Aus dem Institut für Molekulare Medizin I
der Heinrich-Heine-Universität Düsseldorf

Investigation and Analysis of Post-translational Modifications of the SMN Complex in the snRNP Biogenesis

Dissertation

zur Erlangung des Grades eines „Doctor of Philosophy“ (PhD) in Medical Sciences
der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Qiaoping Li
(2026)

Als Inauguraldissertation gedruckt mit der Genehmigung der
Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:
Dekan: Prof. Dr. med. Nikolaj Klöcker
Gutachter: Prof. Dr. rer. nat. Sebastian Wesselborg, Prof. Dr. rer. nat. Jürgen Scheller

Zusammenfassung

Ein gestörter Energiestoffwechsel ist eines der Merkmale von Motoneuron-Erkrankungen (MNDs); sein Einfluss auf die Erkrankungen ist jedoch noch weitgehend ungeklärt. Spinale Muskelatrophie (SMA), eine der MNDs, wird primär durch den Mangel an funktionellem *survival of motor neuron* (SMN)-Protein verursacht. SMN interagiert mit Gemin2–Gemin8 und Unrip, um den SMN-Komplex zu bilden, welcher eine essenzielle Rolle bei der Biogenese von *spliceosomal small nuclear ribonucleoprotein* (snRNP) spielt.

Die vorliegende Studie konzentriert sich auf die SMN-Phosphorylierung und die Gemin3-Proteolyse, um das Verhalten des SMN-Komplexes unter verschiedenen zellulären Energieniveaus aufzuklären. Eine Energiereduktion wurde durch die Behandlung mit *Earle's balanced salt solution* (EBSS) oder *bafilomycin A1* (BafA1) induziert. Zellen, die GFP-SMN oder -Gemin3 Wildtyp sowie entsprechende Mutationen überexprimieren, wurden als primäre Modelle verwendet, um die Auswirkungen auf die Interaktion und die Biogenese von snRNP zu untersuchen. Untersucht wurde dies mittels Größenausschlusschromatographie (SEC), Immunpräzipitation (IP) und Immunfluoreszenz (IF). Die Interaktion wurde darüber hinaus durch Pull-down-Assays unter Verwendung von GST-markierten SMN-Phosphomutanten untersucht. Die durch Calpain vermittelte Spaltung wurde in einem *In-vitro*-Spaltungsassay getestet.

Die vorliegenden Ergebnisse zeigten, dass zwei energiesensorische Kinasen, *unc-51-like autophagy activating kinase 1* (ULK1) und *mammalian target of rapamycin* (mTOR), ungeachtet der Nährstoffbedingungen unabhängig vom SMN-Komplex agieren. Eine Reihe von Interaktionsassays zeigte, dass die einfache SMN-Phosphorylierung an Serin (S) 187 dessen Affinität zu SmB verringert und eine duale Phosphorylierung an Threonin (T) 37 und S187 seine Interaktion mit dem *Protein Arginine N-Methyltransferase 5* (PRMT5)-Komplex drastisch stört. Die IF ergab ferner, dass die duale Phosphorylierung die Bildung von snRNP behindert.

Eine Gemin3-Proteolyse wurde unter basalen Bedingungen beobachtet. Die vorliegende Studie demonstrierte, dass Gemin3 das Substrat von Calpain-1 ist, die basale Spaltung jedoch nicht mit Calpain in Zusammenhang steht. Experimente, die mit Zellen durchgeführt wurden, die trunkiertes GFP-Gemin3 überexprimieren, zeigten unterschiedliche Funktionen der Spaltfragmente auf. Gemin3 wies Eigenschaften von Substraten auf, die über *Chaperone-Mediated Autophagy* (CMA) abgebaut werden, was darauf hindeutet, dass der basale Abbau von Gemin3 über CMA vermittelt werden könnte.

Zusammenfassend lässt sich sagen, dass die Biogenese von snRNP durch die SMN-Phosphorylierung an T37 und S187 sowie durch die Gemin3-Proteolyse beeinflusst wird. Der Einfluss von Stoffwechselstörungen auf diese Prozesse könnte zur Entstehung und Progression von SMA und anderen MNDs beitragen.

Summary

Abnormal energy metabolism is one of the features of motor neuron diseases (MNDs); however, its influence on the diseases remains unclear. Spinal muscular atrophy (SMA), one of MNDs, is primarily caused by the deficiency of functional survival motor neuron (SMN) protein. SMN interacts with Gemin2–Gemin8 and Unrip to form the SMN complex, which plays an essential role in spliceosomal small nuclear ribonucleoprotein (snRNP) biogenesis.

The present study focuses on SMN phosphorylation and Gemin3 proteolysis to elucidate the behaviour of the SMN complex under different cellular energy levels. Energy reduction was induced by treatment with Earle's balanced salt solution (EBSS) or bafilomycin A1 (BafA1). Cells overexpressing GFP-SMN or -Gemin3 wild type and corresponding mutations were used as primary models to examine the impact on interaction and snRNP biogenesis measured via size exclusion chromatography (SEC), immunoprecipitation (IP) and immunofluorescence (IF). The interaction was further investigated via pull-down assays using GST-tagged SMN phospho-mutants. The calpain-mediated cleavage was tested by *in vitro* cleavage assay.

The present findings revealed that two energy-sensing kinases, unc-51-like autophagy activating kinase 1 (ULK1) and mammalian target of rapamycin (mTOR), function independently of the SMN complex regardless of the nutrient conditions. A series of interaction assays showed that SMN single phosphorylation at serine (S)187 reduces its affinity with SmB, and dual phosphorylation at threonine (T) 37 and S187 strikingly disrupts its interaction with the protein arginine N-methyltransferase 5 (PRMT5) complex. IF further revealed that dual phosphorylation impedes the formation of snRNP.

Gemin3 proteolysis was observed under basal conditions. The present study demonstrated that Gemin3 is the substrate of calpain-1, but the basal cleavage is not related to calpain. Assays performing with cells overexpressing truncated GFP-Gemin3 revealed distinct functions of the cleavage fragments. Gemin3 exhibited characteristics of the substrates degraded via chaperone-mediated autophagy (CMA), suggesting that the basal turnover of Gemin3 could be mediated via CMA.

In summary, snRNP biogenesis is affected by SMN phosphorylation at T37 and S187, and Gemin3 proteolysis. The influence of metabolism disorders on these events might contribute to the development and progression of SMA and other MNDs.

List of abbreviations

Biological and medical terms

aa	Amino acids
ALS	Amyotrophic lateral sclerosis
APS	Ammonium persulfate
AR	Autoradiography
ATG	Autophagy-related protein
ATP	Adenosine 5'-triphosphate
BafA1	Bafilomycin A1
BSA	Bovine serum albumin
$CaCl_2$	Calcium chloride
CDK	Cyclin-dependent kinase
CMA	Chaperone-mediated autophagy
co-IP	Co-immunoprecipitation
CS	Coomassie blue staining
DAPI	4',6-Diamidino-2-phenylindole
DDX20	DEAD-box helicase 20
DEAD	The conserved motif Asp-Glu-Ala-Asp
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOX	Doxycycline
DP103	DEAD box protein-103
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EBNA2	Epstein-barr virus nuclear antigen 2
EBNA3C	Epstein-barr virus nuclear antigen 3c
EBSS	Earle's balanced salt solution
EDTA	Ethylenediaminetetra-acetic acid
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
FBS	Foetal bovine serum
FL	Full length
FOXL2	Forkhead box L2
FPR	False positive rate
Fwd	Forward primers
Gemin	Gem-associated protein
GFP	Green fluorescent protein
GO	Gene Ontology
GST	Glutathione S-transferase
HEK293T	Human embryonic kidney 293T cell
HeLa	Henrietta Lacks cell
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesul-fonic acid

HSC70/ HSPA8	Heat shock cognate 71 kDa protein
HspB8/Hsp22	Heat shock protein family B (small) member 8
IB	Immunoblotting
IF	Immunofluorescence
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KCl	Potassium chloride
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knock-out
LAMP2A	Lysosome-associated membrane protein 2a
LB	Lysogeny broth
LC/LC-MS	Liquid chromatography-tandem mass spectrometry
M2H	Mammalian two-hybrid
MEP50/WD45	Methylosome protein 50
METS	Mitogenic Ets transcriptional suppressor
$MgCl_2$	Magnesium chloride
MND	Motor neuron disease
MS	Mass spectrometry
mTOR	Mammalian target of rapamycin
NaCl	Sodium chloride
NANOS1	Nanos homolog 1
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
Olig2	Oligodendrocyte transcription factor 2
PBS	Dulbecco's phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
pICln	Chloride conductance regulatory protein
PKA	Protein kinase A
PKC	Protein kinase C
PL-MS	Proximity labelling mass spectrometry
PRMT5	Protein arginine N-methyltransferase 5
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
Rev	Reverse primers
RNA	Ribonucleic acid
RPS6KB1/ p70S6 kinase	Ribosomal protein S6 kinase beta-1
RT	Room temperature
S.O.C.	Super Optimal broth with Catabolite repression
SD	Standard deviation
sDMA	Symmetric dimethylation
SDS	Sodium dodecyl sulphate

SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SF-1	Steroidogenic factor 1
Sm	Smith antigen
SMA	Spinal muscular atrophy
SMN	Survival motor neuron
snRNA	Small nuclear RNA
snRNP	Spliceosomal small nuclear ribonucleoprotein
TBST	Tris-buffered saline with Tween 20
TEMED	N,N,N',N'-Tetramethylethylene-diamine
Tet-free FBS	Tetracycline-free FBS
TGF-β	Transforming growth factor-beta
ULK1	unc-51-like autophagy activating kinase 1
UPS	Ubiquitin-proteasome system
UV	Ultraviolet
V-ATPase	Vacuolar-type H^+-ATPase
WT	Wild type
Y2H	Yeast two-hybrid

Units

°C	Degree Celsius
g	Gram
h	Hour
kDa	Kilodalton
L	Litre
M	Molar
mA	Milliampere
MDa	Megadalton
mg	Milligram
min	Minute
mL	Millilitre
mM	Millimolar
ng	Nanogram
nM	Nanomolar
rpm	Rounds per minute
U	Enzyme unit
V	Volt
µg	Microgram
µL	Microlitre
µM	Micromolar

Amino acids

A	Alanine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
Y	Tyrosine

Table of Contents

1 Introduction 1

1.1 Spinal muscular atrophy 1

1.2 snRNP biogenesis 1

1.2.1 The process of snRNP assembly 1

1.2.2 The role of SMN and Gemin3 in snRNP biogenesis 2

1.2.3 The energy-dependency of snRNP assembly 4

1.3 SMN phosphorylation 5

1.3.1 The effect of SMN phosphorylation on snRNP biogenesis 5

1.3.2 Phosphorylation at T37 and S187 5

1.4 Cleavage and degradation of Gemin3 6

1.4.1 The structure and function of Gemin3 7

1.4.2 Calpain-mediated cleavage and its implications for snRNP biogenesis 8

1.4.3 The potential interplay between chaperone-mediated autophagy and the SMN complex 9

1.5 Aims of thesis 10

2 Materials and Methods 12

2.1 Materials 12

2.1.1 Reagents 12

2.1.2 Antibodies 16

2.1.3 Buffers 18

2.1.4 Plasmids 20

2.1.5 Oligonucleotides for cloning 21

2.1.6 Bacteria 21

2.1.7 Cell lines 22

2.1.8 Equipment and consumables 22

2.1.9 Computational tools 24

2.2 Methods 25

2.2.1 Molecular biological methods 25

2.2.2 Cellular methods 27

2.2.3 Biochemical methods 29

2.2.4 Bioinformatic analyses 33

3 Results 35

3.1 Interaction between mTOR, ULK1 and SMN protein 35

3.1.1 mTOR, ULK1, and ATG13 are excluded from the SMN complex 35

3.1.2 Ubiquitin-mediated proteolysis might be associated with the turnover of the SMN complex during starvation 39
3.2 Effect of SMN phosphorylation at T37 and S187 on snRNP biogenesis 43
3.2.1 Single phosphorylation at T37 or S187 is dispensable for the interaction with Gemin2 43
3.2.2 S187 phosphorylation affects the binding to SmB 46
3.2.3 Dual phosphorylation at both T37 and S187 distances the PRMT5 complex from the SMN complex and impairs snRNP biogenesis 50
3.3 Gemin3 cleavage 57
3.3.1 Gemin3 is the only Gemin protein sensitive to bafilomycin A1 57
3.3.2 Gemin3 could be cleaved by calpain-1 *in vitro*, but Gemin3 cleavage *in vivo* does not depend on this pathway 60
3.3.3 Frag1-Gemin3 is not included into the SMN complex 66
3.3.4 The C-terminus, rather than the N-terminus, is necessary for Gemin3 to operate in the SMN complex 69
3.4 Gemin3 degradation 75
3.4.1 FL-Gemin3 levels are highly associated with the lysosome 75
3.4.2 FL-Gemin3 levels are regulated by chaperon-mediated autophagy 76
3.4.3 Gemin3 without canonical KFERQ-like motifs remains incorporated in the SMN complex and prompts snRNP biogenesis 79
4 Discussion 86
4.1 Energy-related kinases ULK1 and mTOR are independent of the SMN complex 86
4.2 Energy-mediated phosphorylation modulates the interaction network of the SMN complex 86
4.3 Cleavage of Gemin3 88
4.3.1 Gemin3 is cleaved under basal conditions 88
4.3.2 Gemin3 is the substrate of calpain-1 89
4.3.3 The function of Gemin3 might be dictated by cleavage 90
4.4 Chaperone-mediated autophagy degradation of Gemin3 91
4.5 Cleavage and degradation of the SMN complex 92
4.6 Conclusions and perspectives 94
5 Reference list 95

1 Introduction

1.1 Spinal muscular atrophy

Spinal muscular atrophy (SMA) is an autosomal recessive disease featured with the progressive degeneration of motor neurons and muscle atrophy [1]. SMA is fundamentally driven by mutations in the survival motor neuron (*SMN*) gene that lead to reduced expression of functional SMN protein, which is critical for the generation of spliceosomal small nuclear ribonucleoproteins (snRNPs) [2-4].

SMA belongs to the broad family of motor neuron diseases (MNDs) and shares several underlying cellular dysfunctions with other MNDs, such as amyotrophic lateral sclerosis (ALS) [5, 6]. Motor neurons consume a substantial amount of energy, which makes them more susceptible to energy depletion [7, 8]. Emerging evidence has shown that the dysregulation of autophagy and mitochondria, which drives significant energetic and metabolic deficits, contributes to the development of SMA and ALS [9-12]. These observations highlight the importance of energy homeostasis in MNDs. Research about how the biological processes mediated by cellular energy levels will shed light on the SMA and other MNDs, offering a potential strategy for intervention.

1.2 snRNP biogenesis

1.2.1 The process of snRNP assembly

Disruption of snRNP assembly, caused by SMN deficiency, is widely considered a core pathogenic mechanism of SMA [1, 13]. snRNPs, primarily composed of small nuclear RNA (snRNA) and a heteroheptameric ring of seven Sm proteins (SmD1, D2, D3, B, E, F and G), are the essential component of the spliceosome, a machinery for pre-mRNA splicing [14, 15].

The assembly of snRNA and Sm proteins into snRNPs is orchestrated by the SMN complex in the cytoplasm [16]. The SMN complex, comprising SMN, Gemin2–Gemin8, and Unrip [17], first binds to Sm proteins and subsequently receives snRNA [18]. *In vitro* experiments have shown that Sm proteins can spontaneously assemble on snRNA [19-21]; therefore, the delivery of Sm proteins *in vivo* is tightly regulated by the protein arginine N-methyltransferase 5 (PRMT5) complex, which consists of PRMT5, methylosome protein 50 (WD45; MEP50) and chloride conductance regulatory protein (pICln) [22]. pICln associates with Sm proteins to form two subcomplexes: the 6S complex (pICln-SmD1/D2/F/E/G) and pICln-SmD3/B, which not only sequesters Sm proteins from illicit interaction but also chaperones SmD1, D3, and B to PRMT5/WD45 for symmetric dimethylation (sDMA), an essential modification for subsequent SMN

binding [22]. The SMN complex receives these modified Sm proteins following the dissociation of the chaperone pICln, which is governed by unc-51-like autophagy activating kinase 1 (ULK1)-mediated phosphorylation [23-25]. The target snRNA is specifically recognised and delivered to the SMN complex by Gemin5 [17, 26].

Upon receiving these components, the SMN complex coordinates the proper assembly of Sm proteins and target snRNA into core snRNPs [15]. Chaperoned by the SMN complex, the assembled products are subsequently imported back into the nuclear compartment and localised to Cajal bodies, where they undergo essential modifications and higher-order assembly to form functional spliceosomes [16].

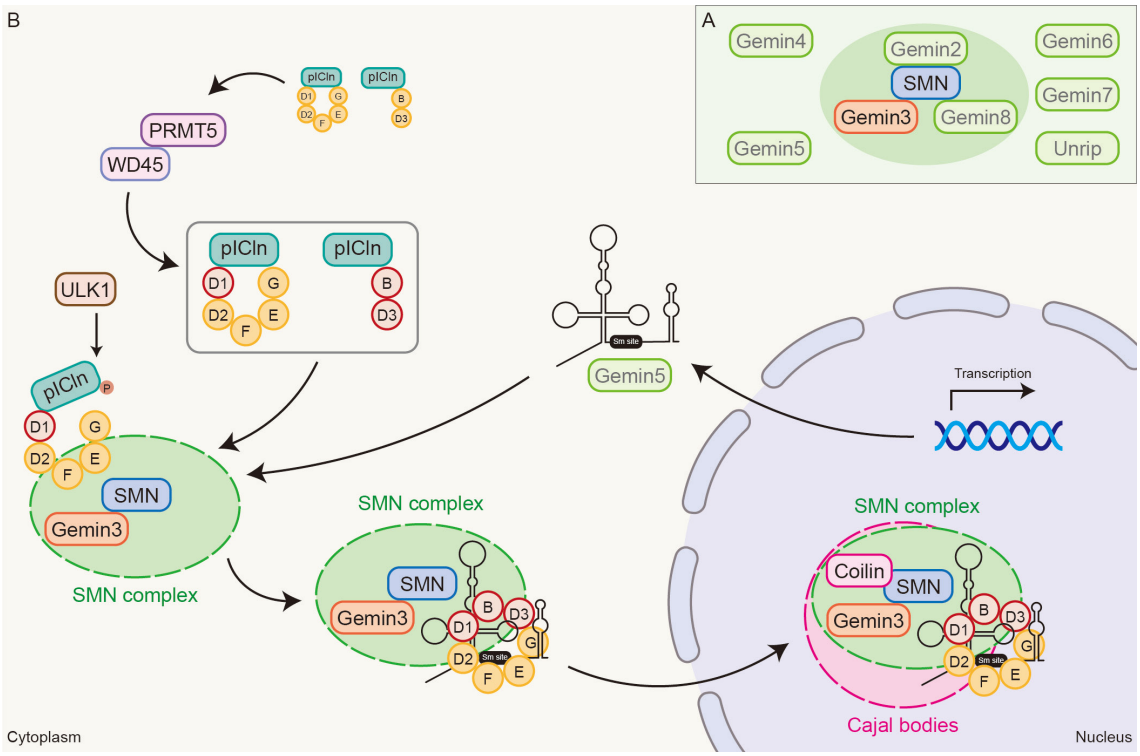


Fig. 1: ***Overview of snRNP biogenesis mediated by the SMN complex.*** *The SMN complex is primarily responsible for the assembly of Sm-class snRNPs, including major spliceosomal U1, U2, U4 and U5 snRNPs [15, 27]. These snRNPs share a similar structure, containing corresponding snRNA and a group of Sm proteins [14]. Here, the U1 snRNA is depicted as a representative snRNA to illustrate the assembly process of snRNPs [28]. (A) The SMN complex consists of SMN, Gemin2–Gemin8, and Unrip [17]. The Gemin2, Gemin3 and Gemin8 which directly bind to SMN are enclosed within the dark green oval, while other proteins are positioned outside [29, 30]. (B) The SMN complex (dashed green oval) receives the Sm proteins methylated by the PRMT5 complex (SmD1, D2 and B with sDMA, red circles) and snRNA via Gemin5 [17, 22, 26]. The 6S complex ring is disrupted by the ULK1 phosphorylation of pICln [23-25]. Regulated by the SMN complex, the Sm proteins are assembled onto snRNAs to form snRNPs which is subsequently delivered into Cajal bodies for further maturation as the spliceosome [15, 16].*

1.2.2 The role of SMN and Gemin3 in snRNP biogenesis

The SMN protein is indispensable for snRNP biogenesis [31]. Compared with other complex proteins, SMN is distributed across a wide range of evolutionary lineages [30]. The absence of the SMN protein is lethal across various species [32]. The clinical

severity of SMA correlates directly with the level of functional SMN protein [4, 33]. This correlation between SMN protein and SMA phenotypes has been utilised to successfully engineer transgenic SMA animal models [34]. The functional significance of SMN stems from its interaction with Sm proteins [34]. SMN interacts directly with all Sm proteins except SmF and SmG primarily through its Tudor domain [35]. Besides, other complex proteins, except Gemin8, also directly bind to specific Sm proteins (Table 1). As the essential core scaffold protein, SMN directly binds to Gemin2, Gemin3, and Gemin8 (Table 1, highlighted in yellow), which in turn serve to incorporate other proteins into the complex [29, 30]. Therefore, SMN also recruits Sm proteins into the complex indirectly through its binding partners.

Table 1: ***Direct interactions between SMN complex proteins and Sm proteins***

Complex components	Interacted Sm proteins	Reference
SMN	B, D1, D2, D3, E	[35]
Gemin2	D1, D2, E, F, G [as a pentamer]	[36], [37]
Gemin3	B, D2, D3	[38]
Gemin4	B, D1, D2, D3, E	[39]
Gemin5	B, D1, D2, D3, E	[40]
Gemin6	B, D2, D3, E	[41]
Gemin7	B, D2, D3, E	[42]
Gemin8	No direct binding	[43]
Unrip	B, (D1), D2, D3, E	[44], [45]

Notes: The direct interaction with Sm proteins was verified by in vitro binding assays. Gemin2 was additionally validated via X-ray crystallography. The notation '()' denotes that the direct interaction between Unrip and SmD1 is only observed in [45]. The direct interaction of SMN protein with Gemin2, Gemin3, and Gemin8 (highlighted in yellow) has been demonstrated, whereas other proteins may be integrated into the complex through an indirect association with SMN [29, 30].

In contrast to SMN, Gemin3 is responsible for snRNA modification during snRNP biogenesis [17, 46]. Gemin3 (also named as DDX20 or DP103) is a DEAD-box RNA helicase [47, 48]. This protein is predominantly found within the metazoan lineage [30], suggesting that Gemin3 might be a late evolutionary acquisition potentially evolved to exert highly specialised functions. The deficiency or mutations of Gemin3 lead to severe degeneration in neuronal and muscular cells that mimics SMA [47, 49]. Gemin3 might

serve as a bridge for the incorporation of Gemin4 and Gemin5 into the SMN complex [29, 38, 39]. Furthermore, recent research has shown that Gemin3 is essential for unwinding the compact RNA structure to expose the Sm site, thereby facilitating the assembly of the Sm protein ring [46].

In summary, SMN recruits Sm proteins to the complex, while Gemin3 mediates the structural rearrangement of the snRNA into the appropriate conformation for subsequent assembly. Both of them are necessary for successful snRNP biogenesis within cellular environments.

1.2.3 The energy-dependency of snRNP assembly

Several *in vitro* assembly assays using cell extracts support the necessity of ATP for the assembly [18, 50-53]. The transfer and unwinding of snRNA during snRNP biogenesis has been demonstrated as ATP-driven processes [46, 54]. These findings indicate that energy is required during the assembly of snRNPs.

Cellular energy-related pathways may play a role in snRNP assembly. Among such mechanisms, the involvement of the ULK1 and mammalian target of rapamycin (mTOR) pathways in snRNP biogenesis has been demonstrated recently. A prior review has detailed the association between the mTOR pathway and snRNP biogenesis, proposing the mTOR pathway as a potential candidate for regulating snRNP biogenesis by SMN phosphorylation [16]. Recent research confirmed this hypothesis by revealing that mTOR-dependent dual phosphorylation of SMN at serine (S) 49 and S63 promotes the liquid-liquid phase separation and efficient condensation of the SMN complex into higher-order structures [55]. Conversely, the ULK1 pathway primarily controls the transfer of Sm proteins during snRNP biogenesis [23-25]. ULK1 phosphorylates the assembly chaperone pICln at three specific residues in its C-terminus: S193, S195, and S197 [23]. This phosphorylation disrupts the binding interface between pICln and the SmG protein, which facilitates the displacement of pICln and the transfer of the Sm proteins onto the SMN complex to assemble functional snRNPs [23-25].

Despite these recent advances, the implications of cellular energy for snRNP biogenesis remain largely incomplete. SMA is characterised by energy disruption [9, 10]. snRNP assembly is an energy-dependent biological process severely compromised in SMA [1, 13]. Thus, further research about the molecular mechanism by which cellular energy levels affect the snRNP biogenesis will be beneficial to the understanding of SMA and other MNDs.

1.3 SMN phosphorylation

Protein phosphorylation serves as a critical molecular switch in signal transduction, which is a potential mechanism coupling energy levels and snRNP biogenesis [56]. The SMN protein is highly phosphorylated in the cytoplasm, where snRNP assembly occurs [16], which suggests that the phosphorylation of SMN plays an essential role in the regulation of snRNP biogenesis.

1.3.1 The effect of SMN phosphorylation on snRNP biogenesis

Several molecular mechanism studies have demonstrated that SMN phosphorylation mediates the composition and stability of the SMN complex, which consequently impacts the snRNP biogenesis [57-61]. Dual phosphorylation at S28 and S31 enhances the association of SMN with Gemin2, Gemin4, and Gemin5, while simultaneously diminishing its affinity for Sm proteins and snRNA [57]. Thus, this dual phosphorylation promotes the release of snRNP, which facilitates assembly *in vitro* [57]. Interaction with Gemin2 and Gemin8 is also mediated by residues threonine (T) 85 and S187, which are phosphorylated by protein kinase A (PKA) [58]. In addition to the Gemin proteins, phosphorylation also regulates the interaction with coilin, a scaffold protein of Cajal bodies where the snRNP undergoes maturation [15, 62, 63]. Research has shown that dephosphorylation at tyrosine (Y) 109, Y127, and Y130 significantly decreases the interaction with coilin and SMN condensation in the nucleus [59]. The condensation of SMN is also governed by two mTOR-dependent phosphorylation sites, S49 and S63 [55]. However, in contrast to the aforementioned tyrosine phosphorylation, these two sites likely modulate condensation by influencing liquid-liquid phase separation instead of interaction [55]. The location of the SMN condensates might be regulated by phosphorylation at S12 [59]. Moreover, the phosphorylation status influences the stability of SMN. Studies have revealed that phosphorylation at S270 and S290 accelerates SMN degradation [60, 61]. S270 might serve as a phosphodegron to trigger the degradation mediated by the ubiquitin-proteasome system (UPS) [60].

1.3.2 Phosphorylation at T37 and S187

T37 and S187 are two phosphorylatable residues on the SMN protein, which have been confirmed under different conditions [55, 58, 64-69]. However, the impact of their phosphorylation remains largely unknown.

T37 exhibited highly conserved evolution. This position is strictly maintained as a threonine across almost all examined species. Although it is replaced by serine in *Drosophila melanogaster*, both residues are phosphorylatable (Fig. 2). Previous

research has suggested that T37 phosphorylation is spleen-specific and related to acute myeloid leukaemia [64, 65].

In contrast, S187 appears to have evolved more recently to support sophisticated regulation. S187 is exclusively conserved in primates (Fig. 2 *Homo sapiens, Macaca fascicularis*, and *Pongo abelii*, highlighted in yellow). S187 is a PKA-specific phosphosite that contributes to the interaction with Gemin2 and Gemin8 [58]. Its phosphorylation is implicated in breast tumours and could be induced by bortezomib, an anti-cancer medication and a proteasome inhibitor [66-68]. Besides, maintenance of S187 phosphorylation levels might require Aurora kinases, Polo-like kinases and ribosomal protein S6 kinase beta-1 (RPS6KB1; p70S6 kinase), as the disruption of any single factor leads to its reduction [55, 69].

While extensive research has underscored the significance of T37 and S187 phosphorylation, their potential involvement in snRNP biogenesis has not yet been fully elucidated.

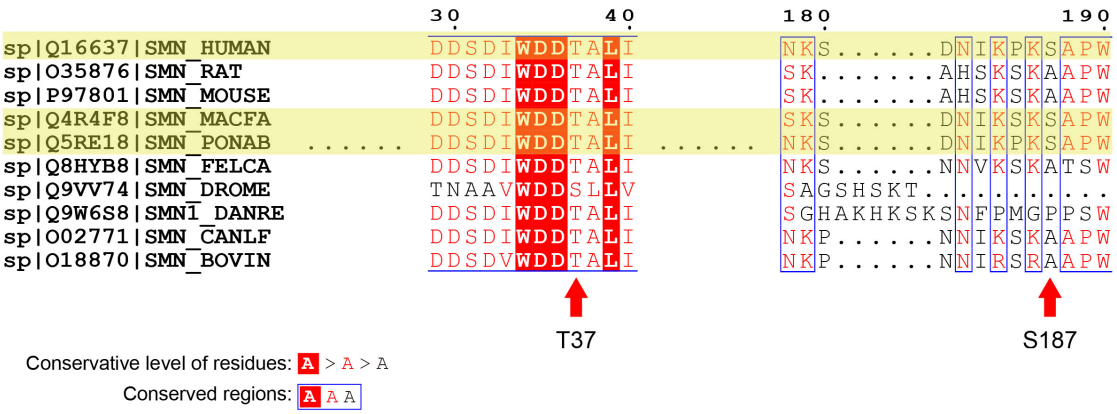


Fig. 2: ***Different evolution patterns of amino acids 37 and 187 on SMN.*** *Protein sequences were retrieved from UniProt, aligned using Clustal Omega 1.2.4 [70, 71], and visualised with ESPript 3.2 [72]. To focus on the 37 and 187 residues, the full-length alignment was truncated to display only the regions including the targeted sites (amino acids 30–40 and 180–190). The specific phosphorylation residues corresponding to residues 37 and 187 in human SMN are indicated by red arrows. Sequences from primates are highlighted in yellow. HUMAN (Homo sapiens); MACFA (Macaca fascicularis, crab-eating macaque); PONAB (Pongo abelii, Sumatran orangutan); MOUSE (Mus musculus); RAT (Rattus norvegicus); BOVIN (Bos taurus, cattle); CANLF (Canis lupus familiaris, dog); FELCA (Felis catus, domestic cat); DANRE (Danio rerio, zebrafish); DROME (Drosophila melanogaster, fruit fly).*

1.4 Cleavage and degradation of Gemin3

Protein cleavage or degradation induced by energy depletion profoundly influences downstream signalling networks [73]. The expression of Gemin3 is highly associated with cellular energy levels (Fig. 18), suggesting that Gemin3 may serve as a target protein for cleavage or degradation to transduce the impact of energy stress on snRNP biogenesis.

1.4.1 The structure and function of Gemin3

Gemin3 is an RNA helicase characterised by a highly conserved N-terminal DEAD-box domain and a divergent C-terminus [47, 48]. The N-terminus endows Gemin3 with the capacity to unwind RNA, and the C-terminus serves as an interaction platform for its binding partners [47, 48]. Various proteins bind with Gemin3 through its C-terminus (Table 2), which allows Gemin3 to mediate multiple biological processes.

Table 2: ***Binding domains on Gemin3***

Processes	Binding proteins	Binding regions (aa)	Experimental method	Reference
snRNP biogenesis	SMN	456–547	*In vitro* binding assay	[38]
Chaperone	HspB8	475–727	Y2H	[74]
Transcription	Olig2	748–825	Y2H	[49]
	SF-1	721–825	M2H *In vitro* binding assay	[75]
	EBNA2; EBNA3C	666–824	Y2H Affinity chromatography	[76]
	FOXL2	542–824	Y2H	[77]
	METS	414–824	*In vitro* binding assay	[78]
Translation	NANOS1	792–825	Y2H	[79]

Notes: This table details the specific regions of Gemin3 required for binding various validated partner proteins. The interacting proteins are categorised according to their primary biological functions. Y2H, yeast two-hybrid; M2H, mammalian two-hybrid; HspB8, heat shock protein family B (small) member 8 (Hsp22); Olig2, oligodendrocyte transcription factor 2; SF-1, steroidogenic factor 1; EBNA2, Epstein-Barr virus nuclear antigen 2; EBNA3C, Epstein-Barr virus nuclear antigen 3C; FOXL2, forkhead box L2; METS, mitogenic Ets transcriptional suppressor; NANOS1, nanos homolog 1; aa, amino acids.

Full length (FL)-Gemin3 is required when it participates in snRNP biogenesis. Gemin3 is incorporated into the SMN complex by interaction with SMN through amino acids (aa) 456–547 [38], and structurally remodels the snRNA to facilitate the transfer of Sm proteins via the helicase domain in the N-terminus [46]. Previous research has shown that the cleaved Gemin3 induced by poliovirus infections is less efficient than FL in assembling snRNPs [80]. Overexpression of C-terminal Gemin3 in *Drosophila* mesoderm and larval muscles results in a destructive dominant-negative effect that causes severe metamorphic abnormalities, massive flight muscle degeneration, and

lethality during the pupal stage [81, 82]. This helicase-lacking mutation has higher affinity with FL-Gemin3 than the wild type (WT), by which endogenous Gemin3 might be sequestered into non-functional oligomers [83]. These observations indicate the necessity for FL-Gemin3 during snRNP biogenesis. Abnormal Gemin3 cleavage or degradation might perturb the assembly of snRNP.

However, some processes might be independent of the helicase property of Gemin3 [75, 84-86], which is conferred by its N-terminus. The C-terminus of Gemin3 is sufficient to interact with steroidogenic factor 1 (SF-1) and repress its transcriptional activity [75, 84, 85]. The helicase-dead Gemin3 mutant retains the ability to interact with transforming growth factor (TGF)-β–activated kinase-1 in the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway and induce breast cancer cell invasion, which is comparable to that of the WT. This suggests that the helicase domain is dispensable for Gemin3 to drive metastasis through the NF-κB pathway [86]. Considering these findings, Gemin3 cleavage fragments might be capable of participating in other biological processes.

1.4.2 Calpain-mediated cleavage and its implications for snRNP biogenesis

Calpains are a highly conserved family of calcium-dependent cysteine proteases that act as modulators in various cellular pathways by performing proteolytic cleavage to alter the activities of their substrates [87]. Calpain-1 (μ-calpain) is one of the most ubiquitous and widely studied conventional members of this family, which could be activated by micromolar (μM) concentrations of calcium [87]. Based on previous research, calpain may mediate snRNP biogenesis through the cleavage of SMN, Gemin3, Gemin5 and coilin [88-92].

The SMN protein is highly associated with calpain-mediated cleavage [88-91]. Recent research has revealed that the calpain system is hyperactivated in SMA motoneurons [88]. Inhibition of calpain activity enhances SMN protein levels and mitigates the symptoms associated with SMA [88, 89]. Indeed, previous studies from A. Gregory Matera's group have demonstrated that SMN could be cleaved by calpain-1 between S192 and phenylalanine (F) 193, generating ~28 kDa N-terminal and ~10 kDa C-terminal fragments [90, 91]. This cleavage is dependent on the structural integrity and specific amino acid sequence of SMN [91]. Interestingly, this cleavage is compartment-specific: the cytoplasmic SMN is readily cleaved, whereas the nuclear one is resistant to this process [91]. Given that the phosphorylation status of SMN is distinct between the cytoplasm and the nucleus [16], phosphorylation is a putative mechanism controlling the calpain-mediated cleavage of SMN [91]. The biological significance of this cleavage

event on SMN is not yet fully understood. In motor neurons, the calpain-mediated cleavage of SMN is regarded as detrimental [88, 89]. However, in muscle tissue, this cleavage might help SMN to remove from Z-disc, the boundary of the sarcomere, to maintain muscle homeostasis [90].

Among the proteins related to snRNP biogenesis, Gemin3, Gemin5 and coilin are the potential targets of calpains as well [90, 92]. While Gemin5 and coilin cleavage could occur in both purified systems and cell lysates, Gemin3 cleavage has only been observed to date in the purified complex [90]. Coilin may be cleaved by calpain-1 within the amino acid range of 475 to 576, consequently producing a fragment of ~28 kDa [92]. This cleavage event might be a mechanism for the translocation of coilin from Cajal bodies to the nucleolus [92].

1.4.3 The potential interplay between chaperone-mediated autophagy and the SMN complex

Autophagy is an intracellular quality control system that uses lysosomes to degrade target proteins [93]. This critical system operates primarily via three distinct pathways: macroautophagy, microautophagy and chaperone-mediated autophagy (CMA) [94]. Unlike macroautophagy and microautophagy, CMA translocates individual cytosolic proteins directly across the lysosomal membrane without requiring vesicles [94, 95]. The KFERQ-like motif (hereafter referred to as KFERQ motif) is the essential feature of CMA substrate proteins, which is specifically recognised by the cytosolic chaperone heat shock cognate 71 kDa protein (HSC70; HSPA8) [95]. Together with other co-chaperones, the substrate-HSC70 complex docks to the cytosolic tail of lysosome-associated membrane protein 2a (LAMP2A) on the lysosomal surface, and the substrate is subsequently internalised into the lysosomal lumen for degradation [95].

Emerging evidence inextricably links CMA to the pathogenesis of SMA. Recent research has identified a missense mutation on HSC70, G470R, which acts as a powerful suppressor of SMA disease by elevating functional SMN protein levels [96, 97]. HSC70 is consistently associated with the SMN complex in several studies [18, 51, 97-99]. The interaction with HSC70 requires the Tudor, Gemin-binding, and proline-rich domains on SMN [97]. Other co-chaperones (Bag3 and HspB8) and autophagy adaptors (p62) are also incorporated into this SMN-HSC70 complex [97]. These findings strongly suggest that the SMN complex is subjected to CMA degradation.

1.5 Aims of thesis

This thesis aims to understand how cellular energy levels regulate snRNP biogenesis through post-translational modification of the SMN complex, particularly of SMN and Gemin3 (Fig. 3). Cells were treated with Earle's balanced salt solution (EBSS) and bafilomycin A1 (BafA1) to reduce the extra- or intracellular energy levels. The interaction of the SMN and Gemin3 was examined by immunoprecipitation (IP), size exclusion chromatography (SEC) and pull-down assay. The capability of snRNP assembly was indicated by the average number of Cajal bodies in the overexpression cells detected via immunofluorescence (IF).

Chapters 3.1 and 3.2 focus on SMN phosphorylation. Recent research suggested that two energy-sensing kinases, ULK1 and mTOR, participate in snRNP biogenesis [23-25, 55]. The association between these two kinases and the SMN complex were examined in Chapter 3.1 under different nutrient conditions. To further understand the energy-regulated phosphorylation events, the regulatory role of SMN phosphorylation at T37 and S187 in the complex interaction and snRNP assembly was studied in Chapter 3.2.

Chapters 3.3 and 3.4 investigate the cleavage and degradation of Gemin3. During the screening, cleaved Gemin3 was observed under basal conditions, and its expression pattern was highly sensitive to energy depletion induced by EBSS and BafA1 (Fig. 18). The mechanisms for Gemin3 cleavage and the biological function of the cleavage fragments in snRNP biogenesis were explored in Chapter 3.3. Since Gemin3 harbours KFERQ motifs, the marker of CMA substrates [95], the association between CMA and Gemin3 degradation was demonstrated in Chapter 3.4.

Disorders in energy metabolism are linked to various MNDs [6], such as SMA and ALS [9-12]. Given that SMN and Gemin3 act as crucial mediators in SMA due to their role in snRNP biogenesis [17], understanding the behaviour of these two proteins under different cellular energy levels will provide insights into the pathogenesis of SMA and other MNDs, thereby offering novel strategies for therapeutic intervention.

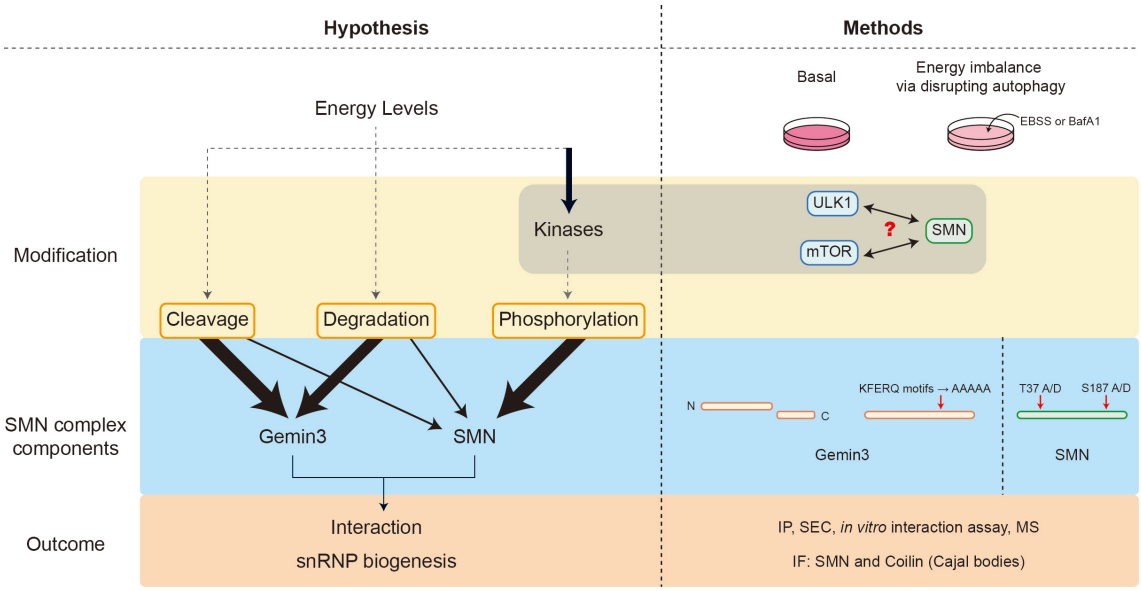


Fig. 3: ***The hypothesis and research methods in this project.*** *Black lines denote the hypotheses examined in the present study, with line thickness proportional to the depth of investigation. Grey dashed lines indicate putative regulatory steps requiring further elucidation.*

2 Materials and Methods

2.1 Materials

2.1.1 Reagents

2.1.1.1 Media and reagents for cell or bacterial culture

Product	*Manufacturer*	*Cat. No.*
Dulbecco's phosphate-buffered saline (PBS)	Gibco	14190144
Dulbecco's modified Eagle medium (DMEM) (4.5 g/L D-glucose)	Gibco	41965-039
Foetal bovine serum (FBS)	Sigma-Aldrich	F9665
Tetracycline-free FBS (Tet-Free FBS)	Gibco	A4736401
Penicillin-Streptomycin (10,000 U/mL)	Gibco	15140-122
Earle's balanced salt solution (EBSS)	Gibco	24010043
Opti-MEM™ reduced serum medium (Opti-MEM)	Gibco	31985062
DharmaFECT™ 1 Transfection Reagent	GE Dharmacon	T-2001-02
FuGENE® 6 Transfection Reagent	Promega	E2691
MISSION® esiRNA targeting human LAMP2 (LAMP2 siRNA)	Sigma-Aldrich	EHU065091-50UG
ON-TARGET*plus*™ Non-targeting control pool	GE Dharmacon	D-001810-10-20
Trypan blue stain, 0.4%	Gibco	15250-061
Lysogeny broth (LB) base powder	Sigma-Aldrich	L3022-1KG
Super Optimal broth with Catabolite repression (S.O.C.) medium	Thermo Fisher Scientific	15544034

2.1.1.2 Chemicals

Product	*Manufacturer*	*Cat. No.*
Bafilomycin A1 (BafA1)	Sigma-Aldrich	B1793
MG132	Selleck Chemicals	S2619
Calpeptin	Santa Cruz Biotechnology	sc-202516

Product	*Manufacturer*	*Cat. No.*
Doxycycline (DOX)	Sigma-Aldrich	05207-5G
Ampicillin	Carl Roth	K029.1
Kanamycin	Carl Roth	T832.2
Hygromycin B Gold	InvivoGen	ant-gh-1
Blasticidin	InvivoGen	ant-bl-05
Dimethyl sulfoxide (DMSO)	AppliChem GmbH	A3672
L-[35S]-Methionine	Hartmann Analytic	SRM-01
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich	I5502-10G
Methanol	VWR International	20847.32
Acetic acid	Carl Roth	3738.5
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich	74255-250G
Sodium deoxycholate	Sigma-Aldrich	30970-100G
Bromophenol blue sodium salt	Carl Roth	A512.1
β-Mercaptoethanol	Carl Roth	4227.1
Paraformaldehyde (PFA)	Carl Roth	335.3
Triton™ X-100	Carl Roth	3051.2
Glycerol	Carl Roth	3783.2
Tween® 20	Sigma-Aldrich	P1379-500ML
IGEPAL® CA-630 (NP-40 substitute)	Sigma-Aldrich	I3021-100ML
Bovine serum albumin (BSA)	Carl Roth	8076.4
Powdered milk (Blotting grade)	Carl Roth	T145.3
Ethylenediaminetetraacetic acid (EDTA) disodium salt dihydrate	Carl Roth	8043.2
Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA)	Amresco (VWR International)	0732-100G

Product	*Manufacturer*	*Cat. No.*
Dithiothreitol (DTT)	Thermo Fisher Scientific	R0861
Adenosine 5'-triphosphate (ATP) solution (100 mM)	Thermo Fisher Scientific	R0441
Glycine	VWR International	101196X
Tris	Carl Roth	4855.2
Sodium chloride (NaCl)	Carl Roth	9265.2
Potassium chloride (KCl)	Carl Roth	P017.1
Magnesium chloride hexahydrate ($MgCl_2 \cdot 6H_2O$)	Carl Roth	2189.1
Calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$)	Carl Roth	HN04.2
ROTIPHORESE® Gel 30 (37.5:1) acrylamide solution	Carl Roth	3029.1
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Carl Roth	2367.1
Ammonium persulfate (APS)	Carl Roth	9592.3
Hydrochloric acid 37%	VWR International	20252.244
Agar	Sigma-Aldrich	A7002
Coomassie® brilliant blue G 250	Carl Roth	9598.2
Amido Black 10B	Carl Roth	9590.1
4’,6-Diamidino-2-phenylindole (DAPI) dihydrochloride	Carl Roth	6335.1

2.1.1.3 Commercial reagents

Product	*Manufacturer*	*Cat. No.*
Gibson Assembly® 2× Master Mix	New England Biolabs	M5510AVIAL
Q5® Hot Start High-Fidelity 2× Master Mix	New England Biolabs	M0494S
KLD Enzyme Mix	New England Biolabs	M0554S

Product	*Manufacturer*	*Cat. No.*
Fluorescein RNA Labelling Mix	Roche	11685619910
TNT® T7 Quick Master Mix	Promega	L520A
cOmplete™ Protease Inhibitor Cocktail	Roche	04693132001
ProLong™ Glass Antifade Mountant	Invitrogen	P36984
MIDORI^{Green} Xtra Stain	NIPPON Genetics	MG10
PageRuler™ prestained protein ladder	Thermo Fisher Scientific	26616
GeneRuler™ 1 kb DNA ladder	Thermo Fisher Scientific	SM0311
GeneRuler™ 100 bp Plus DNA ladder	Thermo Fisher Scientific	SM0321
6× DNA loading dye	Thermo Fisher Scientific	R0611
4-(2-hydroxyethyl)-1-piperazineethanesul-fonic acid (HEPES) solution	Sigma-Aldrich	H0887-100ML
Nuclease-free water	Promega	P119A
G150 X-ray film developer	Agfa	EFBAK
G354 X-ray film fixer	Agfa	2828Q

2.1.1.4 Commercial kits

Kit	*Manufacturer*	*Cat. No.*
NucleoBond® Xtra Maxi Kit	Macherey-Nagel	740424.10
NucleoBond® Xtra Midi Kit	Macherey-Nagel	740420.50
NucleoSpin® Plasmid Kit	Macherey-Nagel	740588.250
Wizard® SV Gel and PCR Clean-Up System	Promega	A9282
Protein Assay Kit II	Bio-Rad	5000002

2.1.1.5 Enzymes

Enzyme	*Manufacturer*	*Cat. No.*
FastDigest Pfl23II	Thermo Fisher Scientific	FD0854
DpnI restriction endonuclease	New England Biolabs	R0176S
Lysozyme	SERVA Electrophoresis	28263
Caspase-3	N/A	In-house
Calpain-1	Sigma-Aldrich	208712-1MG

2.1.2 Antibodies

2.1.2.1 Primary antibodies

Primary antibodies were used at the following dilutions: 1:1000 for immunoblotting (IB), 1:250 for immunofluorescence (IF), and 1:200 for immunoprecipitation (IP).

Primary antibody	*Manufacturer*	*Cat. No.*	*Application*
mTOR	Cell Signaling Technology	2972	IB
ULK1 (D8H5)	Cell Signaling Technology	8054	IB
ATG13	Sigma-Aldrich	SAB4200100	IB
SMN (2B1)	Sigma-Aldrich	05-1532	IB, IP and IF
Gemin2 (E-7)	Santa Cruz Biotechnology	sc-166162	IB
Gemin3 (12H12)	Santa Cruz Biotechnology	sc-57007	IB
Gemin3 (2)	Santa Cruz Biotechnology	sc-135919	IB
Gemin3 (C-5)	Santa Cruz Biotechnology	sc-374373	IB
Gemin3 (D-5)	Santa Cruz Biotechnology	sc-271853	IB
Gemin4 (E-8)	Santa Cruz Biotechnology	sc-365424	IB

Primary antibody	*Manufacturer*	*Cat. No.*	*Application*
Gemin5	Proteintech	24897-1-AP	IB
Gemin6	Proteintech	12307-2-AP	IB
Gemin7	Invitrogen	PA5-116612	IB
Gemin8 (1F8)	Santa Cruz Biotechnology	sc-130669	IB
Unrip		In-house	IB and IP
SmB (Y12)	Invitrogen	MA5-13449	IB
SmD2	Sigma-Aldrich	SAB2102257	IB
SmD3 [EPR7676]	Abcam	ab128857	IB
SmE	Novus Biologicals	NBP2-43792	IB
SmF	Sigma-Aldrich	SAB2102258	IB
SmG	Invitrogen	PA5-49365	IB
pICln (C-5)	Santa Cruz Biotechnology	sc-393525	IB
WD45	Cell Signaling Technology	2823	IB
PRMT5	Cell Signaling Technology	2252	IB
Coilin	Invitrogen	PA5-100038	IF
HSC70	Santa Cruz Biotechnology	sc-7298	IB
LAMP2 (GL2A7)	Abcam	ab13524	IF
LAMP2A [EPR4207(2)]	Abcam	ab125068	IB
GFP (3H9)	ChromoTek	3h9-100	IB
GAPDH (6C5)	Abcam	ab8245	IB
β-actin	Sigma-Aldrich	A5316	IB

2.1.2.2 Secondary antibodies

Secondary antibodies were used at the following dilutions: 1:1000 for immunoblotting (IB) and 1:200 for immunofluorescence (IF).

Secondary Antibody	*Manufacturer*	*Cat. No.*	*Application*
IRDye® 680RD Donkey anti-Rabbit IgG	LI-COR Biosciences	926-68073	IB
IRDye® 800CW Donkey anti-Rabbit IgG	LI-COR Biosciences	926-32213	IB
IRDye® 680RD Donkey anti-Mouse IgG	LI-COR Biosciences	926-68072	IB
IRDye® 800CW Donkey anti-Mouse IgG	LI-COR Biosciences	926-32212	IB
IRDye® 680RD Goat anti-Rat IgG	LI-COR Biosciences	926-68076	IB
IRDye® 800CW Goat anti-Rat IgG	LI-COR Biosciences	926-32219	IB
Alexa Fluor™ 568 Donkey anti-Mouse IgG	Invitrogen	A10037	IF
Alexa Fluor™ 633 Goat anti-Rabbit IgG	Invitrogen	A21070	IF

2.1.3 Buffers

Unless otherwise indicated, all buffers were prepared using ultrapure Milli-Q water. Stock solutions of Tris-buffered saline / Tween20 (TBST), wet transfer buffer, running buffer and TAE were diluted to a 1× working condition using water of at least distilled purity.

2.1.3.1 For immunoblotting

2× Laemmli buffer	125 mM Tris-HCl (pH 7.5), 17.2% (v/v) Glycerol, 4.1% (w/v) SDS, 0.02% (w/v) Bromophenol blue, 2% (v/v) β-Mercaptoethanol; pH 6.8
6× Laemmli buffer	375 mM Tris-HCl (pH 7.5), 25.8% (v/v) Glycerol, 12.3% (w/v) SDS, 0.06% (w/v) Bromophenol blue, 6% (v/v) β-Mercaptoethanol; pH 6.8
10× Laemmli buffer	209 mM Tris-HCl (pH 7.5), 41% (w/v) Glycerol, 7.7% (w/v) SDS, 0.003% (w/v) Bromophenol blue, 10% (v/v) β-Mercaptoethanol; pH 6.8
10× Tris-buffered saline / Tween20 (TBST)	100 mM Tris, 1.5 M NaCl, 0.5% (v/v) Tween® 20; pH 7.6
10× Wet transfer buffer	436 mM Tris, 390 mM Glycine
10× Running buffer	250 mM Tris, 1.9 M Glycine, 1% (w/v) SDS
5% Milk blocking buffer	5% (w/v) Powdered milk in 1× TBST

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels	30% Acrylamide/Bis-acrylamide, Tris-HCl, SDS, APS, TEMED. Proportions adjusted for the required gel percentage based on an established in-house protocol

2.1.3.2 For immunoprecipitation

Washing buffer	10 mM KCl, 10 mM HEPES, 1.5 mM $MgCl_2$, 150 mM NaCl

2.1.3.3 For immunofluorescence

4% PFA fixative solution	4% (w/v) PFA in 1×PBS
0.2% Triton X-100 permeabilization buffer	0.2% (v/v) Triton™ X-100 in 1×PBS
5% BSA blocking buffer	5% (w/v) BSA in 1×PBS

2.1.3.4 For in vitro interaction assay

Interaction buffer	300 mM NaCl, 50 mM Tris-HCl (pH 7.5), 1 mM EGTA, 1 mM EDTA, 1 mM DTT, and 0.01% (v/v) IGEPAL® CA-630 (NP-40)

2.1.3.5 For cellular protein extraction

All extraction buffers were supplemented with the cOmplete™ Protease Inhibitor Cocktail (1 tablet per 50 mL of buffer) immediately prior to use.

Roeder A buffer	10 mM KCl, 10 mM HEPES, 1.5 mM $MgCl_2$
RIPA buffer	150 mM NaCl, 50 mM Tris-HCl (pH 8.0), 1% (v/v) IGEPAL® CA-630 (NP-40), 0.5% (w/v) Sodium deoxycholate, 0.1% (w/v) SDS

2.1.3.6 For size exclusion chromatography

SEC running buffer	150 mM NaCl, 50 mM Tris-HCl (pH 7.5)

2.1.3.7 For protein purification

The lysis buffer was supplemented with the cOmplete™ Protease Inhibitor Cocktail (1 tablet per 50 mL of buffer) immediately prior to use.

Bacterial lysis buffer	300 mM NaCl, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 5 mM EGTA, 0.01% (v/v) IGEPAL® CA-630 (NP-40)

2.1.3.8 For cell culture

DMEM general medium	10% (v/v) FBS, 1% (v/v) Penicillin-Streptomycin in DMEM
DMEM selection medium	10% (v/v) Tet-free FBS, 1% (v/v) Penicillin-Streptomycin, 400 µg/mL Hygromycin, 4 µg/mL Blasticidin in DMEM
Antibiotic-free DMEM medium	10% (v/v) FBS in DMEM

2.1.3.9 Others

Amido Black staining solution	40% (v/v) Methanol, 10% (v/v) Acetic acid, 0.1% (w/v) Amido Black 10B
Coomassie blue staining solution	0.5% (w/v) Coomassie brilliant blue, 50% (v/v) Methanol, 10% (v/v) Acetic acid
Coomassie blue destaining solution	20% (v/v) Methanol, 10% (v/v) Acetic acid
50× TAE buffer	2 M Tris, 1 M Acetic acid, 50 mM EDTA; pH 8.5
Agarose gels	0.5–1% (w/v) agarose and 3–5 µL MIDORI^{Green} Xtra Stain in 1× TAE buffer.

2.1.4 Plasmids

Plasmid	*Source*
pUCBU1-snRNA T7	In-house
pET100-SmD1	synthesised by GeneArt
pET28a-SmB	In-house
pET28a-SmD3	In-house
pET28a-Gemin3	In-house
pGEX-6P-1-SMN	synthesised by GeneArt
pcDNA5-FRT-TO-eGFP-SMN	In-house
pcDNA5-FRT-TO-eGFP	In-house
pOG44	Obtained from Invitrogen (V6005-20)

2.1.5 Oligonucleotides for cloning

Custom oligonucleotide primers utilised for generating desired plasmids were synthesised by Sigma-Aldrich at a 0.025 µmol scale, purified by standard desalting, and delivered as a 100 µM working solution in molecular biology grade water.

Primer name	*Sequence*
SMN T37A	Fwd: 5'-GCAGCACTGATAAAAGCATATGAT-3' Rev: 5'-ATCATCCCAAATGTCAGAATC-3'
SMN T37D	Fwd: 5'-GACGCACTGATAAAAGCATATGAT-3' Rev: 5'-ATCATCCCAAATGTCAGAATC-3'
SMN S187A	Fwd: 5'-GCAGCTCCATGGAACTCTTTTC-3' Rev: 5'-TTTGGGCTTGATGTTATCTGA-3'
SMN S187D	Fwd: 5'-GACGCTCCATGGAACTCTTTTC-3' Rev: 5'-TTTGGGCTTGATGTTATCTGA-3'
Gemin3 Insert	Fwd: 5'-GTCCGGAATGGCGGCGGCATTTGAAG-3' Rev: 5'-TCAGCGGTCACTGGTTACTATGCATCATTTCT TGTAG-3'
Gemin3 Vector	Fwd: 5'-ACCAGTGACCGCTGATCAGCCTCGAC-3' Rev: 5'-GCCGCCATTCCGGACTTGTACAGCTC-3'
Gemin3 N	Fwd: 5'-TGACCGCTGATCAGC-3' Rev: 5'-TTTCTGAATTTGCTTTTTTAAGG-3'
Gemin3 C	Fwd: 5'-ATAGAGAGAACCCTTCAAATTC-3' Rev: 5'-TCCGGACTTGTACAG-3'
Gemin3 MUT-1	Fwd: 5'-CGGCGGCGACCCTTCAAATTCAGAAAG-3' Rev: 5'-CCGCCGCAATTTGCTTTTTTAAGGGTTG-3'
Gemin3 MUT-2	Fwd: 5'-CGGCGGCGAGAGTCCCTGTGTTG-3' Rev: 5'-CCGCCGCATTTCTCACAAAACCATTTTG-3'

Notes: Fwd, forward primers; Rev, reverse primers.

2.1.6 Bacteria

Strain	*Genotype*
ArcticExpress (DE3) competent *Escherichia coli (E. coli)*	B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte [*cpn10 cpn60* $Gent^r$]
DH5α competent *E. coli*	F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG* Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *hsdR17*(r_K^- m_K^+), λ–

2.1.7 Cell lines

Cell lines	*Manufacturer*	*Cat. No.*
Human embryonic kidney (HEK293T) cell	ATCC	CRL-1573™
Henrietta Lacks (HeLa) cell	ATCC	CCL-2™
Flp-In™ T-REx™ 293 cell	Invitrogen	R78007

2.1.8 Equipment and consumables

2.1.8.1 General consumables

All cell culture consumables were supplied by Sarstedt. The gel filtration columns marked with an asterisk (*) have been calibrated with the standard proteins from the calibration kits (GE Healthcare, 17-0442-01 and 17-0441-01) as described in previous studies [23, 24, 100].

Consumable	*Manufacturer*	*Cat. No.*
Protein G Sepharose™ 4 Fast Flow resin	Cytiva	17061801
GFP-Trap® agarose beads	ChromoTek	gta-20
Glutathione Sepharose™ 4B	Cytiva	17-0756-01
Superose™ 6 Increase 10/300 GL column*	Cytiva	29091596
Superdex™ 200 Increase 10/300 GL column*	Cytiva	28990944
WHEATON® dounce tissue grinder (1mL)	DWK life sciences	357538
Millex®-HA syringe filter unit (0.45µm, 33mm)	Merck Millipore	SLHA033SB
Blotting pads	VWR International	732-0594 732-0597
Immobilon®-FL Polyvinylidene difluoride (PVDF) membrane	Merck Millipore	IPFL00010
Amersham™ Hyperfilm™ MP, Autoradiography film	Cytiva	28906844
Microscope slides	Paul Marienfeld	0810000
Cover glasses	Paul Marienfeld	0117520

Consumable	*Manufacturer*	*Cat. No.*
Immersol™ 518 F Fluorescence-free immersion oil	Carl Zeiss	444960-0000-000

2.1.8.2 Equipment

Equipment	*Manufacturer*	*Cat. No. or Model*
Refrigerated microcentrifuge	Eppendorf	5417R
Multifuge™ centrifuge	Thermo Fisher Scientific	3SR+
Avanti high-speed centrifuge	Beckman Coulter	JXN-26
Monochromator-Based Multi-Mode Microplate Reader	BioTek	Synergy™ Mx
Vortex-Genie® 2 mixer	Scientific Industries	Vortex-Genie 2
Tube rotator	neoLab	N/A
Digital block heater HX-1	PEQLAB	91-D1100
Thermomixer comfort	Eppendorf	5355
Innova® 44 incubator shaker	Eppendorf	Innova 44
HERAcell CO_2 Incubator	Thermo Fisher Scientific	VIOS 250i
Class II biological safety cabinet	Thermo Fisher Scientific	Safe 2020
BioPhotometer	Eppendorf	6131
Thermocycler	Biometra	T3000
UV transilluminator	Nippon Genetics	FAS-Digi
Owl™ D2 wide gel electrophoresis system	Thermo Fisher Scientific	D2-BP
NanoDrop™ 1000 spectrophotometer	Thermo Fisher Scientific	ND-1000
ÄKTApurifier UPC 100	Cytiva	28406271
Electrophoresis power supply	Amersham Biosciences	EPS 1001

Equipment	*Manufacturer*	*Cat. No. or Model*
Wide-Mini vertical protein electrophoresis unit	Hoefer	SE640
Trans-Blot® electrophoretic transfer cell	Bio-Rad	1703910
Mini-PROTEAN® tetra vertical electrophoresis cell	Bio-Rad	1658004
Mini Trans-Blot® electrophoretic transfer cell	Bio-Rad	1703930
Odyssey® CLx imaging system	LI-COR Biosciences	9140
Primovert inverted microscope	Carl Zeiss	N/A
Axiocam 208 color microscope camera	Carl Zeiss	N/A
Axio Observer inverted microscope	Carl Zeiss	N/A
ApoTome.2 optical sectioning system	Carl Zeiss	N/A
Plan-APOCHROMAT 40×/1.40 oil DIC objective	Carl Zeiss	N/A
Gel dryer	Bio-Rad	Model 583
Digital camera	Panasonic	DMC-GF7K
Flatbed scanner	Canon	CanoScan 9950F

2.1.9 Computational tools

2.1.9.1 Software and R packages

Software	*Developer/Organization*	*Version*	*Reference*
Adobe illustrator CC 2019	Adobe Inc.	23.0.1	
GraphPad Prism	GraphPad Software	8.0.1	
Fiji (ImageJ)	National Institutes of Health	1.54p	
R software	R Foundation for Statistical Computing	4.5.1	
RStudio	Posit Software, PBC	2025.9.1.401	

Software	*Developer/Organization*	*Version*	*Reference*
tidyverse	N/A	2.0.0	[101]
ComplexUpset	N/A	1.3.3	[102, 103]
clusterProfiler	N/A	4.18.4	[104]

2.1.9.2 Webservers

Webserver	Website	Reference
Clustal Omega 1.2.4	https://www.ebi.ac.uk/jdispatcher/msa/clustalo	[70, 71]
ESPript 3.2	https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi	[72]
DeepCalpain	http://deepcalpain.omicsbio.info/	[105]
KFERQ finder V0.8	https://rshine.einsteinmed.edu/	[106]

2.2 Methods

2.2.1 Molecular biological methods

Throughout this section, commercial reagents and kits are denoted by title case (e.g., Q5® High-Fidelity 2× Master Mix). All procedures involving these reagents and kits were performed according to their respective manufacturer's instructions.

2.2.1.1 Plasmid construction

pcDNA5-FRT-TO-eGFP-Gemin3 wild type (WT) was prepared through Gibson assembly. Both insert encoding Gemin3 WT (cloned from pET28a-Gemin3) and pcDNA5-FRT-TO-eGFP backbone were amplified by polymerase chain reaction (PCR) using the Q5® Hot Start High-Fidelity 2× Master Mix. Following DpnI treatment for 30 minutes (min) at 37°C, the PCR products of the Gemin3 insert and linearised pcDNA5-FRT-TO-eGFP backbone were combined at a 2:1 molar ratio and incubated with the Gibson Assembly® 2× Master Mix for 1 hour (h) at 50°C.

pGEX-6P-1 and pcDNA5-FRT-TO constructs encoding GFP-tagged SMN (T37 A/D, S187 A/D, AA and DD) or Gemin3 (aa 1-478, aa 479-825, MUT-1, MUT-2 and 2MUTs) variants were generated by site-directed mutagenesis. Plasmids containing the desired mutations were cloned from corresponding WT templates using target-specific mutagenic primers (Section 2.1.5) and the Q5® Hot Start High-Fidelity 2× Master Mix.

The resulting PCR products were treated with the KLD Enzyme Mix for at least 5 min at RT.

The desired constructs were selected and amplified on a large scale according to the procedures described in Sections 2.2.1.2–2.2.1.4.

2.2.1.2 Heat shock transformation

Plasmid DNA or PCR/Gibson assembly products were transformed into chemically competent *E. coli* strains using the heat-shock protocol. The cloning strain DH5α was used for plasmid preparation, while the ArcticExpress (DE3) strain was employed for downstream protein expression.

An aliquot of chemically competent cells (50 µL) was gently mixed with approximately 1–5 µL of plasmid DNA or Gibson assembly product, and then heat-shocked at 42°C for 90 seconds. Subsequently, the cell-DNA mixture was supplemented with 600 µL of antibiotic-free S.O.C. medium and incubated for 1–2 h at 37°C with shaking. After the cells were pelleted by centrifugation, 500 µL of the supernatant was discarded and the bacteria were resuspended in the remaining 100 µL of medium. The suspensions were plated onto a LB agar containing appropriate selection antibiotics (100 µg/mL ampicillin or 50 µg/mL kanamycin). The bacteria on the plate were cultured at 37°C overnight.

2.2.1.3 Plasmid preparation

For the initial screening of successful constructs, single colonies from Section 2.2.1.2 were inoculated into the liquid LB medium with appropriate selection antibiotics (100 µg/mL ampicillin or 50 µg/mL kanamycin), and plasmid DNA was isolated using the NucleoSpin® Plasmid Kit. For large-scale plasmid preparation, verified clones were expanded in LB medium, and purified utilising the NucleoBond® Xtra Maxi or Midi Kit.

Plasmid concentrations were measured using the NanoDrop™ 1000 Spectrophotometer. The structural and sequence integrity of the final constructs were verified by agarose gel electrophoresis (Section 2.2.1.4) and standard Sanger sequencing (Microsynth Seqlab GmbH), respectively.

2.2.1.4 Agarose gel electrophoresis

To resolve plasmids or DNA fragments, samples were mixed with 6× DNA loading dye and loaded onto an agarose gel. Electrophoresis was performed at 110 V for 45 min in 1× TAE buffer using a standard horizontal gel electrophoresis system. Following electrophoretic separation, the resolved DNA signals were visualised using a UV transilluminator.

2.2.2 Cellular methods

2.2.2.1 Cell culture and treatment

All cell lines in this thesis were cultured at 37°C with 5% CO_2 in a humidified environment. For routine maintenance, the cells were cultured in DMEM general medium and passaged every 2–3 days at a 1:10 split ratio.

Before treatment, cells seeded in the dishes were washed with PBS. To induce starvation or disrupt proteolysis, the medium was replaced with either EBSS or fresh general medium supplemented with inhibitors (25 nM BafA1, 1–3 μM MG132 or 12.5–50 μM Calpeptin). Corresponding control groups received fresh general medium alone (for EBSS) or the medium containing an equivalent volume of DMSO matched to the highest experimental concentration (for inhibitors). Cells were then incubated for the indicated durations prior to harvesting (Section 2.2.2.2).

2.2.2.2 Cell harvesting and protein extraction

After washing with PBS, cells were harvested by gently scraping into fresh PBS. The cells were pelleted by centrifugation at 1200 rounds per minute (rpm) for 10 min at 4°C.

To prepare cytoplasm extracts, cell pellets were resuspended in Roeder A buffer (3 mL per 1 g of pellet mass) by pipetting up and down. After incubation for 10 min at room temperature (RT), the cell suspensions were homogenized using a dounce homogenizer (10 strokes). These homogenates were adjusted to 150 mM NaCl and then centrifuged at 13500 rpm for 30 min at 4°C.

To extract total protein, cell pellets were resuspended in RIPA buffer (3 mL per 1 g of pellet mass). The lysates were incubated on ice for 30 min with brief vortexing every 10 min, followed by centrifugation at 13500 rpm for 30 min at 4°C.

All harvesting and extraction procedures were performed strictly on ice using pre-chilled tubes and buffers to prevent protein degradation. The resulting supernatants containing cytoplasmic (RoederA) or total soluble (RIPA) proteins were collected for subsequent experiments. Protein concentrations of the cleared extracts were determined by Bradford assay using Protein Assay Kit II.

2.2.2.3 siRNA transfection

To knockdown LAMP2A, HEK293T cells were transfected with 50 nM of LAMP2 siRNA or non-targeting control using DharmaFECT™ 1 Transfection Reagent.

HEK293T cells were seeded into 12-well plates and cultured in 800 μL antibiotic-free general medium till 60–70% confluence. 2.5 μL 20 μM siRNA and 2 μL DharmaFECT1

reagent were diluted separately to 100 µL using Opti-MEM. After incubation for 5 min at RT, two dilutions were combined and incubated for 30 min at RT to allow complex formation. 200 µL of the resulting transfection complex suspension was added dropwise to the cell culture medium. Cells cultured in 1 mL of antibiotic-free general medium served as a negative control.

The cells were incubated with the siRNA mixture/antibiotic-free medium for 24 h, after which they were allowed to recover in general medium for an additional 24 h. Following this incubation, whole-cell lysates were prepared using RIPA buffer (Section 2.2.2.2). The siRNA knockdown efficiency was subsequently determined by assessing LAMP2A levels via immunoblotting (Section 2.2.3.3).

2.2.2.4 Generation of inducible Flp-In™ T-REx™ 293 cell lines

To generate stable cell lines overexpressing SMN or Gemin3 variants, the pcDNA5-FRT-TO plasmids containing the gene of interest (Section 2.2.1.1) was co-transfected with pOG44 into Flp-In™ T-REx™ 293 cells at a 1:9 ratio using the FuGENE® 6 Transfection Reagent.

Cells were grown in a 100-mm culture dish containing 7 mL of antibiotic-free DMEM medium. Transfection was performed when the cells reached 60–70% confluence. A total of 5 µg of plasmid DNA, consisting 0.5 µg of pcDNA5-FRT-TO plasmid containing the gene of interest and 4.5 µg of pOG44, was diluted in 230 µL Opti-MEM. 15 µL of FuGENE® 6 Transfection Reagent was gently mixed with diluted DNA by pipetting up and down, and incubated for 15 min at RT to form DNA-lipid complexes which were subsequently distributed evenly into the antibiotic-free medium. A mock transfection (pOG44 alone) was performed in parallel as a negative control.

The cells were exposed to transfection medium for 24 h and then 10 mL general medium for additional 24 h. After the incubation, the medium was replaced with 10 mL of fresh selection medium. The selection medium was replenished every 3–4 days until antibiotic-resistant colonies emerged. The resulting colonies were pooled and maintained in the general medium as described in Section 2.2.2.1.

To validate the cell lines and optimize induction times, cells were treated with 0.1 µg/mL DOX for various durations, and the levels of the overexpressed GFP-tagged proteins were assessed via immunoblotting (Sections 2.2.2.2 and 2.2.3.3). Based on these optimizations, Flp-In™ T-REx™ 293 cell lines were treated with 0.1 µg/mL DOX overnight (16–18 h) to induce protein overexpression before downstream procedures throughout the present study.

2.2.3 Biochemical methods

2.2.3.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Protein samples were diluted with Laemmli buffer and denatured by heating at 100°C for 10 min. Depending on the molecular weight of the target protein, samples were resolved on 10%, 12%, or 15% SDS-PAGE gels alongside molecular weight markers. Electrophoresis was performed in 1× running buffer at a constant current for ~1 h, until the bromophenol blue dye front reached the bottom of the gel.

2.2.3.2 Coomassie blue staining

After electrophoresis (Section 2.2.3.1), the SDS-PAGE gel was carefully submerged in Coomassie staining solution for 5 min at RT with gentle rocking, and subsequently transferred to destaining solution. Destaining was performed with continuous agitation for a minimum of 2 h until proteins were clearly visible. The destained gel was imaged with a flatbed scanner.

2.2.3.3 Immunoblotting

Following SDS-PAGE (Section 2.2.3.1), resolved proteins were transferred onto a methanol-activated PVDF membrane using a wet transfer system at a constant current (600 mA or 650 mA) or voltage (110 V or 120 V) at 4°C. The membrane was blocked in 5% milk blocking buffer for 1 h at RT, and then probed with primary antibodies overnight at 4°C. After washing three times (1× TBST, 5 min), the membrane was incubated with secondary antibodies for 1 h at RT in the dark. Following three additional washes (1× TBST, 5 min), the membrane was dried between blotting pads. The target proteins visualised by the Odyssey® CLx imaging system.

2.2.3.4 Amido Black staining

After immunoblotting (Section 2.2.3.3), the entire PVDF membrane was incubated with the Amido Black staining solution for 1–5 min at RT. After a brief wash with double-distilled water, the membrane was imaged by a flatbed scanner.

2.2.3.5 Protein expression and purification

pGEX-6P-1 plasmids encoding the recombinant GST-tagged SMN WT and phosphomutant proteins were transformed into ArcticExpress (DE3) competent *E. coli* as described in Section 2.2.1.2. A single colony for each variant was inoculated into liquid LB medium and grown at 37°C to an optical density (OD_{600}) of ~0.6. Protein expression was then induced with 1 mM IPTG at 14°C overnight.

Cells were pelleted by centrifugation at 5000 rpm for 5 min at 4°C, and then resuspended in lysis buffer supplemented with 0.5 mg/mL lysozyme. After incubation for 30 min at 4°C with gentle rotation, the cells were disrupted via sonication on ice, and the resulting homogenates were centrifuged at 36000 rpm for 30 min at 4°C. The clarified lysates were collected and incubated with Glutathione Sepharose™ 4B beads for at least 2 h at 4°C under continuous rotation. The beads were subsequently washed with lysis buffer and pelleted by centrifugation at 4600 rpm for 5 min at 4°C to remove non-specific binding proteins.

To confirm the successful purification of the GST-SMN variants, bead-bound proteins were resolved by SDS-PAGE, followed by Coomassie blue staining (Section 2.2.3.2) and immunoblotting analysis (Section 2.2.3.3). To ensure equal protein loading for each variant across the pull-down assays (Section 2.2.3.8 and 2.2.3.9), the relative concentrations of the GST-SMN variants were semi-quantitatively assessed via Coomassie blue staining (Section 2.2.3.2).

2.2.3.6 Size exclusion chromatography

Cytoplasmic extracts prepared following the protocol in Section 2.2.2.2. were fractionated by the Superose 6 or Superdex 200 increase 10/300 GL column connected to the ÄKTApurifier system. After the column was equilibrated with at least 24 mL SEC running buffer, 1 mL of cytoplasm extract (~4 µg/µL) passed through a 0.45 µm Millex-HA filter was injected onto the column. The separation was performed at a flow rate of 0.5 mL/min. Aliquots from each fraction were either supplemented with 10× Laemmli buffer for immunoblotting (Section 2.2.3.3) or retained for co-immunoprecipitation (co-IP) assays (Section 2.2.3.7).

2.2.3.7 Immunoprecipitation

Following established procedures [23, 24, 100], the co-IP was performed to validate the interactions between the proteins of interest. Cytoplasmic extracts from GFP-tagged overexpression cells or HEK293T cells were prepared as described in Section 2.2.2.2. Extract aliquots containing 25 µg of total protein were supplemented with 2× Laemmli buffer to serve as input controls.

To purified the complexes containing the GFP-tagged proteins (GFP-IP), 1 mg of cytoplasm extract from overexpression cells was incubated with pre-equilibrated 20 µL GFP-Trap® beads for 2 h. To determine the association between the endogenous protein, endogenous co-IP was performed using direct (IP: α-Unrip) or indirect (IP: α-SMN) approach. For the direct method, the Unrip antibody was first conjugated to pre-equilibrated 20 µL Protein G beads by incubation for 2 h with rotation, and the resulting

antibody-conjugated beads were subsequently incubated with 1 mg of HEK293T cell extract for an additional 2 h. For the indirect method, equal amount of cell extract was first incubated with the SMN antibody for 2 h, after which pre-washed 20 µL Protein G beads were added and the mixture was incubated for another 2 h. All incubations during the IP were performed at 4°C with gentle end-over-end rotation.

After incubation, the beads were washed three times by resuspending them in washing buffer and centrifuging at 2000 rpm for 2 min at 4°C. The bound proteins were eluted by 2× Laemmli buffer for subsequent immunoblotting analysis (Section 2.2.3.3).

To conduct GFP-IP with SEC fractions, the GFP-Trap® beads was incubated with a pooled mixture containing 150 µL of each fraction. The remaining procedures were carried out as detailed above.

2.2.3.8 Pull-down assay

Equal amounts of immobilised GST-SMN variants (Section 2.2.3.5) were incubated with 1 mg of cytoplasmic extract from HEK293T cells overnight at 4°C on an end-over-end rotator. After washing twice with cold PBS, the beads were mixed with 2× Laemmli buffer and subjected to immunoblotting (Section 2.2.3.3). Amido Black staining of the membrane served as the loading control (Section 2.2.3.4).

2.2.3.9 In vitro translation and interaction assay

Radiolabelled Sm proteins were synthesised *in vitro* from pET-28a or pET100 constructs (Section 2.1.4) using the TNT® T7 Quick Master Mix in the presence of L-[35S]-Methionine. These *in vitro*-translated radiolabelled proteins were incubated with equivalent amounts of GST-SMN WT and mutations (Section 2.2.3.5) in interaction buffer at 4°C for 2 h with end-over-end rotation. Subsequently, the beads were washed three times with interaction buffer, and the bound proteins were eluted by 2× Laemmli buffer.

The eluted proteins were separated by SDS-PAGE, then fixed and visualised with Coomassie blue staining (Section 2.2.3.2). The gel was dried onto filter paper under vacuum at 80°C for 2 h. To detect the radiolabelled proteins, the dried gel was exposed to autography film at -80°C for indicated times. The resulting autoradiogram was captured by manually developing the film using photographic developer and fixer solutions in a darkroom. The dried gels and resulting autoradiograms were digitised using a flatbed scanner.

2.2.3.10 In vitro cleavage assay

To evaluate the cleavage events among the SMN complex proteins, *in vitro* cleavage assays were conducted using a modified version of the protocols from A. Gregory

Matera's group [90, 91]. Cytoplasmic extracts prepared as described in Section 2.2.2.2 were incubated with 1 mM $CaCl_2$ and enzymes (caspase-3 and calpain-1) at the indicated ratios in a final reaction volume of 100–200 µL at 30°C for 30 min or 2 h. The reactions were terminated by adding 6× Laemmli buffer and subjected to immunoblotting analysis (Section 2.2.3.3).

2.2.3.11 In vitro RNA binding assay

To investigate the capability of the SMN complex to bind with U1 snRNA, *in vitro* RNA binding assay was performed. The plasmid template for *in vitro* transcription was linearised with Pfl23II and purified using the Wizard® SV Gel and PCR Clean-Up System. Fluorescein-labelled U1 snRNA was then synthesised using the Fluorescein RNA Labelling Mix. U1 snRNA-SMN complex formation was initiated by incubating 1.2 mL of cytoplasmic extract (~4 µg/µL) with 60 µL of the labelled U1 snRNA transcript and 2.5 mM ATP. The reaction mixture was protected from light and incubated at 37°C for 40 min.

To isolate the assembled complexes, the reaction was fractionated by SEC using the Superose™ 6 Increase 10/300 GL column (Section 2.2.3.6). Eluted fractions were resolved by SDS-PAGE. To simultaneously track the distribution of U1 snRNA and SMN, the gel was imaged using a UV transilluminator to visualise the fluorescent snRNA before immunoblotting analysis (Section 2.2.3.3).

2.2.3.12 Cajal body measurement by Immunofluorescence

Cells were seeded on coverslips in 24-well plates. After DOX induction or 25 nM BafA1 treatment for 2 h, cells were washed once with PBS and subjected to sequential solution exchanges for immunofluorescence (IF) analysis. With the exception of the primary antibody incubation, all procedures were performed at RT. All steps were separated by two washes with PBS.

Briefly, cells were fixed in 4% PFA fixative solution for 10 min, permeabilized with 0.2% Triton X-100 permeabilization buffer for 15 min, and blocked in 5% BSA blocking buffer for 1 h. Coverslips were then incubated with primary antibodies for at least 2 h at 4°C, followed by incubation with fluorophore-conjugated secondary antibodies for 1 h. Nuclei were stained with DAPI for 1 h in the dark. Finally, coverslips were mounted onto glass slides using ProLong™ Glass Antifade Mountant and allowed to dry overnight prior to imaging.

Fluorescence imaging was performed using the Axio Observer inverted microscope equipped with the ApoTome.2 optical sectioning system and Plan-APOCHROMAT

40×/1.40 oil DIC objective. Acquired images containing at least 1000 cells were analysed by Fiji (ImageJ) as described previously [100]. The number of double-positive puncta (SMN and coilin) within DAPI-stained nuclear region representing Cajal bodies was quantified by the custom automated macro and plotted using GraphPad Prism.

2.2.4 Bioinformatic analyses

2.2.4.1 MS interactome analysis

GFP-IP was performed as described in Section 2.2.3.7 except using RoederA buffer without the protease inhibitor and PBS to wash the beads. The resulting beads were snap frozen in liquid nitrogen and shipped on dry ice to the collaborating laboratory of Prof. Dr. Thilo Kähne (Institute of Experimental Internal Medicine, Otto von Guericke University, Magdeburg, Germany) where the interactors on the beads were identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS) following similar procedures described in previous research [23]. The interactome data were analysed and visualised via RStudio (R version 4.5.1 and RStudio version 2025.9.1.401) using R package including tidyverse (version 2.0.0) [101], ComplexUpset (version 1.3.3) [102, 103] and clusterProfiler (version 4.18.4) [104], along with custom-written scripts.

Proteins with -10lg*P* value ≥ 50 and the number of unique peptides ≥ 3 were considered to be present in the given interactome. Filtered datasets were compared in R to identify shared or unique proteins between groups, which were subsequently represented as an UpSet plot using the ComplexUpset package. The exclusive or intersective proteins of interest were analysed using the compareCluster function within the clusterProfiler package to identify the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched with interactors. To visualise the comparative results, the dot plot was generated to show the top-3 significantly enriched GO terms or top-5 KEGG pathways of each group.

2.2.4.2 SMN conservation analysis

To evaluate the evolutionary conservation of residues 37 and 187 in SMN, multiple sequence alignment across different species was performed. The reviewed Uniprot amino acid sequences of SMN from the species indicated in Fig. 2 were analysed using the Clustal Omega 1.2.4 [70, 71], followed by visualisation using ESPript 3.2 [72]. Default parameters were applied for both web tools. The exported PDF file from ESPript 3.2 was modified by Adobe illustrator CC 2019.

2.2.4.3 Calpain cleavage prediction

To predict the calpain cleavage sites on human Gemin3, the amino sequence of Gemin3 (Uniprot ID: Q9UHI6) was subjected to DeepCalpain [105]. The sequence was queried against all categories of calpain using the medium threshold. The predicted results were sorted in ascending order based on the value of false positive rate and the top-5 cleavage sites were presented in Table 8.

2.2.4.4 KFERQ-like motifs screening

The KFERQ-like motifs in the SMN complex proteins were identified by KFERQ finder V0.8 [106]. SMN, Gemin2–Gemin8 and Unrip (Uniprot ID: Q16637, Q9UHI6, O14893, P57678, Q8WXD5, Q8TEQ6, Q9NWZ8, Q9H840, Q9Y3F4) were screened for standard and advanced motifs in this web server. The numbers of the canonical and post-translational-modification-generated motifs in each protein were visualised as a bar chart (Fig. 40) using GraphPad Prism.

3 Results

3.1 Interaction between mTOR, ULK1 and SMN protein

Spliceosomal small nuclear ribonucleoprotein (snRNP) biogenesis is an ATP-dependent process [18, 50-53]. Recent studies indicate that mammalian target of rapamycin (mTOR) and unc-51-like autophagy activating kinase 1 (ULK1), two energy-sensing kinases, are highly associated with this process [23-25, 55]. Cellular energy conditions dictate the activity of mTOR and ULK1 [107-110]. This chapter was designed to investigate how mTOR and ULK1 involve into the survival motor neuron (SMN) complex under different nutrient conditions.

3.1.1 mTOR, ULK1, and ATG13 are excluded from the SMN complex

To examine how mTOR, ULK1, autophagy-related protein (ATG) 13 and SMN are distributed in the native state under various nutrient conditions, cytoplasmic extracts from GFP-SMN overexpression cells were analysed by size exclusion chromatography (SEC). To simulate cellular energy deprivation, overexpression cell lines were treated with Earle's balanced salt solution (EBSS) for 2 hours (h) after induction. Cytoplasmic proteins were separated by the Superose 6 column (Fig. 4 A). 2-hour EBSS treatment did not alter their cellular distribution of mTOR, ULK1 and ATG13. These proteins partially co-eluted in fractions 11–15, corresponding to 440–660 kDa. To further determine the protein distribution in this molecular weight range, the same samples were subsequently analysed using the Superdex 200 column, which provides optimal separation for proteins between 10 and 600 kDa (Fig. 4 B). Except ATG13, the mTOR, ULK1 and GFP-SMN proteins were co-fractionated in fractions 4–9 (232–668 kDa). No alteration in distribution of mTOR, ULK1, ATG13 and GFP-SMN was found after EBSS treatment.

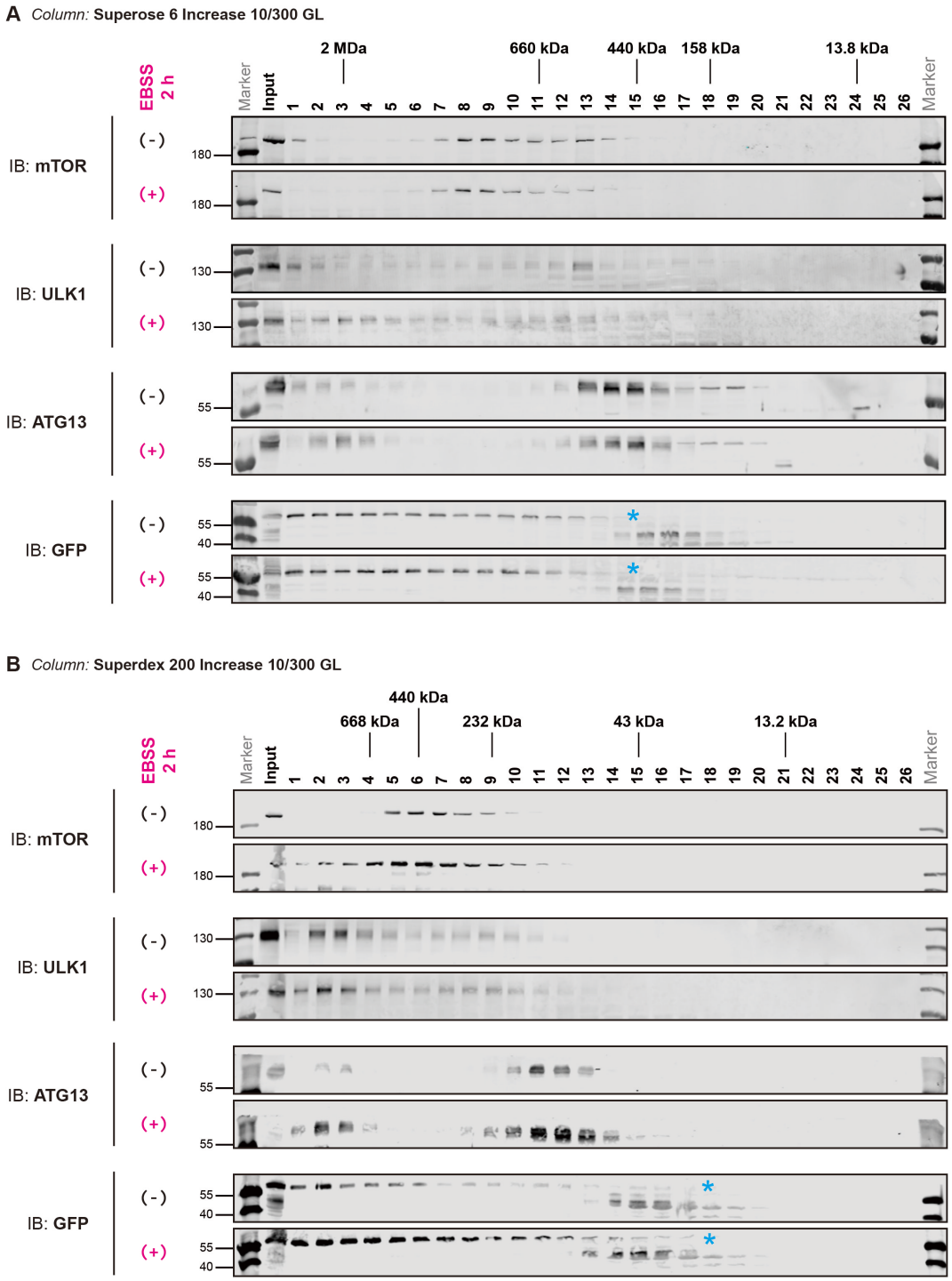


Fig. 4: ***mTOR, ULK1, and ATG13 co-elute with SMN regardless of nutrient conditions.*** *After induction with 0.1 µg/mL doxycycline, GFP-SMN overexpression cells were treated with (+) or without (-) EBSS for 2 h. Their cytoplasm extracts were then loaded onto the (A) Superose 6 and (B) Superdex 200 columns. The distribution of mTOR, ULK1, ATG13 and GFP-SMN was visualised by immunoblotting. The GFP-SMN is marked with asterisks (*). IB, immunoblotting; EBSS, Earle's balanced salt solution.*

To test whether mTOR and ULK1 interact with SMN, co-immunoprecipitation (co-IP) assays were performed. Cytoplasmic extracts from GFP-SMN overexpression cells with or without 2-hour EBSS treatment were incubated with GFP-Trap beads. The isolated binding partners were probed by immunoblotting. As shown in Fig. 5 A, mTOR, ULK1

and ATG13 were not detected in the IP pools of GFP-SMN or GFP-Gemin2 under both nutrient conditions.

To confirm these findings, endogenous IP assays with antibodies against SMN or Unrip using cytoplasmic extracts from HEK293T cells were performed. While these antibodies successfully immunoprecipitated their respective bait proteins and canonical SMN complex components, such as Gemin2 and SMN, the vast majority of mTOR, ULK1 and ATG13 remained free in the supernatants (Fig. 5 B and C).

To preclude the possibility that 2-hour starvation might be insufficient to reduce ATP levels to cause alterations in complex composition, we subsequently conducted GFP-IP using GFP-SMN overexpression cells with extended EBSS treatment (6 or 24 h). In this replicate, weak signals of mTOR were observed on GFP-IP beads regardless of starvation time. However, ULK1 and ATG13 remained undetectable on GFP-IP beads. Gemin and Sm proteins were persistently co-immunoprecipitated by GFP-SMN even under stringent nutrient-deprivation conditions (Fig. 5 D).

These findings suggested that SMN do not interact with ULK1 and mTOR under basal or starvation conditions, and the SMN complex remains stable under metabolic stress.

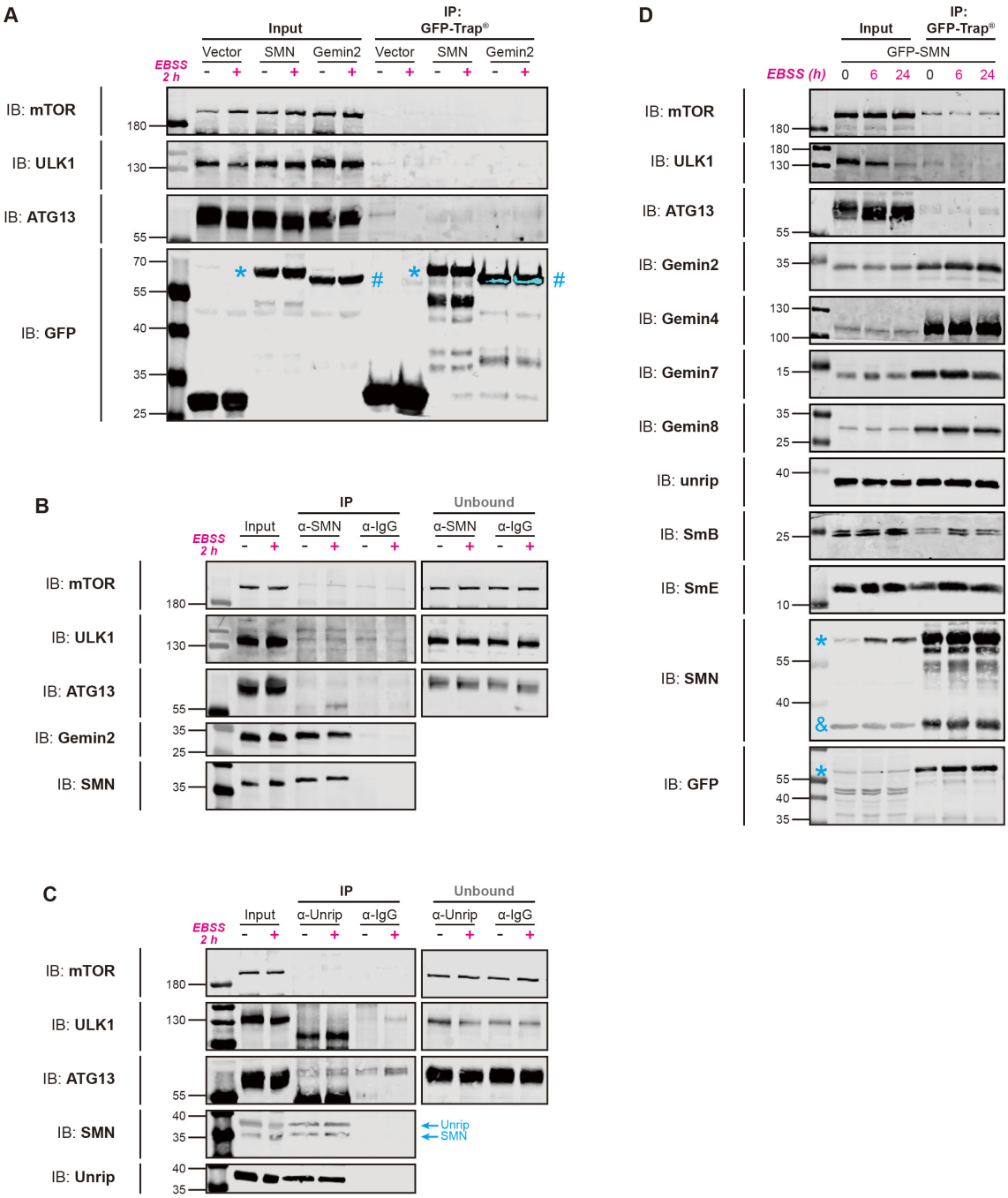


Fig. 5: ***mTOR and ULK1 complexes are absent in the IP pool of GFP-SMN under different nutrient conditions.*** *(A) Following overnight induction with 0.1 μg/mL doxycycline, cells stably transfected with GFP (referred to as "Vector"), GFP-SMN or GFP-Gemin2 were treated (+) with or (-) without EBSS for 2 h. Cytoplasmic fractions were subjected to IP using GFP-Trap beads. Enriched binding proteins were subsequently detected via immunoblotting. Asterisk (*) and hashes (#) indicate the GFP-SMN and GFP-Gemin2, respectively. Cytoplasm extracts from HEK293T cells (+) with or (-) without 2-hour EBSS treatment were incubated with the (B) SMN and (C) Unrip antibodies for co-IP. The symbol "α" represents the prefix "anti-". The eluates from the beads and the supernatant (referred to as "Unbound" in the panels) were analysed by immunoblotting. In (C) immunoblot probed with the SMN antibody, the lower arrow points to newly probed SMN and the upper arrow indicates the residual signal from Unrip after re-incubation. (D) GFP-IP was also performed with GFP-SMN overexpression cells treated with EBSS for 6 and 24 h following the same procedures as described in (A). The signals of GFP-SMN and endogenous SMN served as the bait control, which are marked with asterisks (*) and an ampersand (&), respectively. IB, immunoblotting; IP, immunoprecipitation; EBSS, Earle's balanced salt solution.*

3.1.2 *Ubiquitin-mediated proteolysis might be associated with the turnover of the SMN complex during starvation*

To further explore the interaction network of SMN under different nutrient conditions, a pilot interactome analysis was carried out using GFP-IP beads following the protocol as described in Fig. 5 A. The isolated proteins were identified through mass spectrometry (MS), which was performed by Prof. Dr. Thilo Kähne (Institute of Experimental Internal Medicine, Otto von Guericke University, Magdeburg, Germany). Proteins with high confidence (-10lg*P* ≥ 50, unique peptides ≥ 3) were included in corresponding interactomes.

The MS result further confirmed that the candidate kinases (mTOR, ULK1) and ATG13 were not co-immunoprecipitated by GFP-SMN. However, the difference in SMN interactomes under both nutrient conditions provides preliminary insight into the behaviour of SMN under different energy conditions. As illustrated in the UpSet plot (Fig. 6), 38 proteins invariably interacted with SMN even upon EBSS-induced starvation (yellow bar); while 109 proteins were uniquely bound to SMN after 2-hour EBSS treatment (magenta bar), an almost three-fold increase compared to the 35 proteins identified in the non-treatment group (green bar).

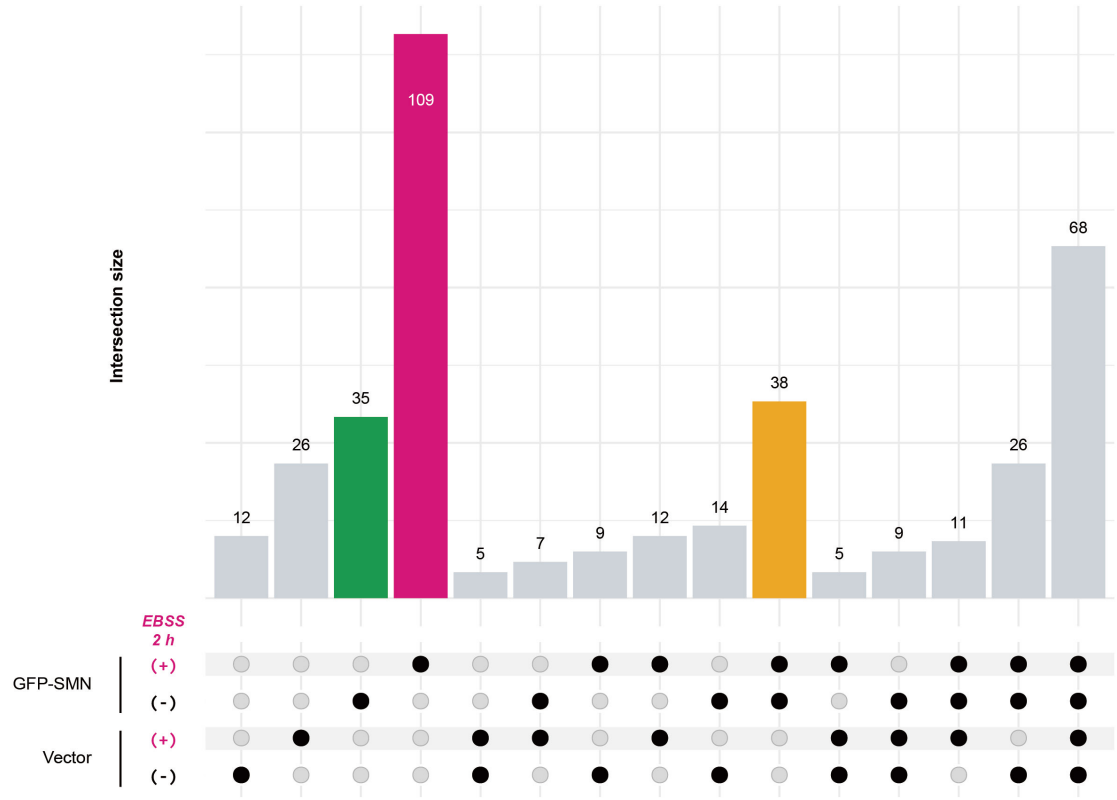


Fig. 6: ***The interaction network of SMN expands after EBSS treatment.*** *In cooperation with Dr. Lea Marie Esser, GFP-IP was performed with GFP (Vector) and GFP-SMN overexpression cells cultured in different nutrient conditions. The resulting beads were subsequently submitted to Prof. Dr. Thilo Kähne for MS identification. Interactors with -10lgP ≥ 50 and unique peptides ≥ 3*

were subjected to the following analysis. The intersection of interactomes under different experiment conditions was visualised by the UpSet plot. The dark dots below represent specific combinations of groups. The number of interactors in each intersection is indicated above the bars. Intersections are colour-coded to differentiate the interactomes of interest from non-specific background (light grey): Green, magenta and yellow highlight the interactome unique to GFP-SMN and GFP-SMN+EBSS, and their intersection, respectively. EBSS, Earle's balanced salt solution.

To gain more mechanistic insights, these protein subsets were annotated using Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway enrichment analyses. GO delineates the biological functions of the interactomes, while KEGG analysis maps the specific pathways in which these interactors are involved. Binding proteins induced by or independent of EBSS treatment were predominantly enriched in the GO terms related to the SMN complex and RNA-associated processes (Fig. 7 A, magenta and yellow), further supporting the stability of SMN complex under different metabolic environments. Proteins that interacted with SMN only under nutrient-rich conditions seemed to be more associated with ribosome (Fig. 7 A, green). Consistent with GO analysis, KEGG analysis revealed that EBSS-specific or unrelational interaction partners of SMN were highly involved in the spliceosome pathway (Fig. 7 B, magenta and yellow), while the binding proteins under the nutrient-rich condition were only enriched in the Covid-19 pathway (Fig. 7 B, green).

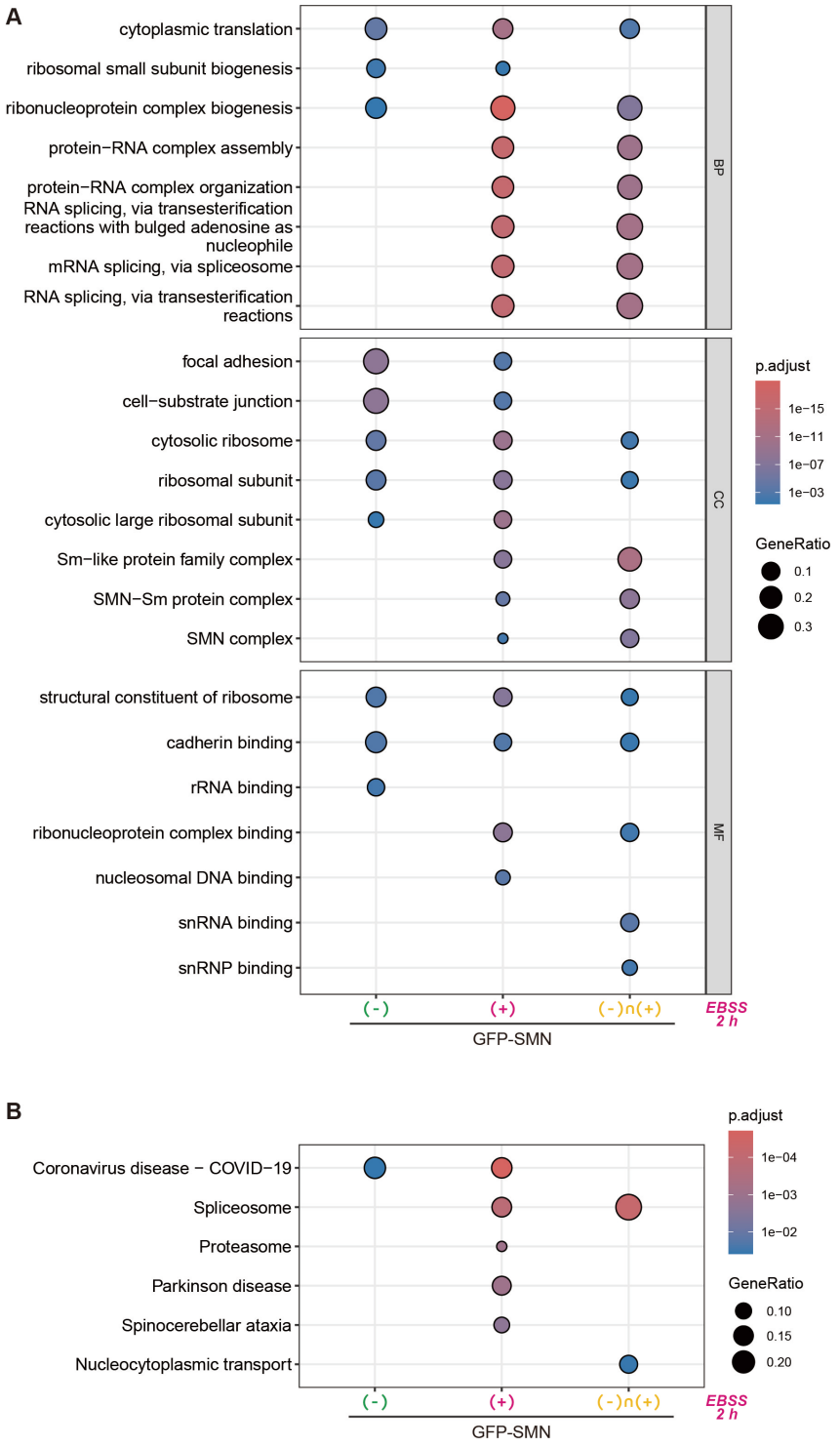


Fig. 7: ***Interaction partners with different functions bind with SMN after EBSS treatment.*** *Specific or shared binding proteins of SMN under different nutrient environments. Subsets highlighted with colours in the UpSet plot (Fig. 6) were subjected to (A) Gene Ontology (GO) and (B) Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analyses. The top-3 GO and top-5 KEGG terms of each group are presented as dot plots. Dot size indicates the Gene Ratio (number of prey proteins from the subset annotated to the corresponding terms / total annotated proteins in the subset), and colour denotes the Benjamini-Hochberg adjusted p-value. Green (-) and magenta (+) indicate the unique interactome of GFP-SMN without or with 2-hour EBSS treatment, respectively; yellow (-)∩(+) refers to the interactome shared by these two groups. BP, biological process; MF, molecular function; CC, cellular component; EBSS, Earle's balanced salt solution.*

Interestingly, compared with the EBSS-independent interactome, the binding proteins induced by starvation enriched in the proteasome pathway (Fig. 7 B, magenta). The enrichment details of this pathway showed that once the nutrient was withdrawn, more proteasomal proteins were recruited to the SMN complex (Table 3). These observations implied that the homeostatic turnover of SMN protein might be disrupted under metabolic stress.

Table 3: ***Proteasomal proteins interact with SMN under different nutrient conditions***

Proteasome part		(-) EBSS	(+) EBSS
19S Regulatory particle	Lid (PA700)	Rpn3	N/A
	Base (PA700)	N/A	Rpn1, Rpn2, Rpt5
20S Core particle (Standard subunits)		N/A	α1, α3

Notes: GFP-IP was performed as described in Fig. 5 A using GFP-SMN overexpression cells treated (+) with or (-) without EBSS for 2 h after induction. The interactors were identified by MS as described in Fig. 6. Proteasomal proteins in the interactomes specific to nutrient-rich ((-) EBSS; Fig. 6, green bar) or -free ((+) EBSS; Fig. 6, magenta bar) conditions are listed in this table.

3.2 Effect of SMN phosphorylation at T37 and S187 on snRNP biogenesis

SMN activity in snRNP biogenesis is governed by its phosphorylation [111]. A SMN peptide with phosphorylated threonine (T) 37 in the IP pool was detected after EBSS treatment during MS identification (Section 3.1.2) (Fig. 39). MS analysis performed by another group revealed that the phosphorylation level of serine (S) 187 decreases after knocking out ribosomal protein S6 kinase beta-1 (RPS6KB1; p70S6 kinase) [55], a downstream target of mTOR [112, 113]. These energy-related phosphorylation events on SMN might bridge the cellular energy levels with snRNP biogenesis. Therefore, the phosphorylation of T37 and S187 was investigated in this chapter to understand their impact on snRNP biogenesis.

3.2.1 Single phosphorylation at T37 or S187 is dispensable for the interaction with Gemin2

T37 is located in the Gemin2-binding domain on SMN [35, 36, 114-118]. Previous research has shown that phosphorylation of S187 and other residues affects interaction with Gemin2 [58]. Therefore, the effect of the phosphorylation at these two sites on interaction with Gemin2 was first examined by GFP-IP using stable cell lines overexpressing SMN variants.

A series of Flp-In™ T-REx™ 293 cell lines overexpressing GFP-tagged SMN with a specific mutation at T37 and S187 was established. Specifically, these sites in overexpressing GFP-SMN were mutated respectively into alanine (A) to simulate dephosphorylation status (T37A, S187A) or into aspartic acid (D) to mimic phosphorylation (T37D, S187D). Cytoplasmic fractions extracted from these cells were incubated with GFP-Trap beads, and then the isolated SMN complex was analysed by immunoblotting. The intensity of Gemin2 captured by GFP-SMN baits was comparable across all SMN variants (Fig. 8 A and B).

GFP-IP reflects the interaction between global SMN and Gemin2. To further investigate the interaction between their subpopulations, the cytoplasm extracts used in GFP-IP were also subjected to the Superose 6 column. SMN and Gemin2 distribution was visualised via immunoblotting. As the core interaction partner of SMN, the Gemin2 protein co-eluted with GFP-SMN from fractions 1 to 14 in a descending gradient, but with a marginal peak from 15 through 19 (Fig. 8 C). This elution pattern was shared among all GFP-SMN cell lines.

Overexpression of SMN could increase the levels of other Gemin proteins [119], which might mask the effect of SMN phospho-mutations on Gemin2 interaction in co-IP. To directly evaluate the affinity between SMN variants and Gemin2, the GST pull-down

assay was conducted. Recombinant GST-tagged SMN variants (wild type (WT) and other four phospho-mutants) were employed as baits to capture Gemin2 from the HEK293T cytoplasmic extracts. Gemin2 was successfully captured by GST-SMN WT and mutations, but no significant difference in their Gemin2 recovery was observed among all GST-SMN variants (Fig. 8 D).

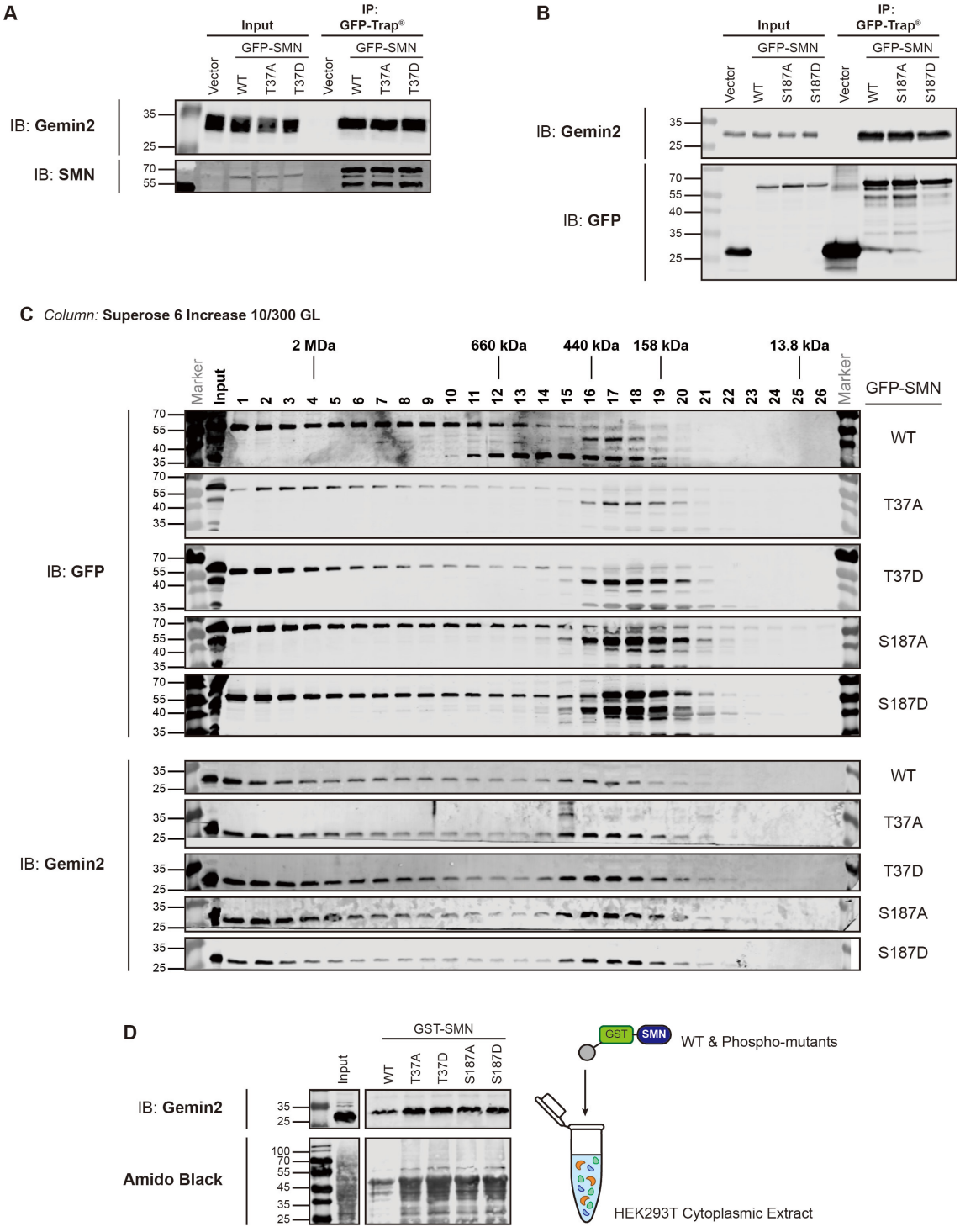


Fig. 8: ***Gemin2 interaction is still maintained after introducing phospho-mutations on SMN at T37 or S187.*** *(A and B) After incubation with 0.1 µg/mL doxycycline overnight, cells stably overexpressing GFP (Vector), GFP-SMN WT and phospho-mutants (T37 A or D; S187 A or D) were lysed to extract the cytoplasmic fractions. These extracts were incubated with GFP-Trap*

beads for 2 h. After washing, the binding proteins on the beads were detected via immunoblotting. SMN or GFP was probed to monitor the amount of immunoprecipitated bait GFP-tagged proteins. (C) The same cytoplasmic extracts used in (A) GFP-IP were subjected to the Superose 6 gel filtration column. The distribution pattern of GFP-SMN and Gemin2 proteins was visualised by immunoblotting. (D) As illustrated on the right of the figure, Gemin2 in HEK293T cytoplasmic fraction was pulled down by different GST-SMN variants (WT, T37A or D, S187A or D). The recovery amount of Gemin2 on GST beads was measured by immunoblotting, and the membrane was stained with Amido Black as the loading control. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.

Other factors in cellular extracts might affect the interaction between SMN and Gemin2. To circumvent this limitation, the cell-free, *in vitro* interaction assay was conducted. L-[35S]-Methionine (^{35}S) radiolabelled Gemin2 was generated via *in vitro* translation and incubated with immobilised GST-only negative control and GST-SMN variants. After washing, the complexes on GST beads were resolved by SDS-PAGE. The loading of bait protein was monitored via Coomassie blue staining, and the radiolabelled prey protein was visualised via autoradiography. No non-specific binding was observed between ^{35}S-Gemin2 and GST protein. With similar GST bait amounts, SMN WT and phospho-mutations were comparable in the radiometric signals corresponding to Gemin2 (Fig. 9 A).

Collectively, these findings suggested that SMN phosphorylation at T37 or S187 is not a dominant driver of Gemin2 interaction.

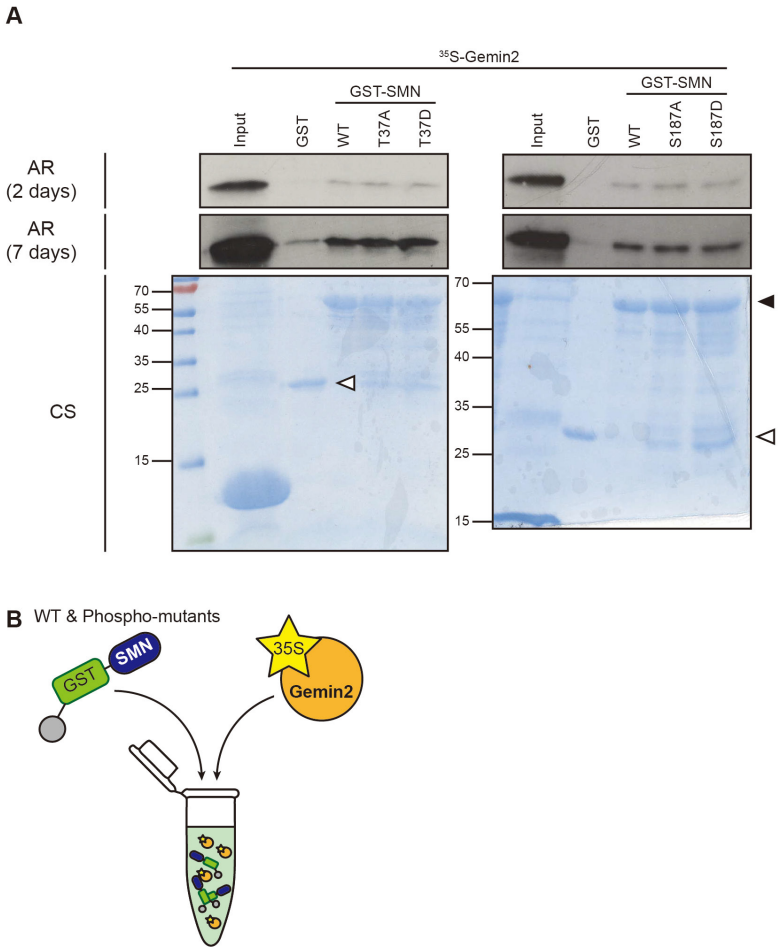


Fig. 9: ***SMN with different phosphorylation states at T37 and S187 shares the same affinity with Gemin2.*** *^{35}S-labelled prey proteins were generated via in vitro translation and pulled down by GST fusion proteins, which were immobilised on beads. (A) ^{35}S-Gemin2 was incubated with GST, GST-SMN WT and its phospho-mutations. The amounts of captured Gemin2 and baits were measured by autoradiography and Coomassie blue, respectively. (B) The corresponding experimental design for bait and prey proteins was illustrated. The GST-SMN variants are indicated by solid arrowheads, while the GST-only protein is marked by unfilled arrowheads. 25% of the ^{35}S-labelled proteins used in assays were loaded as the input. Numbers below AR denote exposure times. WT, wild type; AR, autoradiography; CS, Coomassie blue staining.*

3.2.2 S187 phosphorylation affects the binding to SmB

One of the important functions of the SMN complex is to accept a set of Sm proteins (SmB, D1, D2, D3, E, F and G) [15, 16, 22]. To screen the Sm proteins whose interactions with SMN were disrupted by phosphorylation, GFP-IP was performed following the same protocol as described in Section 3.2.1. Except SmF, other Sm proteins (B, D2, D3, E and G) were successfully co-immunoprecipitated by all GFP-SMN variants. No significant mutation-dependent alterations in the binding capacities with SmE and SmG were observed in this GFP-IP assay. However, compared with WT, four phospho-mutants co-immunoprecipitated less SmB, SmD2 and SmD3 (Fig. 10), which directly interact with SMN [35, 120].

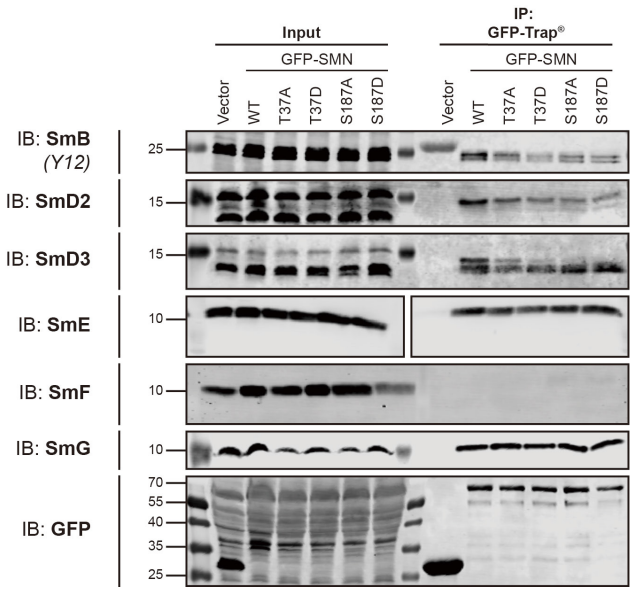


Fig. 10: ***The interaction with Sm proteins might be interfered by SMN phosphorylation.*** *GFP-IP was performed using the cytoplasm extracts from GFP (Vector) and GFP-SMN variants (WT, T37A/D, S187A/D) as previously described (Fig. 8 A and B). Co-purified Sm proteins were detected via immunoblotting with a series of Sm protein antibodies. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.*

To further investigate the impact of T37 or S187 phosphorylation on SMN's capability to accept these Sm proteins, the SmD and SmB proteins in the SEC fractions generated previously in Section 3.2.1 were detected by immunoblotting to delineate their *in vivo* distribution. Due to poor antibody reactivity of SmD group proteins, the membranes were probed by the SmB (Y12) antibody, which recognises symmetrical dimethylarginine (sDMA), a common methylation on the SmB, D1 and D3 proteins [121-123]. Therefore, even though the SmB (Y12) antibody primarily detects SmB [122], it could also monitor the location of methylated D1 and D3 [121]. SmB (~25 kDa) eluted in fractions 10–19 with different distribution patterns among GFP-SMN overexpression cells (Fig. 11, black arrowheads). Specifically, SmB in cells overexpressing GFP-SMN WT and T37A peaked in 11–13 (660 kDa) and 16–18 (158–440 kDa), while in other mutation cell lines, the majority of SmB only resided in fractions 11–13. Methylated SmD1 or SmD3 (~15 kDa) were more clearly visible in fractions 18–20 (158 kDa) in GFP-SMN WT, T37A and S187D overexpression cells (Fig. 11, orange arrowheads and dashed boxes).

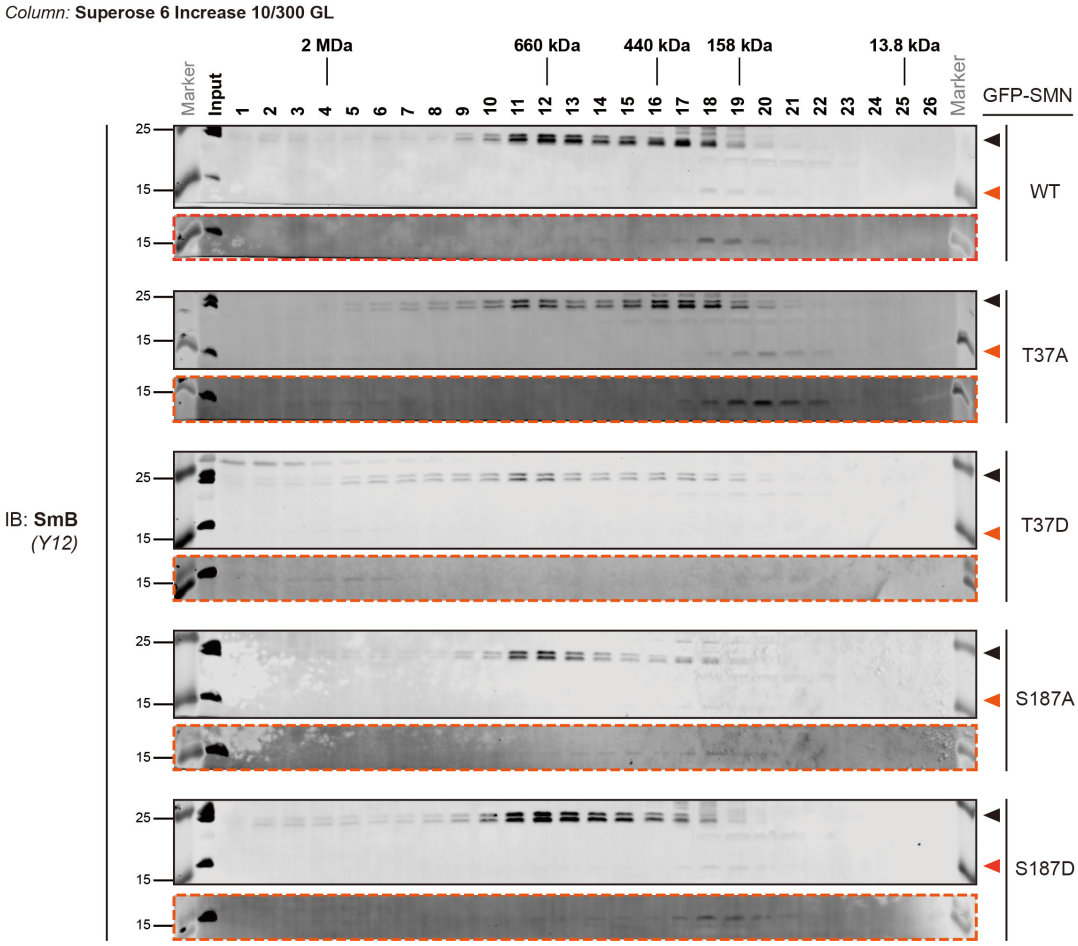


Fig. 11: ***The distribution of SmB, SmD1 and SmD3 varies among cells overexpressing different GFP-SMN variants.*** *Cytoplasmic extracts from induced GFP-SMN variants overexpression cells were fractionated by the Superose 6 column, and the fractions were analysed via immunoblotting. The Y12 antibody was used to primarily detect SmB and partially other proteins with sDMA modification. Black arrowheads point to SmB, whereas orange arrowheads mark methylated SmD1 and D3. To better visualise the faint methylated SmD1 and D3 signals, each blot with higher brightness of the ~15 kDa region is presented under the original blots and indicated by orange dashed boxes. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.*

To more accurately characterise the effect of SMN phosphorylation on the interaction with SmB, SmD1, and SmD3, their direct binding was examined by *in vitro* interaction assays. ^{35}S-labelled Sm proteins were incubated with immobilised GST-tagged SMN variants. The recovery amounts of Sm proteins were detected following the same procedures as described previously in Section 3.2.1. GST did not bind with any tested Sm proteins. All GST-SMN variants pulled down SmD1, SmD3 and SmB. Radiometric signals of SmD group proteins were comparable on all baits (Fig. 12 A and B). However, ^{35}S-SmB was less enriched by the S187D construct (Fig. 12 C).

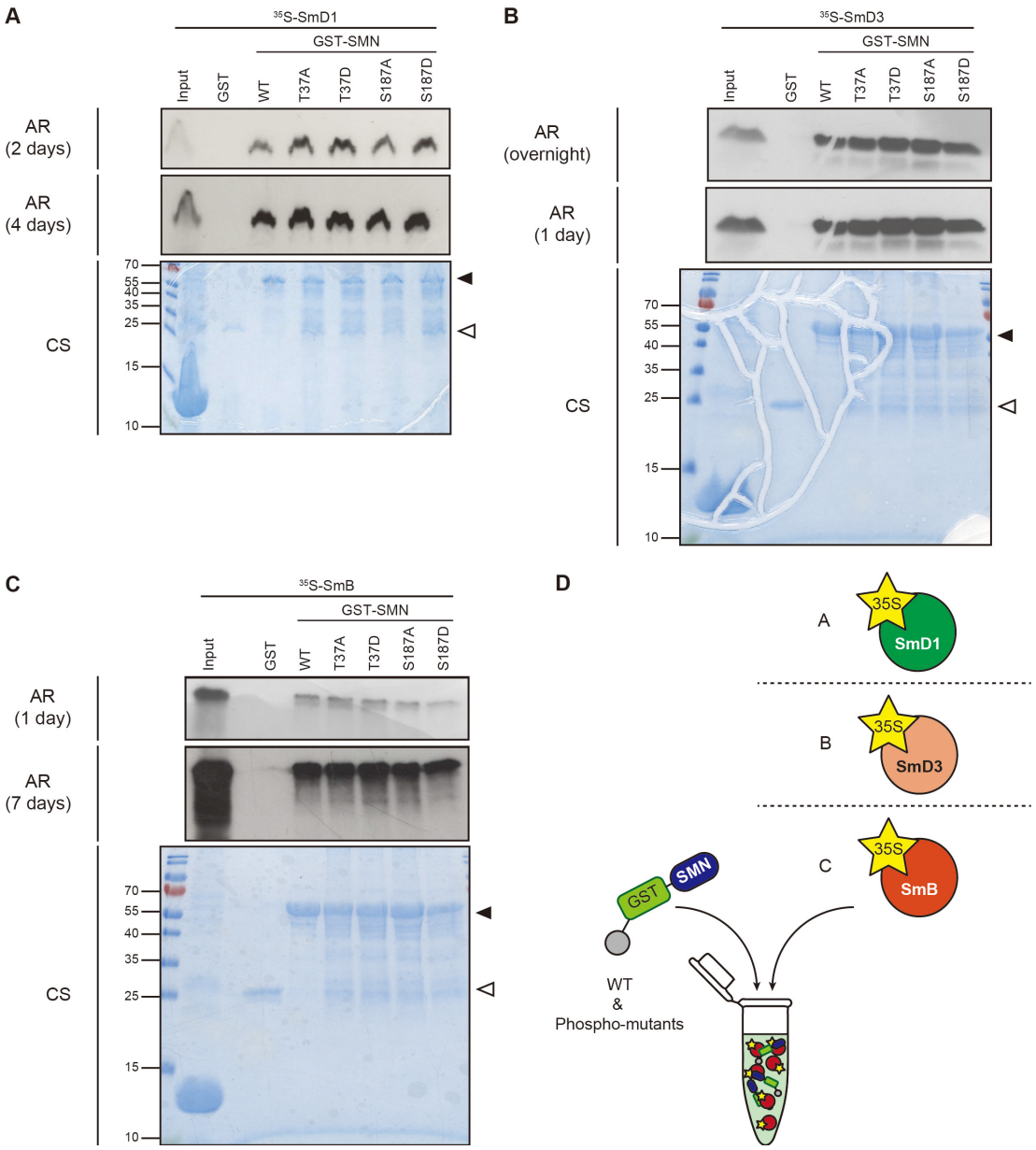


Fig. 12: ***Phosphorylation at S187 on SMN might interfere with SmB interaction.*** *Beads coupled with different GST-SMN fusion proteins (WT and four phospho-mutants) were incubated with (A)^{35}S-SmD1, (B)^{35}S-SmD3, (C)^{35}S-SmB, as illustrated in (D). Bead eluates were resolved by SDS-PAGE, followed by Coomassie blue staining as the loading control and autoradiography as the measurement for recovered prey proteins. The GST-SMN variants are indicated by solid arrowheads, while the GST-only protein is marked by unfilled arrowheads. 25% of the ^{35}S-labelled proteins used in assays were loaded as the input. Numbers below AR denote exposure times. WT, wild type; AR, autoradiography; CS, Coomassie blue staining.*

Since the majority of SmD3 and SmB interact with each other as subcomplexes *in vivo* [15], the interaction with this subcomplex was also evaluated. ^{35}S-labelled SmD3 and SmB proteins were pre-assembled into a complex by incubation before testing their binding with GST-SMN variants. Similarly, the GST-SMN S187D bait captured less SmB and SmD3 (Fig. 13) than WT and S187A variants.

In summary, these findings suggested that SMN phosphorylation at S187 might interfere with SmB interaction but not SmD1 and SmD3.

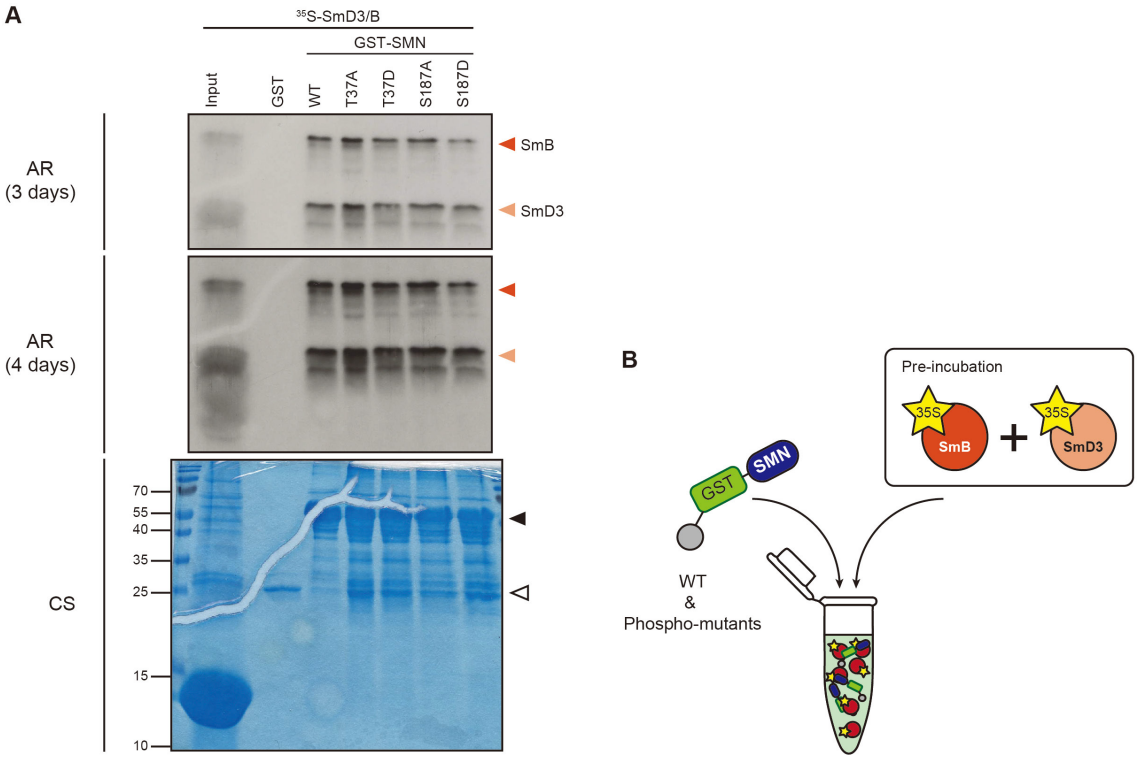


Fig. 13: ***Phosphorylation status at S187 hinders SMN from binding with the SmD3/B subcomplex.*** *(A) An in vitro interaction assay was performed using GST-SMN variants and pre-assembled ^{35}S-SmD3/B subcomplex following the procedures described in Fig. 12. The experimental design for the assay was depicted schematically in (B). The upper autoradiographic signals come from SmB (dark orange arrowheads), while the lower ones are SmD3 (light orange arrowheads) signals. The GST-SMN variants are indicated by solid arrowheads, while the GST-only protein is marked by unfilled arrowheads. 25% of the ^{35}S-labelled proteins used in assays were loaded as the input. Numbers below AR denote exposure times. WT, wild type; AR, autoradiography; CS, Coomassie blue staining.*

3.2.3 Dual phosphorylation at both T37 and S187 distances the PRMT5 complex from the SMN complex and impairs snRNP biogenesis

Sections 3.2.1 and 3.2.2 revealed that the interaction of SMN with Gemin2 and Sm proteins was largely maintained regardless of individual phosphorylation status at T37 or S187. Some studies have shown that the activity of SMN is regulated by combinatorial phosphorylation events [55, 57-59]. Therefore, the biological consequence of dual phosphorylation at these two sites on SMN was investigated. For this, GFP-SMN double phospho-mutant cells were established. GFP-SMN AA cells overexpressed a phosphodefective SMN mutant with dual A substitutions at T37 and S187, mimicking the unphosphorylated state. In contrast, GFP-SMN DD cells overexpressed a phosphomimetic mutant with dual D substitutions to simulate phosphorylation.

To screen the potential interactions governed by this dual phosphorylation, the interactomes of GFP-SMN WT, AA and DD were identified by MS after GFP-IP using cytoplasmic extracts from corresponding overexpression cell lines. The UpSet plot was generated as described in Section 3.1.2 to illustrate the overlap or unique interactors

(Fig. 14 A, various shades of green). The interaction network substantially decreased upon introducing the dual phosphomimic mutation at T37 and S187. The DD variant recovered only eight unique interactors, whereas the WT and AA baits isolated approximately 60 unique interacting proteins each. Eight common proteins were identified in both WT and AA interactomes (Fig. 14 A, red), but there was no overlap between WT and DD interactomes. These findings implied that dual phosphorylation of SMN at T37 and S187 might abrogate certain interactions.

To identify the SMN-related proteins affected by this dual phosphorylation, the binding protein subsets unique to WT, AA and DD were reviewed. Interestingly, protein arginine N-methyltransferase 5 (PRMT5), methylosome protein 50 (WD45; MEP50), and chloride conductance regulatory protein (pICln) were almost undetectable on the GFP-SMN DD bait (Table 4). GFP-IP was performed to further verify these findings. Aligning in part with MS results, PRMT5, WD45 and pICln were readily co-purified by both GFP-SMN AA and WT, but not by the DD mutant (Fig. 14 B). The interaction with other Gemin and Sm proteins was also examined in this GFP-IP assay. Contrary to the methylosome proteins, both phospho-mutant SMN exhibited similar capability in binding with Gemin and Sm proteins compared with WT (Fig. 14 C).

Table 4: ***-10lgP and unique peptide counts for methylosome proteins identified in GFP-IP pools of different GFP-SMN variants***

Proteins		Vector	GFP-SMN		
			WT	AA	DD
PRMT5	-10lgP	19.25 ± 27.22	63.44 ± 89.71	156.05 ± 6.44	14.47 ± 25.07
	unique peptide	0.50 ± 0.71	1.00 ± 1.41	8.33 ± 1.53	0.33 ± 0.58
WD45	-10lgP	0.00 ± 0.00	0.00 ± 0.00	127.55 ± 26.11	0.00 ± 0.00
	unique peptide	0.00 ± 0.00	0.00 ± 0.00	3.67 ± 1.15	0.00 ± 0.00
pICln	-10lgP	0.00 ± 0.00	52.54 ± 74.30	64.54 ± 9.95	0.00 ± 0.00
	unique peptide	0.00 ± 0.00	1.50 ± 2.12	1.00 ± 0.00	0.00 ± 0.00

Notes: The GFP-IP was performed by the author. Subsequent MS identification was performed by Prof. Dr. Thilo Kähne. -10lgP represents 10 times the negative base-10 logarithm of the P-value, indicating the statistical significance of the protein identification. Unique peptide is the number of distinct peptide sequences uniquely mapped to the respective protein, serving as a measure of identification reliability. Higher values of -10lgP and unique peptides indicate greater confidence of identified proteins. Data are presented as the mean ± standard deviation (SD)

across >2 replicates. All final calculated values for both mean and SD are rounded to two decimal places. Zero values are shaded in light grey.

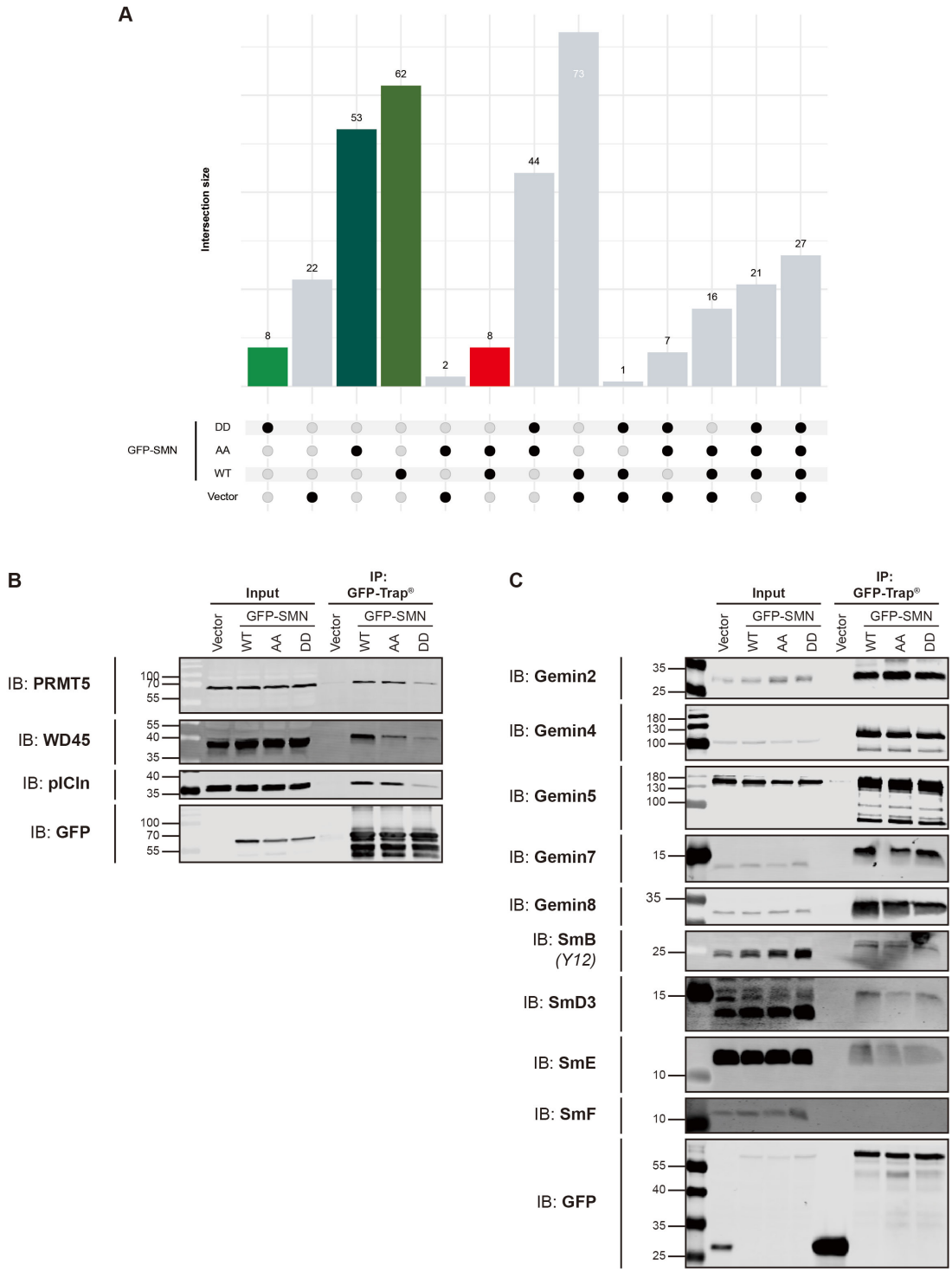


Fig. 14: ***Interaction with PRMT5 complex proteins is modulated by dual phosphorylation at T38 and S187.*** *GFP (Vector), GFP-SMN WT and dual phospho-mutants overexpression cells (AA and DD) were induced by overnight 0.1 µg/mL doxycycline treatment. Cytoplasm fractions from these cells were incubated with GFP-Trap beads. (A) The resulting GFP-Trap beads were submitted to Prof. Dr. Thilo Kähne for MS identification. The identified interactors were filtered with a criterion of -10lgP ≥ 50 and unique peptides ≥ 3. The interactomes of each bait were compared and visualised by the UpSet plot. Groups in the intersections are marked as dark dots in the matrix below, and corresponding intersection sizes are plotted as a bar chart on top. Intersecting sets containing the GFP-only control were classified as background and are indicated in light grey. Specific interacting proteins for GFP-SMN WT, AA, and DD are denoted in olive*

green, dark green, and green, respectively. Proteins shared exclusively between WT and AA, are highlighted in red. (B and C) The co-purified SMN-related proteins by GFP-IP were probed via immunoblotting with the indicated antibodies. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.

To examine whether the distribution of the methylosome was altered after SMN phosphorylation at T37 and S187, PRMT5, WD45 and pICln in GFP-SMN AA and DD overexpression cells were fractionated by the Superose 6 column. Both cell lines shared the same SEC profile of these proteins (Fig. 15 A). To further test their affinity with SMN, GFP-IP was performed using the SEC fractions 11 to 13, which were enriched with PRMT5, WD45 and pICln. Unlike GFP-IP using cytoplasmic extracts, neither SMN phospho-mutants bound with pICln in these fractions corresponding to 660 kDa. In contrast, PRMT5 and WD45 were successfully co-immunoprecipitated by both SMN mutants. But the phosphomimic SMN mutations (DD) captured fewer of these two proteins, even though more baits were used (Fig. 15 B).

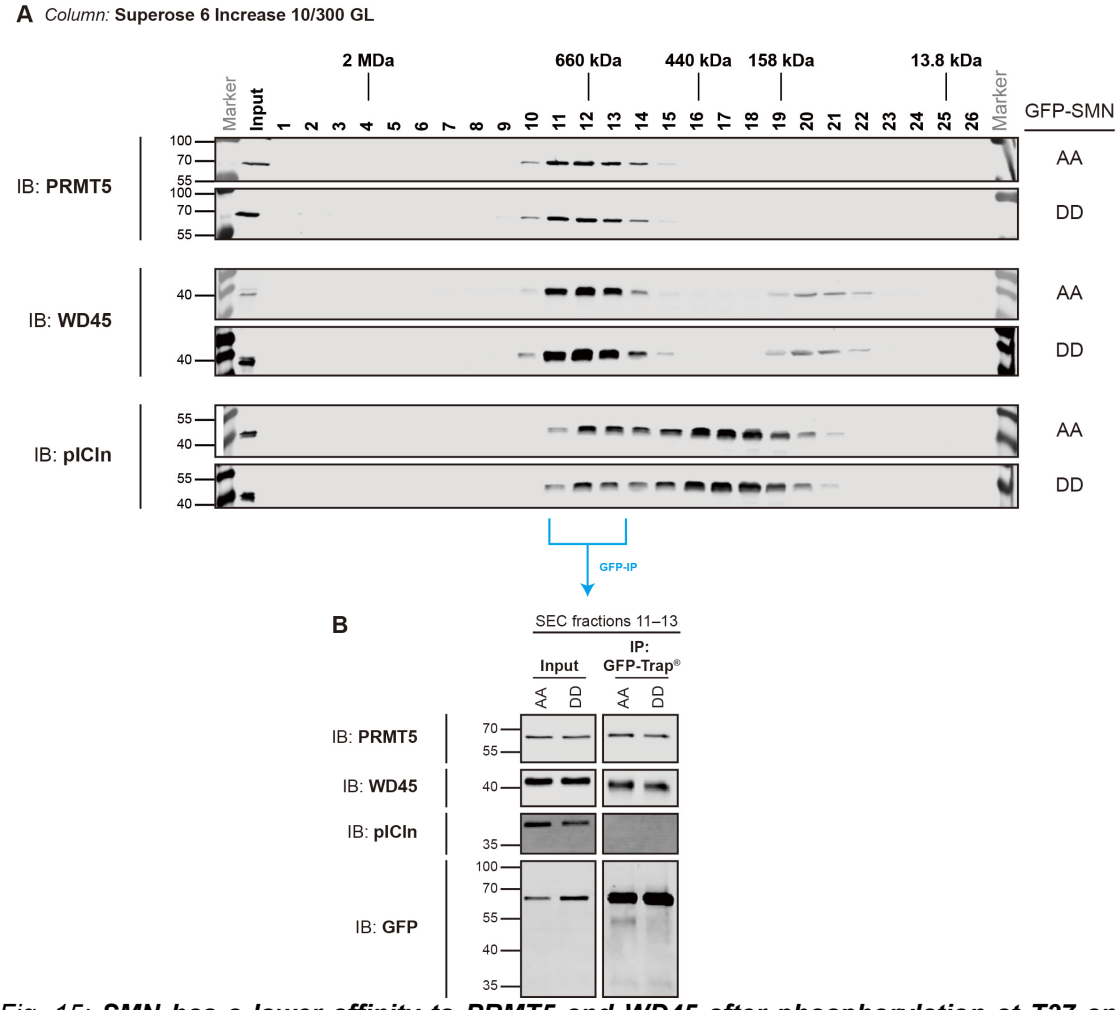


Fig. 15: ***SMN has a lower affinity to PRMT5 and WD45 after phosphorylation at T37 and S187.*** *(A) After overexpression of GFP-SMN AA and DD induced by 0.1 µg/mL doxycycline overnight, cytoplasmic extracts were prepared and then passed through the Superose 6 column. The resulting fractions were analysed via immunoblotting with the PRMT5, WD45 and pICln antibodies. (B) The mixture of the fractions 11 to 13, where the PRMT5 complex predominantly eluted, was incubated with GFP-Trap beads. The co-precipitated PRMT5 complex proteins were*

detected via immunoblotting with the corresponding antibodies. GFP was also probed as the bait control. The protein extraction and subsequent fractionation via SEC were performed by Cand. med. Lilli Albers. The resulting fractions were used by the author for immunoblotting analysis and GFP-IP. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.

PRMT5, WD45 and pICln play an important role in Sm protein methylation and delivery, which is the critical step for successful snRNP assembly [22]. Phosphorylation at T37 and S187, disrupting SMN interaction with these proteins, might have a negative impact on snRNP biogenesis. To examine this, co-immunostaining for coilin and SMN was performed to evaluate the formation of Cajal bodies. Cajal bodies are subnuclear membrane-less organelles for snRNP maturation [16, 124]. Coilin is a marker of Cajal bodies [63, 125]. Accompanied with snRNPs, SMN is localised in Cajal bodies and interacts with coilin [62, 126]. Therefore, the number of Cajal bodies were quantified based on the overlapping signals from SMN and coilin (Fig. 16 A), and its number reflects the efficiency of snRNP assembly *in vivo*.

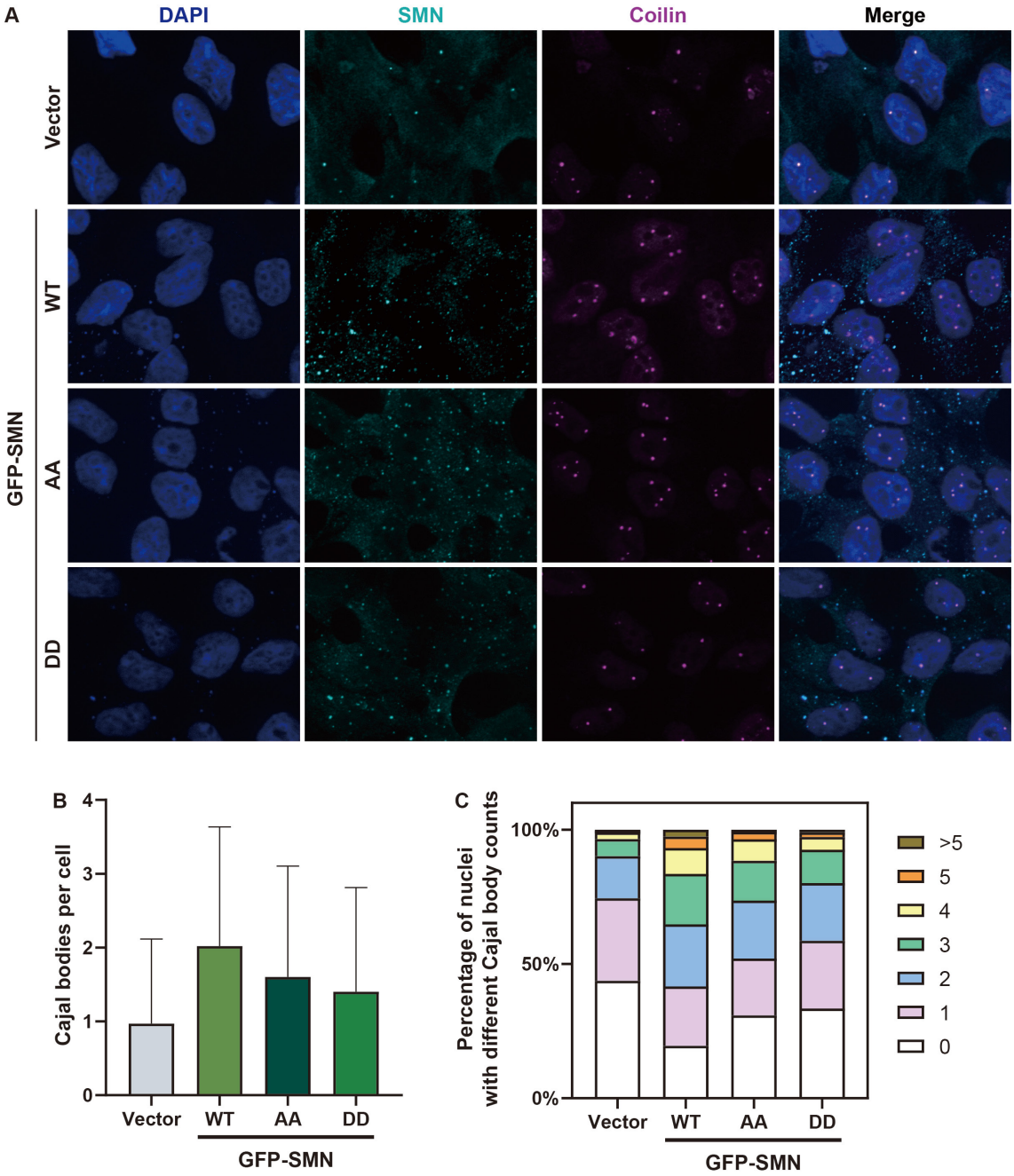


Fig. 16: ***Dual phosphorylation at T37 and S187 might impair capability of SMN to generate snRNPs.*** *(A) Representative immunofluorescence images of GFP (Vector) and GFP-SMN variants (WT, AA and DD) overexpression cell lines immuno-stained with SMN (cyan) and Coilin (magenta). Nuclei were counterstained with DAPI (blue). The colocalization of the SMN and coilin signals represents Cajal bodies, and its number was analysed by Fiji (ImageJ). (B) The total Cajal bodies count was divided by the number of analysed cells to determine the average. Data are presented as mean ± standard deviation (SD). (C) Cells were binned according to the number of Cajal bodies they contained, and the relative frequency of each subset was plotted as a histogram. The corresponding values of (B) and (C) panels are detailed in Table 5 and Table 6. WT, wild type; DAPI, 4',6-Diamidino-2-phenylindole.*

As expected, overexpression of the SMN WT protein significantly increased the number of Cajal bodies, evidenced by an almost two-fold increase in the average number of Cajal bodies per cell compared to the GFP vector control. Overexpression of SMN AA and DD also promoted Cajal body formation to a lesser extent than the WT. However, when

comparing these two mutants, cells overexpressing SMN DD contained slightly fewer Cajal bodies on average than those overexpressing AA mutant (Table 5 and Fig. 16 B).

Table 5: ***Average number of Cajal bodies per cell***

	Vector	GFP-SMN		
		WT	AA	DD
Mean ± SD	0.97 ± 1.15	2.02 ± 1.61	1.60 ± 1.50	1.40 ± 1.41

Notes: All values are rounded to two decimal places. Data are presented as a bar chart in Fig. 16 B. SD, standard deviation.

To further elucidate these differences, cells were categorised based on the number of Cajal bodies they contained, and the frequency distribution was plotted as a histogram. When SMN variants (WT, AA and DD) were overexpressed, more cells contained Cajal bodies. However, among these three cell lines, more GFP-SMN DD overexpression cells failed to generate Cajal bodies (33.38%), and fewer of them contained more than three Cajal bodies (7.45%) (Table 6 and Fig. 16 C). This observation suggested that the phosphorylation of T37 and S187 might impair snRNP biogenesis.

Table 6: ***Percentages of cells containing different numbers of Cajal bodies (%)***

Number of Cajal bodies	Vector	GFP-SMN		
		WT	AA	DD
0	43.72	19.53	30.89	33.38
1	30.76	22.13	21.18	25.23
2	15.68	23.12	21.53	21.53
3	6.36	18.78	14.83	12.40
4	2.46	9.66	8.00	4.79
5	0.54	4.26	2.70	1.65
>5	0.47	2.52	0.86	1.01

Notes: All percentage values are displayed to two decimal places. The corresponding histogram are shown in Fig. 16 C.

3.3 Gemin3 cleavage

In Chapter 3.1, EBSS was used to simulate extracellular nutrient starvation, which induced autophagy in GFP and GFP-SMN overexpression cells. Autophagy maintains cellular energy homeostasis through the degradation of various cellular cargoes [127, 128]. To disrupt the intracellular energy production derived from basal autophagy, BafA1, a widely-used autophagy inhibitor [129], was included in the study. Since autophagy governs protein degradation [130], the levels of the proteins in the SMN complex after EBSS and BafA1 treatment in GFP and GFP-SMN overexpression cells were assessed by immunoblotting. Among all the proteins analysed, Gemin3 was the most significantly affected by BafA1. Under basal conditions, a ~50 kDa Gemin3 fragment was abundant while full length (FL) Gemin3 (~100 kDa) levels were markedly low; however, BafA1 treatment induced the accumulation of FL-Gemin3 and the complete loss of the fragment (Fig. 17). This finding suggested that Gemin3 experiences proteolytic cleavage in the cytoplasm. The following studies in Chapter 3.3 were conducted to identify the mechanism and biological functions of this cleavage.

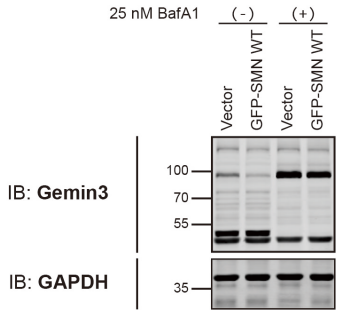


Fig. 17: ***FL-Gemin3 increases in GFP or GFP-SMN overexpression cells after 2-hour 25 nM BafA1 treatment.*** *Cells overexpressing GFP(Vector) or GFP-SMN WT were treated with 25 nM BafA1 for 2 h. Cytoplasmic extracts from these cells were prepared as described in Chapter 3.1, and the cellular level of Geimin3 was measured via immunoblotting. The membrane was cut between 55 kDa and 40 kDa to allow for the separate detection of Gemin3 and GAPDH (loading control), respectively. WT, wild type; IB, immunoblotting; BafA1, bafilomycin A1.*

3.3.1 Gemin3 is the only Gemin protein sensitive to bafilomycin A1

To further confirm that the finding in Fig. 17 was not specific to overexpression cell lines, HEK293T cells were treated with 25 nM BafA1 or EBSS for 2 h. All SMN complex–related proteins in their cytoplasmic extracts were detected via immunoblotting. To monitor the fragments, the full membranes were probed with the indicated antibodies.

The membrane was first incubated with Gemin3 monoclonal antibody (C-5), which specifically recognises the N-terminal of Gemin3 from amino acids (aa) 21–59 (Table 7). Consistent with Fig. 17, the faint FL-Gemin3 signal under basal conditions became prominently visible after BafA1 treatment. Besides, two Gemin3 cleavage fragments at the molecular weights ~50 kDa and ~35 kDa were found in untreated cells (hereafter

referred to as frag1-Gemin3 and frag2-Gemin3, respectively). These fragments were robustly decreased by BafA1, which might explain why more FL-Gemin3 were accumulated in BafA1-treated cells. EBSS also reduced the cleaved Gemin3 levels and correspondingly increased its FL form, but the effect was less pronounced than BafA1 (Fig. 18 A).

To get more information about the amino acid sequence of these cleavage fragments, immunoblotting was also performed with other Gemin3 monoclonal antibodies D-5, 12H12 and (2). As summarised in Table 7, D-5 and 12H12 recognise the internal region of Gemin3, while (2) targets the C-terminus. All three antibodies revealed the same effect of BafA1 on Gemin3 as seen in the C-5 immunoblot. Multiple cleavage fragments were detected by these antibodies as well, but none of them was in the same positions as frag1- and frag2-Gemin3, which were readily probed by C-5 antibody in untreated cells (Fig. 18 A).

Table 7: ***Gemin3 monoclonal antibodies and their target epitopes***

Clone name	Epitope region (aa)
C-5	21–59
D-5	216–360
12H12	368–548
(2)	667–783

Notes: aa, amino acids.

To examine whether the Gemin3 cleavage and its regulation by BafA1 are general, parallel analyses were conducted in other cell lines. Similar to HEK293T cells, FL-Gemin3 in HeLa cells was upregulated, and its cleavage fragments at ~50 kDa and ~35 kDa decreased dramatically after 2 h BafA1 treatment. However, in contrast to HEK293T cells, in which the cleaved Gemin3 was more prominent, the cleavage fragments in untreated HeLa cells were noticeably weaker than the FL form of Gemin3 (Fig. 18 B).

The evaluation was also performed on GFP-Gemin3 WT overexpression cells, which were established for subsequent experiments. GFP-Gemin3 shared the same cleavage pattern as endogenous Gemin3. The cleaved GFP-Gemin3 migrated to molecular weights of ~80 kDa and ~60 kDa (hereafter referred to as frag1-GFP-Gemin3 and frag2-GFP-Gemin3, respectively), aligning with the expected sum of the GFP tag (27 kDa) and

the endogenous cleaved form (~50 kDa and ~35 kDa). In line with HEK293T and HeLa cells, these fragments strikingly decreased after BafA1 treatment (Fig. 18 C).

Collectively, these findings confirmed that Gemin3 is cleaved under basal conditions, and this cleavage is inhibited by BafA1.

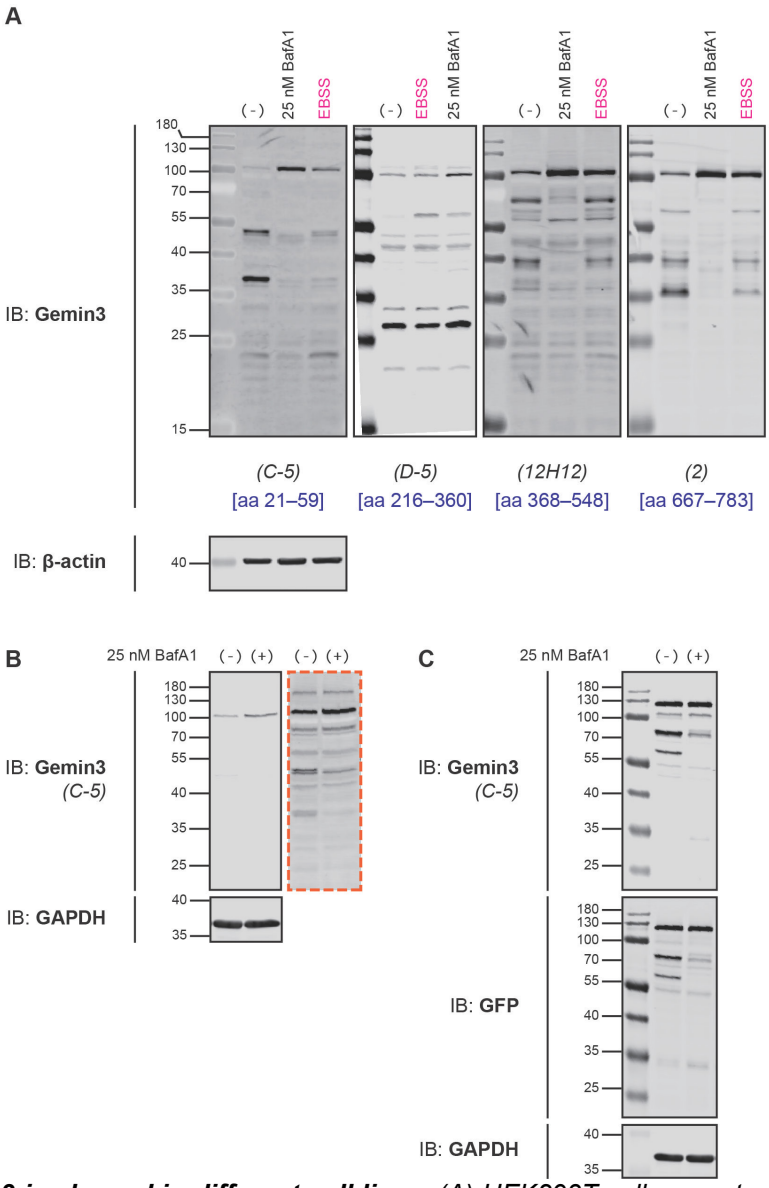


Fig. 18: ***Gemin3 is cleaved in different cell lines.*** *(A) HEK293T cells were treated with EBSS or 25 nM BafA1 for 2 h, and cytoplasm fractions were subsequently prepared and subjected to immunoblotting analysis. Gemin3 was probed by four commercial antibodies with different epitope locations, which are shown under each blot. (B) Gemin3 in HeLa cells with or without 2-hour 25 nM BafA1 treatment was detected following the procedures described in (A). The blot with higher brightness is presented in the right dashed orange boxes. (C) After stimulation with 0.1 μg/mL doxycycline overnight, GFP-Gemin3 overexpression cells were treated with 25 nM BafA1 for 2 h. The protocols of cytoplasmic extraction and immunoblotting are consistent with (A) and (B). GAPDH or β-actin served as the loading control. IB, immunoblotting; BafA1, bafilomycin A1; EBSS, Earle's balanced salt solution.*

Gemin3 is one of the components of the SMN complex [38, 47, 131]. To further explore the effect of BafA1 and EBSS intervention on other SMN complex proteins, HEK293T

cytoplasmic extracts from Fig. 18 A were also probed with SMN, other Gemin proteins and Unrip. Unlike Gemin3, these proteins showed no significant changes in signal distribution and intensity following the treatment (Fig. 19).

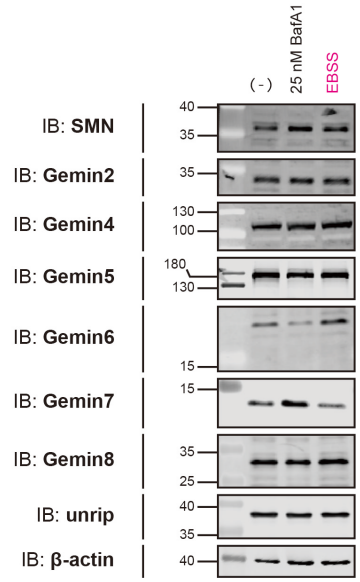


Fig. 19: ***Other SMN complex components are unaffected by BafA1 or EBSS treatment.*** *As described in Fig. 18 A, SMN complex proteins in cytoplasmic extracts from HEK293T cells with 2-hour EBSS or 25 nM BafA1 treatment were detected via immunoblotting. IB, immunoblotting; BafA1, bafilomycin A1; EBSS, Earle's balanced salt solution.*

3.3.2 Gemin3 could be cleaved by calpain-1 in vitro, but Gemin3 cleavage in vivo does not depend on this pathway

Proteolysis is generally regarded as the cause of protein cleavage. Calpain and caspase-3 are two vital cysteine proteases which regulate fundamental cellular processes through proteolysis [87, 132]. To test whether these proteases are related with Gemin3 cleavage, BafA1-treated HEK293T cytoplasmic extracts were subjected to the cleavage assay performed with exogenous caspase-3 or endogenous calpain activated by higher Ca^{2+} concentration. This pilot experiment showed that Ca^{2+} induced the cleavage of FL-Gemin3 instead of caspase-3, suggesting that calpain, a calcium-dependent protease, might be associated with Gemin3 cleavage (Fig. 20, a red arrow). To evaluate the likelihood of Gemin3 being cleaved by calpain, *in silico* prediction was performed by DeepCalpain [105]. Prediction using a general model provided 12 potential calpain cleavage sites on Gemin3. Interestingly, isoform-specific submodel predicted that Gemin3 might be only cleaved by calpain-1 (named as μ-calpain in prediction results), but not calpain-2 (Table 8).

Table 8: ***Top-5 Calpain cleavage site candidates according to FPR value***

Position	Enzyme	FPR	Peptide
311 | 312	μ-calpain	0.05%	FSRIPFNQ|ALVFSNLH
315 | 316	μ-calpain	0.14%	PFNQALVF|SNLHSRAQ
315 | 316	calpain	0.23%	PFNQALVF|SNLHSRAQ
304 | 305	calpain	0.24%	TQHLQELF|SRIPFNQA
309 | 310	μ-calpain	0.27%	ELFSRIPF|NQALVFSN

Notes: Data were downloaded from the DeepCalpain web server (http://deepcalpain.omicsbio.info/) [105]. The amino acid sequence of Gemin3 (Uniprot ID: Q9UHI6) retrieved from Uniprot was uploaded to the web server for analysis. The FPR value reflects the likelihood of prediction errors. Lower FPR values indicate higher prediction confidence. The vertical lines (|) between position numbers and amino acid abbreviations represent the cleavage sites. FPR, false positive rate.

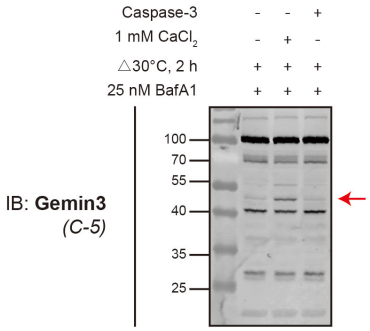


Fig. 20: ***Ca^{2+} induces Gemin3 cleavage.*** *To inhibit the basal cleavage and accumulate FL-Gemin3 in the cytoplasm, HEK293T cells were treated with 25 nM BafA1. Cytoplasmic extracts from these cells were incubated with 1 mM $CaCl_2$ or caspase-3 with a ratio of 0.08 μg of enzyme per 100 μg of total protein. The reactions were incubated at 30°C for 2 h and subsequently subjected to immunoblotting analysis with the Gemin3 (C-5) antibody. The cleavage fragment is marked with a red arrow. IB, immunoblotting; BafA1, bafilomycin A1.*

To further verify the prediction and test whether calpain is responsible for the basal cleavage of Gemin3, the *in vitro* calpain assay was performed as described in Fig. 20, but using native calpain-1. HEK293T cells were pre-treated with BafA1 to eliminate Gemin3 fragments. Cytoplasmic fractions from these cells were incubated with indicated 1 mM $CaCl_2$ or calpain-1 at 30°C for 2 h. By immunoblotting, cleavage products in reactions were analysed along with untreated cells, which served as the reference to determine whether calpain-1 could reproduce the frag1- and frag2-Gemin3.

Following BafA1 treatment, FL-Gemin3 was accumulated substantially, but its cleavage fragments largely disappeared. Consistent with Fig. 20, a faint cleavage band emerged again at a similar position with frag1-Gemin3 (~50 kDa) when cytoplasmic extracts from BafA1-treated HEK293T cells were incubated with 1 mM $CaCl_2$. The cleavage was markedly augmented by calpain-1. When cytoplasmic extracts were exposed to 1.5 U calpain-1, this cleavage fragment became more prominent. Beyond this, another

fragment was detected at around 55 kDa. Reactions with a higher concentration of calpain shared the same cleavage pattern, but the abundance of fragments decreased in a dose-dependent manner (Fig. 21, red arrows). These findings suggested that calpain-1 might be partially responsible for the basal cleavage.

To get more information about the cleavage products by calpain-1, the reactions were also probed with other Gemin3 antibodies targeting different epitopes (Table 7). All antibodies demonstrated that FL-Gemin3 was steadily cleaved with increasing concentration of calpain-1, but they revealed different cleavage fragments (Fig. 21, red arrows). Fragments generated by calpain-1 were also detected in untreated cytoplasmic extract by D-5 and 12H12 antibodies. D-5 revealed a fragment under 25 kDa, slightly increased after incubation with calpain-1. 12H12 antibodies bound with a 55 kDa fragment at the same position as the one observed on the C-5 blot. In contrast to these two antibodies, the (2) antibody recognised a cleavage fragment under 40 kDa, which was not found in untreated sample. These observations demonstrated diverse cleavage sites on Gemin3, which is aligned with the prediction results.

To further explore the effect of calpain-1 on the SMN complex, SMN and other Gemin proteins were also examined in this cleavage assay. Gemin5 is another Gemin protein cleaved by calpain-1. When exposed to 1 mM $CaCl_2$, a faint truncated band appeared at approximately 70 kDa. FL-Gemin5 was further proteolysed into this fragment after incubation with 1.5 U calpain-1 for 2 h. As the concentration of calpain-1 rose to 15 U, ~70 kDa Gemin5 moderately decreased, but an additional band slightly under this fragment became more pronounced (Fig. 21, red arrows). These findings suggested that Gemin5 is also a substrate of calpain-1, and this proteolysis might occur in a stepwise manner.

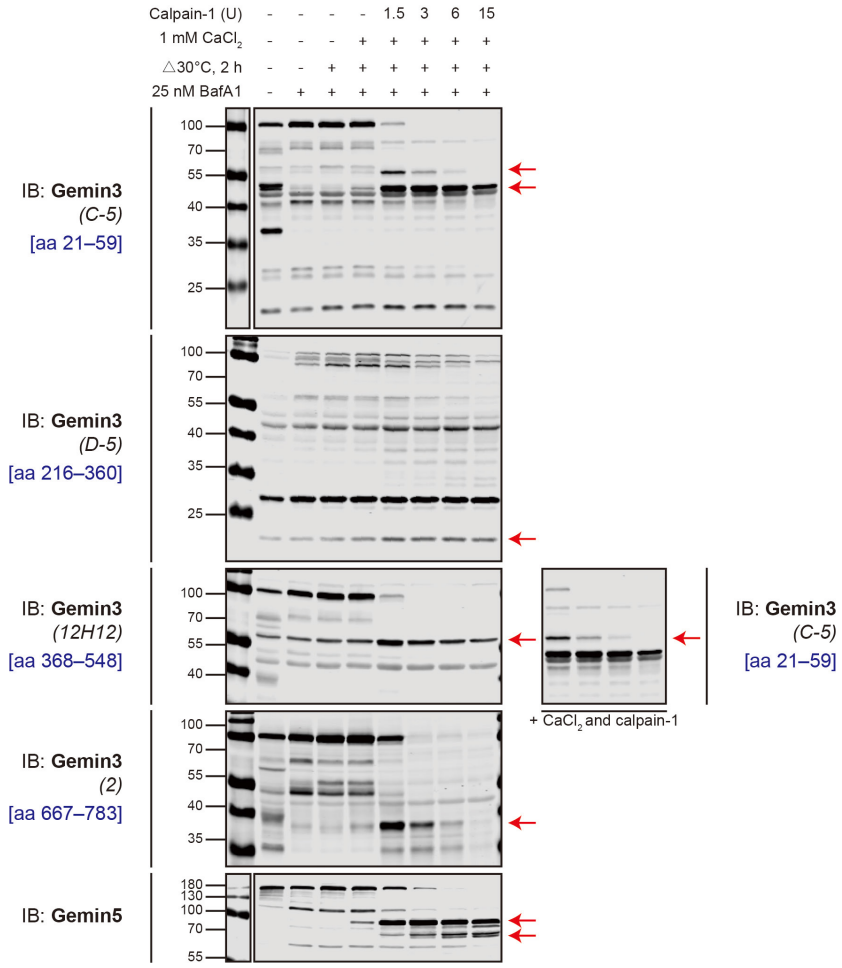


Fig. 21: ***FL-Gemin3 and Gemin5 could be cleaved by calpain-1 in vitro.*** *Cytoplasmic extracts from HEK293T cells pre-treated with 25 nM BafA1 were incubated with 1 mM $CaCl_2$ and calpain-1 at a ratio of 1.5, 3, 6 and 15 U of enzyme per 100 µg of total proteins at 30°C for 2 h. The resulting cleavage products in the assay were analysed via immunoblotting with the Gemin3 and Gemin5 antibodies. To better compare the cleavage bands detected by the 12H12 and C-5 antibodies, an excerpt from the C-5 blot is displayed next to the 12H12 blot with aligned molecular weight markers. Red arrows mark the cleavage fragments. BafA1, bafilomycin A1; IB, immunoblotting.*

No cleavage was observed in other tested proteins; however, they exhibited different sensitivity to calpain-1 (Fig. 22). SMN was most vulnerable to calpain-1, which was completely depleted in calpain-containing reactions. Compared with SMN, Gemin4, Gemin6 and pICln were more stable, but slight degradation could be triggered by high concentration of calpain-1 (15 U). More interestingly, PRMT5 and WD45 were refractory to calpain-1 treatment. They presented robustly under all experimental conditions without any cleavage or degradation.

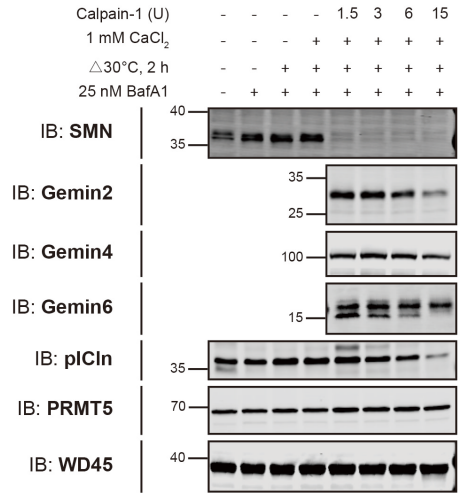


Fig. 22: ***snRNP biogenesis–related proteins have different vulnerability to calpain-1.*** *The reactions from the in vitro calpain-1 cleavage assay (Fig. 21) were probed with the indicated antibodies. BafA1, bafilomycin A1; IB, immunoblotting.*

In this *in vitro* calpain assay, excess calpain (up to 15 U) and prolonged incubation time (2 h) might result in non-specific breakdown. Therefore, to further validate the impact of calpain-1 on SMN complex proteins, especially Gemin3 and Gemin5, a calpain assay was performed again, but with less calpain-1 (1 U and 0.5 U) and shorter incubation (30 minutes (min)). Besides, to exclude the potential effect of BafA1 treatment on the cleavage, cytoplasm extracts from untreated HEK293T cells were also tested in this assay.

Although the amount of enzyme and incubation time were restricted, FL-Gemin3 and Gemin5 were still cleaved by calpain-1 (Fig. 23 A, red arrows). No cleavage fragment produced by calpain-1 was detected in reactions using untreated cells, but the decrease of FL-Gemin3 was still visible (Fig. 23 A, blot in orange dashed box). When cytoplasmic extracts from BafA1-treated cells were incubated with calpain-1, the cleavage fragments of FL-Gemin3 appeared persistently around 55 kDa areas (Fig. 23 A, red arrows). The intensities of these fragments were almost comparable, suggesting that endogenous Ca^{2+} concentration is sufficient for this cleavage. In contrast to Gemin3, FL-Gemin5 was unaffected by BafA1 treatment; therefore, the calpain-1 cleavage fragment around 70 kDa was found in both cytoplasmic extracts upon exposure to calpain-1. These fragment bands were more noticeable if extra Ca^{2+} was supplied into the reactions, implying that Gemin5 cleavage might require certain Ca^{2+} levels.

SMN and other Gemin proteins were also examined. But no cleavage or degradation of them was detected in this assay (Fig. 23 B).

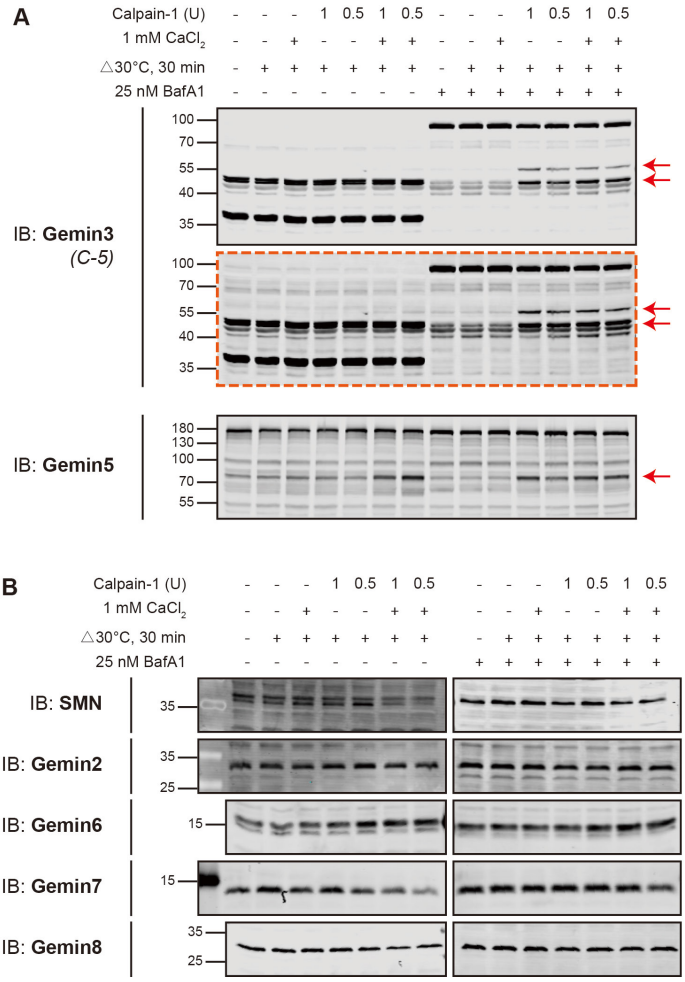


Fig. 23: ***FL-Gemin3 and Gemin5 are sensitive to calpain-1.*** *In vitro calpain assay was performed similarly to Fig. 21 and Fig. 22. HEK293T cells were treated with or without 25 nM BafA1 before extraction. The resulting cytoplasmic extracts were then incubated with 1 mM $CaCl_2$ or calpain-1 at a ratio of 0.5 and 1 U of enzyme per 100 µg of total proteins for 30 min at 30°C. (A) The reactions were subsequently subjected to immunoblotting analysis using the Gemin3 (C-5) and Gemin5 antibodies. (B) Other SMN complex proteins were also probed in the reactions via immunoblotting. Red arrows mark the cleavage fragments. BafA1, bafilomycin A1; IB, immunoblotting.*

The preceding *in vitro* cleavage assays demonstrated that Gemin3 is highly susceptible to calpain-1, and one of the cleavage products around 55 kDa might be frag1-Gemin3. To examine whether this is the case *in vivo*, GFP-Gemin3 overexpression cells and HEK293T cells were treated with 12.5–50 µM of calpeptin (a calpain inhibitor [133]), and then Gemin3 in the cytoplasm was analysed via immunoblotting at various time points. Unlike BafA1, calpeptin failed to prevent the cleavage of Gemin3. Immunoblotting analysis revealed that the signal intensity and distribution of FL and cleaved Gemin3 were unaltered regardless of cell type, calpeptin concentration, or exposure time (Fig. 24).

Taken together, these data demonstrated that calpain-1 is able to cleave Gemin3, but it might not be the primary reason for Gemin3 cleavage under basal conditions.

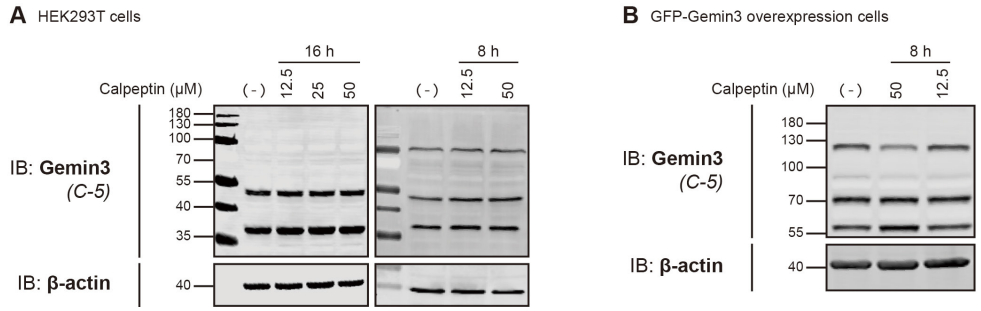


Fig. 24: ***Calpeptin does not inhibit Gemin3 cleavage in vivo.*** *Representative immunoblots of Gemin3 in (A) HEK293T and (B) GFP-Gemin3 overexpression cells. Cells were treated with calpeptin as indicated. Following treatment, cytoplasmic extracts were prepared and then analysed by immunoblotting using the Gemin3 (C-5) antibody. β-actin was used as the loading control to ensure equal protein concentration across samples. IB, immunoblotting.*

3.3.3 Frag1-Gemin3 is not included into the SMN complex

To understand the difference between FL and cleaved Gemin3, the distribution of these three Gemin3 populations was examined by SEC using cytoplasmic extracts from HEK293T and GFP-Gemin3 WT overexpression cells. Although BafA1 significantly reduced the Gemin3 fragments in HEK293T cells, it did not alter the distribution of FL-Gemin3 and other SMN complex proteins. All of them primarily co-eluted in the high molecular fractions from 1 to 5. A portion of frag2-Gemin3 (~35 kDa) eluted in these fractions, but the remainder was co-fractionated with frag1-Gemin3 (~50 kDa), which was exclusively distributed in fractions 17–20 (158–440 kDa) (Fig. 25).

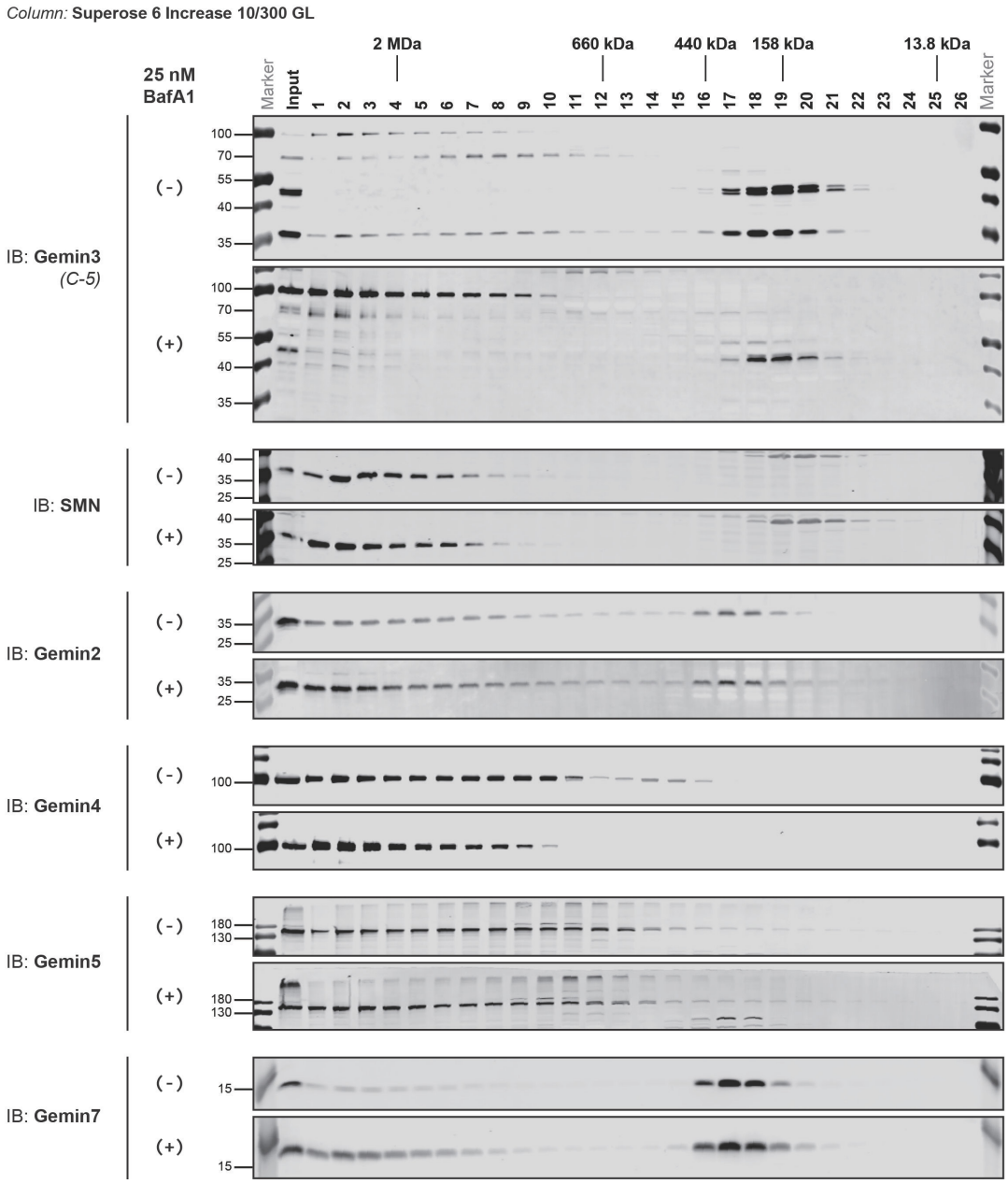


Fig. 25: ***FL-Gemin3, rather than its cleavage forms, co-elutes with other SMN complex components.*** *HEK293T cells were treated with or without 25 nM BafA1 for 2 h. Clarified cytoplasmic extracts were fractionated using the Superose 6 column. Equal volumes of the collected fractions were subjected to immunoblotting using the indicated antibodies. BafA1, bafilomycin A1; IB, immunoblotting.*

GFP-Gemin3 has the same distribution pattern of endogenous FL and fragments. FL- and frag2-GFP-Gemin3 (~60 kDa) were detected early in fractions 1–5, while frag1-GFP-Gemin3 (~80 kDa) eluted much later than them, distributed in 16–18 fractions. SEC suggested that FL and cleaved Gemin3 might be localised into two distinct complexes (Fig. 26 A). To investigate the composition of these complexes, GFP-IP was performed using fractions 1–5 and 16–18. As expected, SMN and Gemin proteins in fractions 1–5 were successfully co-immunoprecipitated with FL- and frag2-GFP-Gemin3. Conversely, only a small amount of Gemin4 was enriched in the IP pool by GFP-Gemin3 fragments

when using fractions 16–18. These findings demonstrated that frag1-Gemin3 was not incorporated into the SMN complex (Fig. 26 B).

To examine the affinity of FL and cleaved Gemin3 with other SMN complex components, one aliquot of cytoplasmic extract from GFP-Gemin3 WT overexpression cells was pre-treated with 2 U calpain-1 at 30°C for 2 h to reduce the level of FL-Gemin3 before GFP-IP. With the exception of SMN, the abundance of Gemin2, Gemin4, Gemin6, Gemin7 and SmB was not altered by calpain-1. Among these unaffected proteins, only Gemin4 was successfully co-precipitated to the same extent in both cytoplasmic extracts. Other proteins in the IP fraction decreased substantially when cytoplasm lacked FL-Gemin3 (Fig. 26 C). This suggested that FL-Gemin3 is necessary for interaction with other SMN complex proteins except Gemin4.

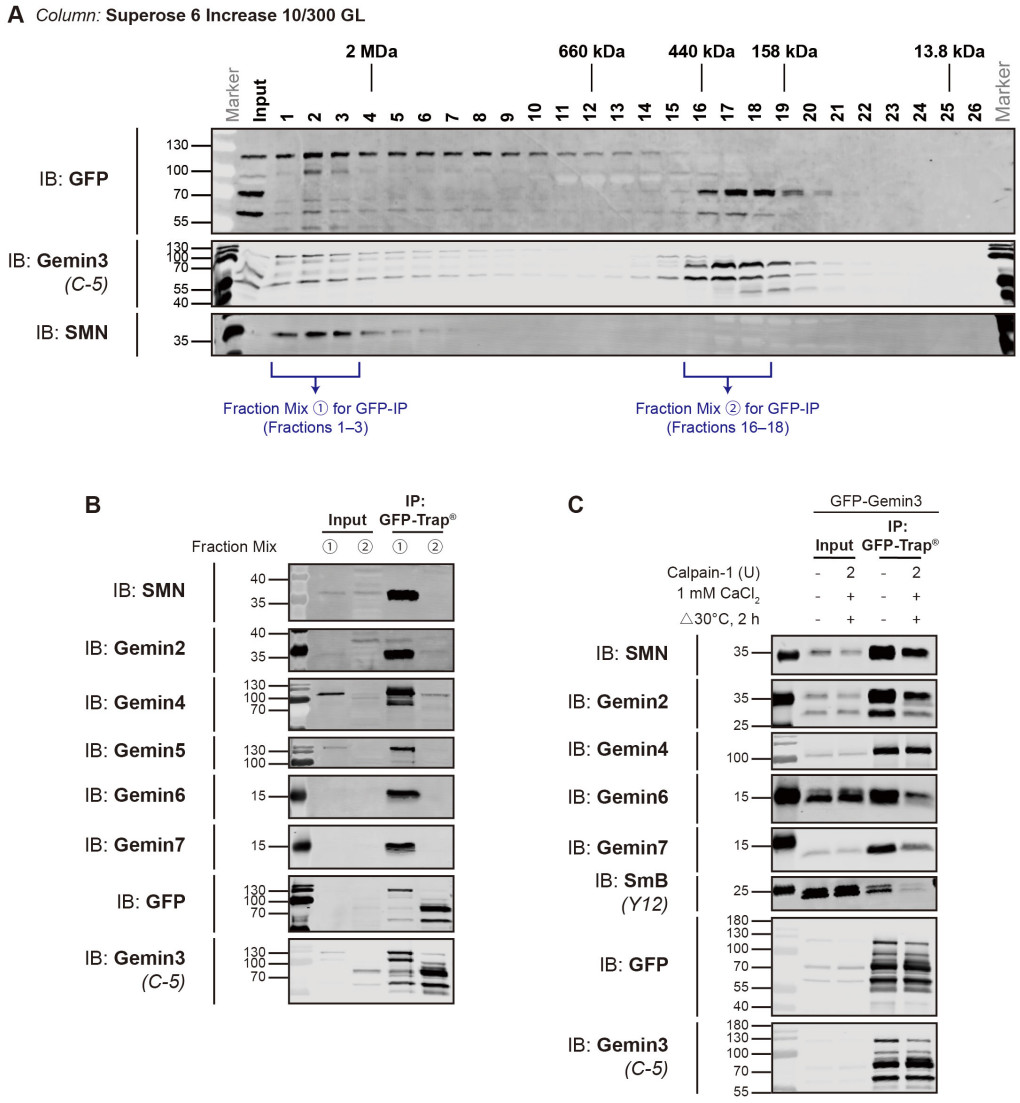


Fig. 26: ***Frag1-Gemin3 is not included in the SMN complex.*** *(A) Immunoblotting analysis of SEC fractions from GFP-Gemin3 WT overexpression cells was performed as described previously. (B) Fractions 1–3 (fraction mix ①) and 16–18 (fraction mix ②) from (A) were pooled and subjected to GFP-IP, followed by immunoblotting for the indicated proteins. (C) The cytoplasmic extract from GFP-Gemin3 overexpression cells was pretreated with or without*

2 U/100 µg of total proteins calpain-1 for 2 h at 30°C before GFP-IP. Co-immunoprecipitated proteins were detected via immunoblotting. IB, immunoblotting; IP, immunoprecipitation.

3.3.4 The C-terminus, rather than the N-terminus, is necessary for Gemin3 to operate in the SMN complex

Section 3.3.3 revealed that frag1-Gemin3 is not incorporated into the SMN complex. To experimentally model this cleavage product and investigate its downstream functions, cell lines overexpressing truncated GFP-Gemin3 were generated. According to Fig. 18 A, the frag-1 fragment should be ~50 kDa containing the N-terminus (C-5 membrane), and its complementary C-terminus might be present in cells as well (12H12 and (2) membranes). Therefore, one GFP-Gemin3 truncation spanned from aa 1 to 478, which closely simulated the size and domain architecture of the endogenous frag1-Gemin3. Another truncation containing the rest aa 479–824 was also included in this study (Fig. 27).

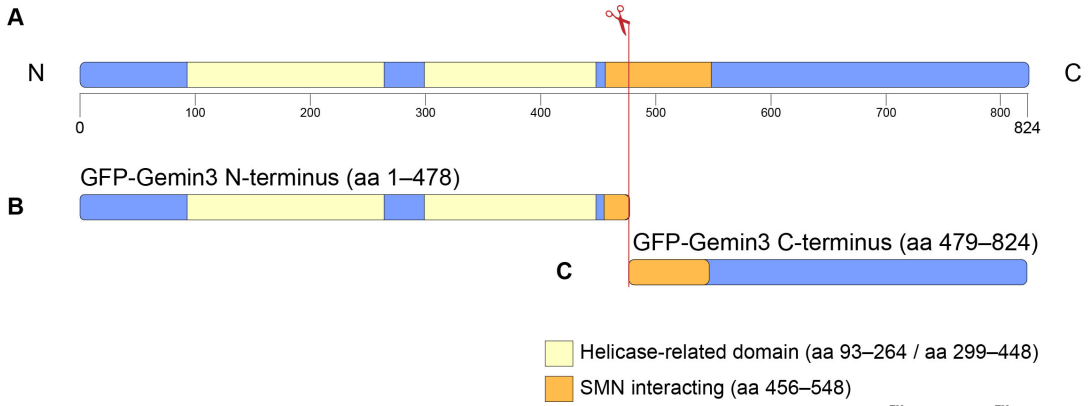


Fig. 27: ***Illustration of truncated GFP-Gemin3 overexpressed in Flp-In™ T-REx™ 293 cell lines.*** *(A) A stylised linear representation of the FL-Gemin3 polypeptide chain, spanning residues 1 to 824. Key functional domains are indicated by coloured blocks. The N-terminal regions, from residues 93–264 and 299–448, are the domains related to its helicase function, which is marked in light yellow. The region from residues 456–548 marked in orange is the domain responsible for SMN interaction. (B) GFP-Gemin3 N-terminus overexpression cells contain GFP-Gemin3 from aa 1–478, whereas (C) GFP-Gemin3 C-terminus cells overexpress the rest of GFP-Gemin3 from aa 479–824. The residue numbers and domain features are based on Uniprot (Q9UHI6, DDX20_HUMAN). aa, amino acids; N, N-terminus; C, C-terminus.*

To determine in which complexes these two GFP-Gemin3 truncations were distributed, SEC was performed using cytoplasmic extracts from GFP-Gemin3 N- and C-terminus overexpression cells. SMN distribution was not affected by the overexpression of these GFP-Gemin3 truncations. GFP-Gemin3 N-terminus has the same distribution pattern as endogenous frag1-Gemin3. When Gemin3 lost its C-terminus, its SEC profile was shifted to lower fractions 16–18. Conversely, Gemin3 without N-terminus still co-eluted with SMN in high-molecular-weight fractions (Fig. 28 A). To test whether the GFP-Gemin3 N-terminus is excluded from the SMN complex, GFP-IP was performed with cell lines overexpressing GFP-Gemin3 truncations. The amount of SMN complex proteins co-

immunoprecipitated by Gemin3 C-terminus was comparable with WT; however, these proteins exhibited significantly reduced affinity with the N-terminus (Fig. 28 B).

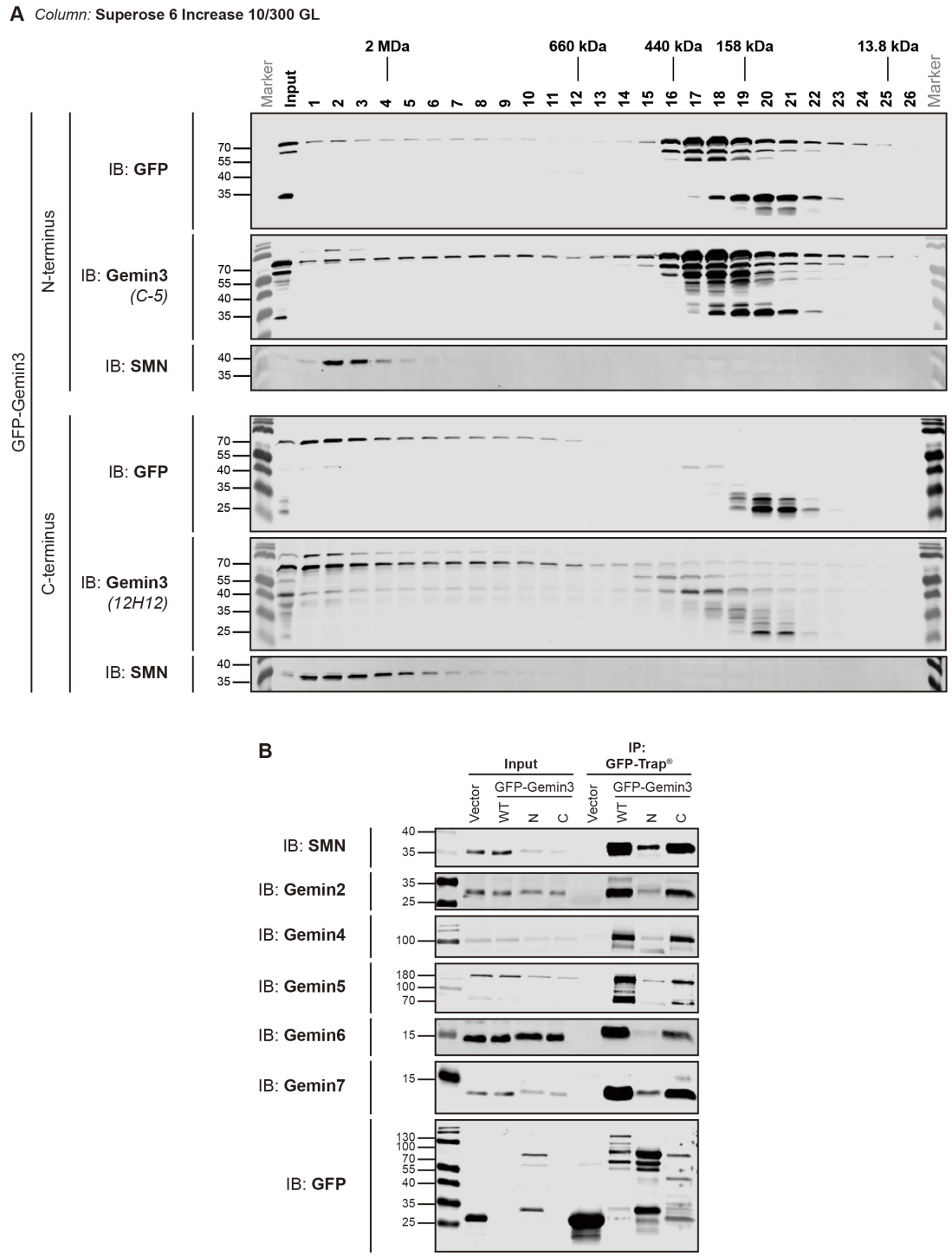


Fig. 28: ***GFP-Gemin3 N-terminus fails to interact with the SMN complex.*** *(A) The cytoplasmic extracts from GFP-Gemin3 N- or C-terminus overexpression cells were subjected to SEC analysis by the Superose 6 column. Protein distribution was visualised via immunoblotting. (B) GFP-IP was performed using cytoplasmic extracts from GFP-Gemin3 variant overexpression cells (WT, N- and C-terminus). Prey proteins were detected by immunoblotting. WT, wild type; N, N-terminus; C, C-terminus; IB, immunoblotting; IP, immunoprecipitation.*

The SMN complex recruits snRNA and Sm proteins for snRNP assembly [15, 16]. Gemin3, as a member of the SMN complex, contains RNA helicase domains in the N-

terminus (Fig. 27 A), which is important for RNA processing [47]. Preceding findings suggested that the Gemin3 N-terminus is not integrated into the SMN complex. To investigate whether the presence of different Gemin3 fragments could reshuffle the snRNA location, cytoplasmic extracts from HEK293T cells and truncated GFP-Gemin3 overexpression cells were pre-incubated with fluorescein-labelled snRNA, and subsequently separated by the Superose 6 column. After resolving the fractions by SDS-PAGE, SMN was detected via immunoblotting to indicate the distribution of the SMN complex, and the fluorescent snRNA was visualised under a UV transilluminator. Two populations of snRNA were observed when snRNA was incubated with HEK293T cytoplasm. One co-eluted with the SMN complex in high-molecular-weight fractions 1 to 5 (2 MDa); another was distributed in fractions 11 to 13 (660 kDa). This distribution pattern of snRNA remained unaltered in the presence of a high concentration of GFP-Gemin3 WT or C-terminus. However, when the GFP-Gemin3 N-terminus was predominant in the reaction, snRNA was only localised to the lower-molecular-weight complex. This finding suggested that the Gemin3 C-terminus is necessary for the SMN complex to accept snRNA (Fig. 29).

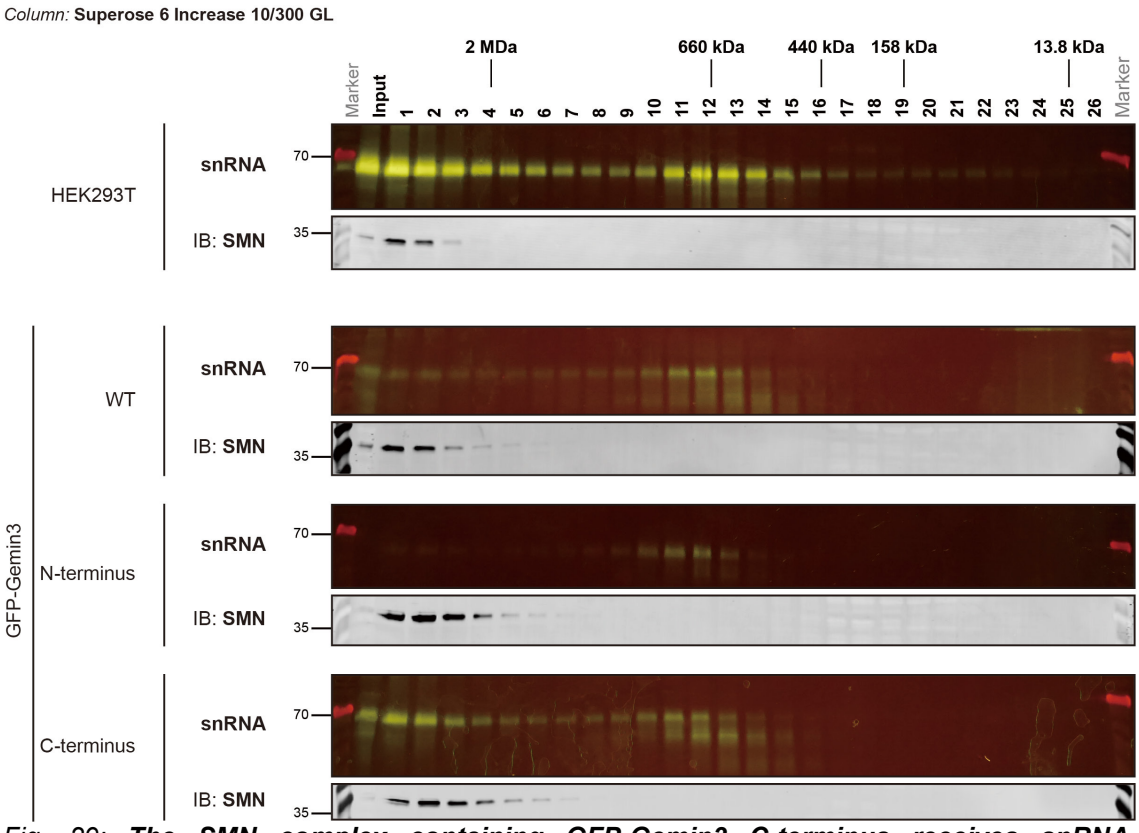


Fig. 29: ***The SMN complex containing GFP-Gemin3 C-terminus receives snRNA.*** *Cytoplasmic extracts from HEK293T and GFP-Gemin3 variant overexpression cells (WT, N- and C-terminus) were incubated with fluorescein-labelled U1 snRNA at 37°C for 30 min before being loaded on the Superose 6 column. The proteins and snRNA in each fraction were separated by SDS-PAGE. The distribution of U1 snRNA was captured under a UV transilluminator. SMN on the same SDS-PAGE gel was detected via immunoblotting as the representative of the SMN complex. IB, immunoblotting; WT, wild type.*

To further examine the capability of snRNP biogenesis when Gemin3 fragments were present in cells, the amount of Cajal bodies were measured by IF as described in Section 3.2.3. The overlap signals of SMN and coilin were considered as Cajal bodies [62, 126]. Unlike SMN, overexpression of Gemin3 WT did not promote Cajal body formation. The average of Cajal bodies was comparable across all cell lines (Table 9, Table 10, Fig. 30 B and C). However, the localisation of the GFP-Gemin3 N-terminus was strikingly different. Whereas both GFP-Gemin3 WT and C-terminus condensed into discrete puncta, the N-terminus protein was diffusely distributed throughout the cells (Fig. 30 A).

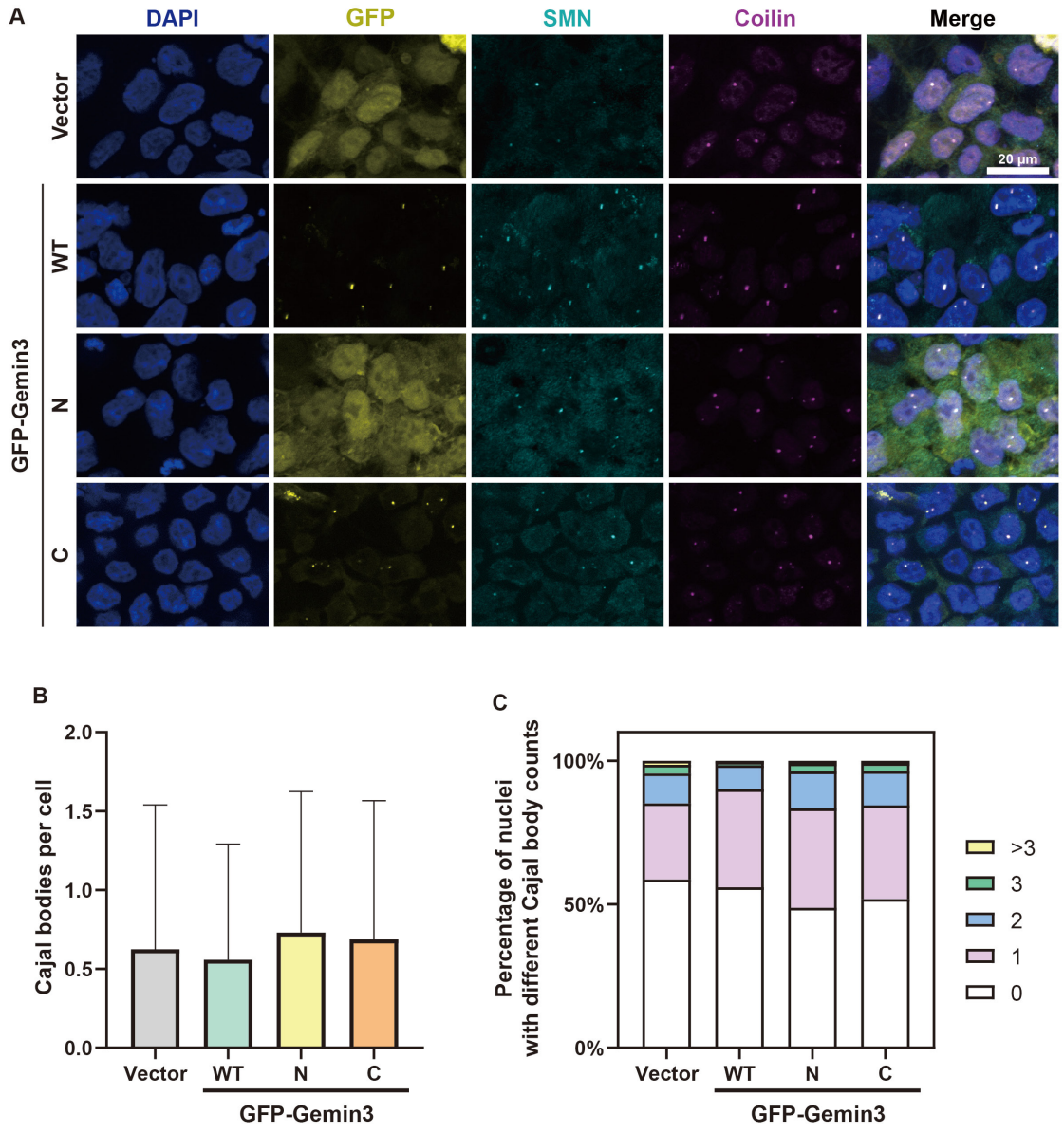


Fig. 30: ***Overexpression of Truncated Gemin3 does not interfere with Cajal body generation.*** *GFP (Vector) and GFP-Gemin3 variants (WT, N- and C-terminus) overexpression cells were subjected to immunofluorescence staining following the same protocol as described in Fig. 16 to visualise and quantify Cajal bodies. (A) Representative immunofluorescence images of cells co-stained with antibodies against SMN (cyan) and coilin (magenta) to locate Cajal bodies. Nuclei were counterstained with DAPI (blue). (B) Average number of Cajal bodies per cell. Data are presented as mean ± standard deviation (SD). (C) Frequency distribution indicating the percentage of cells containing the specified numbers of Cajal bodies. To obtain data for (B) and*

(C), fluorescent images were analysed and quantified using Fiji (ImageJ) software. The corresponding values are detailed in Table 9 and Table 10. WT, wild type; N, N-terminus; C, C-terminus; DAPI, 4',6-Diamidino-2-phenylindole.

Table 9: ***Average number of Cajal bodies per cell***

	Vector	GFP-Gemin3		
		WT	N-terminus	C-terminus
Mean ± SD	0.62 ± 0.92	0.56 ± 0.73	0.73 ± 0.89	0.69 ± 0.88

Notes: All values are rounded to two decimal places. Data are presented as a bar chart in Fig. 30 B. SD, standard deviation.

Table 10: ***Percentages of cells containing different numbers of Cajal bodies (%)***

Number of Cajal bodies	Vector	GFP-Gemin3		
		WT	N-terminus	C-terminus
0	58.66	55.95	48.81	51.81
1	26.52	34.17	34.61	32.69
2	10.44	8.35	12.89	11.86
3	2.98	1.18	2.81	2.83
>3	1.40	0.34	0.88	0.80

Notes: All percentage values are displayed to two decimal places. The corresponding histogram are shown in Fig. 30 C.

To explore the biological functions of these two fragments, interactors in GFP-IP (Fig. 28 B) were identified by MS as described previously (Sections 3.1.2 and 3.2.3), and the interactomes unique to GFP-Gemin3 N- or C- terminus, as well as their respective intersections with WT, were subjected to GO enrichment analysis. Their proteomic networks of N- and C- terminus were highly distinct. Proteins interacting exclusively with the Gemin3 N-terminus were uniquely enriched in cyclin binding and protein kinase C (PKC) binding, whereas the binding partners specific to the C-terminus engaged in diverse biological processes and molecular functions. Notably, only the C-terminus shared an overlapping interactome with WT, which was highly enriched in RNA-related processes, suggesting that the C-terminus preserves more functional properties of WT (Fig. 31).

Collectively, these findings indicated that the C-terminus is more important than the N-terminus for Gemin3 to participate in the SMN complex.

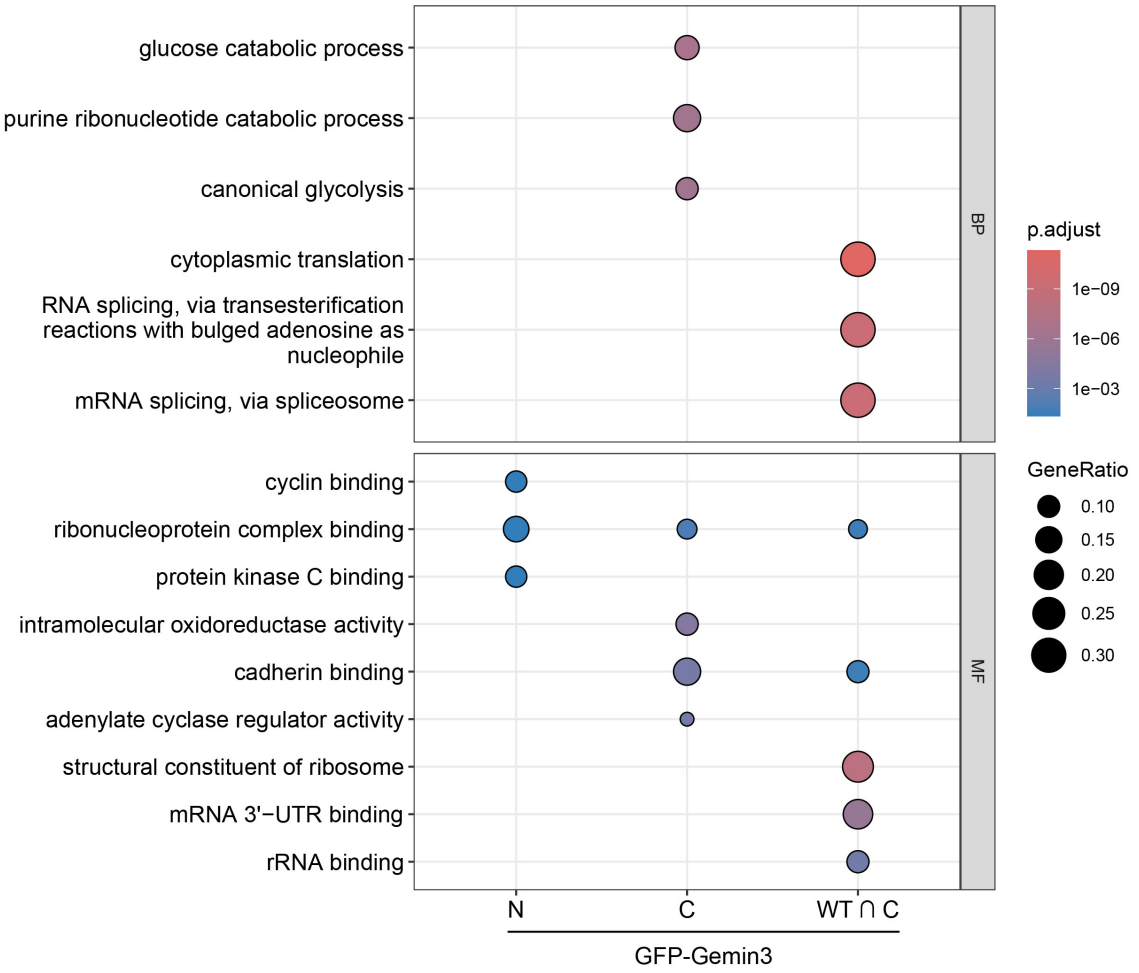


Fig. 31: ***The GFP-Gemin3 N- and C-terminus interactomes are enriched in distinct biological processes and molecular functions.*** *By the analysis method described in Fig. 7, interactomes identified by MS after GFP-IP were subjected to GO enrichment analysis. The top-3 enriched GO terms of biological process and molecular function are shown for the proteins only identified in the GFP-Gemin3 N- or C-terminus networks, as well as the binding proteins shared between WT and C-terminus. The size of each dot represents the number of genes overlapping with the respective GO term (gene ratio), and the colour gradient indicates the statistical significance (adjusted p-value). BP, biological process; MF, molecular function; WT, wild type; N, N-terminus; C, C-terminus.*

3.4 Gemin3 degradation

BafA1 treatment reversed the Gemin3 immunoblotting profile, leading to a dramatic increase in FL-Gemin3 but reduction in fragments (Fig. 18 A). Research about the cleavage inhibition by BafA1 is still in its infancy, but it has been robustly established as a vacuolar-type H^+-ATPase (V-ATPase) inhibitor [134]. V-ATPase within the lysosome membrane plays an important role in lysosomal functions [135]. The alteration of Gemin3 after BafA1 treatment highly implies that Gemin3 might undergo degradation as well. This chapter aimed to explore the potential mechanism for Gemin3 degradation.

3.4.1 FL-Gemin3 levels are highly associated with the lysosome

The ubiquitin-proteasome system (UPS) and lysosomal pathways dominate protein degradation in cells [136]. To test whether the turnover of FL-Gemin3 was regulated by the proteasome, HEK293T cells were treated with MG132, a potent proteasome inhibitor [137]. Following a 24-hour incubation with 1 or 3 μM MG132, cytoplasmic extracts were collected to evaluate Gemin3 levels via immunoblotting analysis. Unlike BafA1, which caused a remarkable increase in FL-Gemin3, no accumulation of Gemin3 was observed in the MG132-treated cells compared to the control (Fig. 32 A). This result indicated that the basal turnover of Gemin3 might not rely on the UPS.

The uptake of proteins into the lysosome primarily relies on the autophagy pathway [138]. The elevation of FL-Gemin3 induced by BafA1, an autophagy inhibitor [129], strongly implies the association between Gemin3 degradation and autophagy. To further understand how the level of FL-Gemin3 was mediated by autophagy, the amount of Gemin3 in ATG5 knock-out (KO) cells was measured via immunoblotting. ATG5 KO cells cannot form autophagosomes, which blocks canonical macroautophagy. FL-Gemin3 did not increase in KO cells, suggesting that the lysosome does not take up Gemin3 via macroautophagy (Fig. 32 B).

Starvation by 2-hour EBSS treatment is a well-established method to induce macroautophagy [129]. However, this treatment had the same effect as BafA1 on FL-Gemin3, which also supported that FL-Gemin3 degradation is not governed by macroautophagy (Fig. 18 A). Since BafA1 inhibits autophagy by targeting the lysosome [129, 134], the increase in FL-Gemin3 caused by EBSS might also be attributed to its impact on the lysosome. To test this hypothesis, HEK293T cells were treated with EBSS or BafA1 for 6, 24 and 72 h, and the levels of Gemin3 and lysosome-associated membrane protein 2A (LAMP2A), a marker of the lysosome, were monitored by immunoblotting. A negative correlation was found between these two proteins. Prolonged exposure to EBSS downregulated LAMP2A and upregulated FL-Gemin3, whereas

treatment with BafA1 for a longer period produced the inverse result (Fig. 32 C). To validate these findings, GFP-Gemin3 overexpression cells were treated with EBSS for 72 h. Consistent with the observation in HEK293T cells, LAMP2A decreased significantly following treatment, whereas both endogenous and GFP-tagged FL-Gemin3 increased markedly (Fig. 32 D).

These findings suggested that FL-Gemin3 degradation is directly sent to the lysosome rather than by vehicles for degradation, which implied that the FL-Gemin3 might be involved in chaperone-mediated autophagy (CMA).

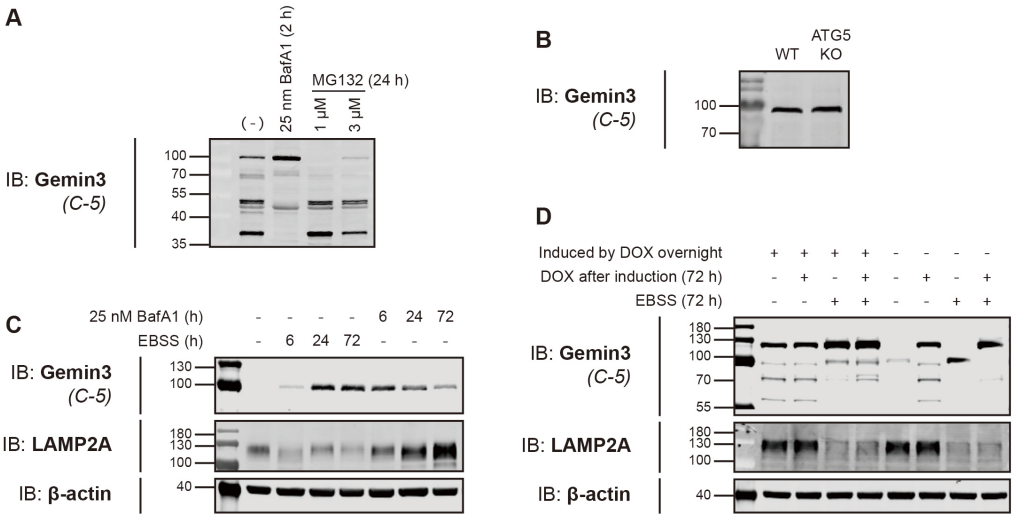


Fig. 32: ***Gemin3 abundance is negatively correlated with LAMP2A.*** *(A) HEK293T cells were treated with BafA1 and MG132 as indicated. Cytoplasmic extracts were prepared subsequently, and Gemin3 was measured via immunoblotting. (B) WT and ATG5 KO HeLa cells were prepared by Dr. María José Mendiburo. (C) HEK293T cells were treated with 25 nM BafA1 or EBSS for the indicated time. (D) GFP-Gemin3 overexpression cells were cultured with or without 0.1 μg/mL doxycycline overnight. Following the pre-treatment, cells were treated with EBSS alone or co-treated with EBSS and doxycycline for an additional 72 h. For (B–D) panels, whole-cell lysates prepared with RIPA buffer were analysed by immunoblotting using the indicated antibodies. β-actin was used to verify equal protein loading. IB, immunoblotting; WT, wild type; KO, knock-out; DOX, doxycycline; BafA1, bafilomycin A1; EBSS, Earle's balanced salt solution.*

3.4.2 FL-Gemin3 levels are regulated by chaperon-mediated autophagy

The presence of KFERQ-like motifs (hereafter referred to as KFERQ motifs) is a critical prerequisite for a protein degraded by CMA [95, 139]. Heat shock cognate 71 kDa protein (HSC70; HSPA8) specifically binds to this motif and sends the substrates to the lysosome [95]. To evaluate the potential of Gemin3 to be a substrate for CMA, its amino acid sequence was analysed by KFERQ finder V0.8 [106]. Nine KFERQ motifs were found on the sequence of Gemin3 by the computational screening. Two of them were canonical KFERQ motifs (Table 11, highlighted in yellow and Fig. 33, red lines with asterisk), which inherently bind HSC70. The rest were currently incomplete but could serve as functional targeting sites after post-translational modification (Table 11, unhighlighted and Fig. 33, green lines with asterisk).

Table 11: ***KFERQ motifs on Gemin3 (Uniprot ID: Q9UHI6)***

Motif	Motif starts at aa	Motif type
KTRLK	175	acetyl and phos. act.
TRLKK	176	acetyl and phos. act.
YKVVN	280	asparagine and phos. act.
QKIER	477	canonical
VKSKN	506	asparagine and phos. act.
NKVIE	633	asparagine
KVIEQ	634	canonical
KTRLK	720	acetyl and phos. act.
SFRLQ	743	phos. act.

Notes: The data were retrieved from KFERQ finder V0.8 (https://rshine.einsteinmed.edu/)[106]. acetyl.act. and phos. act. indicate that the sequences could be activated as KFERQ motifs after acetylation and phosphorylation. Asparagine motif type allows the sequences to be flanked by asparagine (N) rather than glutamine (Q).

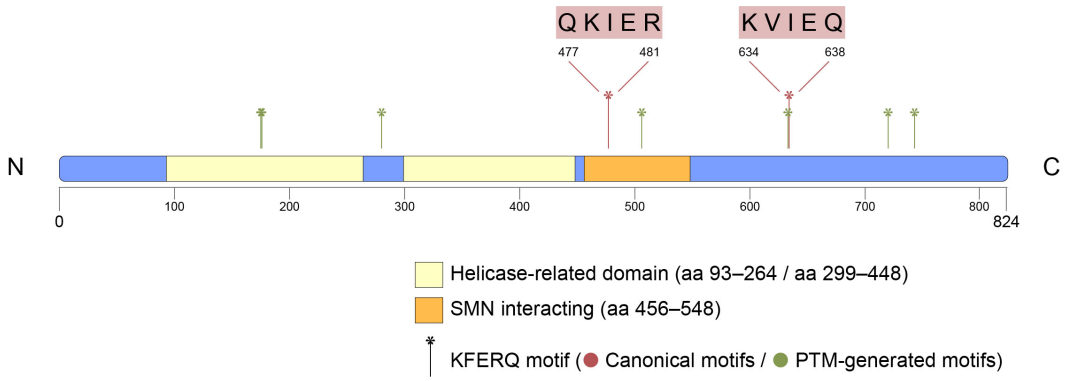


Fig. 33: ***KFERQ motifs on Gemin3.*** *The Gemin3 polypeptide chain is illustrated as Fig. 27 A. The position of canonical or PTM-generated KFERQ motifs on Gemin3 is indicated by lines with asterisks in red and green, respectively. The sequences of two canonical motifs are shown at the top. aa, amino acids; N, N-terminus; C, C-terminus.*

To experimentally validate the *in silico* predictions, GFP-IP was performed using GFP-SMN and GFP-Gemin3 overexpression cells. Cytoplasmic fractions of these two cell lines were incubated with GFP-trap beads, and immunoprecipitated proteins were visualised via immunoblotting. HSC70 was successfully captured by GFP-SMN and GFP-Gemin3 (Fig. 34 A). The modest intensity of HSC70 in IP fractions suggested that their bindings may be restricted to a small subpopulation of the protein under basal conditions. Therefore, to investigate which populations of Gemin3 interact with HSC70, the cytoplasmic extract from GFP-Gemin3 WT overexpression cells was subjected to SEC analysis. HSC70 (red) and GFP-Gemin3 (green) in fractions were detected via

immunoblotting. HSC70, as a highly expressed protein, spanned from 1 to 22 fractions, but the majority co-eluted with Gemin3 fragments in fractions 16–18 (Fig. 34 B). To verify whether HSC70 have affinity with these Gemin3 fragments, GFP-IP was performed using SEC fractions 1–3 and 16–18. HSC70 was successfully precipitated by FL or GFP-Gemin3 fragments, but slightly more HSC70 was detected in the IP using 16–18 fractions (Fig. 34 C).

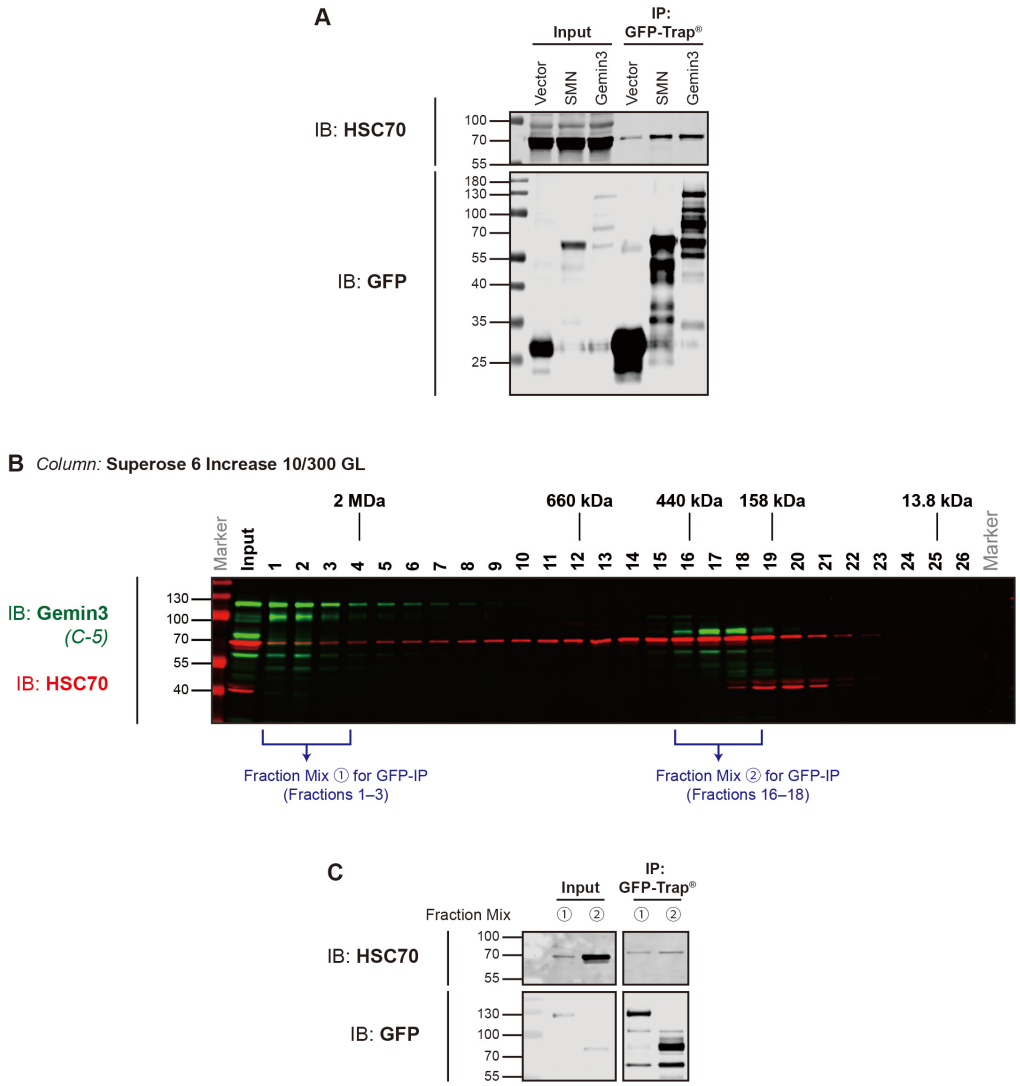


Fig. 34: ***HSC70 is co-immunoprecipitated by Gemin3.*** *(A) Cytoplasmic extracts from cells overexpressing GFP (Vector), GFP-SMN and GFP-Gemin3 WT were subjected to GFP-IP. Pull-down proteins were probed with the indicated antibodies. (B) The cytoplasmic extract from GFP-Gemin3 WT overexpression cells was resolved by the Superose 6 column. Fractions were probed with antibodies against Gemin3 (green) and HSC70 (red). (C) Fractions 1–3 (fraction mix①) and 16–18 (fraction mix ②) from (B) were pooled and incubated with GFP-Trap beads. HSC70 in the IP pools was identified by immunoblotting. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.*

HSC70-substrate complex docks at the cytosolic tail of the CMA receptor LAMP2A on the lysosomal surface, which is a requisite step for degradation [95]. To evaluate whether Gemin3 was escorted to the lysosome, IF was performed on HEK293T cells using the

Gemin3 (C-5) and LAMP2A antibodies. The IF imaging revealed that partial Gemin3 signal (cyan) co-localised with LAMP2A (red) regardless of BafA1 treatment (Fig. 35 A, black arrows), which supported that Gemin3 could be successfully recruited to the lysosome. To further examine whether the degradation of Gemin3 relies on LAMP2A, Gemin3 levels were measured via immunoblotting after LAMP2A was knocked down in HEK293T cells. LAMP2 siRNA successfully reduced LAMP2A expression, but did not cause a significant increase in FL-Gemin3 (Fig. 35 B).

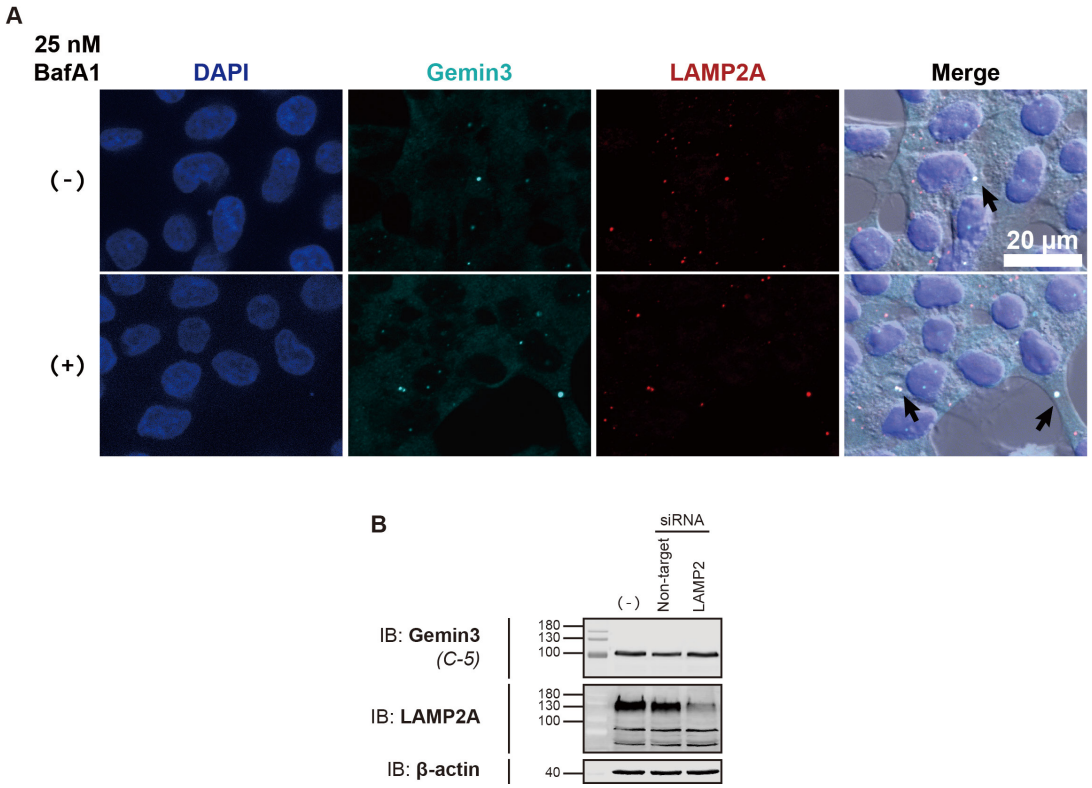


Fig. 35: ***Gemin3 co-locates with LAMP2A.*** *(A) Representative immunofluorescence images of HEK293T cells cultured in the absence (top row) or presence (bottom row) of 25 nM BafA1 for 2 h. Cells were fixed and co-stained with antibodies against Gemin3 (C-5) (cyan) and LAMP2A (red). Nuclei were counterstained with DAPI (blue). Merged images indicate the relative spatial distribution and colocalization of both proteins (black arrows). (B) Representative immunoblotting analysis of Gemin3 and LAMP2A expression following RNA interference. HEK293T cells were transfected with 50 nM of either a non-targeting control or the specific siRNA against LAMP2. Cells were harvested to prepare whole cell lysates following a 24 h transfection and a 24 h recovery phase. Proteins were resolved by SDS-PAGE and immunoblotted with the Gemin3 and LAMP2A antibodies. β-actin was used as the endogenous loading control. BafA1, bafilomycin A1; DAPI, 4',6-Diamidino-2-phenylindole; IB, immunoblotting.*

3.4.3 Gemin3 without canonical KFERQ-like motifs remains incorporated in the SMN complex and prompts snRNP biogenesis

To further investigate the CMA mechanism for fine-tuning Gemin3 levels in cells and its consequence on snRNP biogenesis, three other cell lines overexpressing GFP-Gemin3 without KFERQ motifs were established. Specifically, two canonical KFERQ motifs on

GFP-Gemin3 were abolished individually (MUT-1, MUT-2) or simultaneously (2MUTs) by substituting motif residues with alanine (A) (Fig. 36).

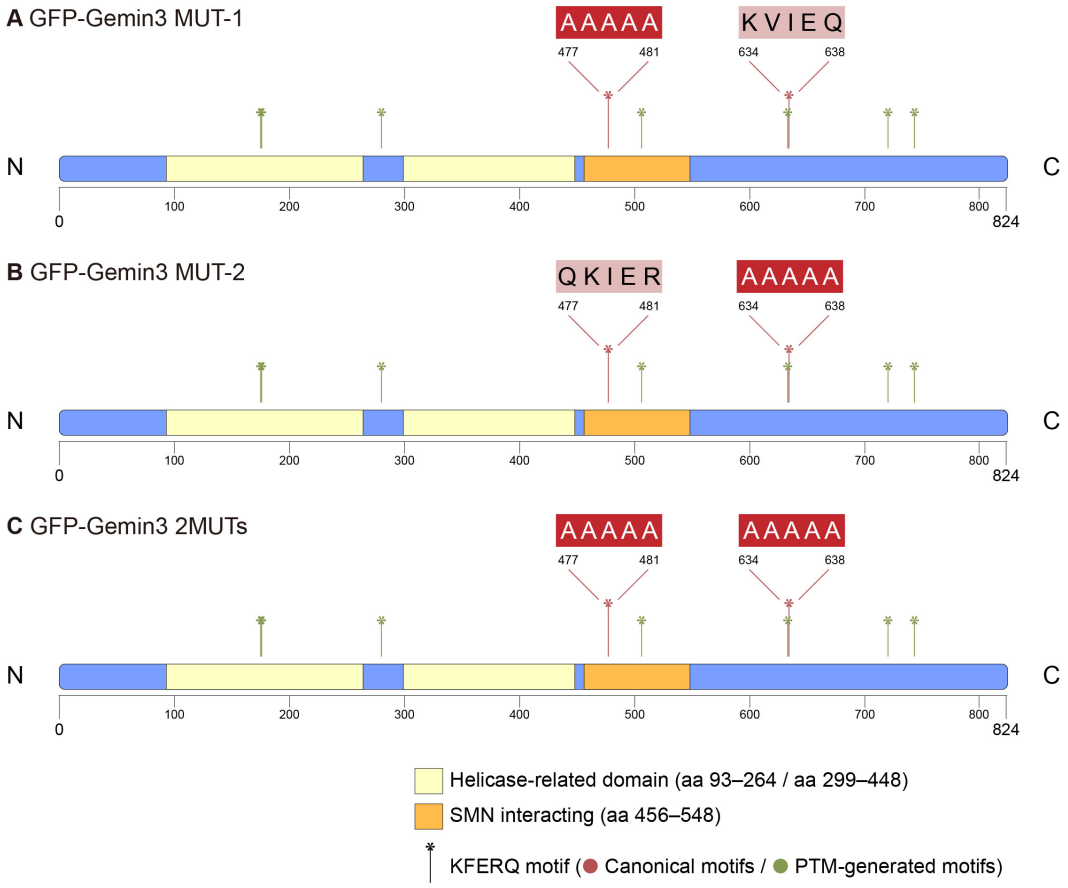


Fig. 36: ***Schematic representation of KFERQ motif–abolished GFP-Gemin3 mutants overexpressed in the Flp-In™ T-REx™ 293 cell lines.*** *The Gemin3 WT backbone is depicted as Fig. 27 A and Fig. 33. Two canonical KFERQ motifs were substituted with five alanine residues in GFP-Gemin3 mutation cell lines. Specifically, (A) MUT-1 cell overexpresses a Gemin3 with a substitution at residues 477–481 and (B) MUT-2 at residues 634–638. (C) GFP-Gemin3 2MUTs contains both substitutions. N, N-terminus; C, C-terminus.*

GFP-IP was performed using these cell lines to examine whether these GFP-Gemin3 mutants without KFERQ motifs lost the capability to interact with HSC70. Immunoblotting analysis revealed that HSC70 was persistently immunoprecipitated by MUT-1 and 2MUTs GFP-Gemin3 variants (Fig. 37 A). This interaction might be derived from the remaining KFERQ motifs or hydrophobic patches on these GFP-Gemin3 mutations (Table 11, Fig. 33 and Fig. 36). To test whether abolishing KFERQ motifs alleviates the CMA degradation, the abundance of each GFP-Gemin3 mutant in the cytoplasm was measured via immunoblotting. Compared to the GFP-Gemin3 WT overexpression cells, a moderate accumulation of FL-GFP-Gemin3 was observed in cells overexpressing GFP-Gemin3 MUT-1 and 2MUTs. Interestingly, Gemin3 fragments decreased slightly if both canonical motifs were abolished (Fig. 37 B). These findings indicated that the clearance of Gemin3 is associated with CMA.

To investigate the impact on the SMN complex when Gemin3 degradation was attenuated, SEC and GFP-IP were performed using cytoplasmic extracts from motif-abolished Gemin3 overexpression cells. Gemin3 without KFERQ motifs shared the same SEC profile with WT. SMN from all cell lines still co-eluted with FL-GFP-Gemin3 variants in high-molecular-weight fractions (Fig. 37 C). GFP-IP analysis revealed that all the Gemin3 mutants without KFERQ motifs robustly co-precipitated with SMN and other Gemin proteins (Fig. 37 D).

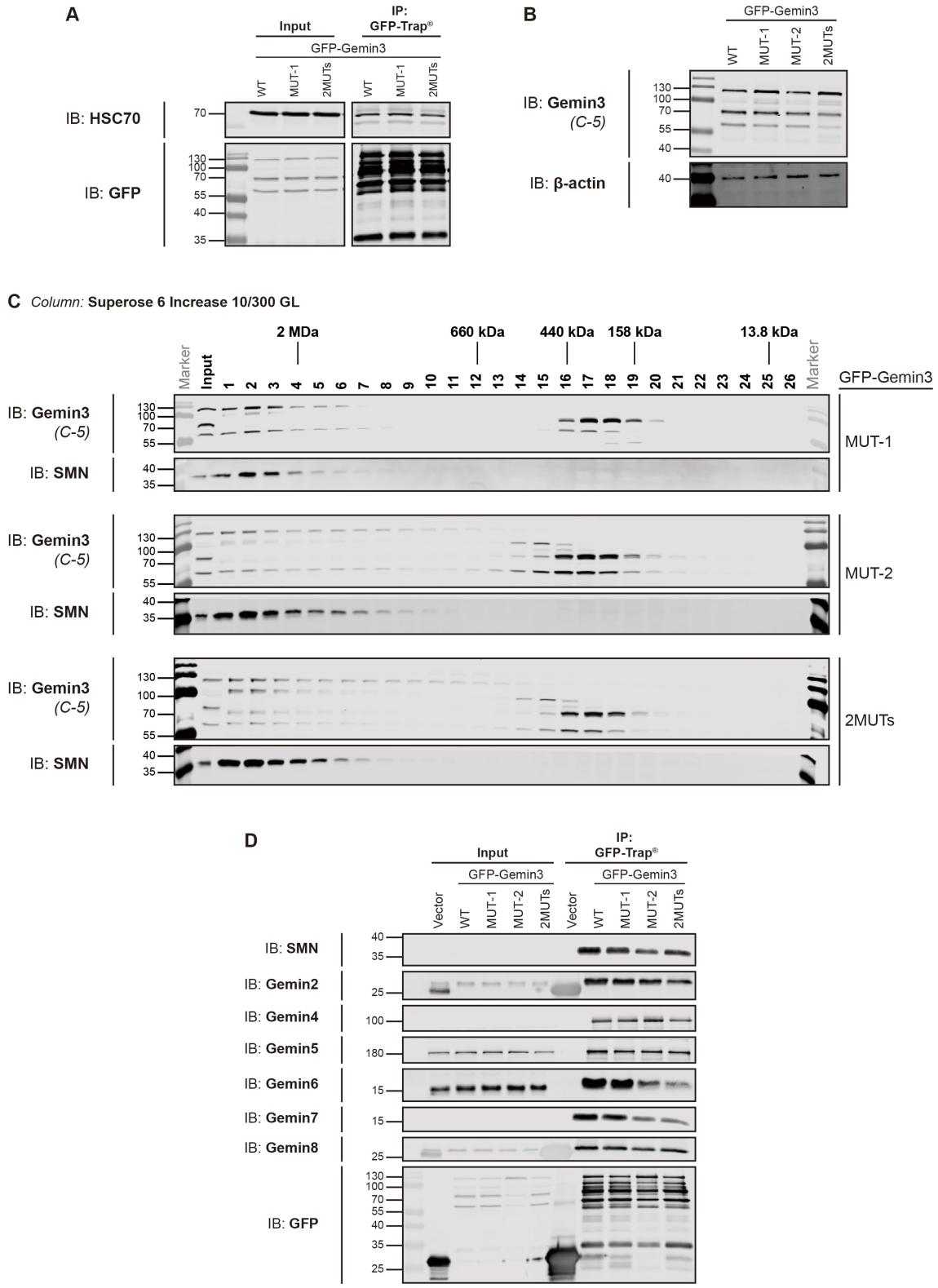


Fig. 37: ***Gemin3 without KFERQ motifs does not affect its interaction with the SMN complex.*** *(A) HSC70 was immunoprecipitated by GFP-Trap beads in cytoplasmic extracts from GFP-Gemin3 variant overexpression cells (WT, MUT-1 and 2MUTs) and detected via immunoblotting. (B) The basal level of Gemin3 in cytoplasmic extracts from cells overexpressing GFP-Gemin3 without the KFERQ motif (WT, MUT-1, MUT-2 and 2MUTs) was analysed by immunoblotting. β-actin served as the endogenous loading control. (C) Cytoplasmic extracts from KFERQ motif–abolished GFP-Gemin3 were resolved on the Superose 6 column. Eluted fractions were collected, and equivalent volumes were probed for Gemin3 and SMN by immunoblotting. (D) Cytoplasmic extracts from GFP and GFP-Gemin3 mutation overexpression cells were subjected to GFP-IP.*

Co-precipitated SMN complex proteins were probed via immunoblotting. WT, wild type; IB, immunoblotting; IP, immunoprecipitation.

Although the SMN complex largely remained intact, its activity might be altered. To compare the capability to assemble snRNP in the presence of Gemin3 without KFERQ motifs, the amount of Cajal bodies *in vivo* was measured via IF as described previously in Sections 3.2.3 and 3.3.4. IF analysis revealed that more Cajal bodies were produced when cells overexpressing GFP-Gemin3 MUT-2 or 2MUTs (Fig. 38 A). These two cell lines contained around one Cajal body per cell on average, which is over 1.5 times greater than the remaining three cell lines (Table 12 and Fig. 38 B). The histogram showed that half of the cells overexpressing the GFP vector, GFP-Gemin3 WT and MUT-1 failed to generate Cajal bodies; however, when introducing MUT-2 and 2MUTs mutations into Gemin3, the proportion of cells without Cajal bodies decreased markedly to approximately 30% (Table 13 and Fig. 38 C). These findings suggested that snRNP biogenesis might be facilitated when the CMA-mediated degradation of Gemin3 was hindered.

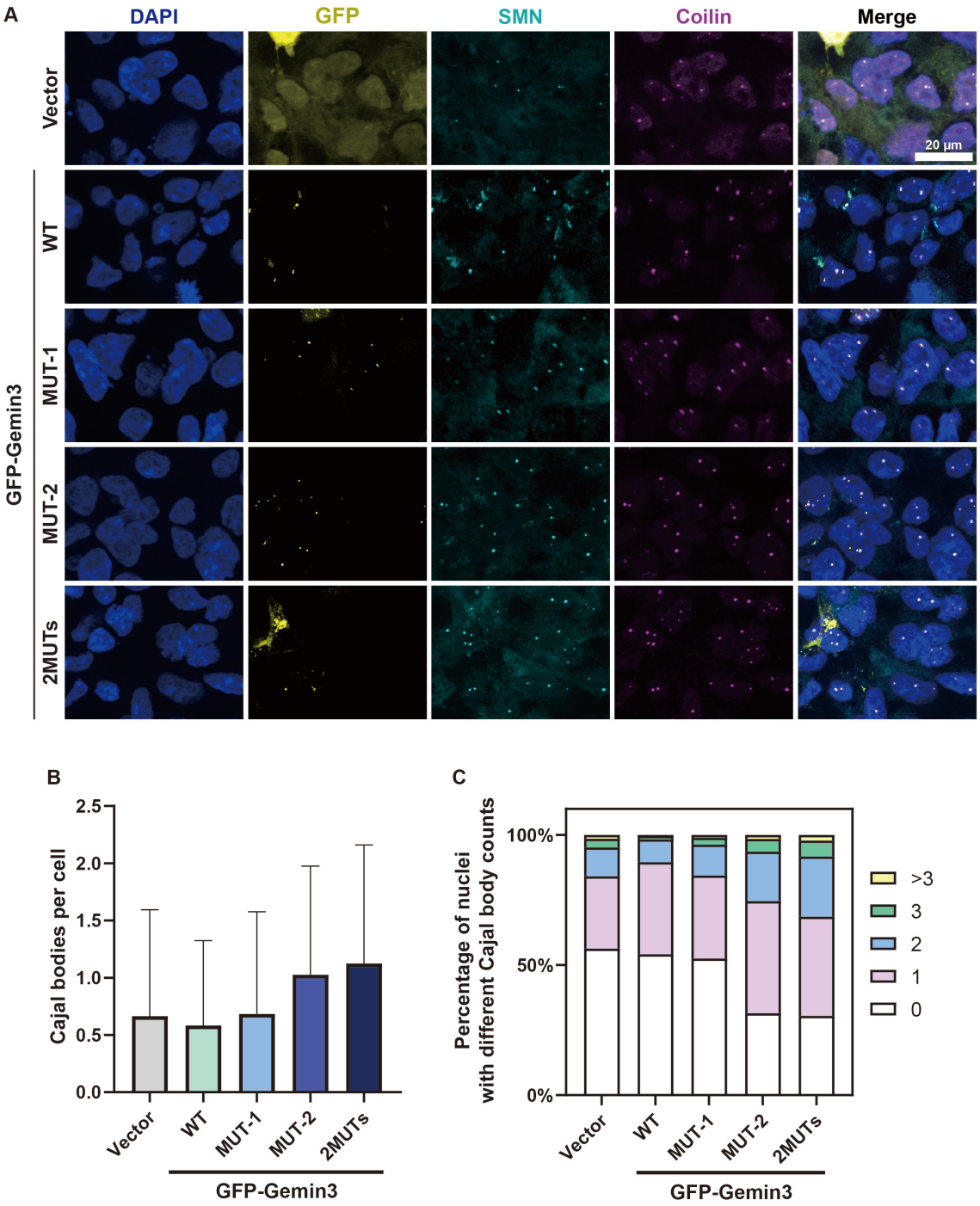


Fig. 38: ***Gemin3 without KFERQ motifs benefits Cajal body formation.*** *GFP (Vector) and GFP-Gemin3 with or without KFERQ motifs overexpression cells were subjected to immunofluorescence analysis following the procedures used in Fig. 16 and Fig. 30. (A) Representative immunofluorescence images of cells co-stained with the SMN (cyan) and coilin (magenta) antibodies to locate the Cajal bodies. Nuclei were visualised with DAPI (blue). (B) Average number of Cajal bodies per cell in cells overexpressing GFP-Gemin3 WT or various mutations. Data are presented as mean ± standard deviation (SD). (C) Frequency distribution showing the percentage of cells containing different numbers of Cajal bodies. The corresponding values are detailed in Table 12 and Table 13. WT, wild type; DAPI, 4',6-Diamidino-2-phenylindole.*

Table 12: ***Average number of Cajal bodies per cell***

	Vector	GFP-Gemin3			
		WT	MUT-1	MUT-2	2MUTs
Mean ± SD	0.66 ± 0.93	0.58 ± 0.74	0.68 ± 0.89	1.03 ± 0.95	1.13 ± 1.04

Notes: All values are rounded to two decimal places. Data are presented as a bar chart in Fig. 38 B. SD, standard deviation.

Table 13: ***Percentages of cells containing different numbers of Cajal bodies (%)***

Number of Cajal bodies	Vector	GFP-Gemin3			
		WT	MUT-1	MUT-2	2MUTs
0	56.32	54.18	52.55	31.51	30.48
1	27.75	35.37	31.84	43.02	38.14
2	11.17	8.80	11.89	19.05	23.14
3	3.30	1.30	2.63	4.86	6.11
>3	1.45	0.35	1.09	1.56	2.13

Notes: All percentage values are displayed to two decimal places. The corresponding histogram are shown in Fig. 38 C.

4 Discussion

4.1 Energy-related kinases ULK1 and mTOR are independent of the SMN complex

Early studies have indicated that spliceosomal small nuclear ribonucleoprotein (snRNP) biogenesis in cell extracts is dependent on survival motor neuron (SMN) protein and ATP [18, 50-53]. However, the cellular mechanism coupling energy levels and snRNP biogenesis remains an open question. Recent research has revealed that unc-51-like autophagy activating kinase 1 (ULK1) and mammalian target of rapamycin (mTOR) pathways are involved in the snRNP biogenesis [23-25, 55]. The core of these two pathways, ULK1 and mTOR, are two energy-sensing kinases [140]. Therefore, the interaction of ULK1 and mTOR with the SMN complex likely functions as a bridge between upstream energy levels and downstream snRNP biogenesis. A co-localisation of SMN and phosphorylated mTOR in the nucleus was observed previously by immunofluorescence (IF) when ribosomal protein S6 kinase beta-1 (RPS6KB1; p70S6 kinase) was depleted [55]. However, co-immunoprecipitation (co-IP) assays in the present study indicated that the SMN complex, ULK1 and mTOR were independent in the cytoplasm regardless of nutrient conditions (Fig. 5). Similarly, proximity labelling mass spectrometry (PL-MS) analysis using whole cell lysates has shown that ULK1 and mTOR were absent in the immediate vicinity of SMN [141], which perfectly complements the present findings by further demonstrating that ULK1 and mTOR neither transiently interact with SMN, nor do they associate in the nucleus. These findings collectively suggest that while the interaction might be triggered by specific stimuli, the SMN complex is separated from ULK1 and mTOR in the steady state.

4.2 Energy-mediated phosphorylation modulates the interaction network of the SMN complex

Beyond affecting the SMN-kinase interaction, cellular energy levels might alter SMN phosphorylation. Threonine (T) 37 phosphorylation of SMN was observed after 2-hour starvation (Fig. 39), suggesting this phosphorylation might depend on cellular energy levels. Serine (S) 187 was postulated to be another energy-associated phosphorylation site based on previous studies [55, 58]. S187 is phosphorylated by protein kinase A (PKA) [58], and its phosphorylation level decreases following the knockout of RPS6KB1 [55], a crucial player in the mTOR signalling [113]. The activities of PKA and mTOR pathways are tightly coupled to cellular metabolic states and energy availability [113, 142]. The phosphorylation at T37 and S187 has been identified in several high-throughput mass spectrometry (MS) analyses [55, 58, 64-69]. Based on these observations, the effect of

these two energy-mediated phosphorylations of SMN was investigated in the present study to provide deeper insights into the energy-dependence of snRNP biogenesis.

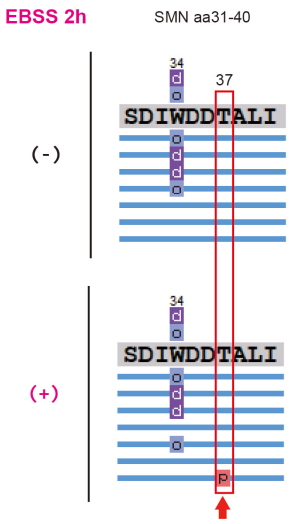


Fig. 39: ***SMN is phosphorylated at T37 after 2-hour EBSS treatment.*** *The amino acid sequence of SMN from 31 to 40 is displayed. Blue horizontal lines beneath the sequence represent individual peptide fragments identified in IP pools of GFP-SMN by MS performed in Section 3.1.2. Phosphorylation is indicated by pink-coloured square markers containing the P letter. A red box indicates T37 in the identified fragments under both experimental conditions. The phosphorylation in the sample with EBSS treatment is pointed out by a red arrow.*

Previous work has implied that T37 and S187 phosphorylation might affect the interaction with Gemin2 [35, 36, 58, 115, 118]. T37 is located in the Gemin2-binding domain of SMN, which spans from amino acids (aa) 26 to 51 [35, 36, 115, 118]. S187 is located in an intrinsically disordered region, but its phosphorylation status is one of the factors in the interaction between SMN and Gemin2.[58]. However, this study showed that SMN preserved the full binding capacity with Gemin2 regardless of the individual or dual phosphorylation status at T37 and S187 (Fig. 8, Fig. 9 and Fig. 14 C), indicating that phosphorylation at these two sites is dispensable for Gemin2 interaction.

Beyond Gemin proteins, the direct interaction of SMN with SmD1, B and D3 has been demonstrated in several studies [35, 120, 143-148], which is fundamentally mediated by the SMN's central Tudor domain around aa 90–160 [143, 147, 149, 150]. Consistently, SmB, SmD1 and SmD3 directly bound with all GST-SMN constructs containing the intact Tudor domain in this study (Fig. 12 and Fig. 13). Interestingly, although S187 is located outside the Tudor domain, the phosphomimic mutation at this site impaired the binding affinity with SmB, thereby diminishing the interaction with SmB/D3 subcomplex (Fig. 12 and Fig. 13). An early study has shown that a SMN C-terminal fragment (aa 240–267) competitively displaced full length (FL) SMN from SmB [35]. These findings suggest that the interaction between SmB and SMN might not rely solely on the Tudor domain.

Unlike Gemin and Sm proteins, observations regarding the interaction between the SMN and protein arginine N-methyltransferase 5 (PRMT5) complexes are inconsistent within the current literature [18, 141, 151-153]. A reciprocal co-IP assay of HeLa cytoplasmic fractions has revealed that neither the SMN complex nor PRMT5/chloride conductance regulatory protein (pICln) were detected in corresponding IP pools by immunoblotting [151]. However, in another study using the same extracts, both SMN complex and PRMT5 complex were identified by MS in the eluted fraction from an anti-pICln affinity column [18]. Their association could be observed in other cell lines through MS or co-IP as well [141, 152, 153]. The latter opinion is also supported by the present study in which the PRMT5 complex proteins were co-immunoprecipitated by GFP-SMN in the cytoplasmic extracts (Table 4 and Fig. 14 B) or SEC fraction pool (Fig. 15 B). This conflict might be attributed to the sensitivity of the analytical techniques employed. As shown in the present study (Fig. 14 B) and previous research [18, 153], the co-immunoprecipitated PRMT5 complex proteins were less abundant than the bait protein SMN, indicating their interaction *in vivo* is modest. Because of this, the co-precipitated SMN or PRMT5 complex proteins might fall below the limit of detection of conventional immunoblotting [151], but could be readily detected by highly sensitive MS [18, 141, 152].

In addition to corroborating the interaction of the SMN-PRMT5 complex, the present study further reveals that this interaction is regulated by dual phosphorylation at T37 and S187 of SMN (Table 5, Table 6, Fig. 14 B, Fig. 15 B and Fig. 16). The interaction with PRMT5 complex proteins was strikingly reduced in the phosphomimic GFP-SMN DD cell line (Fig. 14 B and Fig. 15 B). An early study has shown that the SMN-PRMT5 complex assembled snRNP more efficiently compared with the SMN complex alone [18], suggesting the interaction of the SMN-PRMT5 complex facilitates snRNP biogenesis. In agreement with this, cells overexpressing GFP-SMN DD exhibited a lower average number of Cajal bodies (Table 5, Table 6 and Fig. 16), indicating that concurrent phosphorylation at T37 and S187 of SMN, which results in the dissociation with the PRMT5 complex, impairs its capacity for snRNP assembly.

Collectively, these data suggest that energy-associated phosphorylation at T37 and S187 governs the interaction of the SMN complex with the PRMT5 complex, by which the impact of cellular energy levels might be transmitted to snRNP biogenesis.

4.3 Cleavage of Gemin3

4.3.1 Gemin3 is cleaved under basal conditions

Beyond the interaction, the expression of each SMN complex protein was evaluated under different nutrient conditions. Interestingly, among all complex components, only

Gemin3 was sensitive to the extra- or intracellular energy reduction caused by Earle's balanced salt solution (EBSS) and bafilomycin A1 (BafA1), most notably by BafA1 (Fig. 17–Fig. 19). Upon BafA1 treatment, the levels of additional <100 kDa Gemin3 decreased dramatically, whereas the FL form (~100 kDa) was notably elevated in different cell lines (Fig. 18). The response to BafA1 suggests that these extra Gemin3 fragments are more likely to be products from proteolytic cleavage rather than isoforms produced by alternative splicing. Gemin3 fragments in unstimulated cells have been reported in other studies [38, 154]. An early study has shown that the Gemin3 (12H12) antibody weakly detected a protein of approximately 55 kDa in an immunoblot of total HeLa cell extracts [38]. In another study, the FL and a <70 kDa signal probed by the Gemin3 antibody were abolished following transient lentiviral shRNA knockdown of Gemin3, confirming that the lower signal also represents Gemin3 rather than a non-specific artifact [154]. These findings indicate that Gemin3 undergoes extensive proteolytic cleavage under basal conditions.

4.3.2 Gemin3 is the substrate of calpain-1

Calpain has been implied as a protease for Gemin3 cleavage by prior research [90]. However, the cleavage has been observed only in the purified native SMN complex, but not in total HeLa cell lysate [90]. In contrast, the present study demonstrated that Gemin3 in HEK293T cytoplasmic extracts was successfully cleaved by calpain-1 (Fig. 20, Fig. 21 and Fig. 23 A), even under milder conditions comparable to those in the former research [90] (Fig. 23 A). Calpain-1 produced two N-terminal fragments (~50 kDa and ~55 kDa) detected by the Gemin3 (C-5) antibody *in vitro* (Fig. 21 and Fig. 23 A). The ~55 kDa N-terminal fragment was also detected by the Gemin3 (12H12) antibody, suggesting this fragment might include the internal region from aa 368–548 (Fig. 21). This N-terminal fragment and another C-terminal fragment ~40 kDa, probed by the Gemin3 (2), might be the products from the same cleavage event since the sum of their molecular weight is roughly equivalent to that of FL-Gemin3 (Fig. 21). These observations substantiate that Gemin3 is the target of calpain-1.

There are two possible explanations for the discrepancy between the present work and the previous report [90]. First, the previous study drew its conclusion primarily based on the alteration in FL-Gemin3 levels in the assay [90]. However, as shown in Fig. 23 A, the level of FL-Gemin3 does not reduce significantly under mild conditions. Second, BafA1-treated HEK293T extract was used in the present study. This extract contains a high concentration of FL-Gemin3, making it easier to detect the cleavage products. The present findings do not necessarily contradict the prior study [90], but rather extend the

current understanding about how calpain-1 cleaves Gemin3 by providing the profile of cleavage fragments.

Although Gemin3 is the substrate of calpain-1, calpeptin (a calpain-1 inhibitor [133]) failed to abrogate the cleavage in untreated cells (Fig. 24). This finding rules calpain-1 out as the protease responsible for the cleavage observed under basal conditions. However, under pathological conditions, Gemin3 could be cleaved by calpain-1. Spinal muscular atrophy (SMA) and amyotrophic lateral sclerosis (ALS) are characterised by energy deficiency in motor neurons [9-12]. A computational model has shown that the reduction of ATP in motor neurons might upregulate the cellular concentration of Ca^{2+}, which could elevate the activity of calpain [8]. Indeed, the calpain system has been demonstrated to be overactivated during SMA and ALS [89, 155]. The phenotypes of both diseases are ameliorated by inhibiting the calpain pathway [88, 89, 156]. As the substrate of calpain, Gemin3 is highly vulnerable to be cleaved by calpain under SMA and ALS conditions, which could be one of the contributors to the disease development.

4.3.3 The function of Gemin3 might be dictated by cleavage

The RNA helicase–related domains are localised to the N-terminus of Gemin3, whereas the binding domains of multiple proteins are clustered within its C-terminus [47, 48]. In agreement with these structural features, the endogenous/GFP-tagged Gemin3 fragments and GFP-Gemin3 N-terminus (aa 1-478), which are devoid of the C-terminus, were expelled from the SMN complex (Fig. 25, Fig. 26 and Fig. 28). Consistently, Gene Ontology (GO) enrichment analysis revealed that the interactome of N-terminal GFP-Gemin3 shifted away from the canonical processes driven by Gemin3 wild type (WT) to the binding with cyclin and protein kinase C (PKC) (Fig. 31). This finding implies that the N-terminus of Gemin3 is a potential target for phosphorylation by cyclin-dependent kinases (CDKs) and PKC.

Gemin5, the only component binding to snRNA in the SMN complex [157], is integrated into the SMN complex through its interaction with Gemin3 [29, 38]. Due to the departure of the N-terminal Gemin3 from the SMN complex, the snRNA failed to be loaded on the complex through Gemin3-Gemin5 interaction (Fig. 29). Furthermore, unlike the C-terminus and WT, which condensed into punctate signals, the GFP-Gemin3 N-terminus diffused in the whole cells (Fig. 30 A). Similar IF patterns have been observed in other truncated Gemin3 which is unable to interact with SMN [38, 81]. These results indicate that the spatial distribution of Gemin3 is dependent on its Interaction with the SMN complex.

In contrast to the N-terminus, the C-terminus of Gemin3 largely preserves the interaction of FL (Fig. 28–Fig. 31). Although C-terminal GFP-Gemin3 (aa 479–824) contains a partial SMN binding domain (aa 456–547) [38], it was still incorporated into the SMN complex (Fig. 28). Besides, the C-terminal fragment retained certain interactors of FL that are related to RNA processing (Fig. 31, WT ∩ C). Similar to the complex with FL-Gemin3, this incomplete SMN complex recruited snRNA (Fig. 29) and was situated in Cajal bodies (Fig. 30 A).

Accumulating evidence has validated that Gemin3 fragments are insufficient for the proper functioning of the SMN complex [80-82, 158]. The cleavage of Gemin3 induced by poliovirus infection disrupts snRNP assembly [80]. The yeast two-hybrid assay has demonstrated that the Gemin3 C-terminus preferentially binds with its FL form [83]; therefore, overexpressing this fragment in the *Drosophila* model might impede endogenous FL-Gemin3 from functional interactions, leading to SMA-like disability [81, 82, 158]. In contrast to these previous results, overexpression of the Gemin3 fragments in the HEK293T cells did not adversely affect the formation of Cajal bodies (Table 9, Table 10, Fig. 30 B and C). This tolerance to Gemin3 fragments might be the result of the intrinsic proteolytic mechanism in HEK293T cells, which could fine-tune the amount of Gemin3.

Even though snRNP biogenesis requires FL-Gemin3, certain activities are independent of its N-terminal RNA helicase domain [75, 86]. The abundance of Gemin3 fragments under basal conditions suggests that they play specific roles in cellular processes (Fig. 18). In the present study, only the interaction with Gemin4 was unaffected by the level of FL-Gemin3 (Fig. 26 C), and the Gemin3 C-terminus co-immunoprecipitated a comparable amount of Gemin4 with WT (Fig. 28 B). These findings imply that the C-terminal fragment is sufficient for Gemin4 interaction. However, the biological function of this subcomplex needs further investigation.

4.4 Chaperone-mediated autophagy degradation of Gemin3

Several studies have reported that the level of Gemin3 was insensitive to the inhibitors that interfere with the proteasome (MG132, bortezomib) or autophagosome (3-Methyladenine and chloroquine) [159-161]. Consistent with these observations, no significant alteration in Gemin3 levels was caused by MG132 treatment and ATG5 knockout in the present research (Fig. 32 A and B). These results exclude the proteasome and autophagosome as the machinery for Gemin3 degradation. Conversely, during prolonged BafA1 and EBSS treatment, the abundance of Gemin3 exhibited a strong inverse correlation with that of lysosome-associated membrane protein type 2a

(LAMP2A), a marker of the lysosome, suggesting that the degradation of Gemin3 highly depends on the lysosome (Fig. 32 C and D). Based on these findings, it was reasonable to hypothesise that Gemin3 might be directly sent to the lysosome for degradation.

Chaperone-mediated autophagy (CMA) is one of the autophagy pathways which satisfy the criteria of this hypothesis [95]. The present findings validated the interaction of Gemin3 with heat shock cognate 71 kDa protein (HSC70; HSPA8) and LAMP2A (Fig. 34 and Fig. 35), which is the hallmark of CMA-degraded substrates [95]. Gemin3 possesses nine KFERQ-like motifs (hereafter referred to as KFERQ motifs), two of which are canonical motifs (Table 11, Fig. 33 and Fig. 40). The interaction of HSC70 with Gemin3 has been reported recently by PL-MS [98]. In line with this, GFP-Gemin3 co-immunoprecipitated a modest amount of HSC70 in cytoplasmic extracts and SEC fractions (Fig. 34 A and C), suggesting that a pool of Gemin3 could be the substrate of CMA. Recently, the defective HSC70 chaperone has been proposed as the common aetiology between SMA and ALS [162]. As the client protein for HSC70 in CMA, Gemin3 might be implicated in this chaperone dysregulation, driving the progression of SMA and ALS.

Further IF analysis in HEK293T cells consistently revealed a partial co-localisation between endogenous Gemin3 and LAMP2A within the cytoplasm (Fig. 35 A), further verifying that a fraction of Gemin3 is delivered to the lysosome for degradation. Cells were immuno-stained by the Gemin3 (C-5) antibody which recognises the N-terminus of Gemin3. As shown in Fig. 30 A, the C-terminus is necessary for the condensation of Gemin3. Therefore, the puncta signal of Gemin3 in the present IF suggests that it is FL-Gemin3 that is transported to the lysosome.

Since CMA only processes partial FL-Gemin3, LAMP2A knockdown did not lead to a marked increase in FL-Gemin3 (Fig. 35 B). However, the level of FL-Gemin3 (Fig. 37 B) and the average of Cajal bodies (Table 12, Table 13, Fig. 38 B and C) were upregulated after the suppression of CMA by abolishing canonical KFERQ motifs.

Collectively, these lines of evidence support that the turnover of FL-Gemin3 is mediated via CMA.

4.5 Cleavage and degradation of the SMN complex

Beyond the Gemin3 protein, Gemin5 and SMN have been identified as the target of calpain-1 as well [88-91]. This conclusion is further validated by the evidence from the present cleavage assays (Fig. 21–Fig. 23). Gemin5 was cleaved into the fragments around 70 kDa (Fig. 21 and Fig. 23 A). However, unlike previous reports [90, 91], no

SMN cleavage fragment was observed in the present study (Fig. 22 and Fig. 23 B). This might be attributed to the differences in the antibodies and cell lines employed in the assays. The fragments were probed by antibodies targeting the N-terminus (Clone 8) and C-terminus (9F2) of SMN in previous studies [90, 91], whereas the SMN monoclonal antibody (2B1) used in the present study was raised against purified recombinant His-tagged SMN protein [126]. Cleavage might destroy the target epitope, rendering the SMN (2B1) antibody unable to detect the fragments. Besides, the C-terminal SMN fragment could be detected in U2-OS cells [91], but not in HeLa cells [90], which suggests that the stability of fragments might be different among different cell lines.

Substantial studies have demonstrated that the level of SMN is mediated by the ubiquitin-proteasome system (UPS) pathway [60, 163-166]. Consistently, MS analysis of GFP-IP fractions revealed that more proteasomal proteins interacted with SMN upon EBSS treatment (Table 3 and Fig. 7 B), suggesting that starvation might induce UPS degradation of SMN. However, no significant change in SMN levels was observed after EBSS treatment (Fig. 19) presumably because most of the SMN were shielded in the complex, which is less vulnerable to UPS degradation than the free monomer SMN [60, 163, 167].

In addition to UPS, the SMN complex is subjected to CMA degradation, as evidenced by several co-IP assays demonstrating the presence of HSC70 in the SMN interaction network [18, 51, 97-99]. Consistently, HSC70 was successfully co-purified by GFP-SMN in the present study (Fig. 34 A). The KFERQ motif is required for HSC70 binding during CMA [95, 139]. Given the distribution of this motif on each complex protein (Fig. 40), the regulation of CMA on the SMN complex under basal conditions might stem from Gemin2–6 and Unrip rather than SMN.

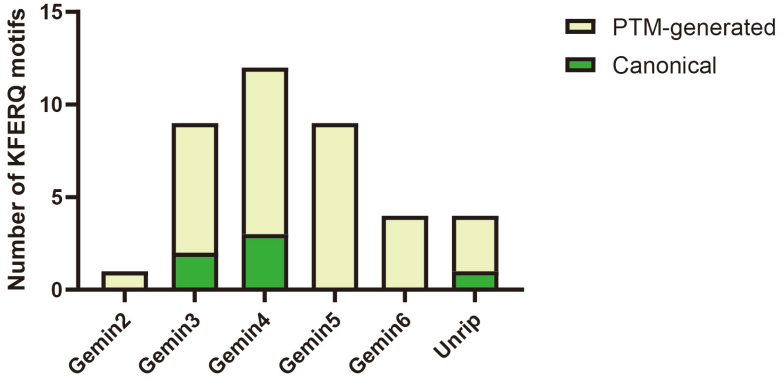


Fig. 40: ***Gemin2–Gemin6 and Unrip are the components containing KFERQ motifs in the SMN complex.*** *The Uniprot entries of SMN, Gemin2–Gemin8 and Unrip (Q16637, Q9UHI6, O14893, P57678, Q8WXD5, Q8TEQ6, Q9NWZ8, Q9H840, Q9Y3F4) were submitted to KFERQ finder V0.8 (https://rshine.einsteinmed.edu/) for motif screening [106]. Green represents the number of canonical motifs, and cream represents motifs requiring activation by post-translational*

modification. SMN, Gemin7 and Gemin8 without the KFERQ motif were excluded from this visualisation. PTM, post-translational modification.

4.6 Conclusions and perspectives

In summary, the present study elucidates the impact of cellular energy fluctuation on snRNP biogenesis through SMN phosphorylation and Gemin3 proteolysis.

Although cellular energy levels do not affect the interaction of SMN with ULK1 and mTOR (Chapter 3.1), it might alter the phosphorylation status at T37 and S187 which modulates the affinity of SMN with SmB and the PRMT5 complex (Chapter 3.2). Notably, the dissociation with the PRMT5 complex induced by dual phosphorylation comprises the snRNP biogenesis *in vivo* (Chapter 3.2). In future research, *in vitro* binding and snRNP assembly assays are required to further illustrate the impact of SMN phosphorylation at T37 and S187 on its direct interaction with the PRMT5 complex proteins and activity in snRNP formation.

Extensive Gemin3 proteolysis occurs under nutrient-rich conditions, which could be blocked by BafA1 treatment (Chapter 3.3). The basal degradation is partially carried out via CMA (Chapter 3.4). Although Gemin3 is the target of calpain-1 (Chapter 3.3), the specific cytoplasmic proteases responsible for the basal cleavage remain to be identified.

The study about the truncated Gemin3 further substantiates the necessity of its C-terminus in the interaction with the SMN complex and suggests the N-terminus as the target region for CDKs and PKC phosphorylation (Chapter 3.3). As discussed above (Section 4.3.3), both termini of Gemin3 are required in snRNP biogenesis [80-82, 158]; however, certain activities are not dependent on the N-terminus [75, 86]. Given their stability and abundance, the Gemin3 fragments might perform specific tasks, which warrants further investigation.

Motor neurons and the muscular tissue experience severe metabolic disorders in several motor neuron diseases (MNDs) [6], especially SMA and ALS [9-12]. However, little is known about the implications of energy reduction for these diseases. The present research provides a molecular mechanism about this question from the aspect of snRNP biogenesis. The current therapies remain insufficient to ALS and the later stage of SMA [168-170]. The restoration of SMN phosphorylation or Gemin3 proteolysis could be alternative strategies to enhance treatment efficacy.

5 Reference list

1 Mercuri E, Sumner CJ, Muntoni F, et al. Spinal muscular atrophy. *Nat Rev Dis Primers*. 2022;8(1):52.

2 Lefebvre S, Burglen L, Reboullet S, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. 1995;80(1):155-65.

3 Melki J, Lefebvre S, Burglen L, et al. De novo and inherited deletions of the 5q13 region in spinal muscular atrophies. *Science*. 1994;264(5164):1474-7.

4 Lefebvre S, Burlet P, Liu Q, et al. Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet*. 1997;16(3):265-9.

5 Mathis S, Beauvais D, Duval F, et al. The various forms of hereditary motor neuron disorders and their historical descriptions. *J Neurol*. 2024;271(7):3978-90.

6 Barone C, Qi X. Altered Metabolism in Motor Neuron Diseases: Mechanism and Potential Therapeutic Target. *Cells*. 2023;12(11).

7 Harris JJ, Jolivet R, Attwell D. Synaptic energy use and supply. *Neuron*. 2012;75(5):762-77.

8 Le Masson G, Przedborski S, Abbott LF. A computational model of motor neuron degeneration. *Neuron*. 2014;83(4):975-88.

9 Rashid S, Dimitriadi M. Autophagy in spinal muscular atrophy: from pathogenic mechanisms to therapeutic approaches. *Front Cell Neurosci*. 2023;17:1307636.

10 James R, Chaytow H, Ledahawsky LM, et al. Revisiting the role of mitochondria in spinal muscular atrophy. *Cell Mol Life Sci*. 2021;78(10):4785-804.

11 Vandoorne T, De Bock K, Van Den Bosch L. Energy metabolism in ALS: an underappreciated opportunity? *Acta Neuropathol*. 2018;135(4):489-509.

12 Zhang X, Yang K, Le W. Autophagy and Motor Neuron Diseases. *Adv Exp Med Biol*. 2020;1207:53-74.

13 Nishio H, Niba ETE, Saito T, et al. Spinal Muscular Atrophy: The Past, Present, and Future of Diagnosis and Treatment. *Int J Mol Sci*. 2023;24(15).

14 Patel SB, Bellini M. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res*. 2008;36(20):6482-93.

15 Matera AG, Wang Z. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol*. 2014;15(2):108-21.

16 Gruss OJ, Meduri R, Schilling M, et al. UsnRNP biogenesis: mechanisms and regulation. *Chromosoma*. 2017;126(5):577-93.

17 So BR, Zhang Z, Dreyfuss G. The Function of Survival Motor Neuron Complex and Its Role in Spinal Muscular Atrophy Pathogenesis. Spinal Muscular Atrophy2017. p. 99-111.

18 Meister G, Fischer U. Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *Embo j*. 2002;21(21):5853-63.

19 Raker VA, Hartmuth K, Kastner B, et al. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol Cell Biol*. 1999;19(10):6554-65.

20 Raker VA, Plessel G, Lührmann R. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *The EMBO Journal*. 1996;15(9):2256-69.

21 Sumpter V, Kahrs A, Fischer U, et al. In vitro reconstitution of U1 and U2 snRNPs from isolated proteins and snRNA. *Mol Biol Rep*. 1992;16(4):229-40.

22 Fischer U, Englbrecht C, Chari A. Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdiscip Rev RNA*. 2011;2(5):718-31.

23 Schmitz K, Cox J, Esser LM, et al. An essential role of the autophagy activating kinase ULK1 in snRNP biogenesis. *Nucleic Acids Res*. 2021;49(11):6437-55.

24 Esser LM, Schmitz K, Hillebrand F, et al. Phosphorylation of pICln by the autophagy activating kinase ULK1 regulates snRNP biogenesis and splice activity of the cell. *Comput Struct Biotechnol J*. 2023;21:2100-9.

25 Grimmler M, Reinhart M, Alers S, et al. Spliceosomal Sm core assembly: AlphaFold 3 predicted structure and phosphorylation-dependent regulation of the human 6S complex. *Comput Struct Biotechnol J*. 2026;31:51-60.

26 Pineiro D, Fernandez-Chamorro J, Francisco-Velilla R, et al. Gemin5: A Multitasking RNA-Binding Protein Involved in Translation Control. *Biomolecules*. 2015;5(2):528-44.

27 Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol*. 2007;8(3):209-20.

28 Stark H, Dube P, Luhrmann R, et al. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature*. 2001;409(6819):539-42.

29 Otter S, Grimmler M, Neuenkirchen N, et al. A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *J Biol Chem*. 2007;282(8):5825-33.

30 Cauchi RJ. SMN and Gemins: 'We are family' … or are we? *BioEssays*. 2010;32(12):1077-89.

31 Coady TH, Lorson CL. SMN in spinal muscular atrophy and snRNP biogenesis. *Wiley Interdiscip Rev RNA*. 2011;2(4):546-64.

32 Burghes AH, Beattie CE. Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat Rev Neurosci*. 2009;10(8):597-609.

33 Kolb SJ, Kissel JT. Spinal Muscular Atrophy. *Neurol Clin*. 2015;33(4):831-46.

34 Chaytow H, Huang YT, Gillingwater TH, et al. The role of survival motor neuron protein (SMN) in protein homeostasis. *Cell Mol Life Sci*. 2018;75(21):3877-94.

35 Liu Q, Fischer U, Wang F, et al. The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell*. 1997;90(6):1013-21.

36 Zhang R, So BR, Li P, et al. Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell*. 2011;146(3):384-95.

37 Yi H, Mu L, Shen C, et al. Negative cooperativity between Gemin2 and RNA provides insights into RNA selection and the SMN complex's release in snRNP assembly. *Nucleic Acids Res*. 2020;48(2):895-911.

38 Charroux B, Pellizzoni L, Perkinson RA, et al. Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J Cell Biol*. 1999;147(6):1181-94.

39 Charroux B, Pellizzoni L, Perkinson RA, et al. Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J Cell Biol*. 2000;148(6):1177-86.

40 Gubitz AK, Mourelatos Z, Abel L, et al. Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J Biol Chem*. 2002;277(7):5631-6.

41 Pellizzoni L, Baccon J, Rappsilber J, et al. Purification of native survival of motor neurons complexes and identification of Gemin6 as a novel component. *J Biol Chem*. 2002;277(9):7540-5.

42 Baccon J, Pellizzoni L, Rappsilber J, et al. Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J Biol Chem*. 2002;277(35):31957-62.

43 Carissimi C, Saieva L, Baccon J, et al. Gemin8 is a novel component of the survival motor neuron complex and functions in small nuclear ribonucleoprotein assembly. *J Biol Chem*. 2006;281(12):8126-34.

44 Carissimi C, Baccon J, Straccia M, et al. Unrip is a component of SMN complexes active in snRNP assembly. *FEBS Lett*. 2005;579(11):2348-54.

45 Grimmler M, Otter S, Peter C, et al. Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Hum Mol Genet*. 2005;14(20):3099-111.

46 Panek J, Roithova A, Radivojevic N, et al. The SMN complex drives structural changes in human snRNAs to enable snRNP assembly. *Nat Commun*. 2023;14(1):6580.

47 Curmi F, Cauchi RJ. The multiple lives of DEAD-box RNA helicase DP103/DDX20/Gemin3. *Biochem Soc Trans*. 2018;46(2):329-41.

48 He L, Yang J, Hao Y, et al. DDX20: A Multifunctional Complex Protein. *Molecules*. 2023;28(20).

49 Bizen N, Bepari AK, Zhou L, et al. Ddx20, an Olig2 binding factor, governs the survival of neural and oligodendrocyte progenitor cells via proper Mdm2 splicing and p53 suppression. *Cell Death Differ*. 2022;29(5):1028-41.

50 Pellizzoni L, Yong J, Dreyfuss G. Essential role for the SMN complex in the specificity of snRNP assembly. *Science*. 2002;298(5599):1775-9.

51 Meister G, Bühler D, Pillai R, et al. A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. *Nat Cell Biol*. 2001;3(11):945-9.

52 Kleinschmidt AM, Patton JR, Pederson T. U2 small nuclear RNP assembly in vitro. *Nucleic Acids Res*. 1989;17(12):4817-28.

53 Temsamani J, Rhoadhouse M, Pederson T. The U2 small nuclear ribonucleoprotein particle associates with nuclear factors in a pre-mRNA independent reaction. *J Biol Chem*. 1991;266(30):20356-62.

54 Battle DJ, Kasim M, Wang J, et al. SMN-independent subunits of the SMN complex. Identification of a small nuclear ribonucleoprotein assembly intermediate. *J Biol Chem*. 2007;282(38):27953-9.

55 Schilling M, Prusty AB, Boysen B, et al. TOR signaling regulates liquid phase separation of the SMN complex governing snRNP biogenesis. *Cell Rep*. 2021;35(12):109277.

56 Lee MJ, Yaffe MB. Protein Regulation in Signal Transduction. *Cold Spring Harb Perspect Biol*. 2016;8(6).

57 Grimmler M, Bauer L, Nousiainen M, et al. Phosphorylation regulates the activity of the SMN complex during assembly of spliceosomal U snRNPs. *EMBO Rep*. 2005;6(1):70-6.

58 Wu CY, Curtis A, Choi YS, et al. Identification of the phosphorylation sites in the survival motor neuron protein by protein kinase A. *Biochim Biophys Acta*. 2011;1814(9):1134-9.

59 Husedzinovic A, Oppermann F, Draeger-Meurer S, et al. Phosphoregulation of the human SMN complex. *Eur J Cell Biol*. 2014;93(3):106-17.

60 Gray KM, Kaifer KA, Baillat D, et al. Self-oligomerization regulates stability of survival motor neuron protein isoforms by sequestering an SCF(Slmb) degron. *Mol Biol Cell*. 2018;29(2):96-110.

61 Rademacher S, Detering NT, Schuning T, et al. A Single Amino Acid Residue Regulates PTEN-Binding and Stability of the Spinal Muscular Atrophy Protein SMN. *Cells*. 2020;9(11).

62 Hebert MD, Szymczyk PW, Shpargel KB, et al. Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes Dev*. 2001;15(20):2720-9.

63 Machyna M, Heyn P, Neugebauer KM. Cajal bodies: where form meets function. *Wiley Interdiscip Rev RNA*. 2013;4(1):17-34.

64 Weber C, Schreiber TB, Daub H. Dual phosphoproteomics and chemical proteomics analysis of erlotinib and gefitinib interference in acute myeloid leukemia cells. *J Proteomics*. 2012;75(4):1343-56.

65 Huttlin EL, Jedrychowski MP, Elias JE, et al. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell*. 2010;143(7):1174-89.

66 Mertins P, Yang F, Liu T, et al. Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Mol Cell Proteomics*. 2014;13(7):1690-704.

67 Mertins P, Qiao JW, Patel J, et al. Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nat Methods*. 2013;10(7):634-7.

68 Mertins P, Mani DR, Ruggles KV, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature*. 2016;534(7605):55-62.

69 Kettenbach AN, Schweppe DK, Faherty BK, et al. Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Sci Signal*. 2011;4(179):rs5.

70 Madeira F, Madhusoodanan N, Lee J, et al. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res*. 2024;52(W1):W521-W5.

71 Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011;7:539.

72 Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42(Web Server issue):W320-4.

73 McShane E, Selbach M. Physiological Functions of Intracellular Protein Degradation. *Annu Rev Cell Dev Biol*. 2022;38:241-62.

74 Sun X, Fontaine JM, Hoppe AD, et al. Abnormal interaction of motor neuropathy-associated mutant HspB8 (Hsp22) forms with the RNA helicase Ddx20 (gemin3). *Cell Stress Chaperones*. 2010;15(5):567-82.

75 Yan X, Mouillet JF, Ou Q, et al. A novel domain within the DEAD-box protein DP103 is essential for transcriptional repression and helicase activity. *Mol Cell Biol*. 2003;23(1):414-23.

76 Grundhoff AT, Kremmer E, Tureci O, et al. Characterization of DP103, a novel DEAD box protein that binds to the Epstein-Barr virus nuclear proteins EBNA2 and EBNA3C. *J Biol Chem*. 1999;274(27):19136-44.

77 Lee K, Pisarska MD, Ko JJ, et al. Transcriptional factor FOXL2 interacts with DP103 and induces apoptosis. *Biochem Biophys Res Commun*. 2005;336(3):876-81.

78 Klappacher GW, Lunyak VV, Sykes DB, et al. An induced Ets repressor complex regulates growth arrest during terminal macrophage differentiation. *Cell*. 2002;109(2):169-80.

79 Ginter-Matuszewska B, Kusz K, Spik A, et al. NANOS1 and PUMILIO2 bind microRNA biogenesis factor GEMIN3, within chromatoid body in human germ cells. *Histochem Cell Biol*. 2011;136(3):279-87.

80 Almstead LL, Sarnow P. Inhibition of U snRNP assembly by a virus-encoded proteinase. *Genes Dev*. 2007;21(9):1086-97.

81 Borg R, Cauchi RJ. The Gemin associates of survival motor neuron are required for motor function in Drosophila. *PLoS One*. 2013;8(12):e83878.

82 Cauchi RJ, Davies KE, Liu JL. A motor function for the DEAD-box RNA helicase, Gemin3, in Drosophila. *PLoS Genet*. 2008;4(11):e1000265.

83 Cacciottolo R, Ciantar J, Lanfranco M, et al. SMN complex member Gemin3 self-interacts and has a functional relationship with ALS-linked proteins TDP-43, FUS and Sod1. *Sci Rep*. 2019;9(1):18666.

84 Ou Q, Mouillet JF, Yan X, et al. The DEAD box protein DP103 is a regulator of steroidogenic factor-1. *Mol Endocrinol*. 2001;15(1):69-79.

85 Lee MB, Lebedeva LA, Suzawa M, et al. The DEAD-box protein DP103 (Ddx20 or Gemin-3) represses orphan nuclear receptor activity via SUMO modification. *Mol Cell Biol*. 2005;25(5):1879-90.

86 Shin EM, Hay HS, Lee MH, et al. DEAD-box helicase DP103 defines metastatic potential of human breast cancers. *J Clin Invest*. 2014;124(9):3807-24.

87 Ono Y, Saido TC, Sorimachi H. Calpain research for drug discovery: challenges and potential. *Nat Rev Drug Discov*. 2016;15(12):854-76.

88 de la Fuente S, Sansa A, Periyakaruppiah A, et al. Calpain Inhibition Increases SMN Protein in Spinal Cord Motoneurons and Ameliorates the Spinal Muscular Atrophy Phenotype in Mice. *Mol Neurobiol*. 2019;56(6):4414-27.

89 de la Fuente S, Sansa A, Hidalgo I, et al. Calpain system is altered in survival motor neuron-reduced cells from in vitro and in vivo spinal muscular atrophy models. *Cell Death Dis*. 2020;11(6):487.

90 Walker MP, Rajendra TK, Saieva L, et al. SMN complex localizes to the sarcomeric Z-disc and is a proteolytic target of calpain. *Hum Mol Genet*. 2008;17(21):3399-410.

91 Fuentes JL, Strayer MS, Matera AG. Molecular determinants of survival motor neuron (SMN) protein cleavage by the calcium-activated protease, calpain. *PLoS One*. 2010;5(12):e15769.

92 Velma V, Broome HJ, Hebert MD. Regulated specific proteolysis of the Cajal body marker protein coilin. *Chromosoma*. 2012;121(6):629-42.

93 Kenney DL, Benarroch EE. The autophagy-lysosomal pathway: General concepts and clinical implications. *Neurology*. 2015;85(7):634-45.

94 Yamamoto H, Matsui T. Molecular Mechanisms of Macroautophagy, Microautophagy, and Chaperone-Mediated Autophagy. *J Nippon Med Sch*. 2024;91(1):2-9.

95 Kaushik S, Cuervo AM. The coming of age of chaperone-mediated autophagy. *Nat Rev Mol Cell Biol*. 2018;19(6):365-81.

96 Kim JK, Jha NN, Awano T, et al. A spinal muscular atrophy modifier implicates the SMN protein in SNARE complex assembly at neuromuscular synapses. *Neuron*. 2023;111(9):1423-39 e4.

97 Her YR, Fuentes-Moliz A, Kothary R, et al. Protein-stabilizing and neurotransmission-potentiating activities of a synaptic chaperone modify spinal muscular atrophy in model mice. *bioRxiv*. 2026.

98 Ryu SW, Stewart R, Pectol DC, et al. Proteome-wide identification of HSP70/HSC70 chaperone clients in human cells. *PLoS Biol*. 2020;18(7):e3000606.

99 Matera AG, Steiner RE, Mills CA, et al. Proteomic analysis of the SMN complex reveals conserved and etiologic connections to the proteostasis network. *Front RNA Res*. 2024;2.

100 Esser LM, Li Q, Judt M, et al. The Impact of p70S6 Kinase-Dependent Phosphorylation of Gemin2 in UsnRNP Biogenesis. *Int J Mol Sci*. 2023;24(21).

101 Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *Journal of open source software*. 2019;4(43):1686.

102 Lex A, Gehlenborg N, Strobelt H, et al. UpSet: Visualization of Intersecting Sets. *IEEE Transactions on Visualization and Computer Graphics*. 2014;20(12):1983-92.

103 Krassowski M, Arts M, Lagger C. krassowski/complex-upset: v1. 3.5. *Zenodo*. 2022.

104 Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-7.

105 Liu ZX, Yu K, Dong J, et al. Precise Prediction of Calpain Cleavage Sites and Their Aberrance Caused by Mutations in Cancer. *Front Genet*. 2019;10:715.

106 Kirchner P, Bourdenx M, Madrigal-Matute J, et al. Proteome-wide analysis of chaperone-mediated autophagy targeting motifs. *PLoS Biol*. 2019;17(5):e3000301.

107 Egan DF, Shackelford DB, Mihaylova MM, et al. Phosphorylation of ULK1 (hATG1) by AMP-activated protein kinase connects energy sensing to mitophagy. *Science*. 2011;331(6016):456-61.

108 Gwinn DM, Shackelford DB, Egan DF, et al. AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol Cell*. 2008;30(2):214-26.

109 Inoki K, Zhu T, Guan KL. TSC2 mediates cellular energy response to control cell growth and survival. *Cell*. 2003;115(5):577-90.

110 Kim J, Kundu M, Viollet B, et al. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nat Cell Biol*. 2011;13(2):132-41.

111 Detering NT, Schuning T, Hensel N, et al. The phospho-landscape of the survival of motoneuron protein (SMN) protein: relevance for spinal muscular atrophy (SMA). *Cell Mol Life Sci*. 2022;79(9):497.

112 Laplante M, Sabatini DM. mTOR signaling in growth control and disease. *Cell*. 2012;149(2):274-93.

113 Liu GY, Sabatini DM. mTOR at the nexus of nutrition, growth, ageing and disease. *Nat Rev Mol Cell Biol*. 2020;21(4):183-203.

114 Gupta K, Martin R, Sharp R, et al. Oligomeric Properties of Survival Motor Neuron.Gemin2 Complexes. *J Biol Chem*. 2015;290(33):20185-99.

115 Sarachan KL, Valentine KG, Gupta K, et al. Solution structure of the core SMN-Gemin2 complex. *Biochem J*. 2012;445(3):361-70.

116 Singh NN, Seo J, Rahn SJ, et al. A multi-exon-skipping detection assay reveals surprising diversity of splice isoforms of spinal muscular atrophy genes. *PLoS One*. 2012;7(11):e49595.

117 Singh RN, Howell MD, Ottesen EW, et al. Diverse role of survival motor neuron protein. *Biochim Biophys Acta Gene Regul Mech*. 2017;1860(3):299-315.

118 Wang J, Dreyfuss G. Characterization of functional domains of the SMN protein in vivo. *J Biol Chem*. 2001;276(48):45387-93.

119 Miralles MP, Sansa A, Beltran M, et al. Survival motor neuron protein and neurite degeneration are regulated by Gemin3 in spinal muscular atrophy motoneurons. *Front Cell Neurosci*. 2022;16:1054270.

120 Pellizzoni L, Charroux B, Dreyfuss G. SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. *Proc Natl Acad Sci U S A*. 1999;96(20):11167-72.

121 Brahms H, Raymackers J, Union A, et al. The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies. *J Biol Chem*. 2000;275(22):17122-9.

122 Hirakata M, Craft J, Hardin JA. Autoantigenic epitopes of the B and D polypeptides of the U1 snRNP. Analysis of domains recognized by the Y12 monoclonal anti-Sm antibody and by patient sera. *J Immunol*. 1993;150(8 Pt 1):3592-601.

123 Lerner EA, Lerner MR, Janeway CA, Jr., et al. Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. *Proc Natl Acad Sci U S A*. 1981;78(5):2737-41.

124 Ry C. Un sencillo metodo de coloracion selectiva del reticulo protoplasmico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. *Lab Invest Biol Univ (Madrid)*. 1903;1292143.

125 Andrade LE, Chan EK, Raska I, et al. Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. *J Exp Med*. 1991;173(6):1407-19.

126 Liu Q, Dreyfuss G. A novel nuclear structure containing the survival of motor neurons protein. *Embo j*. 1996;15(14):3555-65.

127 Mizushima N, Komatsu M. Autophagy: renovation of cells and tissues. *Cell*. 2011;147(4):728-41.

128 He C. Balancing nutrient and energy demand and supply via autophagy. *Curr Biol*. 2022;32(12):R684-R96.

129 Klionsky DJ, Abdel-Aziz AK, Abdelfatah S, et al. Guidelines for the use and interpretation of assays for monitoring autophagy (4th edition)(1). *Autophagy*. 2021;17(1):1-382.

130 Mizushima N, Klionsky DJ. Protein turnover via autophagy: implications for metabolism. *Annu Rev Nutr*. 2007;27:19-40.

131 Gubitz AK, Feng W, Dreyfuss G. The SMN complex. *Exp Cell Res*. 2004;296(1):51-6.

132 Asadi M, Taghizadeh S, Kaviani E, et al. Caspase ‐ 3: structure, function, and biotechnological aspects. *Biotechnology and applied biochemistry*. 2022;69(4):1633-45.

133 Tsujinaka T, Kajiwara Y, Kambayashi J, et al. Synthesis of a new cell penetrating calpain inhibitor (calpeptin). *Biochem Biophys Res Commun*. 1988;153(3):1201-8.

134 Bowman EJ, Siebers A, Altendorf K. Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells. *Proc Natl Acad Sci U S A*. 1988;85(21):7972-6.

135 Mindell JA. Lysosomal acidification mechanisms. *Annu Rev Physiol*. 2012;74:69-86.

136 Pohl C, Dikic I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science*. 2019;366(6467):818-22.

137 Lee DH, Goldberg AL. Proteasome inhibitors: valuable new tools for cell biologists. *Trends Cell Biol*. 1998;8(10):397-403.

138 Mizushima N. Autophagy: process and function. *Genes Dev*. 2007;21(22):2861-73.

139 Wing SS, Chiang HL, Goldberg AL, et al. Proteins containing peptide sequences related to Lys-Phe-Glu-Arg-Gln are selectively depleted in liver and heart, but not skeletal muscle, of fasted rats. *Biochem J*. 1991;275 (Pt 1)(Pt 1):165-9.

140 Dunlop EA, Tee AR. The kinase triad, AMPK, mTORC1 and ULK1, maintains energy and nutrient homoeostasis. *Biochem Soc Trans*. 2013;41(4):939-43.

141 Binda O, Juillard F, Ducassou JN, et al. SMA-linked SMN mutants prevent phase separation properties and SMN interactions with FMRP family members. *Life Sci Alliance*. 2023;6(1).

142 London E, Bloyd M, Stratakis CA. PKA functions in metabolism and resistance to obesity: lessons from mouse and human studies. *J Endocrinol*. 2020;246(3):R51-r64.

143 Bühler D, Raker V, Lührmann R, et al. Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. *Hum Mol Genet*. 1999;8(13):2351-7.

144 Brahms H, Meheus L, de Brabandere V, et al. Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *Rna*. 2001;7(11):1531-42.

145 Friesen WJ, Massenet S, Paushkin S, et al. SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol Cell*. 2001;7(5):1111-7.

146 Friesen WJ, Dreyfuss G. Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN). *J Biol Chem*. 2000;275(34):26370-5.

147 Selenko P, Sprangers R, Stier G, et al. SMN tudor domain structure and its interaction with the Sm proteins. *Nat Struct Biol*. 2001;8(1):27-31.

148 Meister G, Eggert C, Buhler D, et al. Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol*. 2001;11(24):1990-4.

149 Tripsianes K, Madl T, Machyna M, et al. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat Struct Mol Biol*. 2011;18(12):1414-20.

150 Sprangers R, Selenko P, Sattler M, et al. Definition of domain boundaries and crystallization of the SMN Tudor domain. *Acta Crystallogr D Biol Crystallogr*. 2003;59(Pt 2):366-8.

151 Friesen WJ, Paushkin S, Wyce A, et al. The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol*. 2001;21(24):8289-300.

152 Cho NH, Cheveralls KC, Brunner AD, et al. OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science*. 2022;375(6585):eabi6983.

153 Zhao DY, Gish G, Braunschweig U, et al. SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature*. 2016;529(7584):48-53.

154 Tsalikis J, Tattoli I, Ling A, et al. Intracellular Bacterial Pathogens Trigger the Formation of U Small Nuclear RNA Bodies (U Bodies) through Metabolic Stress Induction. *J Biol Chem*. 2015;290(34):20904-18.

155 Metwally E, Al-Abbadi HA, Hussain T, et al. Calpain signaling: from biology to therapeutic opportunities in neurodegenerative disorders. *Front Vet Sci*. 2023;10:1235163.

156 Lescouzères L, Butti Z, Chaineau M, et al. Targeting calpastatin pharmacologically restores synaptic proteolysis and preserves motor neurons survival and function in C9orf72 ALS. *bioRxiv*. 2026:2025.12.08.692907.

157 Battle DJ, Lau CK, Wan L, et al. The Gemin5 protein of the SMN complex identifies snRNAs. *Mol Cell*. 2006;23(2):273-9.

158 Cacciottolo R, Cauchi RJ. A critical genetic interaction between Gemin3/Ddx20 and translation initiation factor NAT1/eIF4G2 drives development. *Dev Biol*. 2025;521:37-51.

159 Huang Y, Wang C, Li K, et al. Death-associated protein kinase 1 suppresses hepatocellular carcinoma cell migration and invasion by upregulation of DEAD-box helicase 20. *Cancer Sci*. 2020;111(8):2803-13.

160 He L, Fan X, Zhu Z, et al. African swine fever virus MGF360-9L degrades DDX20 through the Rab1A-dependent autophagy pathway to antagonize its antiviral effect. *Virol Sin*. 2025;40(5):822-34.

161 Bushman JW, Deng W, Samarasinghe KTG, et al. Discovery of a VHL molecular glue degrader of GEMIN3 by Picowell RNA-seq. 2025.

162 Matera AG. Chaperone dysfunction in motor neuron disease: new insights from studies of the SMN complex. *Genetics*. 2025;229(3).

163 Burnett BG, Munoz E, Tandon A, et al. Regulation of SMN protein stability. *Mol Cell Biol*. 2009;29(5):1107-15.

164 Locatelli D, Terao M, Kurosaki M, et al. Different Stability and Proteasome-Mediated Degradation Rate of SMN Protein Isoforms. *PLoS One*. 2015;10(7):e0134163.

165 Chang HC, Hung WC, Chuang YJ, et al. Degradation of survival motor neuron (SMN) protein is mediated via the ubiquitin/proteasome pathway. *Neurochem Int*. 2004;45(7):1107-12.

166 Hsu SH, Lai MC, Er TK, et al. Ubiquitin carboxyl-terminal hydrolase L1 (UCHL1) regulates the level of SMN expression through ubiquitination in primary spinal muscular atrophy fibroblasts. *Clin Chim Acta*. 2010;411(23-24):1920-8.

167 Cho S, Dreyfuss G. A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes Dev*. 2010;24(5):438-42.

168 Hensel N, Kubinski S, Claus P. The Need for SMN-Independent Treatments of Spinal Muscular Atrophy (SMA) to Complement SMN-Enhancing Drugs. *Front Neurol*. 2020;11:45.

169 Wen X, Lan T, Su W, et al. Latest progress and challenges in drug development for degenerative motor neuron diseases. *Neural Regen Res*. 2026;21(5):1849-63.

170 Yesbek Kaymaz A, Bora-Akoglu G, Erdem Yurter H, et al. Gene Targeted Therapies for Neurodegenerative Disorders: Strategies and Implications in ALS and SMA. *Genes (Basel)*. 2026;17(4).

Acknowledgements

First and foremost, I would like to thank my supervisors, Prof. Sebastian Wesselborg, Prof. Jürgen Scheller and Dr. Christoph Peter for their endless support in every aspect of my work. I am really grateful to Prof. Sebastian Wesselborg and Dr. Christoph Peter for providing me with the opportunity to join the Institute of Molecular Medicine I institute and perform this research project. I specifically want to thank Dr. Christoph Peter for his academic guidance. This simple acknowledgement is hardly enough to express my profound gratitude. Thank you for everything!

Many thanks to my thoughtful and helpful colleagues at the Institute of Molecular Medicine I. Their patience, collaboration and shared expertise made the daily work of this project not only easier, but more enjoyable. Special thanks go to Dr. Lea Marie Esser and Dr. Jan Cox who have generously taught me many valuable experiment techniques at the beginning of this project. I genuinely appreciate the help from my close colleague M. Sc. Maximilian Jüdt. From managing the paperwork of product orders and MS samples deliveries, to helping with heavy tasks like changing the CO_2 bottles, his assistance made everything run more smoothly. I would also like to extend my thanks to Cand. med. Lilli Albers and M.Sc. Isabel Michelmann. Although they were here for only a short time, their hard work and positive energy made my journey more colourful.

This study was further bolstered by the cooperation with Prof. Thilo Kähne and Dr. María José Mendiburo. I am grateful to Prof. Thilo Kähne for conducting MS and Dr. María José Mendiburo for providing ATG5 KO cells.

I would like to thank the staff at our university and the local municipal offices. Their prompt responses and invaluable administrative support allowed me to focus more fully on my research. I would especially like to thank our institution coordinator, Annette Eder-Martin, for her assistance with various administrative tasks.

The research was made possible by a generous scholarship from the China Scholarship Council. I am deeply grateful for their financial support, which allowed me to start the research.

Some people in my personal life deserve to be mentioned here. I want to thank my friend, Dr. Mingming Cui, a brave and kind person who is a role model for me. I am truly thankful for her encouragement and family-like companionship throughout this journey. Besides, thank you to my parents for giving me freedom to do whatever I am interested.

Last, thank you to this beautiful city Düsseldorf for giving me a lovely neighbourhood for living and diverse cuisines to comfort my stomach.

This journey of PhD study likes a holiday to me, full of blessings and full of serendipities. God is good to me everyday. *Deo gratias*.