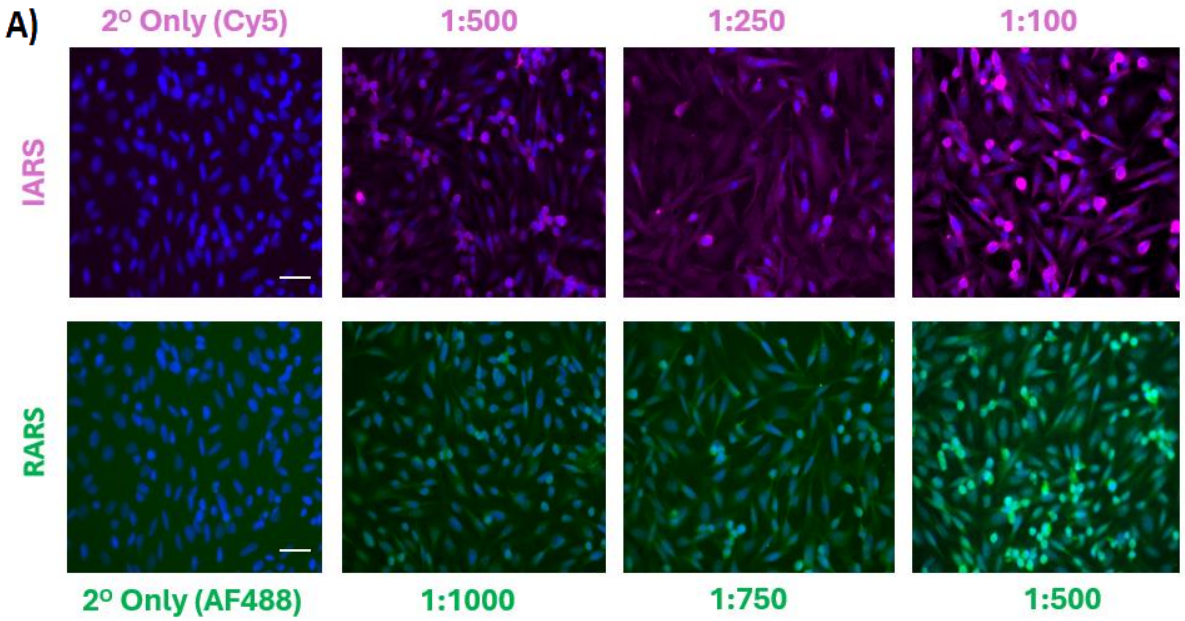
Zilio, E., Piano, V., & Wirth, B. (2022). Mitochondrial Dysfunction in Spinal Muscular Atrophy. *International Journal of Molecular Sciences*, *23*(18), 10878. https://doi.org/10.3390/ijms231810878

Supplementary Information

S1: Antibody Titrations for Optimal Immunofluorescence and Western Blotting

To investigate any potential interaction of the RARS and IARS proteins with SMN, it would be essential to find antibodies fit for purpose in the use of both immunofluorescence and western blotting. Once selected, these antibodies were titrated to determine the optimal concentration for use in each technique.

For immunofluorescent staining, Hela cells were fixed on coverslips and stained with varying concentrations of IARS (1:500, 1:250 and 1:100 dilutions) and RARS (1:1000, 1:750, and 1:500 dilutions) antibodies, using the antibody product sheets to determine a sufficient starting concentration. Images were obtained using the EVOS^{tm} M5000 Imaging System at x40 objective, in either the AF488 (Green – RARS) and Cy5 (Magenta – IARS) channels, along with the DAPI channel as counterstain (Figure 5.3A).



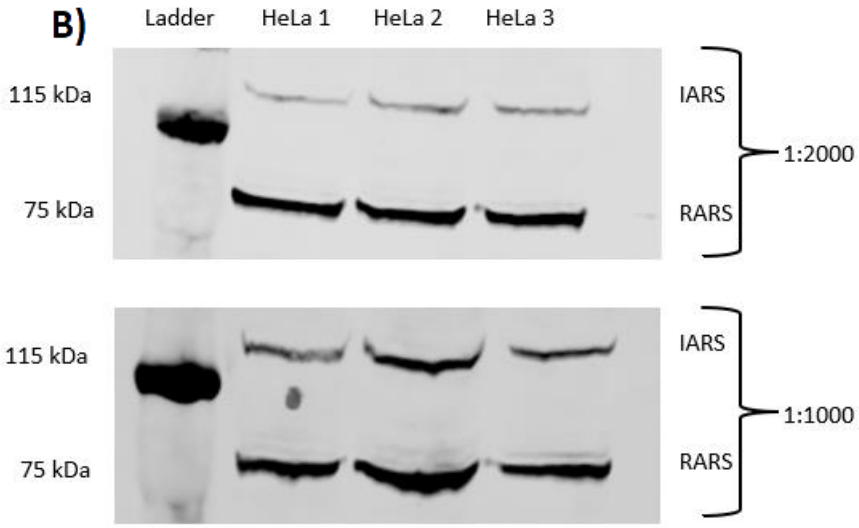


FIGURE S1: TITRATING ANTIBODY CONCENTRATIONS FOR OPTIMAL IMMUNOFLUORESCENT STAINING AND WESTERN BLOTTING.

A) Hela cells stained with increasing concentrations of the IARS/RARS antibodies (concentrations denoted above/below in either magenta/green). Relevant secondary only staining was performed as a control. Cells counterstained with DAPI. Scale = 75 microns. B) Western blots of a low concentration (top blot – 1:2000) and higher concentration (bottom blot – 1:1000) of both IARS and RARS antibodies in Hela lysates.

For western blot optimization, two concentrations (1:2000 and 1:1000) were selected for both RARS and IARS optimisation. Hela whole cell lysates were utilised and blots were subsequently imaged on the Li-Cor Odyssey CLx (Figure 5.3B)

From these results, it was decided to use the higher antibody concentrations of both antibodies in future immunofluorescence and western blotting experiments.

S2: Proteins in TurboID-SMNΔ7 with Significantly Increased or Decreased Interactivity in FLAG IP or Streptavidin Pulldown Datasets

Protein lists – including Gene Name, Protein Name and Canonical Functions - of differentially interacting TurboID-SMNΔ7 interactors which meet the stringency cutoff (Fold Change <0.5 and P Value <0.05).

FLAG IP TurboID-SMNΔ7 **Decreased** Interactors – sorted by Fold Change (largest to smallest). P<0.05.

Gene Name	Protein Name	Canonical Functions	Fold Change (Δ7/SMN)
CCT7	T-complex protein 1 subunit eta	Protein folding, chaperone-mediated protein complex assembly	0.25
ARF4	ADP-ribosylation factor 4	Vesicle-mediated transport, regulation of intracellular protein transport, small GTPase-mediated signal transduction	0.27
HNRNPK	Heterogeneous nuclear ribonucleoprotein K	mRNA splicing, regulation of alternative mRNA splicing via spliceosome, mRNA export from nucleus	0.28
GART	Trifunctional purine biosynthetic protein adenosine-3	Purine nucleotide biosynthetic process, ATP binding, magnesium ion binding	0.29
FAM35A	Protein FAM35A	Unknown function	0.29
CPS1	Carbamoyl-phosphate synthase [ammonia], mitochondrial	Urea cycle, metabolic process, carbamoyl phosphate biosynthetic process	0.29
CCT3	T-complex protein 1 subunit gamma	Protein folding, chaperone-mediated protein complex assembly	0.30
HK1	Hexokinase-1	Glucose metabolism, catalysis of ATP binding, phosphorylation of glucose	0.34
YARS	Tyrosine--tRNA ligase, cytoplasmic	Aminoacyl-tRNA ligase activity, tRNA aminoacylation for protein translation, ATP binding	0.35
RAB1B	Ras-related protein Rab-1B	Intracellular protein transport, vesicle-mediated transport, regulation of small GTPase-mediated signal transduction	0.36
SFXN1	Sideroflexin-1	Iron ion transport, mitochondrial outer membrane, iron ion transmembrane transporter activity	0.39
ATP5B	ATP synthase subunit beta, mitochondrial	ATP synthesis coupled proton transport, proton-transporting ATP synthase complex, mitochondrial ATP synthesis coupled proton transport	0.40

RPL27	60S ribosomal protein L27	Structural constituent of ribosome, cytoplasmic translation, ribosomal large subunit assembly	0.42
RPL4	60S ribosomal protein L4	Structural constituent of ribosome, cytoplasmic translation, ribosomal large subunit assembly	0.42
MRPL16	39S ribosomal protein L16, mitochondrial	Mitochondrial translation, ribosomal large subunit assembly, structural constituent of ribosome	0.43
RPS13	40S ribosomal protein S13	Structural constituent of ribosome, cytoplasmic translation, ribosomal small subunit assembly	0.44
HNRNPH1	Heterogeneous nuclear ribonucleoprotein H	mRNA processing, mRNA splicing, regulation of mRNA metabolic process	0.44
OXCT1	Succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	Ketone body metabolic process, mitochondrial matrix, transferase activity	0.45
CCT5	T-complex protein 1 subunit epsilon	Protein folding, chaperone-mediated protein complex assembly	0.47
RPS2	40S ribosomal protein S2	Structural constituent of ribosome, cytoplasmic translation, ribosomal small subunit assembly	0.47
NQO1	NAD(P)H dehydrogenase [quinone] 1	Electron transport chain, NADH dehydrogenase (quinone) activity, mitochondrial respiratory chain complex I	0.49

Streptavidin Pulldown TurboID-SMNΔ7 **Increased** Interactors – sorted by Fold Change (largest to smallest). P<0.05.

Gene Name	**Protein Name**	**Canonical Functions**	**Fold Change (Δ7/SMN)**
GEMIN7	Gem-associated protein 7	Regulation of mRNA stability, mRNA processing, RNA splicing	5.81
DDX20	Probable ATP-dependent RNA helicase DDX20	RNA helicase activity, ATP-dependent RNA helicase activity, regulation of RNA metabolic process	5.21
PSMG2	Proteasome assembly chaperone 2	Proteasome assembly, chaperone-mediated protein folding, protein refolding	4.45
TUBGCP2	Gamma-tubulin complex component 2	Microtubule organizing center organization, gamma-tubulin complex assembly, microtubule nucleation	3.06
SON	Protein SON	mRNA splicing, regulation of alternative mRNA splicing via spliceosome, mRNA export from nucleus	3.03

GEMIN4	Gem-associated protein 4	Regulation of mRNA stability, mRNA processing, RNA splicing	2.58
PDS5A	Sister chromatid cohesion protein PDS5 homolog A	Sister chromatid cohesion, regulation of DNA replication, DNA repair	2.45
KIF5B	Kinesin-1 heavy chain	Microtubule-based movement, ATPase activity, microtubule motor activity	1.90
SAFB	Scaffold attachment factor B1	Chromatin binding, DNA binding, regulation of gene expression	1.80
SUGP2	SURP and G-patch domain-containing protein 2	mRNA splicing, regulation of alternative mRNA splicing via spliceosome, mRNA export from nucleus	1.66
PRKACA	cAMP-dependent protein kinase catalytic subunit alpha	Protein phosphorylation, cAMP-mediated signaling, regulation of protein kinase activity	1.65
ECM29	Proteasome-associated protein ECM29 homolog	Regulation of proteasomal ubiquitin-dependent protein catabolic process, protein deubiquitination, protein binding	1.65
DNAJC13	DnaJ homolog subfamily C member 13	Protein folding, chaperone-mediated protein folding, heat shock protein binding	1.63
TAF6	Transcription initiation factor TFIID subunit 6	Transcription initiation, RNA polymerase II transcription factor activity, TFIID-class transcription factor binding	1.60
PFAS	Phosphoribosylformylglycinamidine synthase	Purine nucleotide biosynthetic process, ATP binding, transferase activity	1.57
UBAP1	Ubiquitin-associated protein 1	Ubiquitin-protein transferase activity, protein ubiquitination, ubiquitin-dependent protein catabolic process	1.54
S100A10	Protein S100-A10	Calcium ion binding, protein homodimerization activity, regulation of protein stability	1.53

Streptavidin Pulldown TurboID-SMNΔ7 **Decreased** Interactors – sorted by Fold Change (largest to smallest)

Gene Name	**Protein Name**	**Canonical Functions**	**Fold Change (Δ7/SMN)**
CALML3	Calmodulin-like protein 3	Calcium ion binding, regulation of muscle contraction, protein homodimerization activity	0.28
DCTN1	Dynactin subunit 1	Intracellular transport, cytoplasmic dynein complex, regulation of microtubule-based process	0.31

RBM12B	RNA-binding protein 12B	mRNA processing, mRNA splicing, regulation of mRNA metabolic process	0.31
ERAL1	GTPase Era, mitochondrial	GTP binding, GTPase activity, regulation of ribosome assembly	0.32
CUL2	Cullin-2	Ubiquitin-protein transferase activity, protein neddylation, cullin-RING ubiquitin ligase complex	0.32
AVL9	Late secretory pathway protein AVL9 homolog	Golgi apparatus organization, protein transport, regulation of vesicle-mediated transport	0.33
GIT1	ARF GTPase-activating protein GIT1	Small GTPase-mediated signal transduction, regulation of ARF protein signal transduction, GTPase activator activity	0.33
KIAA0196	WASH complex subunit strumpellin	Actin cytoskeleton organization, intracellular protein transport, regulation of endocytosis	0.33
GNL1	Guanine nucleotide-binding protein-like 1	Small GTPase-mediated signal transduction, regulation of GTPase activity, GTP binding	0.33
ATP13A1	Manganese-transporting ATPase 13A1	Metal ion transport, manganese ion transmembrane transporter activity, ATP binding	0.33
FANCI	Fanconi anemia group I protein	DNA repair, interstrand cross-link repair, Fanconi anemia pathway	0.33
SRGAP2	SLIT-ROBO Rho GTPase-activating protein 2	Regulation of small GTPase-mediated signal transduction, axon guidance, GTPase activator activity	0.34
CDK12	Cyclin-dependent kinase 12	Cell cycle, regulation of cyclin-dependent protein serine/threonine kinase activity, protein phosphorylation	0.34
NAV1	Neuron navigator 1	Axon guidance, neuron projection development, cell adhesion molecule binding	0.35
PSMD3	26S proteasome non-ATPase regulatory subunit 3	Proteasome-mediated ubiquitin-dependent protein catabolic process, protein deubiquitination, protein binding	0.35
SPATA5L1	Spermatogenesis-associated protein 5-like protein 1	Spermatogenesis, regulation of protein localization to cell surface, protein homodimerization activity	0.35
DDX42	ATP-dependent RNA helicase DDX42	RNA helicase activity, ATP-dependent RNA helicase activity, regulation of RNA metabolic process	0.35

DCUN1D1	DCN1-like protein 1	Protein neddylation, protein neddylation via Cullin-RING ubiquitin ligase complex, protein binding	0.36
PANK4	Pantothenate kinase 4	Coenzyme A biosynthetic process, ATP binding, magnesium ion binding	0.36
CHTF18	Chromosome transmission fidelity protein 18 homolog	Chromosome segregation, sister chromatid cohesion, regulation of mitotic cell cycle	0.36
EIF3I	Eukaryotic translation initiation factor 3 subunit I	Translation initiation, regulation of translational initiation, eukaryotic translation initiation factor activity	0.36
XPOT	Exportin-T	Nuclear protein export, regulation of mRNA export from nucleus, nuclear export signal receptor activity	0.36
HSPA9	Stress-70 protein, mitochondrial	Response to stress, chaperone-mediated protein folding, protein folding within mitochondrion	0.37
TFIP11	Tuftelin-interacting protein 11	Tooth development, enamel mineralization, protein binding	0.37
TRRAP	Transformation/transcription domain-associated protein	Regulation of transcription, RNA polymerase II transcription cofactor activity, transcription corepressor activity	0.37
GSTT1	Glutathione S-transferase theta-1	Glutathione transferase activity, detoxification of electrophilic compounds, conjugation with glutathione	0.37
SLC25A22	Mitochondrial glutamate carrier 1	Mitochondrial glutamate import, transmembrane transport, mitochondrial inner membrane	0.37
MSH6	DNA mismatch repair protein Msh6	Mismatch repair, DNA binding, ATPase activity	0.38
DDX39A	ATP-dependent RNA helicase DDX39A	RNA helicase activity, ATP-dependent RNA helicase activity, regulation of RNA metabolic process	0.38
CAPZB	F-actin-capping protein subunit beta	Actin filament capping, F-actin binding, cytoskeletal protein binding	0.38
TTC4	Tetratricopeptide repeat protein 4	Protein binding, regulation of protein stability, protein homodimerization activity	0.38
CCAR2	Cell cycle and apoptosis regulator protein 2	Regulation of cell cycle, regulation of apoptotic process, protein kinase activity	0.38
QIL1	Protein QIL1	Mitochondrial respiratory chain complex assembly,	0.38

		mitochondrion organization, mitochondrial membrane	
HSP90AB2P	Putative heat shock protein HSP 90-beta 2	Protein folding, chaperone-mediated protein folding, heat shock protein binding	0.38
TTC37	Tetratricopeptide repeat protein 37	Protein binding, regulation of protein stability, molecular chaperone	0.39
COG1	Conserved oligomeric Golgi complex subunit 1	Golgi apparatus organization, protein transport, regulation of vesicle-mediated transport	0.39
GTPBP4	Nucleolar GTP-binding protein 1	Ribosome biogenesis, regulation of rRNA processing, GTP binding	0.39
DDX19A	ATP-dependent RNA helicase DDX19A	RNA helicase activity, ATP-dependent RNA helicase activity, regulation of RNA metabolic process	0.39
HSPA1A	Heat shock 70 kDa protein 1A	Response to heat, protein folding, chaperone-mediated protein folding	0.39
DENND4C	DENN domain-containing protein 4C	Small GTPase-mediated signal transduction, regulation of GTPase activity, GTPase activator activity	0.39
PPP6R3	Serine/threonine-protein phosphatase 6 regulatory subunit 3	Protein dephosphorylation, regulation of protein phosphatase activity, protein phosphatase inhibitor activity	0.39
PRPF31	U4/U6 small nuclear ribonucleoprotein Prp31	mRNA splicing, U4/U6 small nuclear ribonucleoprotein complex assembly, U4 snRNA binding	0.40
PHLDB2	Pleckstrin homology-like domain family B member 2	Phosphatidylinositol binding, protein kinase B signaling, regulation of small GTPase mediated signal transduction	0.40
UFSP2	Ufm1-specific protease 2	Ubiquitin-specific protease activity, protein deubiquitination, negative regulation of ubiquitin-protein ligase activity	0.40
SPAG5	Sperm-associated antigen 5	Spermatogenesis, fertilization, protein binding	0.40
IDH3A	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	Citric acid cycle, NAD binding, isocitrate dehydrogenase activity	0.40
NDC80	Kinetochore protein NDC80 homolog	Chromosome segregation, kinetochore organization, microtubule binding	0.40
INPPL1	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2	Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase activity, regulation of	0.40

		phosphatidylinositol metabolic process	
SORD	Sorbitol dehydrogenase	Sorbitol metabolic process, zinc ion binding, zinc-dependent alcohol dehydrogenase activity	0.41
TCEB1	Transcription elongation factor B polypeptide 1	Transcription elongation, RNA polymerase II transcription cofactor activity, protein binding	0.41
FUS	RNA-binding protein FUS	RNA binding, mRNA processing, regulation of alternative mRNA splicing	0.41
HSPA8	Heat shock cognate 71 kDa protein	Protein folding, chaperone-mediated protein folding, response to heat stress	0.41
RBM25	RNA-binding protein 25	mRNA splicing, regulation of alternative mRNA splicing via spliceosome, RNA binding	0.41
ABCC2	Canalicular multispecific organic anion transporter 1	Organic anion transport, bile acid transport, integral component of plasma membrane	0.42
SAMD9	Sterile alpha motif domain-containing protein 9	RNA binding, mRNA splicing, regulation of mRNA processing	0.42
MAP7	Ensconsin	Microtubule binding, cytoskeleton organization, regulation of microtubule polymerization	0.42
NCKAP1	Nck-associated protein 1	Protein binding, cytoskeletal protein binding, actin filament binding	0.42
ACAD9	Acyl-CoA dehydrogenase family member 9, mitochondrial	Fatty acid beta-oxidation, acyl-CoA dehydrogenase activity, mitochondrial matrix	0.42
NGDN	Neuroguidin	Axon guidance, regulation of dendrite development, protein binding	0.42
CCDC6	Coiled-coil domain-containing protein 6	Protein binding, cytoskeletal protein binding, microtubule binding	0.42
SF3B6	Splicing factor 3B subunit 6	mRNA splicing, spliceosome assembly, regulation of mRNA splicing	0.42
NOB1	RNA-binding protein NOB1	mRNA processing, ribosome biogenesis, endoribonuclease activity	0.43
RPS29	40S ribosomal protein S29	Structural constituent of ribosome, cytoplasmic translation, ribosomal small subunit assembly	0.43
ZC3H14	Zinc finger CCCH domain-containing protein 14	RNA binding, mRNA stability, regulation of mRNA metabolic process	0.43

MKI67	Antigen KI-67	Regulation of cell cycle, chromosome organization, DNA replication	0.43
RNF20	E3 ubiquitin-protein ligase BRE1A	Protein ubiquitination, histone H2B monoubiquitination, E3 ubiquitin ligase activity	0.44
MCM4	DNA replication licensing factor MCM4	DNA replication initiation, DNA replication preinitiation complex, ATP binding	0.44
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1	Translation initiation, regulation of translational initiation, eukaryotic translation initiation factor activity	0.44
CYB5R3	NADH-cytochrome b5 reductase 3	Electron transport chain, NADH dehydrogenase activity, cytochrome-c oxidase activity	0.44
AATF	Protein AATF	Regulation of transcription, apoptosis, protein binding	0.44
ATP2B3	Plasma membrane calcium-transporting ATPase 3	Calcium ion transport, ATPase activity, plasma membrane	0.44
TM9SF3	Transmembrane 9 superfamily member 3	Endosomal transport, transmembrane receptor protein tyrosine kinase signaling pathway, transmembrane transporter activity	0.45
MYO1E	Unconventional myosin-Ie	Actin filament binding, ATP binding, motor activity	0.45
NOP56	Nucleolar protein 56	Ribosome biogenesis, nucleolus organization, rRNA binding	0.45
SRSF3	Serine/arginine-rich splicing factor 3	mRNA splicing, regulation of alternative mRNA splicing, RNA binding	0.45
PSMA4	Proteasome subunit alpha type-4	Proteasome-mediated ubiquitin-dependent protein catabolic process, ATPase activity, proteasome core complex	0.46
MMS19	MMS19 nucleotide excision repair protein homolog	DNA repair, nucleotide-excision repair, zinc ion binding	0.46
PDE1B	Calcium/calmodulin-dependent 3,5-cyclic nucleotide phosphodiesterase 1B	cAMP catabolic process, metal ion binding, phosphoric diester hydrolase activity	0.46
KNTC1	Kinetochore-associated protein 1	Chromosome segregation, kinetochore organization, microtubule binding	0.46
VPS37A	Vacuolar protein sorting-associated protein 37A	Endosomal transport, vacuolar protein sorting, protein binding	0.46
YTHDC2	Probable ATP-dependent RNA helicase YTHDC2	RNA helicase activity, ATP-dependent RNA helicase activity, regulation of RNA metabolic process	0.46

SEC31A	Protein transport protein Sec31A	COPII vesicle coating, ER to Golgi vesicle-mediated transport, protein binding	0.47
HNRNPUL2	Heterogeneous nuclear ribonucleoprotein U-like protein 2	mRNA splicing, regulation of alternative mRNA splicing via spliceosome, mRNA export from nucleus	0.47
TOR4A	Torsin-4A	Regulation of cellular protein localization, ATP binding, hydrolase activity	0.47
SPATS2L	SPATS2-like protein	mRNA splicing, regulation of alternative mRNA splicing via spliceosome, mRNA export from nucleus	0.47
ZAK	Mitogen-activated protein kinase kinase kinase MLT	MAP kinase kinase kinase activity, intracellular signal transduction, protein serine/threonine kinase activity	0.47
MGEA5	Protein O-GlcNAcase	O-GlcNAcase activity, protein deglycosylation, protein binding	0.47
PGM2	Phosphoglucomutase-2	Glucose metabolic process, phosphoglucomutase activity, magnesium ion binding	0.47
ATG9A	Autophagy-related protein 9A	Autophagy, regulation of autophagy, protein binding	0.48
BCLAF1	Bcl-2-associated transcription factor 1	Regulation of transcription, transcription factor activity, apoptosis	0.49
ZNF598	Zinc finger protein 598	DNA binding, metal ion binding, zinc ion binding	0.49
G3BP1	Ras GTPase-activating protein-binding protein 1	Regulation of small GTPase-mediated signal transduction, protein binding, Ras GTPase activator activity	0.50
BRIX1	Ribosome biogenesis protein BRX1 homolog	Ribosome biogenesis, rRNA processing, snoRNA binding	0.50