



**University of
Sheffield**

**Novel Gene Therapy Approaches for
The Viral and Non-viral Delivery of
Cell-permeable Peptides Targeting
the SRSF1-dependent Nuclear Export
of Pathological *C9ORF72*-ALS/FTD
Repeat Transcripts**

Submitted for the degree of Doctor of Philosophy (PhD) October 2023

Aytac GUL 180285554

Supervisors: Professor Guillaume Hautbergue, Professor Mimoun Azzouz and
Professor Dame Pamela Shaw

Department of Neuroscience

Sheffield Institute for Translational Neuroscience, University of Sheffield

04/10/2023

Dedication

This thesis is dedicated to those who have suffered from neurodegenerative diseases to their families and loved ones, and to the dedicated scientists and healthcare professionals working tirelessly towards uncovering the mysteries of the human brain and finding a cure for these devastating diseases.

I would also like to dedicate this thesis to my beloved relatives and friends who were taken from us too soon in the devastating earthquake which happened on 6th February in 2023, in my hometown in Turkey. Their sudden departure has left a void that can never be filled. I am grateful for the time we shared together and will carry their love and kindness with me always. I hope that through my work and dedication in this thesis, I can honour their memory. May Allah (God) grant them peace, and may their legacy continue to inspire and motivate us to strive for a better tomorrow.

Acknowledgements

I would like to express my deepest gratitude to my thesis advisors, Prof Guillaume Hautbergue, Prof Mimoun Azzouz and Prof Dame Pamela Shaw for their unwavering support, guidance, and encouragement throughout my research journey. Their invaluable advice, expertise, and insightful suggestions have been instrumental in shaping the direction and success of my work. I am so grateful for everything I have learned from you.

I would also like to thank Prof. Dr. Abdullah Arpacı, my scholarship (YLSY) supervisor at Hatay Mustafa Kemal University in Turkey, for his invaluable advice, encouragement and motivation which have significantly contributed my academic and personal development.

This PhD would not have been as successful without the help of some very important people at SITraN especially Dr. Lydia Castelli, Dr. Ya-Hui Lin, Dr. Tobias Moll, Dr. Adrian Higginbottom, Dr. David Burrows, Dr. Kamallia Mohd Imran, Ms. Chloé Moutin. They spent time to help me to gain technical laboratory skills, the ability to think critically and provide valuable advice. I am grateful to all the technical staff in the department, but especially to Mr Ian Coldicott. Thank you for helping me to perform all the animal work in this thesis. I would also like to thank present and past members of the Hautbergue RNA Biology especially Ms Nikita Soni and Mr. Sajid Shah.

I'd like to extend my heartfelt gratitude to my dear brother, Mr. Tunahan Kirktepelı, to my dear friends, Ms. Neslihan Kirktepelı, Mr. Onur Ozcelik, Ms. Melis Kahyaoglu, and to my little brother, Burak Kirktepelı. Your unwavering support and constant encouragement have been indispensable to my journey. Thank you for being more than just friends and family during this process; you've been pillars of strength.

A special thank you to the Republic of Turkiye Ministry of National Education for making this PhD opportunity possible for me by providing financial support throughout my PhD, giving special attention to sort out the difficulties I have experienced. I will always be thankful for that.

Lastly, I extend my heartfelt thanks to my parents Hatice Oya Gul and Asim Gul who have played a significant role in shaping me into the person I am today. Their love, support, encouragement, and all that they have done for me made it possible for me to complete this PhD. Also, I would like to express my deepest gratitude to my wife, Dr. Bahar Unlu Gul, for her unwavering love, encouragement, and support throughout my PhD journey. Her presence in my life every single day with unconditional support and holding my hand every step of the way was the driving force behind the success of this thesis. This thesis also belongs to you.

Declaration

I, Aytac Gul, declare that the work contained in this thesis entitled “Novel Gene Therapy Approaches for The Viral and Non-viral Delivery of Cell-permeable Peptides Targeting the SRSF1-dependent Nuclear Export of Pathological *C9ORF72*-ALS/FTD Repeat Transcripts”, is my own original work and that any assistance received has been fully acknowledged. I also declare that to the best of my knowledge and belief, the work has not been submitted elsewhere for a degree or award at any other university.

As specified, part of this thesis is consisting of following publication:

- Chapter 4: **Functional Characterization of Recombinant SRSF1-V5-TAT CPP in *in vitro* and *in vivo* models of *C9ORF72*-ALS/FTD** – Castelli et al, 2023. A cell-penetrant peptide blocking *C9ORF72*-repeat RNA nuclear export suppresses neurodegeneration. *Science Translational Medicine* DOI: [10.1126/scitranslmed.abo3823](https://doi.org/10.1126/scitranslmed.abo3823).

Research Output

Conferences, Presentations and Course Attended

- The University of Sheffield, The Medical School, 1st year Oral platform presentation (2019) - ‘Novel therapeutic strategies for motor neurone disease: moving towards gene therapy approaches’.
- American Society of Gene and Cell Therapy Conference (2020). Online
- Sheffield Institute for Translational Neuroscience Oral platform presentation (2021)- ‘Non-viral and Viral Gene Therapy Approaches to target nuclear export of pathological *C9orf72* repeat transcripts in ALS/FTD’.
- The University of Sheffield, The Medical School, Annual Medical Research Meeting (2021)- Oral platform presentation ‘Novel Gene Therapy Approaches to target nuclear export of pathological *C9orf72* repeat transcripts’.
- The University of Sheffield, The Medical School, 3rd year (2021). Oral platform presentation ‘A potential novel gene therapy approach to target the nuclear export of pathological *C9orf72* repeat transcripts in ALS/FTD’.
- Medical School Research Meeting of the University of Sheffield (2021) - Oral platform presentation ‘A potential novel gene therapy approach to target the nuclear export of pathological *C9orf72* repeat transcripts in ALS/FTD’.
- 32nd International Symposium on ALS/MND (2021)- Online.
- The University of Sheffield, The Faculty of Medicine, Dentistry and Health Research Staff Association (2022) Oral platform presentation ‘A potential novel gene therapy approach to target the nuclear export of pathological *C9orf72* repeat transcripts in ALS/FTD’ Best Oral Talk Award.
- The University of Sheffield, Biological Service Unit, Home Office Personal License Course. 22/4/21.

- 33rd International Symposium in ALS/MND (2022). Poster presented ‘Development of a methodology for the bio-synthesis of a novel “drug-like” cell permeable peptide inhibiting the SRSF1-dependent nuclear export of *C9orf72*-repeat transcripts in ALS/FTD’.
- Translational UK Conference in Cambridge (2023). Poster presented ‘Development of a methodology for the bio-synthesis of a novel “drug-like” cell-permeable peptide inhibiting the SRSF1-dependent nuclear export of C9ORF72-repeat transcripts in ALS/FTD’.

Publications

- Simeon R. Mihaylov, Lydia M. Castelli, Ya-Hui Lin, **Aytac Gül**, Nikita Soni, Christopher Hastings, Helen R. Flynn, Mark J. Dickman, Ambrosius P. Snijders, Oliver Bandmann, Heather Mortiboys, Sila K. Ultanir, Guillaume M. Hautbergue. The master energy homeostasis regulator PGC-1alpha couples transcriptional co-activation and mRNA nuclear export. *Nature Communication* 2023; 14: 5496.
- Lydia M. Castelli, Alvaro Sanchez-Martinez, Ya-Hui Lin, **Aytac Gül** Santosh Kumar Upadhyay, Adrian Higginbottom, Johnathan Cooper-Knock, , Amy Walton, Claire Montmasson, Rebecca Cohen, Claudia S. Bauer, Kurt J. De Vos, Mimoun Azzouz, Pamela J. Shaw, Cyril Dominguez, Laura Ferraiuolo, Alexander J. Whitworth, Guillaume M. Hautbergue. A cell-penetrant peptide blocking C9ORF72-repeat RNA nuclear export reduces the neurotoxic effects of dipeptide repeat proteins. *Science Translational Medicine* 2023; 15, eabo3823.
- **Aytac Gül**, Wang-Ping Huan, Mimoun Azzouz, Pamela J. Shaw, Ya-Hui Lin, Guillaume M. Hautbergue. Expression and Purification of Recombinant Cell-Penetrant Peptides (In Preparation).

Abstract

Background: Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are an incurable neurodegenerative diseases involving selective death of upper and lower motor neurons and loss of neurons in the frontal and temporal lobes of the brain respectively. Hexanucleotide (GGGGCC) repeat expansions in *C9orf72* are the most common cause of ALS and FTD. Our lab has reported that the nuclear export adaptor protein SRSF1 (Serine/arginine-rich splicing factor 1) dependent nuclear export of abnormal *C9orf72* repeat transcripts causes cytoplasmic production of dipeptide-repeat proteins (DPRs), which is one of the main driving pathogenic mechanisms. Antagonizing the function of SRSF1 was further shown to confer neuroprotection in reporter cell models, patient-derived neurons and *Drosophila* models of *C9ORF72*-ALS/FTD.

Aims and objectives: In this context, I aimed to: **(i)** Develop a robust methodology for biosynthesis of a novel, drug like SRSF1 cell permeable peptide (CPP); **(ii)** Evaluate functionality and efficacy of SRSF1-linker CPP in proof-of-concept experiments using relevant *in vitro* and *in vivo* models of *C9ORF72*-ALS/FTD; **(iii)** Design novel viral gene therapy treatments consisting of self-complementary adeno-associated virus (scAAV9) SRSF1-linker CPP and scAAV9 SRSF1- RRM2 α 1 CPP; **(iv)** Validate functionality of generated scAAV9 viral gene therapy tools using *in vitro* and animal models of *C9ORF72*-ALS/FTD.

Results: I performed detailed optimization experiments for the expression and purification of the SRSF1-linker CPP. The functionality of recombinant SRSF1-linker CPP was further validated in a human embryonic kidney cell (HEK) model of *C9ORF72*-ALS/FTD where it mediated inhibition of the nuclear export of pathogenic *C9orf72* repeat transcripts and concomitantly reduced the production of sense and antisense DPRs and their related toxicity. Then, I reported the testing of SRSF1-linker CPP in C9-500 BAC mice which resulted in a substantial decrease of 70 % poly (GP) accumulation in the brain. In parallel, I described a rational design of gene therapy approaches involving scAAV9 mediated expression of SRSF1-linker CPP and SRSF1-RRM2 α 1 CPP. Subsequently, I demonstrated that transfecting scAAV9 viral vectors into cell cultures led to expression of V5 tagged CPPs, which resulted in a decrease of both sense and antisense DPRs and their associated toxicity. Lastly, our viral method was effective *in vivo*, shown by the reduction of poly (GP) expression in the CNS.

Conclusion: This thesis provided proof-of-concept for SRSF1 mediated peptide as a drug and viral gene therapy approaches with therapeutic potential for the treatment *C9ORF72*-ALS/FTD.

Table of Contents

Dedication	i
Acknowledgements	ii
Declaration	iv
Research Output	v
Abstract	vii
Table of Contents	viii
List of Figures	xiii
List of Tables.....	xvii
Abbreviations	xix
1. Chapter 1 - Introduction	1
1.1. Background Information Relating to Amyotrophic Lateral Sclerosis.....	2
1.2. Clinical Presentation of ALS	2
1.3. Epidemiology.....	2
1.4. Genetics of ALS	3
1.4.1. <i>SOD1</i>	3
1.4.2. <i>TARDBP</i>	4
1.4.3. <i>FUS</i>	4
1.4.4. <i>C9ORF72</i>	5
1.5. The Pathophysiological Mechanisms of ALS	9
1.5.1. RNA Metabolism Defects in ALS	9
1.6. <i>C9ORF72</i> -ALS/FTD	11
1.7. <i>C9ORF72</i> -Mediated ALS Mechanisms.....	13
1.7.1. Loss of <i>C9ORF72</i> Function Through Haploinsufficiency.....	13
1.7.2. <i>C9ORF72</i> Toxic Gain of Function Mechanisms	14

1.7.3. SRSF1-Dependent Nuclear Export of Pathological <i>C9ORF72</i> -Repeat Transcripts.....	17
1.8. <i>C9ORF72</i> -ALS/FTD Mouse Models	25
1.9. Therapeutic Interventions for ALS/FTD	27
1.9.1. Drug Treatment.....	27
1.9.2. Stem Cell Therapy	28
1.9.3. Gene Therapy Approach in ALS/FTD.....	29
1.9.4. Therapeutic Strategies for <i>C9ORF72</i> -ALS/FTD	39
1.10. Limitation of Therapeutic Approaches	46
1.11. Cell-Penetrating Peptides.....	47
1.11.1. Classification of CPPs.....	47
1.11.2. Intracellular Uptake Mechanisms of CPPs	50
1.11.3. Pre-clinical and Clinical Studies	51
1.12. Hypothesis and Aims	54
2. Chapter 2 - Materials and Method.....	55
2.1. Materials	56
2.1.1. Strains and Cell Lines	56
2.1.2. Plasmids and Vectors	57
2.1.3. Primers	59
2.1.4. Media and Antibiotics for Bacterial and Mammalian Cell Lines	63
2.1.5. Restriction and DNA Modifying Enzymes.....	65
2.1.6. Antibodies	66
2.1.7. Molecular Biology Kits.....	67
2.1.8. Equipment.....	68
2.2. Methods	69
2.2.1. Maintenance of Plasmids	69
2.2.2. Molecular Cloning	74

2.2.3.	Recombinant Protein Expression and Purification	78
2.2.4.	Mammalian Protein Expression	84
2.2.5.	Analysis of Protein by SDS-PAGE Gel Electrophoresis	86
2.2.6.	Immunocytochemistry	89
2.2.7.	GST-NXF1, NXF2 and NXF3 p15 Pull-Down Assays	90
2.2.8.	Co-Immunoprecipitation (CoIP) Assay	91
2.2.9.	Total and Cytoplasmic and Nuclear RNA Extraction.....	91
2.2.10.	MTT Cell Proliferation Assay.....	95
2.2.11.	Animal Work.....	95
2.2.12.	Statistical Analysis of Data	107
3.	Chapter 3 - Expression and Purification of Cell Permeable Peptides	108
3.1.	Introduction	109
3.2.	Expression and Purification of Recombinant Peptides.....	110
3.2.1.	Recombinant Peptide Expression	110
3.2.2.	Methodologies Used for Purification of Peptides	112
3.3.	Aims and Objectives.....	121
3.4.	Results	122
3.4.1.	Designing SRSF1(89-120)-V5-TAT PTD CPPs	122
3.4.2.	Production of Recombinant SRSF1(89-120)-V5-W-TAT (His) ₆ CPP ...	126
3.4.3.	Optimisation of SRSF1(89-120)-V5-W-TAT PTD-(His) ₆ CPP Expression (with the presence of tryptophan)	135
3.4.4.	Optimisation of SRSF1(89-120)-V5-TAT PTD-(His) ₆ CPP Expression (without the presence of tryptophan)	137
3.4.5.	Production of Recombinant GST-SRSF1(89-120)-V5-TAT PTD CPP .	139
3.4.6.	Expression and Purification of GST-control V5-TAT PTD CPP	168
3.5.	Discussion.....	177
3.6.	Experimental Conclusion	189

4. Chapter 4 - Functional Characterization of Recombinant SRSF1-V5-TAT CPP in <i>In Vitro</i> and <i>In Vivo</i> models of <i>C9ORF72</i> -ALS/FTD	192
4.1. Introduction	193
4.2. Aim and Objectives	194
4.3. Results	195
4.3.1. Recombinant CPPs are delivered into the nuclear and cytoplasmic compartments.....	195
4.3.2. Evaluation of Intracellular Distribution of Synthetic CPPs.....	198
4.3.3. Investigating The Potential <i>In Vitro</i> Interaction of Recombinant SRSF1-CPP with Tap/NXF1	200
4.3.4. Investigating The Potential Direct Interactions of Recombinant SRSF1-CPP with TAP/NXF1, NXF2 and NXF3	202
4.3.5. Recombinant SRSF1-CPP Inhibits SRSF1:NXF1 Dependent Nuclear Export of Sense and Antisense- <i>C9orf72</i> -Repeat RNA.....	206
4.3.6. Recombinant SRSF1-CPP Inhibits Dose-dependent Sense (poly-GA, GP, GR) and Antisense (poly-GP, PR, PA) DPR Expression	208
4.3.7. Recombinant SRSF1-CPP Inhibits Dose-dependent Reduction of Sense (poly-GA, GP, GR) and Antisense (poly-GP, PR, PA) DPR-associated Cytotoxicity...	210
4.3.8. Therapeutic Efficiency of SRSF1-CPP in an Animal Model of <i>C9ORF72</i> -ALS/FTD.....	212
4.4. Discussion.....	218
4.5. Experimental Conclusion	227
5. Chapter 5 - Engineering a Viral Vector Expressing SRSF1 Cell Permeable Peptides for The Development of Novel Gene Therapy Approaches	228
5.1. Introduction	229
5.2. Aims and Objectives.....	230
5.3. Results	231
5.3.1. Stability of SRSF1-CPP	231
5.3.2. Cloning of SRSF1-linker and RRM2α1 CPPs Into scAAV Constructs .	232

5.3.3. In vitro Validation of The Functionality of The Constructs	239
5.3.4. In Vivo Validation of The Functionality of scAAV9 Viruses	247
5.4. Discussion.....	261
5.5. Experimental Conclusion	277
6. Chapter 6 - General Conclusions and Discussion	278
6.1. Project Aims and Main Findings	279
6.2. Developing a Robust Protocol to Produce Recombinant SRSF1-linker CPP.....	280
6.3. Functionality of SRSF1-linker CPP in Human Cell Models of <i>C9ORF72</i> - ALS/FTD.....	281
6.4. <i>In vivo</i> Functionality of SRSF1-linker CPP	283
6.4.1. Drosophila Model of <i>C9ORF72</i> -ALS/FTD	283
6.4.2. Mouse Model of <i>C9ORF72</i> -ALS/FTD	283
6.5. Clinical Applications of CPP	288
6.6. Limitation and Future Work	291
6.7. Conclusions	293
7. References	295

List of Figures

Figure 1.1. Widespread dysregulation of components of RNA metabolism in ALS	9
Figure 1.2. Diagram representing the proposed <i>C9ORF72</i> hexanucleotide-repeat-mediated mechanisms of neurotoxicity	12
Figure 1.3. Common principles of NXF1 dependent nuclear export of eukaryotic bulk mRNA.	20
Figure 1.4. The identified two therapeutic strategies to inhibit the nuclear export of pathological <i>C9ORF72</i> repeat transcripts.	23
Figure 2.1. The representative image of southern blot transfer steps.....	98
Figure 2.2. QPCR plate design for viral biodistribution across several organs.	105
Figure 3.1. Schematic representation of typical anion exchange chromatography separation steps.....	115
Figure 3.2. Schematic representation of general affinity chromatography purification steps.	117
Figure 3.3. Flow diagram illustrating steps for size exclusion chromatography.....	120
Figure 3.4. Plasmid maps of engineered SRSF1(89-120)-V5-TAT PTD-His6 and SRSF1 (89-120)-V5-W-TAT PTD-His6.....	124
Figure 3.5. Induction time course of SRSF1(89120)-V5-W-TAT PTD-(His)6 in BL21-RP, BL-21 Rosetta and BL21-RIL <i>E. coli</i> strains.....	127
Figure 3.6. Two-step purification of SRSF1(89-120)-V5-W-TAT PTD-(His)6 CPP with Cobalt affinity and Mono-S strong cationic ion exchange chromatography.	130
Figure 3.7. Two-step purification of SRSF1(89-120)-V5-W-TAT PTD-(His)6 CPP with Cobalt affinity and SP weak cationic ion exchange chromatography.	132
Figure 3.8. Concentrating SRSF1(89-120)-V5-W-TAT PTD-(His)6 CPP fractions with Cobalt affinity chromatography.	134
Figure 3.9. Induction time course of SRSF1(89120)-V5-W-TAT PTD-(His)6 CPP expressed in BL21-RP <i>E. coli</i> strains.....	136
Figure 3.10. Induction time course of SRSF1(89-120)-V5-TAT PTD-(His)6 CPP expressed in BL21-RP <i>E. coli</i> strains.....	138
Figure 3.11. Induction time course of GST-SRSF1(89120)-V5-TAT PTD in BL21-RP, BL-21 Rosetta <i>E. coli</i> strains.	140
Figure 3.12. Small scale purification of GST-SRSF1-V5-TAT PTD CPP with GSH affinity chromatography.....	143

Figure 3.13. Large scale two-step purification of GST-SRSF1-V5-TAT PTD CPP including GSH affinity and SP ion exchange chromatography.	146
Figure 3.14. Effect of mobile phase with a range of salt concentrations (200 mM – 500 mM) in the presence of urea on the binding ability of cleaved SRSF1-CPP to SP chromatography resin.	148
Figure 3.15. Effect of three different buffer compositions on the binding ability of cleaved SRSF1-CPP and salt concentrations of elution buffer to SP chromatography resin.....	151
Figure 3.16. Investigating binding ability of cleaved SRSF1-CPP onto CM chromatography resin at range of salt concentrations (200 mM – 500 mM) and in presence of urea in mobile phase.....	153
Figure 3.17. Investigating binding ability of cleaved SRSF1-CPP onto CM chromatography resin at low range of salt concentrations (50 mM – 150 mM) and in the presence of urea in mobile phase.....	155
Figure 3.18. Investigating binding ability of cleaved SRSF1-CPP onto Heparin adsorption chromatography resin at low range of salt concentrations (50 mM – 150 mM) and in the presence of urea in mobile phase.	157
Figure 3.19. Investigating the binding ability of cleaved SRSF1-CPP onto Heparin adsorption chromatography resin at higher range of salt concentrations (200 mM, 250 mM and 300 mM) and in the presence of urea in mobile phase.....	158
Figure 3.20. Investigating total salt concentrations in elution buffer for desorption of SRSF1-CPP at 100 mM and 150 mM NaCl from Heparin adsorption chromatography.....	160
Figure 3.21. Investigating total salt concentrations in the Elution buffer for desorption of SRSF1-CPP at 150 mM and 200 mM NaCl from Heparin adsorption chromatography.....	161
Figure 3.22. Large scale two-steps purification of GST-SRSF1-V5-TAT PTD CPP with GSH affinity and Heparin adsorption chromatography.	163
Figure 3.23. Size exclusion chromatography of SRSF1-CPP.	165
Figure 3.24. Quantification and concentration of SRSF1-CPP.....	167
Figure 3.25. Cloning strategy of Control-V5-TAT PTD CPP.	169
Figure 3.26. Generation of GST-Control-V5-TAT PTD CPP plasmid.....	170
Figure 3.27. Induction time course of GST-V5-TAT PTD in BL-21 Rosetta <i>E. coli</i> strains.	171
Figure 3.28. Large scale purification of GST-Ctrl-V5-TAT PTD CPP with GSH affinity and Heparin adsorption chromatography.....	172

Figure 3.29. Size exclusion chromatography of Ctrl CPP.	174
Figure 3.30. Quantification and concentration of Ctrl CPP.	176
Figure 3.31. Schematic representation of the optimization strategies for expression and purification of SRSF1(89-120)-V5-TAT PTD-(His) ₆ CPP (with or without tryptophan).	189
Figure 3.32. Schematic representation of the optimization strategies for expression and purification of GST SRSF1(89-120)-V5-TAT PTD.	190
Figure 3.33. The optimised purification method for Recombinant SRSF1 and Ctrl-CPP ...	191
Figure 4.1. The recombinant CPPs are detected in the nucleus and cytoplasm of > 90 % of HEK cells.....	198
Figure 4.2. The synthetic CPPs are taken up by > 90 % cells in concentrations which were detected in both the cytoplasm and nucleus of HEK cells.	200
Figure 4.3. The recombinant SRSF1-CPP inhibits interaction of endogenous SRSF1 with NXF1 in human cells.....	201
Figure 4.4. The recombinant SRSF1-CPP specifically interacts with TAP/NXF1: p15 in a dose dependent manner, but not with NXF2 and NXF3 which are expressed in only mouse testis.	205
Figure 4.5. The recombinant SRSF1-CPP inhibits nuclear export of sense and antisense C9orf72-repeat transcripts.....	207
Figure 4.6. The recombinant and synthetic SRSF1-CPP inhibits expression of sense antisense DPRs in dose dependent manner.....	209
Figure 4.7. The recombinant and synthetic SRSF1-CPP rescue sense antisense DPRs associated toxicity in dose dependent manner.....	211
Figure 4.8. The synthetic SRSF1-CPP inhibits RAN-translation of DPRs and reduce motor deficits in vivo in <i>C9ORF72</i> -ALS/FTD <i>Drosophila</i>	214
Figure 4.9. The synthetic SRSF1-CPP inhibits RAN-translation of DPRs in vivo in <i>C9ORF72</i> -ALS/FTD mouse brains.	217
Figure 5.1. Stability assessment of SRSF1-CPP in HEK cell culture medium over a 12-day period.....	231
Figure 5.2. Plasmid maps of engineered scAAV SRSF1(89-120)-V5-TAT PTD and scAAV SRSF1-RRM2α1(132-144) V5-TAT PTD.....	234
Figure 5.3. Cloning of SRSF1-linker and RRM2α1 CPPs into scAAV _Efs1α GFP.....	236
Figure 5.4. Cloning of SRSF1-linker and RRM2α1-CPPs into scAAV_CBh-SRSF1-m4...	238
Figure 5.5. scAAV_H1 transfection facilitates expression of V5-tagged SRSF1-linker and SRSF1-RRM2α1 CPPs in HEK293T cells.	241

Figure 5.6. scAAV_H1 SRSF1-linker and scAAV_H1 SRSF1-RRM2 α 1 CPPs inhibits expression of sense and antisense DPRs.	244
Figure 5.7. scAAV_H1 SRSF1-linker and scAAV_H1 SRSF1-RRM2 α 1 CPPs rescue sense and antisense DPRs-associated toxicity.	246
Figure 5.8. PCR-based sexing analysis using sex 1 sry_Fw and sex 1 sry_Rev primers for gender identification of neonatal mice at P1 using DNA extracted from the tail.	247
Figure 5.9. PCR-based genotyping analysis using C9-GT-F, C9-GT-R, Vgll4-F, Vgll4-R primers for genotyping of neonatal mice at P1 using DNA extracted from tail.	248
Figure 5.10. Southern blot analysis for identification of <i>C9ORF72</i> repeat expansions length in C9-500 BAC mice.	250
Figure 5.11. Viral genome quantification development assay and validation in <i>C9ORF72</i> -ALS/FTD mouse brains.	252
Figure 5.12. Biodistribution analysis of C9-500 BAC mice injected with scAAV9_H1 SRSF1-linker and scAAV9_H1 SRSF1-RRM2 α 1 CPP.	254
Figure 5.13. <i>C9ORF72</i> -ALS/FTD mouse brain biodistribution of scAAV9_H1 SRSF1-linker and scAAV9_H1 SRSF1-RRM2 α 1 CPP one month post injection in neonatal mice.	258
Figure 5.14. Quantification of poly-GP DPRs in mouse brains at 1-month post-injection at P1 C9-500 BAC Mice.	260
Figure 6.1. Project outcomes and highlights: the promise of therapeutic in vivo non-viral and viral gene therapy approaches for the treatment of <i>C9ORF72</i> -ALS/FTD.	287

List of Tables

Table 1.1. Some of the key genes and functions implicated in pathophysiology of ALS.....	7
Table 1.2. Dysregulation of RNA export in neurodegenerative diseases.....	24
Table 1.3. Examples of various CPPs with their detailed information.....	49
Table 1.4. Pre-clinical studies of CPPs in animal models of various diseases.....	52
Table 1.5. Clinical studies of CPP based therapies in various diseases.	53
Table 2.1. Plasmid source and description.	57
Table 2.2. PCR primer sequence for molecular cloning.....	59
Table 2.3. Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) primers.	61
Table 2.4. Antibiotics used for bacterial cell culture.....	64
Table 2.5. Restriction and DNA modifying enzymes.	65
Table 2.6. Primary antibodies used in western Blotting and immunofluorescence microscopy.	66
Table 2.7. Secondary antibodies used in western blotting and immunofluorescence microscopy.	67
Table 2.8. Molecular biology assay kit details.	67
Table 2.9. The compositions of TfbI and TfbII buffers.....	69
Table 2.10. Restriction digest reaction for PCR product.....	73
Table 2.11. Inverse PCR reaction assembly.	74
Table 2.12. Herculase II DNA fusion polymerase PCR programme.	74
Table 2.13. Taq PCR reaction assembly.....	75
Table 2.14. Taq DNA polymerase PCR programme.....	75
Table 2.15. The loading, lysis, washing, and elution buffers used in affinity and ion exchange chromatography systems for purification of His ₆ -tagged SRSF1-linker CPP.	79
Table 2.16. The loading, lysis, washing, and elution buffers used in affinity and ion exchange chromatography systems for purification of GST-tagged SRSF1-linker and Ctrl CPPs.	80
Table 2.17. Transfection condition for HEK293T cell line.....	85
Table 2.18. Composition of SDS polyacrylamide gels.	86
Table 2.19. Composition of Coomassie blue staining solution.	87
Table 2.20. Composition of de-staining solution.	87
Table 2.21. Composition of Ponceau staining solution.	88
Table 2.22. The name of reagents, volumes and times for probe detection and imaging.	100

Table 3.1. Ion-exchange chromatography resins and their functional groups.....	113
Table 3.2. The conditions used in Strategy 1, Strategy 2, and Strategy 3.	141
Table 3.3. The composition of tested loading and elution buffers conditions.....	149

Abbreviations

aa	Amino acids
AAV	Adeno-Associated Virus
AC	Affinity chromatography
AD	Autosomal Dominant
ADs	Alzheimer Disease
ALS	Amyotrophic Lateral Sclerosis
ALS2	Alsin
ALYREF	Aly/REF export factor
AMD	Alpha-mannosidosis
AMO	Antisense morpholino oligonucleotide
AMPs	Antimicrobial peptides
ANOVA	Analysis of variance
ANG	Angiogenin
AR	Autosomal Recessive
ASO	Antisense Oligonucleotide
ATP	Adenosine Triphosphate
ATXN2	Ataxin 2
BAC	Bacterial artificial chromosome
BBB	Blood-Brain Barrier
BIND	Beta-structured inhibitor
Bcl-2	B-cell Lymphoma 2
bp	Base pair
BS	Brainstem
BSA	Bovine serum albumin
C ₄ G ₂	CCCCGG repeat expansions
C9orf72	Chromosome 9 open reading frame 72
C9orf72-L	C9orf72 long isoform
C9orf72-S	C9orf72 short isoform
C9-500	500 C9ORF72 Repeat Transcripts
CAG	Cytosine-adenine-guanine
Cb	Cerebellum

CBC	Cap binding complex
CBh	Chicken- β -actin hybrid
cDNA	Complementary DNA
CDP	Chemiluminescent substrate for alkaline phosphatase
Cas9	CRISPR-associated protein 9
CHMP2B	Charged multivesicular body protein 2B
CHCHD10	Coiled-coil-helix-coiled-coil-helix domain-containing 10
CIAP	Calf Intestinal Alkaline Phosphatase
CM	Carboxymethyl
CM	Cisterna Magna
CNS	Central Nervous System
CRISPR	Clustered regularly interspaced short palindromic repeats
CPPs	Cell-penetrating peptides
CSF	Cerebrospinal fluid
Cx	Cortex
Ct	Cycle threshold
DAPI	4', 6-diamidino-2-phenylindole
Dbp5	DEAD box RNA helicase
DCTN1	Dynactin subunit 1
DEAE	Diethylaminoethyl
DENN	Differentially expressed in normal and neoplastic cells
DEPC	Diethyl Pyrocarbonate
DFNB	Autosomal recessive
DIG	Digoxigenin
DNA	Deoxyribonucleic Acid
DM	Myotonic dystrophy
DM1	Myotonic dystrophy type 1
DM2	Myotonic dystrophy type 2
DMD	Duchenne muscular dystrophy
DMEM	Dulbecco's Modified Eagle's Medium
dNTPs	Deoxynucleotide Triphosphates
DTT	Dithiothreitol
DPRs	Dipeptide repeat proteins

DRG	Dorsal root ganglia
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediamine Tetra acetic Acid
EF1 α	Elongation Factor 1 alpha
eIF4A	Eukaryotic translation initiation factor-4A
eIF4B	Eukaryotic translation initiation factor-4B
eIF4H	Eukaryotic translation initiation factor-4H
ELISA	Enzyme-linked immunosorbent assay
ERBB4	Receptor tyrosine-protein kinase erbB 4
EVs	Extracellular vesicles
fALS	Familial ALS
FBS	Fetal Bovine Serum
FCS	Foetal Calf Serum
FDA	Food and Drug Administration
FG	Phenylalanine and glycine
FIG4	Polyphosphoinositide phosphatase
FPLC	Fast protein liquid chromatography
FT	Flow through
FTD	Frontotemporal Dementia
fFTD	Familial FTD
FUS	Fused in Sarcoma RNA-binding Protein
FVB/N	Friend Virus B NIH
FVB/NJ	Friend Virus B NIH Jackson
G4C2	GGGGCC repeat expansions
GA	Glycine-Alanine
GAA	Alpha-glucosidase
GB1	Immunoglobulin-binding domain B1 of streptococcal protein G
GEF	Guanine exchange factors
GLP-1	Glucagon-like peptide-1
GLT8D1	Glycosyltransferase 8 domain-containing protein 1
GFP	Green fluorescence protein
GP	Glycine-Arginine

GR	Glycine-Arginine
GSH	Glutathione-sepharose (GSH)
GST	Glutathione S-transferases
HA	Haemagglutinin
HD	Huntington Disease
HEK293T	Human embryonic kidney 293 cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic Acid
H1	Histone promoter 1
HIC	Hydrophobic interaction chromatography
HIV	Human Immunodeficiency Virus
HI	Hypoxic–ischaemic
hnRNP	Heterogeneous Nuclear Ribonucleoproteins
HNRNPA1	Heterogeneous Ribonucleoprotein A1
HPLC	High-Performance Liquid Chromatography
HRP	Horseradish peroxidase
His ₆	Hexa-histidine
hESC	Human embryonic stem cells
HS	Human serum
ICV	Intracerebroventricular
IMAC	Immobilized metal ion affinity chromatography
Inverso	D- amino acids
INP	Input
iPSCs	Induced pluripotent stem cells
IEX	Ion exchange chromatography
IP6	Inositol hexakisphosphate
ITRs	Inverted terminal repeats
JNK	c-Jun N-terminal kinase
Kb	Kilobases
kDa	Kilo Dalton
KIF5A	Kinesin heavy chain isoform 5A
LCA	Leber’s congenital amaurosis
LB	Luria-Bertani broth

LMN	Lower motor neuron
LP	Lumbar puncture
LV	Lentivirus
MATR3	Matrin 3
MCAO	Middle cerebral artery occlusion
microRNA	miRNA
mRNA	Messenger RNA molecules
mRNPs	Messenger ribonucleoproteins
MSC	Mesenchymal stem cells
MSD	Meso Scale Discovery
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Mono-S	Methyl sulfonate
mSOD1	Mutant Superoxide Dismutase 1
N2A	Neuro2a cells
NASH	Non-alcoholic steatohepatitis
NaCl	Sodium Chloride
NaOAc	Sodium Acetate
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NEA	Nuclear export receptors
NEK1	Serine/threonine-protein kinase Nek1
NEFH	Neurofilament heavy polypeptide
NF-H	Neurofilament heavy subunit
NF-kB	Nuclear factor kappa B
NF-L	Neurofilament light subunit
NF-M	Neurofilament medium subunit
NMN	Nicotinamide mononucleotide molecules
NSC	Foetal neural progenitors
NTg	Nontransgenic
NTF	Neurotrophic factor
NXF1	Nuclear RNA Export Factor 1
OCT	Optimal cutting temperature
P1	Post-natal Day 1
P34	Post-natal Day 34

PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PCI	Phenol/chloroform/isoamyl alcohol
PEI	Polyethylenimine
PFA	Paraformaldehyde
PIC	Protease Inhibitor Cocktail
PKR	RNA-dependent protein K
PMSF	Phenylmethanesulphonyl fluoride
PNK	Polynucleotide Kinase
polyQ	Polyglutamine
PMOs	Phosphorodiamidate morpholino oligomers
PrP	Prion protein
PRPH	Peripherin
PTD	Protein transduction domain
pI	Isoelectric point
Penstrep	Penicillin/Streptomycin
rAAV	Recombinant AAV
RAB	Proteins Ras-associated binding Protein
RAN Translation	Repeat associated non-AUG translation
REF	RNA export factor
RPE65	Retinal pigment epithelium-specific 65-kDa protein gene
RNA	Ribonucleic Acid
RNAse	Ribonuclease
RNAi	RNA interference
RNP	Ribonucleoprotein
ROS	Reactive oxygen species
RRM RNA	Recognition Motif
RT-qPCR	Quantitative reverse transcription PCR
RSC	Royal college of surgeons
r-hGH	Recombinant human growth hormone
ROCK	Potent rho kinase
Rt	Recombinant
sALS	Sporadic ALS
sFTD	Sporadic Frontotemporal Dementia

SAP	Sweet arrow Peptide
SB	Super Broth
SC	spinal cord
SCA8	Spinocerebellar Ataxia 8
SEC	Size exclusion chromatography
SETX	Senataxin
SIGMAR1	Sigma non-opioid intracellular receptor 1
siRNAs	Small interfering RNAs
SMA	Spinal muscular atrophy
SMN	Survival motor neuron protein
SMN1	Survival of Motor Neuron 1
SOD1	Cu/Zn super oxide dismutase 1
SP	Sulphopropyl
SPG11	Spatacsin
SRSF1	Serine and Arginine-Rich Splicing Factor 1
ssAAV	Single stranded AAV vectors
scAAV	Self-complementary AAV vectors
SQSTM1	Sequestosome 1
SS1b	Splice site 1b
SSC buffer	Saline-sodium citrate buffer
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SSTR2	Somatostatin receptor subtype
shRNA	Short Hairpin RNA
TAP	Tip associating protein
TAP/NXF1	Nuclear RNA export factor 1
TAT	Trans-activator of transcription
TB	Terrific Broth
TBK1	Serine/threonine-protein kinase TBK1
TBST	Tris-buffered saline, Tween 20
TEMED	Tetramethylethylenediamine
TDP-43	Transactive response DNA binding protein 43 kDa
TrxA	Thioredoxin
TREX	Transcription export complex
TUBA4A	Tubulin α 4A

TUJ1	Class III beta-tubulin
TMPyP4	5,10,15,20-tetra(N-methyl-4-pyridyl) porphyrin
tRNA	Transfer RNA
Tg	Transgenic
UBQLN2	Ubiquilin-2
UMN	Upper motor neuron
UTR	Untranslated region
UPS	Ubiquitin-proteasome system
V1	Variant 1
V2	Variant 2
V3	Variant 3
VCN	Vector copy number
VAPB	Vesicle-associated membrane protein-associated protein B/C
VCP	Valosin-containing protein
VEGF	Vascular endothelial growth factor
WT	Wild type
XD	X-Linked Dominant

1. Chapter 1 - Introduction

1.1. Background Information Relating to Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ALS) is an incurable neurodegenerative disease involving selective death of upper and lower motor neurons, progressive paralysis and death usually within 3-5 years from symptom onset (Brown and Al-Chalabi, 2017). The French neurologist Jean-Martin Charcot discovered ALS in 1869 (Wijesekera and Nigel Leigh, 2009). ALS usually begins with a focal onset of motor problems with contiguous spread to anatomically adjacent areas. The progressive loss of motor neurons causes progressive muscle weakness, atrophy and spasticity and failure of the neuromuscular system including muscles involved in limb, bulbar and respiratory function (Al-Chalabi et al., 2016a; Rowland and Shneider, 2001).

1.2. Clinical Presentation of ALS

ALS is a clinically heterogeneous disease (Hardiman et al., 2017). ALS symptoms mainly are categorized as upper motor neuron (UMN) and lower motor neuron dysfunctions (LMN). Approximately, one third of ALS patients present with weakness in the bulbar muscles causing difficulties with speech and swallowing often with associated emotional lability (Kiernan et al., 2011). Limb-onset ALS cases are responsible for 60-70% of ALS cases (Gordon, 2013). In addition, cognitive and behavioural impairments have been observed in up to 50% of ALS cases and 13% of patients experience behavioural changes in keeping with frontotemporal dementia (FTD) (Al-Chalabi et al., 2016b). There is a relationship between age and the disease because it is adult-onset disease. The average age of disease onset varies between 63 and 65 years in Europe. The disease age for North America and East Asia is approximately 59 years old. A younger age of disease onset can be observed, including some cases with juvenile onset (Shaw, 2005).

1.3. Epidemiology

The incidence rate for ALS cases is 2-3 per 100,000 people in Europe (Hardiman et al., 2017). In contrast, the incidence of ALS is 0.8 and 0.7 cases per 100,000 in East Asia and South Asia, though this may be partially due to under-reporting (Marin et al., 2014).

The incidence and prevalence of the disease as well as the symptomatic deficits are altered depending on geographical factors. For example, ALS has a relatively low incidence in non-European countries (Chiò et al., 2013). However, some regions in the Western Pacific and Japan have higher incidence rates such as 7.0 per 100,000 in Guam (Waring et al., 2004) and 9.5 per 100,000 in the Kii peninsula of Japan (Kihira et al., 2012). ALS has two subtypes recognised as a familial and sporadic. Sporadic ALS (sALS) represents approximately 90 percent of all ALS cases. The rest of the cases are inherited and described as familial ALS (fALS) (Talbot et al., 2016). The incidence of ALS varies between males and females. The ratio of male to female is approximately 1.5/1 (Manjaly et al., 2010).

1.4. Genetics of ALS

New emerging technologies and studies for gene mapping and DNA analysis have made it easier to identify genes which predispose to ALS (van Rheenen et al., 2016). Until now, mutations in approximately 30 genes associated with ALS have been identified (Al-Chalabi et al., 2016b) and key genes are listed in **Table 1.1**. Among these, mutations in 4 genes underlie the majority of fALS cases (Chiò et al., 2014).

1.4.1. *SOD1*

Mutations in *SOD1* (superoxide dismutase 1) were discovered as the first genetic cause of ALS in 1993 and this represents second commonest cause of fALS after *C9orf72*. *SOD1* mutations are responsible for 20% fALS and 1-2% sALS (Rosen et al., 1993). *SOD1* is a metalloenzyme encoding 153 amino acids, and its normal function is to catalyse the conversion of superoxide radicals into molecular oxygen and hydrogen peroxide (Rakhit and Chakrabarty, 2006). This reaction occurs through the oxidation of the copper atom at the active site of the *SOD1* enzyme. The *SOD1* enzyme contains a zinc atom that helps maintain its protein structure. *SOD1* is widely distributed in the CNS and makes up approximately 1% of brain protein. Indeed, it is present in all other body tissues (Okado-Matsumoto and Fridovich, 2001; Shaw, 2005). So far, more than 185 disease related mutations in *SOD1* across the five exons have been discovered. These are mainly missense mutations with some additional small gene deletions, or insertions (Yamashita and Ando, 2015). The pathogenesis was initially considered to occur through a loss-of-function (LOF) mechanism due to a reduction in the *SOD1* enzyme activity by 50 to 89% (Rosen et al., 1993). However, it was later found that there is no correlation between severity of disease phenotypes and enzyme activity.

This suggests that a toxic gain-of-function (GOF) may be responsible for motor neuron injury. Evidence for the toxic GOF mechanism in ALS was supported by the results of the study that investigated neuropathological and behavioural features of a mouse model lacking the SOD1 enzyme. They reported that these mice did not develop the ALS specific symptoms (Reaume et al., 1996). There are several proposed pathophysiological mechanisms including RNA dysfunction, mitochondrial damage, oxidative stress, excitotoxicity and disruptions of axonal and cellular transport that conferred toxicity and motor neuron injury and cell death due to mutations in the *SOD1* gene (Hayashi et al., 2016; Shaw, 2005).

1.4.2. *TARDBP*

In 2008, *TARDBP*, which encodes transactive response DNA binding protein of 43 kDa (TDP-43), was identified as a cause of ALS (Schmid et al., 2013). *TARDBP* mutations account for 4% fALS and 1% sALS cases (Schmid et al., 2013; Sreedharan et al., 2008). So far more than 48 mutations in the *TARDBP* gene have been identified which are mainly located in exon 6 within the C terminal glycine rich domain. TDP-43 binds to RNA / DNA and interacts with other ribonuclear proteins. TDP-43 can act as a controller of gene expression and is involved in RNA metabolism (Ratti and Buratti, 2016). Nuclear clearance of TDP-43 and the formation of cytoplasmic TDP-43 inclusions across motor cortex, brainstem, spinal cord and motor neurons are defined as the pathological hallmark of ALS (Neumann et al., 2006). However, it has not been fully identified that this cytoplasmic accumulation is causing either a loss-of-function or toxic gain-of-function of TDP-43. Several *in vitro* and *in vivo* models have been generated to test these hypotheses and it was reported that neurodegenerative phenotypes, DNA damage and change in the cellular transcriptome were observed involving both mechanisms across several these models (Mejzini et al., 2019).

1.4.3. *FUS*

Similar to TDP-43, *FUS* (Fused in Sarcoma) encodes an RNA/DNA binding FET family protein (Kwiatkowski et al., 2009). *FUS* has been linked to various cellular processes such as transcription, pre-mRNA splicing, RNA transport, regulation of translation, and DNA repair mechanisms (Ratti and Buratti, 2016; Wang et al., 2013). *FUS* is primarily located in the nucleus of cells, but in patients with ALS who have *FUS* mutations, mislocalization to the cytoplasm and the formation of *FUS*-positive inclusions have been observed.

More than 50 mutations in *FUS* have been reported and these mutations are responsible for 4% fALS and 1% sALS cases (Kwiatkowski et al., 2009; Vance et al., 2009). Similar to TDP-43, it is still not fully identified how protein inclusions may contribute to neurodegeneration. Two mechanisms have been proposed to explain the toxicity of *FUS* mutations. Loss-of-function (LOF) mutations were reported to promote cytoplasmic aggregation and nuclear loss of *FUS*. However, LOF mutations within *FUS* are unlikely to cause ALS as knockdown of *FUS* in mice does not lead to neuronal loss (Kino et al., 2015). Nevertheless, similar research using fruit flies showed that removing *FUS* led to the degeneration of neurons and disruptions in movement (Sharma et al., 2016). There is considerable evidence that *FUS*-ALS is caused by a gain of function (GOF) rather than LOF. For example, it was found that overexpression of normal human *FUS* in mice caused significant motor neuron degeneration and accumulation of *FUS* in the cytoplasm (Mitchell et al., 2013).

1.4.4. *C9ORF72*

Among the genes involved in ALS, hexanucleotide (GGGGCC) repeat expansions in intron 1 of the *C9orf72* (chromosome 9 open reading frame 72) gene were identified as the most common genetic cause of ALS and frontotemporal dementia in 2011, accounting for 40-50% fALS and 7-10% sALS and 30% familial frontotemporal dementia (fFTD) and 7% sporadic FTD (sFTD) cases (DeJesus-Hernandez et al., 2011; Majounie et al., 2012; Renton et al., 2011). Most healthy individuals have between 2 to 24 GGGGCC repeat expansions in the *C9orf72* gene. Currently, the minimum number of repeat expansions that cause disease is considered to be ≥ 30 (Simón-Sánchez et al., 2012). It also corresponds to the length that leads to RNA foci and dipeptide repeat protein (DPR) production (Gijssels et al., 2016; Hautbergue et al., 2017).

For many years, it was assumed that ALS and FTD represented a single spectrum of diseases, but with the discovery of *C9orf72* it became apparent that ALS and FTD share commonalities where for example ~ 15-20% of patients with ALS also have certain cognitive and behavioural problems similar in keeping with FTD. The precise pathophysiological mechanisms to explain how the repeat expansions cause ALS/FTD have not yet fully defined, but three non-mutually exclusive mechanisms of neurotoxicity have been proposed: **(i)** haploinsufficiency (see **section 1.6.1**); **(ii)** the formation RNA foci with sequestration of RNA binding proteins and **(iii)** dipeptide repeat mediated toxicity (see **section 1.6.2**).

The *C9orf72* gene encodes three main variants (variant 1 (V1), variant 2 (V2) and variant 3 (V3)) which are detected in most human tissues, including all brain regions and the spinal cord, with high expression in myeloid cells such as CD14⁺ monocytes and eosinophils and lower expression in lymphoid cells and other tissues (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Rizzu et al., 2016). Variant 1 encodes a short C9orf72 protein which has 222 amino acids whereas variants 2 and 3 encode the long C9orf72 protein which has 481-amino acids (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Examination of the distribution of isoforms in human brain tissue has shown that the long isoform is dispersed in the cytoplasm while the short isoform is positioned at the nuclear membrane. In addition, it was found that the C9orf72 protein is highly expressed in the brain, spinal cord, and immune system compared to other organs such as the lung, heart, liver, kidney, and skeletal muscle. This expression pattern also aligns well with the transcript expression profile (Laflamme et al., 2019; Smeyers et al., 2021; Xiao et al., 2015). The cellular function of C9orf72 protein was unknown when it was discovered in 2011; however, studies have since revealed several roles including interacting with a protein that takes part in nuclear import, the formation and degradation of stress granules, regulation of autophagy as well as axonal growth in primary mouse embryonic motor neurons (Chitiprolu et al., 2018; Sivadasan et al., 2016; Webster et al., 2016; Xiao et al., 2015).

Table 1.1. Some of the key genes and functions implicated in pathophysiology of ALS. The table is adapted from (Hardiman et al., 2017).

Locus	Gene (Protein)	Inheritance	Disease Mechanism	Reference
21q22.1	<i>SOD1</i>	AD or AR	Oxidative stress	(Rosen et al., 1993)
14q11	<i>ANG</i> (angiogenin)	AD	RNA metabolism	(Greenway et al., 2006)
q36	<i>TARDBP</i>	AD	RNA metabolism	(Kabashi et al., 2008)
16p11.2	<i>FUS</i>	AD or AR	RNA metabolism	(Kwiatkowski et al., 2009)
9p13.3	<i>VCP</i> (valosin-containing protein)	AD	Autophagy	(Johnson et al., 2010)
10p15-p14	<i>OPTN</i> (optineurin)	AD or AR	Autophagy	(Maruyama et al., 2010)
9p21-22	<i>C9orf72</i> (guanine nucleotide exchange <i>C9orf72</i>)	AD	RNA metabolism and Autophagy	(DeJesus-Hernandez et al., 2011; Renton et al., 2011)
Xp11.23-Xp13.1	<i>UBQLN2</i> (ubiquilin-2)	XD	Ubiquitin proteasome (UPS) and Autophagy	(Deng et al., 2011)
5q35	<i>SQSTM1</i> (sequestosome 1)	AD	Autophagy	(Fecto et al., 2011)
17p13.2	<i>PFN1</i> (<i>profilin 1</i>)	AD	Cytoskeleton	(Wu et al., 2012)
12q13.1	<i>HNRNPA1</i> (HnRNP A1)	AD	RNA metabolism	(Kim et al., 2013)
5q31.2	<i>MATR3</i> (matrin 3)	AD	RNA metabolism	(Johnson et al., 2014)
4q33	<i>NEK1</i> (Serine/threonine-protein kinase Nek1)	AD	DNA-damage response	(Kenna et al., 2016)
12q13.3	<i>KIF5A</i> (Kinesin heavy chain isoform 5A)	AD	Cytoskeletal defects	(Nicolas et al., 2018)
2p13.1	<i>DCTN1</i> (Dynactin subunit 1)	AD or AR	Unknown	(Münch et al., 2004)
22q12.2	<i>NEFH</i> (Neurofilament heavy polypeptide)	AD or AR	Neurofilamentous Swellings	(Figlewicz et al., 1994)

AD: autosomal dominant ****AR:** Autosomal Recessive, ****XD:** X-Linked Dominant

2q33.1	<i>ALS2</i> (alsin)	AR	Endosomal trafficking	(Hadano et al., 2001; Yang et al., 2001)
18q21	Unknown	AD	Unknown	(Hand et al., 2002)
9q34.13	<i>SETX</i> (senataxin)	AD	RNA metabolism	(Chen et al., 2004)
15q21.1	<i>SPG11</i> (Spatacsin)	AR	DNA damage repair and axon growth	(Hentati et al., 1994)
20p13	Unknown	AD	Unknown	(Sapp et al., 2003)
20q13.32	<i>VAPB</i> (vesicle-associated membrane protein-associated protein B/C)	AD	Endoplasmic reticulum stress	(Nishimura et al., 2004)
6q21	<i>FIG4</i> (polyphosphoinositide phosphatase)	AD	Endosomal trafficking	(Chow et al., 2009)
12q24.12	<i>ATXN2</i> (ataxin 2)	AD	RNA metabolism	(Elden et al., 2010)
9p13.3	<i>SIGMAR1</i> (sigma non-opioid intracellular receptor 1)	AD	UPS and autophagy	(Al-Saif et al., 2011; Luty et al., 2010)
3p11.2	<i>CHMP2B</i> (charged multivesicular body protein 2B)	AD	Endosomal trafficking	(Parkinson et al., 2006)
2q34	<i>ERBB4</i> (receptor tyrosine-protein kinase erbB 4)	AD	Neuronal development	(Takahashi et al., 2013)
2q35	<i>TUBA4A</i> (tubulin α 4A)	AD	Cytoskeleton	(Sama et al., 2014)
5q35.3	<i>CHCHD10</i> (coiled-coil-helix-coiled-coil-helix domain-containing 10)	AD	Mitochondrial maintenance	(Bannwarth et al., 2014)
12q14.2	<i>TBK1</i> (serine/threonine-protein kinase TBK1)	AD	Autophagy	(Matsumoto et al., 2015; Weidberg and Elazar, 2011)
3p21.1	<i>GLT8D1</i> (Glycosyltransferase 8 domain-containing protein 1)	AD	Glycosyltransferase activity	(Cooper-Knock et al., 2019)
12q13.12	<i>PRPH</i> (Peripherin)	AD or AR	Neurofilament disorganization	(Gros-Louis et al., 2004)

1.5. The Pathophysiological Mechanisms of ALS

The pathophysiological mechanisms of ALS are not fully understood since they are very complex. The proposed pathological mechanisms which are not mutually exclusive include; glutamate excitotoxicity, structural and functional dysfunctions of mitochondria, impaired axonal structure and transport, oxidative stress, RNA metabolism defects, protein and neurofilament aggregation, impaired protein clearance and autophagy, inflammatory dysfunction and altered astrocyte-motor neuron cross-talk (Ferraiuolo et al., 2011).

1.5.1. RNA Metabolism Defects in ALS

The widespread alteration of RNA processing (**Figure 1.1**) is now recognized as one of the main mechanisms which contributes to ALS pathogenesis because many of the genes which have been implicated in familial ALS have a role in RNA metabolism. Studying RNA metabolism involves not only focusing on one mechanism. Since disease pathways are encoded by proteins and proteins are encoded by RNA, RNA metabolism is involved in all pathways which are altered in the disease. Our therapeutic target is a protein involved in gene expression. Changing the gene expression of SRSF1 has been shown to prevent DPR production (Hautbergue et al., 2017). The details of some major contributory mutations in genes involved in RNA processing are given below (Butti and Patten, 2019).

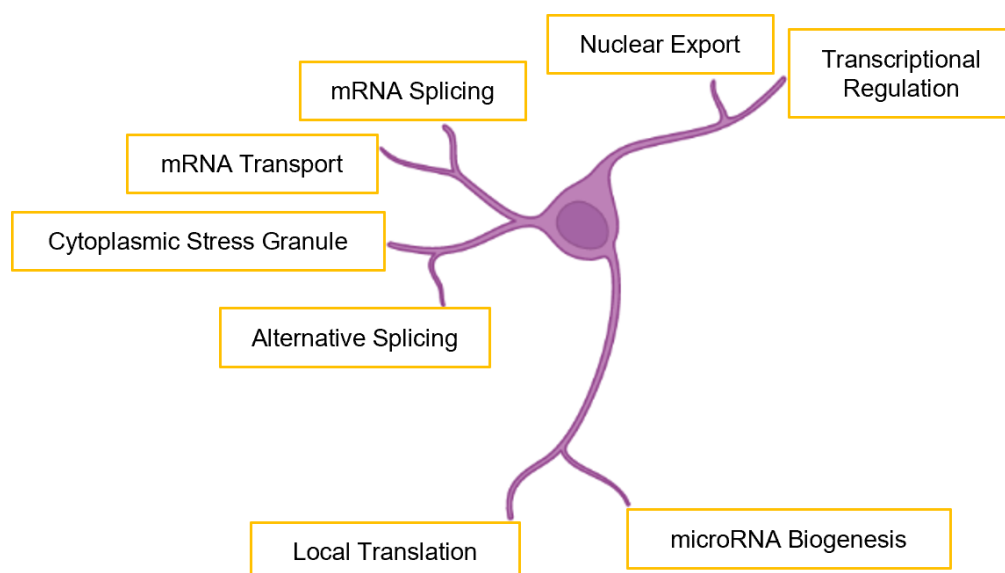


Figure 1.1. Widespread dysregulation of components of RNA metabolism in ALS (Ferraiuolo et al., 2011).

There is a widespread alteration of RNA metabolism in ALS, but it is very difficult to understand what are the upstream changes that cause disease versus changes secondary to the primary insult. For example, there are widespread transcriptome alterations in *C9ORF72*-ALS. It has been reported that one third of the transcriptome is altered in ALS, but due to this very large number of changes, it is not possible to precisely identify the alterations that cause disease from the changes that are consequences of the neurodegenerative process (Prudencio et al., 2015). Mutations in *TARDBP* involve dysregulation of many RNA processing functions such as transcriptional regulation, alternative splicing, microRNA (miRNA) biogenesis, axonal transport of mRNAs, mRNA splicing and RNA transport (Mackenzie and Rademakers, 2008; Neumann et al., 2006). Post-mortem studies of the CNS show that there is nuclear clearance of TDP-43 and cytoplasmic mislocalisation/aggregation of TDP-43 within affected neurons which is known as TDP-43 proteinopathy. Interestingly, the proteinopathy occurs in 97% of ALS cases (both in familial and sporadic cases) except in *SOD1* and *FUS* linked ALS cases (Mackenzie et al., 2007). TDP-43 aggregation is the hallmark of ALS and FTD (Prasad et al., 2019). The mechanisms that induce TDP-43 proteinopathy are still unknown and it is also remains controversial whether this generates loss or gain of toxic functions in ALS (or both) (French et al., 2019).

Mutations in the *FUS* gene alter the transportin-mediated nuclear localization of FUS and cause the occurrence of the formation of FUS-containing cytoplasmic stress granules. The FUS protein is also associated with RNA and DNA processing such as mRNA transport and mRNA alternative splicing (Dormann et al., 2010) and these functions may be impaired in the presence of *FUS* mutations. Dysregulation of gene expression was observed in the presence of mutations in *C9orf72*. The production of dipeptide repeat proteins (DPRs) interferes with multiple biological processes and steps in RNA metabolism including altered nucleolar processing, splicing and mRNA transport. The GGGGCC (G₄C₂) repeat expansion mutation in *C9orf72* forms G-quadruplex secondary structures which are likely to affect RNA processing (Haeusler et al., 2014). Also, this expansion results in the production of nuclear and cytoplasmic RNA foci (Lee et al., 2013) and may cause RNA toxicity arising from sequestration of RNA binding proteins and DPRs (Cooper-Knock et al., 2014b). Furthermore, impairment of DNA repair and gene specific instability arise from repeat expansions in *C9orf72* through the formation of R-loops. (Reddy et al., 2014; Walker et al., 2017).

In addition to *TARDBP*, *FUS* and *C9orf72*, there are other genes associated with ALS which are involved in RNA metabolism such as *angiogenin (ANG)*, *senataxin (SETX)*, *TAF15*, *EWSR1*, *ELP3*, *HNRNPA1*, *HNRNPA2B1* and *MATR3* (Hardiman et al., 2017). The presence of multiple mutations in genes involved in RNA metabolism highlights that dysfunction of RNA metabolism is an important mechanism in the pathophysiology of ALS.

1.6. C9ORF72-ALS/FTD

We focused on *C9orf72* mediated ALS/FTD. The rationale for this was firstly that the hexanucleotide expansion in *C9orf72* has been identified as the most common cause of ALS and FTD in 2011 (DeJesus-Hernandez et al., 2011; Renton et al., 2011). This repeat transcript in intron 1 and this abnormal intron containing mRNA is exported into the cytoplasm which is an abnormal process and therefore we are interested in investigating the abnormal nuclear export of the pathological repeat transcripts. The size of the repeat expansion varies among ALS patients, with the larger expansions ranging from 200 to 5000 repeats (Cooper-Knock et al., 2014a). *C9ORF72* mediated ALS is different from other ALS subtypes because it involves a microsatellite repeat expansion and repeat associated non-ATG (RAN) translation. RAN translation is a non-conventional form of translation that occurs in all 3 frames in the absence of canonical ATG codons. It involves microsatellite repeat expansions and secondary structures such as DNA–RNA heteroduplexes (Zamiri et al., 2018), RNA duplexes (Vatovec et al., 2014) and I- motifs (Kovanda et al., 2015). It is not precisely understood how the repeat expansions cause ALS/FTD but three non-mutually exclusive mechanisms of neurotoxicity have been proposed (**Figure 1.2**): **(A)** haploinsufficiency; **(B)** sequestration of RNA-binding proteins by RNA foci and **(C)** RAN translation resulting in the production of dipeptide repeat proteins (DPRs) (Balendra and Isaacs, 2018).

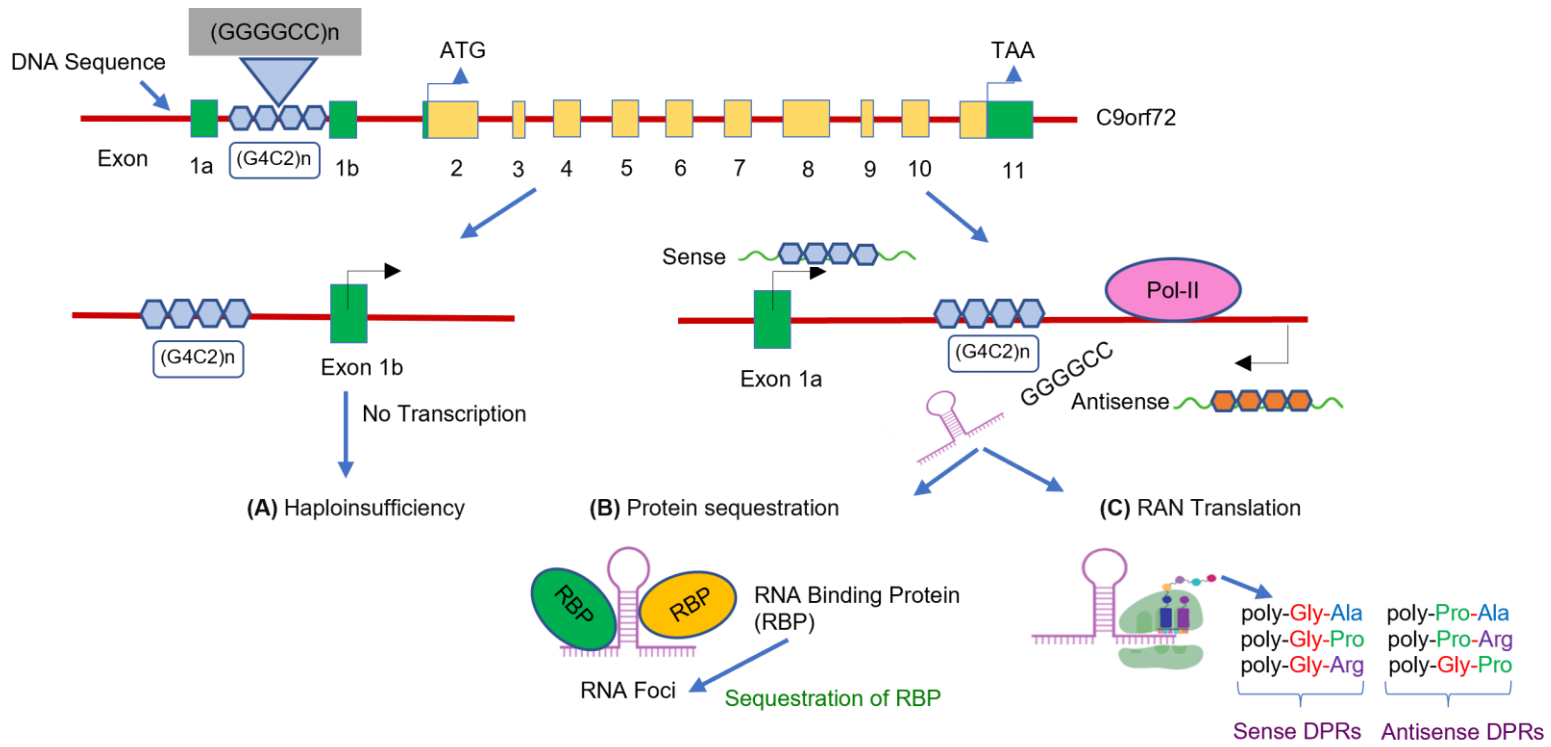


Figure 1.2. Diagram representing the proposed C9orf72 hexanucleotide-repeat-mediated mechanisms of neurotoxicity (Walsh et al., 2015).

1.7. C9ORF72-Mediated ALS Mechanisms

1.7.1. Loss of C9ORF72 Function Through Haploinsufficiency

In the presence of a pathological GGGGCC intron 1 expansion *C9orf72* has reduced expression levels of three alternatively spliced mRNA variants (variants 1, 2 and 3). This leads to a decrease in the production of the C9orf72 protein and potential haploinsufficiency. The hexanucleotide repeat expansion is found in intron 1 of variant 1 and 3 or located in the promoter region of variant 2 (Donnelly et al., 2013). The C9orf72 protein is structurally related to the DENN (differentially expressed in normal and neoplastic cells) family of proteins which act as guanine exchange factors (GEF) to activate RAB proteins (Levine et al., 2013). The function of RAB proteins is to regulate membrane trafficking. Indeed, the C9orf72 protein has also recently been shown to play roles in activation of autophagy, nucleocytoplasmic trafficking and endosomal membrane trafficking (Farg et al., 2014; Sellier et al., 2016; Webster et al., 2016).

The underlying loss of function might arise from several factors, due to secondary structures formed by repeat expansions in the DNA and RNA, impaired polymerase activity and abortive transcription (Gendron and Petrucelli, 2018). Also, epigenetic changes have been reported to affect the expression of *C9orf72* mRNAs and hypermethylation of histones in the promoter is one of the epigenetic processes that inhibits expression of the *C9orf72* transcripts (Belzil et al., 2013; Haeusler et al., 2014). In addition, it was found in *C9ROF72* patients that RNA foci and DPR production are decreased by hypermethylation (Bauer, 2016). A loss of C9orf72 function mechanism was corroborated in *C. elegans* and zebrafish models. Knockdown of the *C. elegans* *C9orf72* ortholog resulted in age-related motility deficits causing paralysis. Also axonal and behavioural deficits and morphological changes in motor neurons were observed following knockdown of *C9orf72* in zebrafish (Ciura et al., 2013). However, mouse studies dedicated to downregulation of *C9orf72* did not cause any neuronal deficits (Wen et al., 2014). Some studies in mouse models demonstrated that knockout of *C9orf72* led to a reduction in lifespan, and weight loss but no studies showed neural loss or behavioural deficits (Lagier-Tourenne et al., 2013). Also, knock out of *C9orf72* led to immunological problems, liposomal accumulation, lymphadenopathy and splenomegaly (O'Rourke et al., 2016). Overall, these findings clearly suggest that loss of C9orf72 function is likely to be a contributor to ALS progression and pathogenesis but is likely not to be the main causative factor of the disease.

1.7.2. *C9ORF72* Toxic Gain of Function Mechanisms

1.7.2.1. Potential Gain of Toxic Function From Repeat-Containing RNAs

The second proposed mechanism is that repeat expansion of *C9orf72* causes ALS/FTD via gain of an RNA toxic mechanism which arises from sequestration of RNA binding proteins (RBPs) by expanded repeat RNA transcripts causing potential loss of functions of the sequestered proteins (Gendron et al., 2013) and a resulting toxic gain-of-function (Hautbergue et al., 2017). Sense (G₄C₂)_n and antisense (C₄G₂)_n repeats are transcribed bidirectionally to produce sense and antisense RNA foci (DeJesus-Hernandez et al., 2011). The formation of RNA foci is a common feature of *C9ORF72*-ALS and other repeat expansion-based diseases. RNA foci are aggregates of RNA (Wojciechowska and Krzyzosiak, 2011). So far, there is limited experimental evidence showing that the RNA foci cause cellular dysfunction and toxicity, although RNA foci are highlighted in some reports to be toxic to cells (Balendra and Isaacs, 2018). In fact, it has been shown that increased levels of RNA foci do cause increased neurotoxicity (Tran et al., 2015). Several papers have reported the use of plasmids with or without a stop codon. When the plasmid containing the stop codon was employed, it didn't produce the DPR, rendering it non-toxic, even with an increased presence of RNA foci in *Drosophila*. Conversely, without the stop codon, the resulting DPR product was highly toxic (Mizielinska et al., 2014). Additionally, in 2017, we found that artificially augmenting RNA foci by retaining transcripts in the nucleus did not induce toxicity in either *Drosophila* or patient-derived neurons (Hautbergue et al., 2017).

RNA foci have been reported to sequester RNA binding proteins (RBPs) which regulate mRNA processing, splicing, translational regulation and RNA transport and degradation. Several RNA-binding proteins, such as ADARB2, ALYREF, hnRNP H, hnRNP A1, nucleolin, Pur- α , SRSF1, specifically bind to the *C9orf72* repeat expansion containing RNA (Babi Leko et al., 2019). HnRNPs are a family of RNA binding proteins. These proteins may cause either a toxic gain or loss of function because they are sequestered (Mori et al., 2013). Among them, SRSF1 (serine/arginine- rich splicing factor 1) underpins a gain function through abnormal nuclear mRNA export which leads to RAN translation of DPR proteins (Hautbergue et al., 2017).

Both sense and antisense RNA foci are mainly found in the nucleus, but also cytoplasmic RNA foci are observed in mitotic cells (DeJesus-Hernandez et al., 2011). (G₄C₂)_n and (C₄G₂)_n repeats have been detected mainly in neurons of the central nervous system (CNS) but they also exist in astrocytes in the spinal cord and cerebellum and in oligodendrocytes and microglia in motor cortex, frontal cortex and hippocampus (Mizielinska et al., 2013). Overall, the current evidence suggests that RNA foci are likely not toxic and that RAN translation of DPRs (below) is the main mechanism causing neurodegeneration in cell and animal models of *C9ORF72*-ALS/FTD.

1.7.2.2. Gain-of-Function Through RAN Translation of Dipeptide Repeat Proteins

The third pathological mechanism of *C9orf72* repeat-mediated neurodegeneration implicates repeat associated non-AUG (RAN) translation of dipeptide repeat proteins with neurotoxic aggregating properties. RAN translation is an unconventional form of translation that occurs in microsatellite repeat expansion disorders. It has been discovered in other repeat expansion diseases such SCA8 (Spinocerebellar Ataxia 8) and DM1 (myotonic dystrophy type 1) (Zu et al., 2011) and more recently in DM2 (myotonic dystrophy type 2) (Zu et al., 2017). RAN translation can occur in different conditions, for example it can require poly (A) tails or cis- regulatory elements such as promoters and repressors (Bañez-Coronel et al., 2015) and ALS papers that show cap dependent/independent RAN translation can occur (Green et al., 2017; Kearse et al., 2016). RAN translation occurs in multiple frames to generate various polymeric protein products from expanded repeat transcripts. Similar to the formation of RNA foci, both sense and antisense *C9orf72* repeat transcripts are RAN translated to produce five neurotoxic dipeptide-repeat proteins (DPRs) (Gendron et al., 2013; Mann et al., 2013; Zu et al., 2013).

Poly(GP), poly(GA) and poly(GR) are synthesized by the sense (G₄C₂)_n RNA and poly(PG), poly(PR) and poly(PA) are derived from the antisense (C₄G₂)_n RNA which localize in cerebellum, thalamus, hippocampus, frontal cortex, spinal cord and motor cortex (Gendron et al., 2013; Mori et al., 2013). Among those DPRs, inclusion of poly(GP) is the most common form of DPR in *C9ORF72*-ALS/FTD neurons due to translation of both sense and antisense expanded RNA transcripts (Gendron et al., 2017; Zu et al., 2013). Poly-Arginine DPRs have been reported as the most toxic (Hao et al., 2019; Mizielinska et al., 2014).

The DPR production has been proposed to involve both RAN translation through ribosomal scanning for a near-cognate CUG start codon and potentially ribosomal frameshifting (Tabet et al., 2018) although the later has not yet been clearly demonstrated. RAN translation of sense (G₄C₂)_n RNA repeats use a near-cognate CUG start codon inserted in Kozak consensus sequence and the poly-(GA). DPR production is also sensitive to the inhibition or depletion of general translation factors including eIF4A (Green et al., 2017) eIF4B and eIF4H (Goodman et al., 2019). Among these factors, eIF4A is capable of unwinding (G₄C₂)_n repeat RNA which is important for translation. Depletion of these canonical translation factors alleviates G₄C₂ based toxicity. In addition to reducing in G₄C₂ toxicity, depletion of eIF4B and eIF4H1 also reduced poly (GR)-GFP levels (Goodman et al., 2019).

Several groups have demonstrated different types of DPR toxicity. For example, incubation of either poly (GR)₂₀ or poly (PR)₂₀ with U2OS cells cause a decrease in cell viability (Kwon et al., 2014). Another study demonstrated that treatment of human astrocytes with the same form of DPRs results in morphological changes (Xu et al., 2013). The literature also shows that poly (GA), poly (GP) and poly (PA) are involved in inducing stress granule formation. Among them, poly (GA) has less toxic effects compared to arginine-rich DPRs, but expression of poly (GA) in neurons, cell models and *Drosophila* neurons causes mild toxicity. Furthermore, in vivo *Drosophila* studies showed that poly(GR)_{36 or 100} and poly(PR)_{36 or 100} led to eye degeneration and reduced survival, but poly(GA)_{36 or 100} and poly(PA)_{36 or 100} did not cause any adverse effects under the same conditions (Mizielinska et al., 2014).

DPRs also play a role in nucleolar disruption, DNA damage, and impairment of ubiquitin proteasome function (Freibaum and Taylor, 2017). In conclusion, a growing body of evidence implicates the production of DPRs through RAN-translation as one of the main components of neuronal injury at least in cell and animal models of C9ORF72-ALS/FTD (Moens et al., 2017).

1.7.3. SRSF1-Dependent Nuclear Export of Pathological C9ORF72-Repeat Transcripts

1.7.3.1. Physiological mRNA Nuclear Export Pathway

Eukaryotic cells have a nucleus separated from the cytoplasm with a nuclear membrane. The nuclei harbour the chromosomes which carry the genetic information in the form of DNA-protein complexes known as chromatin (Jackson, 2011). Regulation and control of many vital cellular process such as DNA replication and transcription of genes takes place in the nucleus, but translation of messenger RNA molecules (mRNA), which is another step of gene expression, occurs in ribosome organelles to build protein. Since ribosomes are present in the cytoplasm, mature mRNA molecules need to be exported from the nucleus to the cytoplasm to be translated into proteins (Katahira, 2015). The nuclear export of mRNA consists of 3 coupled steps mainly involving the co-transcriptional and splicing formation of messenger ribonucleoproteins (mRNPs), the translocation of mRNPs through the nuclear pores and the release of mRNPs into the cytoplasm (Köhler and Hurt, 2007). First, mRNPs are generated by adding nuclear export proteins onto pre-mRNAs through recruitment of the human TREX complex to mRNA during splicing. (Masuda et al., 2005; Walsh et al., 2010).

Gene expression relies on the co-transcriptional formation of mature mRNPs which involves key processing steps to allow for nuclear export-competent mRNPs to be transported from the nucleus to the cytoplasm (Carmody and Wente, 2009). First modification 5'capping which is adding of a 7-methylguanosine cap to the 5'end of newly synthesized pre-mRNA is required. This modification prevents degradation of pre-mRNA and also is needed for splicing to take place (Shatkin and Manley, 2000). The cap binding complex (CBC) binds to 5'capping which provides a binding platform to nuclear export proteins (Cheng et al., 2006). The next step is splicing of pre-mRNA transcripts. Splicing is basically the removing of introns from transcripts to allow the release of mRNA from the nucleus (Legrain and Rosbash, 1989). During the splicing process, two important classes of nuclear export adaptor proteins are recruited to the spliced mRNAs. These are either part of the TREX (Transcription-Export complex is an evolutionarily conserved multiprotein complex that plays a major role in the functional coupling of different steps during mRNA biogenesis, including mRNA transcription, processing, decay, and nuclear export) complex such as ALYREF (Masuda et al., 2005) or are serine and arginine-rich splicing factors (SRSF1, 3, 7) (Huang and Steitz, 2005).

The TREX complex harbours the RNA export adaptor (ALYREF) for binding nuclear export receptor NXF1, RNA helicase (UAP56 or DDX39B) which is component of the spliceosome and THO (THO is a multi-protein complex that promotes coupling between transcription and mRNA processing), which is composed of many proteins and is responsible for transcriptional elongation and RNA export (Masuda et al., 2005; Strässer et al., 2002). In humans, the nuclear export of mRNAs is coupled to splicing (Masuda et al., 2005). Other mRNA nuclear adaptor proteins are part of the SRSF protein family and involve SRSF1, SRSF3 and SRSF7. The SRSF proteins are recruited to the nascent RNAs for alternative splicing in the phosphorylated forms (Huang et al., 2004), thereby working as a component of the spliceosome (Lareau et al., 2007). However, when these proteins work as an mRNA nuclear adaptor protein, they undergo dephosphorylation to interact with TAP/NXF1 (Nuclear RNA export factor 1) which in turn interacts with nucleoporins composing the nuclear pore complexes. The next step is cleavage/polyadenylation which is coupled to the nuclear export of mRNAs in yeast (where the vast majority of gene are intronless) but not in humans (where the vast majority of genes are spliced). The polyadenylation process is required for releasing the mRNA from the gene template. Once mRNPs are released from the gene template, they start moving from the gene site to the nuclear pore. In 2004, it was reported that mRNAs are moving randomly which is called also the 'zigzag' model (Shav-Tal et al., 2004). When they reach nuclear pore, the nuclear export receptor NXF1 in complex with p15 binds respectively to TREX or the SRSF proteins to license transport of mRNPs through the nuclear pore to the cytoplasm.

The interaction of ALYREF or SRSF1, 3, 7 adaptors with NXF1 triggers a remodelling of NXF1 in a high RNA-binding affinity mode which triggers handover of the mature mRNA molecules from the adaptor proteins to NXF1, thereby inducing the nuclear export process (Hautbergue et al., 2008; Viphakone et al., 2012). The next stage is passing of mRNAs through the nuclear pore complex (NPC). The NXF1 proteins are responsible for motion of mRNAs to the inside of the nuclear pore through transient hydrophobic interactions with the FG (phenylalanine and glycine) repeats of nucleoporins, which constitute one third of the total NPC nucleoporins. These FG repeats form a gel-like structure in the centre of the NPC which blocks passing molecules larger than 30 kDa and provides a binding platform (Alber et al., 2007; Cole and Scarcelli, 2006). In this step, the importance of nuclear export receptor proteins is clearly understood because without their specific interaction with the FG repeat, it will be difficult move through the interior of the NPC.

The final step is the release of the mRNPs into the cytoplasm. This process is performed by nuclear export factors; Gle1, Dbp5 (DEAD box RNA helicase) and cofactor; IP₆ (inositol hexakisphosphate) (Alcázar-Román et al., 2006). To exit successfully from the nuclear pore, the mRNP structure needs to be disassembled by removing the NXF1 protein, otherwise mRNPs will fall back into nucleus during translocation because NXF1 onto mRNPs is still active to interact with nucleoporins (Stewart, 2007). This interaction is prevented by removing NXF1 with the help of Dbp5 which is activated by Gle1 and along with IP₆ (Tran et al., 2007). In this way, stripped mRNPs cannot interact with the NPC anymore. It is important to note that activation of Dbp5 triggers disassembling of other bound proteins onto the mRNP complex into the cytoplasm. Some of these proteins removed by Dbp5, such as adaptor proteins, move back into the nucleus (Lund and Guthrie, 2005). A diagram of nuclear export dependent bulk mRNA is given in **Figure 1.3**.

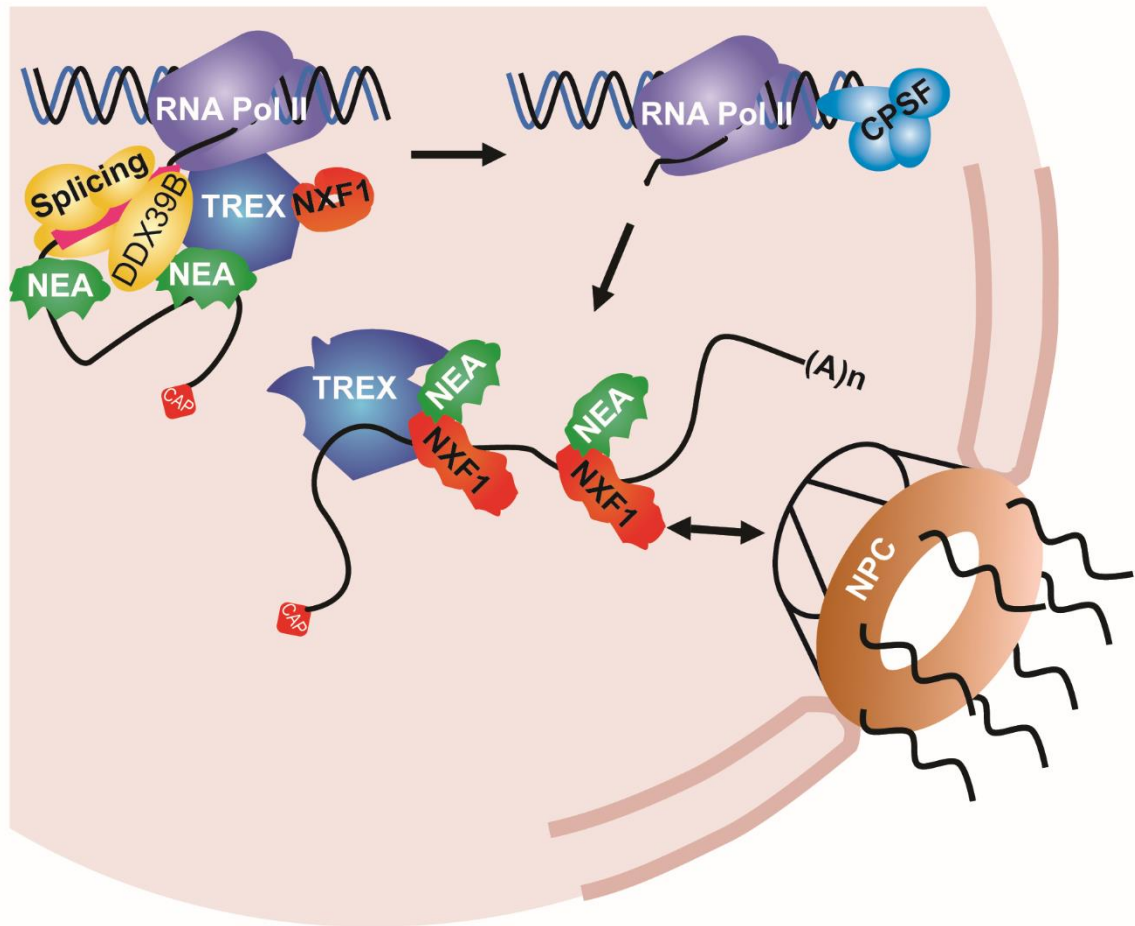


Figure 1.3. Common principles of NXF1 dependent nuclear export of eukaryotic bulk mRNA. The process consists of three steps. The first step is pre-initiation which is required for triggering transcription of DNA by recruiting TBP (TATA box binding protein) and transcription start site (TSS). After initiation of transcription, capping enzyme (CE), CAP are involved. The next step is elongation which involves recruitment of the nuclear adaptor protein (TREX complex) by the spliceosome during splicing. During this step, NXF1 is recruited to the active transport of mRNP transcripts. The process then follows cleavage of the 3' end and polyadenylation. The nuclear export ends with transport of transcripts through the NPC and then release into the cytoplasm. **The figure is adapted from** (Hautbergue et al., 2017).

1.7.3.2. Alteration of Nucleocytoplasmic Transport in *C9ORF72*-ALS/FTD

The nucleocytoplasmic transport of proteins and mRNA has been shown to be altered in neurodegenerative diseases such as Huntington and Poly-Q disease. However, this does not depend on SRSF1 (Walsh et al., 2015). The identified dysregulation of nucleocytoplasmic transport of RNA in neurodegenerative diseases are described in **Table 1.2**. *C9ORF72*-related ALS is subset of this disease group which is caused by hexanucleotide repeat expansion of sense (G_4C_2) and antisense (G_2C_4) *C9orf72* repeats (DeJesus-Hernandez et al., 2011). In *C9ORF72* mediated ALS, unspliced *C9orf72* pre-mRNAs are exported to the cytoplasm where they are subjected to RAN translation which leads to the production of toxic dipeptide repeat proteins as translation occurs in all the frames (Hautbergue et al., 2017). It has been reported that the nuclear export protein, SRSF1 specifically binds to the pathogenic *C9orf72* repeat transcripts (Hautbergue et al., 2017). Following the sequestration of SRSF1 onto the repeat transcript, interaction with the amino-terminal domain of the nuclear export receptor NXF1, which is also part of nuclear export machinery, is promoted (Huang et al., 2004). Aminoacids 89-120 of SRSF1 were shown by my supervisor to interact with NXF1. Indeed, NXF1 binds to the nucleopore and allows the process of nuclear export by itself binding to RNA. The NXF1:SRSF1 interaction allows for nuclear export of these abnormal pre-mRNA transcripts and this subsequently leads to the production dipeptide repeat transcripts by translation within the cytoplasm (Hautbergue et al., 2017).

Also, it should be noted that SRSF1 does not bind the non-pathogenic transcript of *C9orf72* and it is not involved in the nuclear export of the non-pathological repeat transcripts (Hautbergue et al., 2017). Since SRSF1 is specific to the pathological species, it provides an opportunity to manipulate its function which is associated with nucleocytoplasmic transport, thereby inhibiting the nuclear export of pathological *C9orf72* repeat transcripts and preventing the production of pathological DPRs (Hautbergue, 2016). In this respect, my research group in collaboration with others, has recently shown two therapeutic strategies that inhibit SRSF1-dependent nuclear export of both sense and antisense *C9orf72* repeat transcripts (shown in **Figure 1.4**).

These strategies consist of either depleting SRSF1 or expression of SRSF1 m4 mutant protein, a protein engineered by replacing 4 arginine residues, which are required for the interaction with NXF1 by alanines. It has been shown that if you deplete SRSF1 partially within a cell, then this will lead to a reduction in the sequestration of the protein onto the pathogenic pre-mRNA transcript. Therefore, this will inhibit the interaction with nuclear export machinery and nuclear export of the pathogenic transcripts. Genome wide investigation has shown that SRSF1 is involved in nuclear export of very small subgroup of RNA (225 transcripts) in a mammalian cell model, due to the fact that other SRSF 2 to 7 factors play redundant and cooperative roles in the nuclear export pathway (Müller-McNicoll et al., 2016). In another paper published from my lab has shown that 50 % depletion SRSF1 leads to alteration of nuclear export of only 177 transcripts (137 mRNAs) out of 45,000 genes in C9ORF72-ALS patient-derived neurons. Therefore, partial depletion of SRSF1 has very limited off-target effects (less than 0.4% cell transcriptome) (Castelli et al., 2021).

Consequently, less DPRs will be produced in the cytoplasm of the cell, preventing neurotoxicity. The second strategy involves the mutant SRSF1 m4 protein. The mutant SRSF1-m4 protein retains ability to bind *C9orf72* repeat transcripts but fails to interact with NXF1 which leads to the inhibition of the nuclear export and inhibition of DPR production. Those two strategies have been shown to confer neuroprotection by inhibiting the nuclear export of *C9orf72* repeat transcripts in both patient derived neurons and in a *Drosophila* model of ALS (Hautbergue et al., 2017).

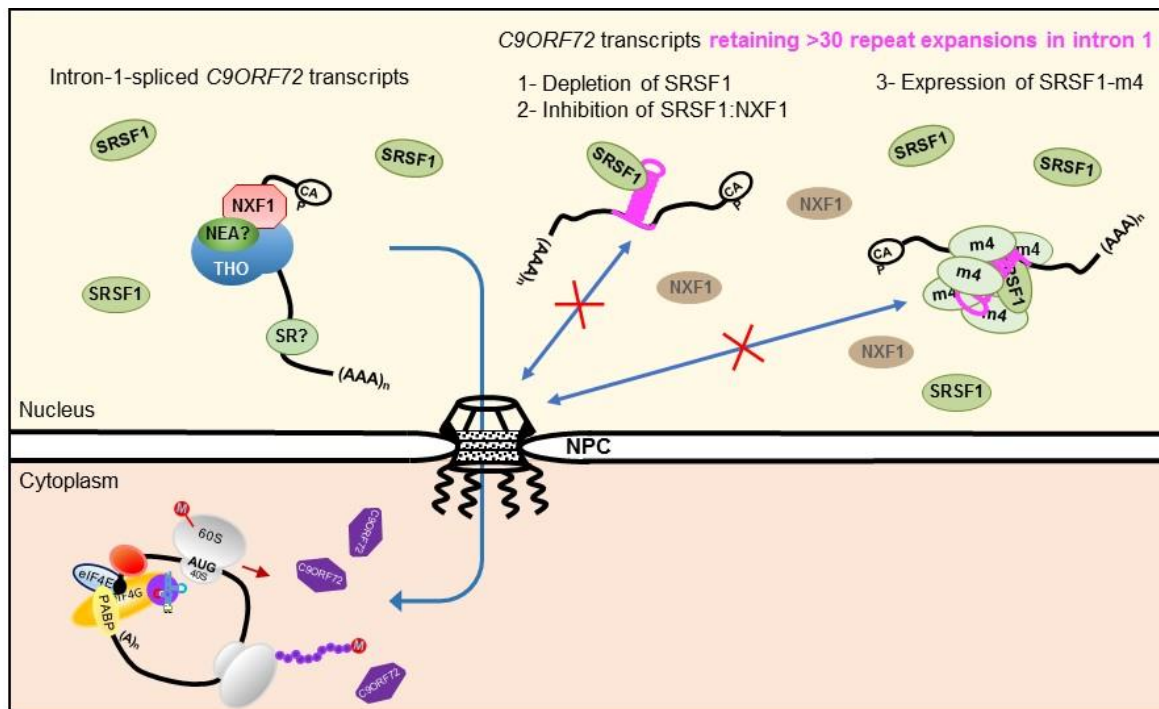


Figure 1.4. The identified two therapeutic strategies to inhibit the nuclear export of pathological *C9ORF72* repeat transcripts. These strategies involve either depletion of SRSF1 or expression of engineered SRSF1-m4 protein on pathogenic *C9ORF72* repeat transcripts. Both strategies do not interfere with the production of the wild type (WT) *C9orf72* protein but they prevent nucleocytoplasmic transport of pathogenic *C9ORF72* repeat transcripts by inhibiting NXF1:SRSF1 interaction. **The figure is adapted from** (Lydia M. Castelli, 2018).

Table 1.2. Dysregulation of RNA export in neurodegenerative diseases. The table is adapted from (Boehringer and Bowser, 2018).

Defects	Disease Model	Reference
Nuclear membrane structural abnormalities	AD: Identified by Nup62 immunostaining in patient hippocampal tissue	(Sheffield et al., 2006)
Nuclear membrane structural abnormalities	ALS: Nuclear envelope abnormalities denoted by Nup62, Nup88, and Nup153 immunoreactivity in spinal cord tissue from mutant SOD1 G93A mice and ALS patients	(Kinoshita et al., 2009)
Increased NPC protein immunoreactivity	ALS: SOD1 G93A mice showed increased immunoreactivity for GP210 and Nup205 in spinal cord tissue and patient tissue showed increased GP210 immunoreactivity in the nuclear envelope and cytoplasm	(Shang et al., 2017)
NPC and RNA export proteins identified as disease modifiers in <i>Drosophila</i> genetic screens	ALS: Loss of function of Nup50 and dominant negative Ran enhanced an eye phenotype in flies expressing the <i>C9orf72</i> repeat expansion, loss of function of Nup107, and Nup160 suppressed phenotype	(Freibaum et al., 2015)
Altered levels or localization of Ran and Ran binding protein	ALS: Knockdown of TDP-43 in SH-SY5Y cells led to decreased levels of Ranbp1 and TDP-43 knockdown in Neuro2a cells led to decreased levels of Ran mRNA and protein	(Ward et al., 2014)
Altered levels or localization of Ran and Ran binding protein	HD: iPSC derived neurons from Huntington's patients exhibited decreased nuclear to cytoplasmic Ran ratio	(Grima et al., 2017)
Nuclear retention of RNA and mRNA	ALS: An increased nuclear to cytoplasmic ratio of RNA was identified by RNA-seq when G93A mutant SOD1 was expressed in NSC-34 cells	(Kim et al., 2017)
Nuclear retention of RNA and mRNA	ALS: mRNA was retained within the nucleus of cells expressing a C-terminal fragment of TDP-43 or an artificial aggregation prone β -sheet construct	(Woerner et al., 2016)
RNA export proteins interact with disease linked proteins or dipeptide repeats	ALS: CRM1 and Aly identified as interactors of the DPRs PR and GR (from <i>C9orf72</i> repeat expansion) expressed in HEK293 cells	(Lee et al., 2016a)
Mutations of Gle1 in ALS patients	ALS: mutations in Gle1 which likely lead to haploinsufficiency are a rare genetic cause of ALS	(Kaneb et al., 2014)
Polymerization of FG nucleoporins	ALS: PR DPR formed from <i>C9orf72</i> repeat expansion bound to FG repeat of the central channel of the nuclear pore keeping them in a polymerized state	(Shi et al., 2017)

1.8. *C9ORF72*-ALS/FTD Mouse Models

Generating a mouse model that is able recapitulate pathological mechanisms and clinical features of *C9ORF72* mediated ALS/FTD is highly important to tease out disease mechanisms and investigate the efficacy of treatments. Up to now, several research groups have developed *C9ORF72*-ALS/FTD mouse models. The first animal model was generated by Chew et al, 2015. To develop this model, AAV2/9 harbouring (G₄C₂)₆₆ repeat DNA was administered via bilateral intra-cerebroventricular injection into the CNS of C57BL/6J mice at post-natal day 0 to obtain efficient and wide-spread transduction. Behavioural, pathological, and biochemical features of mice were assessed after 6 months. They found that sense repeat transcripts were present in around 50% of cells in specific brain areas such as motor cortex, hippocampus of AAV-(G₄C₂)₆₆ mice, but not in the control mice. Additionally, poly (GA), poly (GP), and poly (GR) DPRs were found in particular areas of the brain (cortex, hippocampus, and Purkinje cells). Furthermore, the study revealed that at 6 months, mice injected with repeat DNA old displayed elevated anxiety levels and hyperactivity, and a decline in social behaviour compared to control mice. Indeed, the treated mice showed substantial difficulties with their motor coordination. However, no motor neuron loss was detected in the spinal cord of the mice (Chew et al., 2015).

More recent mouse models for *C9ORF72*-ALS/FTD used BAC (bacterial artificial chromosome) transgenic models. The researchers created transgenic mice by using four different C9-BAC DNA clones. These clones had different G₄C₂-repeat lengths, varied sizes of flanking genomic regions and contained either a complete or partial *C9orf72* coding region. Two mouse models of *C9ORF72*-ALS/FTD were generated using similar methods by incorporating partial regions of the *C9orf72* gene into BAC DNA clones in different mouse strains. Peters *et al.* used a BAC clone containing a partial human *C9orf72* gene region (exon 1 to 6) with a (G₄C₂)₅₀₀ repeat region in SJL/B6 mouse strain (Peters et al., 2015), whereas (Jiang et al., 2016) generated a mouse model using a BAC clone along with a (G₄C₂)₄₅₀ repeat region (exon 1 to 5) in C57BL6/C3H mouse strain. In both mouse strains, the presence of sense and antisense G₄C₂-repeat RNA foci was observed in the CNS and spinal cord. The number of antisense RNA foci was significantly less than the number of sense RNA foci. An MSD assay was used to analyse DPR-mediated pathology in both BAC mouse models.

The assay revealed that the highest amount of expressed poly (GP) DPR was detected in cerebellum in comparison to cortex, hippocampus, midbrain, hindbrain and spinal cord of SJL/B6 mouse strain at 4 months (Peters et al., 2015). As for the C57BL6/C3H mouse strain, poly (GA), poly (GP) and poly (GR) DPRs were detected in retro-splenial cortex at 3 months. The authors also investigated an age-dependent increase in DPR expression in the BAC-C57BL6/C3H mice and found that the amount and size of Poly (GA) increased in the dentate gyrus from 6 to 22 months, with the size doubling by 22 months (Jiang et al., 2016). Although both mouse models display pathological features similar to those observed in patients with *C9ORF72*-associated ALS/FTD, they did not show significant behavioural abnormalities or neuronal loss even after 12 months.

O'Rourke *et al.* utilized a BAC clone containing the complete *C9orf72* gene, including a (G4C2)₈₀₀ repeat region on C57BL/6J background. Like other BAC transgenic models, this model exhibited both sense and antisense RNA foci, as well as poly (GP) toxic DPRs in various CNS regions, such as the spinal cord, frontal cortex, and cerebellum. The researchers then examined the mice for any behavioural changes after 18 months. However, they did not observe significant alterations between transgenic and the age-matched control group (O'Rourke et al., 2015).

In 2016, Liu and colleagues developed a BAC-*C9orf72* mouse model by inserting BAC clone containing the *C9orf72* gene along with (G₄C₂)₈₅₀ repeat region into FVB/NJ mouse strain (Liu et al., 2016; Nguyen et al., 2020). This transgenic mouse model was used in the present project as it most closely mirrors human *C9ORF72*-ALS/FTD disease related phenotypes and pathological features. This model exhibits a number of key features that are similar to those seen in human patients, including expression of sense and antisense repeat transcripts, accumulation of both sense and antisense DPRs, and RNA foci formation in CNS regions. Indeed, they also observed cytoplasmic and nuclear TDP-43 inclusions specifically in degenerating neurons across the brain of transgenic mice. In terms of behavioural and functional features, this model has been shown to exhibit progressive motor neuron loss and dysfunction, muscle weakness, tremors, paralysis, and reduced lifespan which are relevant to ALS and FTD. However, they also reported that some mice developed severe motor deficits, while others displayed a slower progressive disease similar to human ALS and FTD which constitute a disadvantage for this model in terms of predicting which mice will develop phenotype (Liu et al., 2016).

Mice with severe symptoms tend to die earlier than those with milder symptoms. This unpredictability makes it difficult to evaluate treatments, as it may be hard to tell if any improvement is due to the intervention or the natural progression of the disease in the mouse. However, this mouse model still currently is considered the most appropriate model that exhibits the phenotypes and pathological features of human *C9ORF72*-ALS/FTD. Therefore, we decided to use this model to evaluate the efficacy of our SRSF1-mediated non-viral and viral gene therapy approaches.

1.9. Therapeutic Interventions for ALS/FTD

1.9.1. Drug Treatment

There is currently no effective treatment to cure the motor deficits and improve survival of patients due to the complexity of ALS pathophysiology. However, some devices such as neck collars and feeding tubes have been developed to improve the quality of life of patients (Pancani et al., 2016). There are only few approved drug treatments for patients. One is Riluzole which was approved in 1996 and modestly extends lifespan by 3 or 4 months. It is thought to modulate glutamate excitotoxicity (Mitumoto et al., 2014) but does not correct the motor deficits that lead to progressive paralysis. The second drug is Edaravone, a free radical scavenger that was approved in 2017 by the US Food and Drug Administration (FDA) (Ito et al., 2008). Studies suggest that Edaravone reduces weight loss and the progression of motor neuron degeneration in *SOD1^{G93A}* mice. Additionally, it has been demonstrated that edaravone is effective in decreasing *SOD1* protein accumulation. Another drug is Relyvrio which was approved in USA and Canada. More recently, QALSODY (formerly known as tofersen) was approved on 25/04/23 by the FDA for *SOD1*-ALS. There is also a potential treatment using the tyrosine kinase inhibitor masitinib. Administration of masitinib in the *SOD1^{G93A}* mice resulted in slowing down microgliosis, motor neuron loss and extending lifespan of animal. However, clinical trial investigation is still ongoing. (Trias et al., 2016). A similar observation was also made with administration of Fasudil which is potent rho kinase (ROCK) inhibitor (Takata et al., 2013). In addition, there are currently 47 potential drug treatments now under clinical trial (Mead et al., 2023).

1.9.2. Stem Cell Therapy

Stem cell-based therapy is being investigated for its potential to treat ALS/FTD due to the ability of stem cells to differentiate into human embryonic stem cells (hESC), induced pluripotent stem cells (iPSC), mesenchymal stem cells (MSC) and foetal neural progenitors (NSC). It has been shown that they prevent motor neuron loss by releasing neuroprotective trophic factors as well as preventing toxicity (Ciervo et al., 2017). One example of a study investigated the transplantation of hESC into the spinal cord of SOD1^{G93A} mice model. The results revealed that motor neuron loss close to injection site was reduced in transplanted animals (Wyatt et al., 2011). Although hESC exhibits potentially promising results, clinical studies were blocked due to ethical issues, potential tumour formation within a living organism and the risk of transplant rejection (Toma et al., 2015).

The use of NSC for ALS/FTD treatment is another area of research that has received some attention. Several studies have been performed to evaluate the therapeutic potential of NSC transplantation into ALS mice. For example, a study investigated the efficiency of NSI-566RSC cell injection in the ventral horn of SOD1^{G93A} mice. Transplantation of NSI-566RSC cells resulted in delaying disease onset and progression in comparison to control group (Yan et al., 2006). Similar observations were also made by the same research when they investigated therapeutic efficiency of NSI-566RSC cells in pre-symptomatic SOD1^{G93A} mice. MSCs are a type of adult stem cell that have shown promise in the treatment ALS. In preclinical studies, MSCs have been shown to have a beneficial effect on motor function and survival in animal models of ALS. For example, 14 days delay in disease onset and extended life span were observed in SOD1^{G93A} ALS mice that received human bone marrow mesenchymal stem cells (hBMMSCs) intravenously (Uccelli et al., 2012). MSCs also have been studied into clinical trial, NurOWN, where the efficacy and safety were assessed. In this trial, ALS patients received intrathecal injection of MSC- NTF (neurotrophic factors), which resulted in decrease in CSF inflammatory biomarkers and was deemed safe (Berry et al., 2019).

1.9.3. Gene Therapy Approach in ALS/FTD

Gene therapy is also considered another promising therapeutic approach in ALS/FTD. Gene therapy is defined as a medical approach that involves modifying the gene expression or correcting a defective gene responsible for the disease by introducing new genetic material such as DNA or RNA into the target cells and tissue. The genetic material can be delivered to the target cells using different methods, including viral vectors and non-viral vectors (Bulcha et al., 2021). Gene therapy is an emerging approach and there are only a limited number of gene therapies approved in the UK at present.

1.9.3.1. Viral Vector Mediated Gene Therapy

Viral vectors have been widely used in gene therapy research, accounting for more than 75% of global gene therapy applications (Chaitanya, 2022). These vectors can efficiently deliver transgenes to the nuclei of human cells. Viral gene therapy mediated expression of a transgene of interest, either a protein or a shRNA cassette, is currently considered a promising therapeutic approach in ALS given the recent successful results of Phase 2 clinical trials and FDA approval for the adeno-associated viral (AAV) replacement strategy of SMN (survival motor neuron protein) in spinal muscular atrophy. In addition, gene therapy approaches have entered clinical trials for ALS. The main viral vectors commonly used for gene therapy in the CNS are based on: lentiviruses and adeno-associated viruses (Chandran, 2017). AAVs or self-complementary scAAVs are the preferred virus targeting the CNS because of their ability to deliver genes with high efficiency and promote long-term stable gene expression (McCarty et al., 2004). They are also non-pathogenic, present a very low rate of genomic integration, and are able to transduce both dividing and non-dividing post-mitotic cells (Munis, 2020).

1.9.3.1.1. Lentiviral Vectors

Lentiviral vectors are a type of retrovirus vectors containing single-stranded RNA that usually ranges in length from 7 kilobases (kb) to 12 kb. These vectors are recognized for their distinct ability to stably integrate their viral DNA into the host genome (Lundstrom, 2019). Indeed, they can efficiently transfer genes to non-dividing cells (Liu and Berkhout, 2014). Additionally, the integration of lentiviral vectors into genomic sites is evenly distributed, which diminishes the risk of mutagenesis or activation of deleterious genes.

The use of lentiviral-mediated gene therapy is prevalent in the treatment of different types of cancers like melanoma and ovarian cancer. It is also commonly used for treating genetic disorders such as adenosine deaminase deficiency, severe combined immune deficiency, and Gaucher's disease, as well as ophthalmic diseases (Ghosh et al., 2020). For example, the efficacy of an HIV-1 based lentiviral vector harboring a murine *Mertk* cDNA was evaluated in a Royal College of Surgeons (RCS) rat model which develops severe retinal degeneration. The lentivirus expressing *Mertk* was injected subretinally into the right eye resulting in a reduction in photoreceptor loss and subsequent improvement in retinal function by 27 weeks (Tschernutter et al., 2005).

Additionally, vectors derived from HIV-1 have been used to overexpress various therapeutic molecules in animal models of motor neuron diseases. For example, lentiviral mediated SOD1 depletion through expression of shRNAs, via intramuscular injection with retrograde transport to the spinal cord, has been successfully tested in mouse models resulting in a reduction in the expression levels of wild-type and mutant SOD1 proteins and leading to delay in disease progression and prolonged survival of SOD1^{G93A} ALS mice (Ralph et al., 2005; Raoul et al., 2005). A similar observation was made in another study investigating the therapeutic efficiency of intramuscular injection of EIAV lentivirus expressing VEGF in SOD1^{G93A} ALS mice. The authors reported that mice treated with EIAV-VEGF showed decreased motor neuron degeneration, delayed disease onset and extended life span by 38 days in comparison to an age matched control group (Azzouz et al., 2004).

Despite being a useful tool in gene therapy, lentiviral vectors have some limitations. One such concern is insertional mutagenesis, where the vector's insertion into the genome can disrupt essential genes, especially in actively transcribed genes. It has been observed that even those vectors with strong promoter and enhancer elements can activate neighbouring genes (Bulcha et al., 2021; Zufferey et al., 1998). Lentiviral vectors can also form chimeric gene fusions consisting of host and proviral sequences arising from vector genome integration (Almarza et al., 2011; Moiani et al., 2012). Another potential issue with lentiviral vectors is their limited cargo capability and ability to cause abnormal splicing of cellular transcripts (Cesana et al., 2014). Despite these challenges, lentiviral vectors are commonly used in gene therapy research and clinical trials, representing about 22% of all trials conducted globally (Zhao et al., 2022).

1.9.3.1.2. Adeno-associated Virus

AAVs are a small (18-25 nm) helper-dependent virus type derived from a non-harmful parvovirus which are utilized as carriers for gene expression while posing minimal risk of side effects. These viruses are associated with liver carcinoma in humans. AAVs have a capsid made of 60 protein structures enclosed in a non-enveloped icosahedral shape (Meyer et al., 2019; Wu et al., 2006). The single-stranded DNA genome of AAVs is approximately 4.7 kb and contains two open reading frames which are rep and cap, surrounded by inverted terminal repeats (ITRs). These ITRs create hairpin structures that play a significant role in the replication and packaging processes (Meyer et al., 2019). AAVs have been demonstrated as an effective delivery vector for the brain due to their ability to transduce both dividing and non-dividing cells, their reduced risk of genetic alterations and silencing of the transgene and their ability to provide long-term expression (Li and Samulski, 2020). So far, 12 different AAV serotypes have been identified, along with the isolation of over 100 variants from humans and primates (Kotterman and Schaffer, 2014). Different AAV vectors have been identified that preferentially transduce many different cell types, there are still cell types for which AAV has proven difficult to transduce (Naso et al., 2017).

Furthermore, the capability of AAV serotypes to penetrate the blood-brain barrier has further boosted the utilization of AAVs in neurological diseases by targeting the CNS (Kuzmin et al., 2021). Recently, an FDA approved gene therapy drug under the name Zolgensma was designed to treat SMA type I and II. SMA is caused by mutations in the *SMN1* gene and is characterised by progressive degeneration of anterior horn cells and subsequent muscle weakness and permanent difficulty with breathing by the age of two. Zolgensma is a single-dose intravenous injection of an AAV9 vector to introduce a functional copy of the *SMN1* gene into the patient's cells to replace the missing or non-functional *SMN1* gene. The aim of the therapy is to slow or stop the progression of SMA by restoring the function of motor neurons, which can improve muscle control, strength, and movement in affected children. Zolgensma is the second approved genetic therapy for SMA and has shown promising results in clinical trials, including significant improvements in motor function and milestones in infants treated before the age of two years (Mendell et al., 2017).

In the context of ALS treatment, the frequently employed AAV serotypes for gene therapy purposes are AAVs 2, 5, 6, 9 and 10 which allow passage through the blood-brain barrier following intravenous or intrathecal injection (Naso et al., 2017). Among the commonly used AAV serotypes, AAV9 has received a lot of attention due to its high ability to efficiently transduce the brain in various species such as mice, rats, cats, and monkeys following systemic administration (Hunter et al., 2022). There is a great deal of literature highlighting the therapeutic success of AAVs in gene therapy in ALS models. A single intravenous injection of AAV9 encoding SOD1 shRNA in SOD1^{G93A} mice before disease onset resulted in a reduction in mutant SOD1, delayed disease onset, extended survival, and slower disease progression. In the same study, the efficiency of AAV9-SOD1 shRNA on SOD1^{G37R} mice already displaying the disease phenotypes was investigated. They found that single injection of virus slowed down disease progression and extended survival by 86.5 days. After these promising results, they transitioned into non-human primate testing of AAV9 with shRNA and found that intrathecal injection of AAV9 via lumbar puncture significantly reduced SOD1 protein and SOD1 RNA levels (Foust et al., 2013).

Similar observations were made in another study investigating the efficiency of a single intrathecal injection of AAV9 encoding RNAi against hSOD1 into SOD1^{G93A} mice (Iannitti et al., 2018). In a separate study, delivery of shRNAs via AAV2 to spinal motor neurons of SOD1^{G93A} mice showed that intramuscular injection of siRNA reduced the functional properties and levels of SOD1. Additionally, intrathecal and peripheral delivery of shRNA with AAV9 resulted in delayed disease onset in mice by approximately 40% and an increase in retrograde axonal transport (Miller et al., 2005). In another study, administration of AAV10-U7-hSOD1 (sSOD1 inserted in AAV10) to SOD1^{G93A} mice prevented weight loss, rescued neuromuscular function, and extended survival (Biferi et al., 2017). Besides the use of AAV in ALS, expression of SMN using AAV9 is very effective in treating SMA and has recently been approved for a treatment of this childhood form of motor neuron disease (Meyer et al., 2015).

An AAV mediated gene therapy strategy was also applied for the treatment of other diseases. For example, In 2008, two independent groups reported the successful gene therapy outcomes for the treatment of Leber's congenital amaurosis (LCA) which is group of childhood onset inherited retinal degenerative diseases. One form of LCA is caused by mutations in the retinal pigment epithelium specific 65-kDa protein gene (*RPE65*). Subretinal injection of recombinant AAV2 viral vector expressing *RPE65* complementary DNA (cDNA) to patients with *RPE65* deficiency enhanced vision of patients with inherited blindness (Bainbridge et al., 2008; Maguire et al., 2008). Both studies provided hope for future potential AAV approaches for the treatment LCA. This led in 2017, to Luxturna developed by Spark Therapeutics to treat *RPE65* deficiency, becoming the first FDA-approved AAV based gene therapy product.

Pompe disease is rare neuromuscular disorder caused by mutation in the *GAA* gene. The therapeutic efficacy of AAV8 mediated delivery of the *GAA* transgene was evaluated in *GAA* knockout mouse model. The study found that intravenous injection of an AAV8 viral vector resulted in a reduction of glycogen accumulation in both muscle and CNS, as well as a decrease in cardiac hypertrophy and an increase in survival of the mice (Puzzo et al., 2017). A recent study involving *GAA* knockout mice and non-human primates investigated the effects of the drug AT845. This drug is an AAV8 serotype vector containing the murine MCK promoter/enhancer, which drives a codon-optimized human *GAA*. Systemic administration of AT845 led to enzyme activity levels exceeding physiological norms. This, in turn, resulted in notable functional improvements and efficient glycogen removal in crucial target tissues, all in a dose-dependent manner (Eggers et al., 2022).

Alpha-mannosidosis (AMD) disease is a neurodegenerative lysosomal storage disorder which is caused by mutation in the *MANB21* (alpha mannosidase) gene which leads to deficiency in alpha mannosidase enzyme activity (Malm and Nilssen, 2008). Researchers administered AAV.hu32 carrying the feline *MANB* gene to cats with AMD via intracarotid injection. The treatment resulted in widespread transduction across multiple brain regions, reversed storage lesions, improved lifespan, and produced higher *MANB* activity in the cerebrospinal fluid and serum, and histological alleviation of the pathological changes (Yoon et al., 2020).

Gene therapy has emerged as a powerful approach for treating microsatellite repeat expansions for example in Huntington disease (HD) caused by a CAG repeat expansion within exon 1 of the *HTT* gene. One of the most promising results in gene therapy was seen in a study spearheaded by Uniqure. They are developing an AAV-based gene therapy, AMT-130. This therapy employs a single brain surgery procedure to administer the adeno-associated viral vector serotype 5 (AAV5), which subsequently results in a continuous reduction of Htt expression. The introduction of AAV5 into the brain tissue is achieved through MRI-guided techniques (Rodrigues and Wild, 2020). The miRNA will be loaded into the nucleus of the cell and bind to the Htt mRNA, marking it for degradation. *In vivo* experiments conducted on Q175 mouse models, in which AAV5 was surgically injected and demonstrated a notable dose-correlated decline in mHTT protein, observing reductions of as much as 39% in the striatum and 13% in the cortex. The AMT-130-01 study is currently evaluating the treatment's efficacy and safety in humans. Designed as a double-blind trial, it includes a sham control group and is at the Phase I/II investigation stage. Recruitment is active in the US. Once the 26 participants undergo the brain surgery for the therapy, they'll be closely monitored at dedicated HD clinics worldwide (Martinez and Peplow, 2021).

AAV has some limitations that must be considered when designing and implementing gene therapy protocols. The most significant limitation of AAV vectors is the small cloning capacity, which restricts the transgene and ITRs to be cloned within a 4.9 kb size. To address the limitation, several methods have been proposed, such as decreasing the transgene size or utilizing fragmented or dual vector approaches (Au et al., 2022; Chamberlain et al., 2016; Naso et al., 2017). However, these strategies still face efficacy problems that prevent their widespread application. Another challenge is the delay in transgene expression due to the single-stranded nature of the AAV genome. A possible solution involves using self-complementary AAV vectors (scAAV), which generate double-stranded DNA without requiring additional synthesis or pairing. However, this approach reduces the packaging capacity of the AAV vector to 2.3 kb, which is half of its original size limit (Ail et al., 2023).

1.9.3.2. Non-viral Gene Therapy Approach

Non-viral gene therapy approaches involve using varying combinations of DNA or RNA which is easier to produce than viral vectors but less efficient for gene delivery. Efforts are underway to improve gene delivery using non-viral approaches such as cationic lipids, antisense oligonucleotides (ASO), cell-penetrating peptides and others. The use of these methods alone or in combination may eventually improve the efficacy of non-viral gene therapy approaches (Castro and Kipps, 2018).

1.9.3.2.1. Antisense Oligonucleotide (ASO) Therapy

ASOs are short synthetic, unmodified or chemically modified single-stranded DNA sequences ranging from 8 to 50 nucleotides that are designed to bind to an mRNA target. They can inactivate the expression of a gene by binding to a promoter (Dhuri et al., 2020) stimulating the degradation of a target mRNA through RNA-interference and modulation of pre-mRNA splicing (Boros et al., 2022). These properties make ASOs suitable for a range of neurodegenerative diseases and they can provide advantages over viral gene therapy approaches in the case of toxic protein gain of function mechanisms. For example, ASOs can be tailored to precisely target and degrade this aberrant mutant mRNA, preserving the unaffected wild-type mRNA—a degree of specificity that might be challenging with some viral gene therapies (Boros et al., 2022).

If the pathological mechanism of neurodegeneration is caused by a loss of protein activity, using an ASO may not be safe (Mis et al., 2017). Take the *C9ORF72*-ALS/FTD case as an example. Degradation of *C9ORF72* mRNA would reduce the expression of the C9orf72 protein, which is involved in activating autophagy, nucleocytoplasmic trafficking, and endosomal membrane trafficking. As C9orf72 is crucial for initiating autophagy, its reduction could impair this process. Still, it's important to note that this concern doesn't always apply; ASOs are utilized for certain conditions characterized by a loss of function, such as SMA.

In SMA, Spinraza® is a type of ASO therapy that has received approval for usage. It is administered through intrathecal delivery and is composed of a 2-MOE-PS oligonucleotide developed by Ionis Pharmaceuticals. Its functional mechanism involves correcting *SMN2* splicing to produce the full length SMN protein. ASOs were also tested on polyglutamine (poly Q) diseases which are a group of nine uncommon neurological conditions caused by abnormal expansion of CAG repeat expansions resulting in the formation of insoluble poly (Q) protein aggregates through a toxic GOF. Pre-clinical studies have demonstrated that ASOs have the ability to decrease the levels of mutant poly Q-expanded protein and related disease abnormalities (Matos et al., 2018). Furthermore, in the current clinical trial for Huntington's disease, the use of ASOs from Ionis Pharmaceuticals was found to be safe and has resulted in a dose-dependent reduction of mutant huntingtin in the cerebrospinal fluid. These outcomes are encouraging, and hold promise for future clinical research (Rodrigues and Wild, 2018).

Duchenne muscular dystrophy (DMD) is another rare neurodegenerative disease that leads to the gradual degeneration and weakening of muscles due to mutations in the *DMD* gene (Yiu and Kornberg, 2015). Sarepta Therapeutics has developed an ASO-based therapy called Exondys 51™ (eteplirsen) that was approved by the FDA in 2016. ASO treatment corrected the genetic reading frame of the abnormal *DMD* gene producing a functional dystrophin protein. Clinical trials have demonstrated that the ASO therapy has led to an increase in muscle fibres that test positive for dystrophin and an improvement in walking distance in treated patients, although the precise impact of these results is not yet fully understood (Lim et al., 2017).

Several clinical trials using ASO technology are currently being conducted for treatment ALS and FTD. An ASO targeting SOD1 led to an improvement in the lifespan of SOD1^{G93A} mice. A Phase I human clinical trial of ASO targeting SOD1 named ISIS 333611 was completed in 2012 (NCT01041222) (Miller et al., 2013). No adverse effects were observed in the CNS in both the treated and control groups. Also, the authors checked the SOD1 protein levels in the cerebrospinal fluid of patients and no noticeable decrease was found (Miller et al., 2013).

A recent ASO-based therapy called tofersen, BIIB067 (Biogen) was investigated in phase I and II clinical trials (NCT02623699). 50 ALS patients with *SOD1* mutations received intrathecal doses of tofersen ranging from 20 mg to 100 mg. The patients were followed for 31 weeks post-injection and the dose-dependent efficiency of tofersen on the *SOD1* concentration in the CSF was evaluated. They found that patients receiving the highest dose (100 mg) displayed a reduction in the level of *SOD1* protein in CSF by 36% (Miller et al., 2020). Additionally, administration of 100 mg of tofersen resulted in a reduction of neurofilament concentrations in CSF and plasma (Miller et al., 2020). The initial trial of BIIB067 had a small number of participants, so the study was extended to a phase 3 clinical trial. The phase 3 clinical trial failed to achieve statistical significance in measuring the progression of ALS in fast progressors group after 28 weeks of treatment. However, significant clinical benefits were observed 12 months after treatment initiation (Miller et al., 2022). Recently, the ASO treatment was approved by the FDA on April 25, 2023.

Although ASOs are a promising therapeutic approach for targeting neurodegenerative disease, they have several limitations. The main challenge is that ASOs are not able to pass the blood brain barrier (BBB) when delivered systemically compared to delivery of RNAi within a viral vector (Scarrott et al., 2015). In addition, the patients must currently undertake regular lumbar punctures for intrathecal delivery of the ASO due to the relatively short half-life.

1.9.3.3. Limitations of Gene Therapy Approaches

Non-viral and viral gene therapies have recently provided promising results. In particular, this approach has recently achieved significant success as demonstrated by the FDA approval of Zolgensma or Luxturna for the treatment of SMA and LCA respectively. These significant breakthroughs have opened new avenues for the development of gene therapy based therapeutic tools for other genetic diseases. The ongoing clinical trials for gene therapy in the UK is consistently rising annually. In 2022, there were 178 trials currently underway, reflecting an increase from the 168 trials reported the previous year. AAV and lentiviral vectors continue to be the prevailing options for gene delivery. In clinical trials focusing on in vivo gene therapy, AAV-derived vectors maintain their status as the primary choice (38% of cases), closely pursued by lentiviral vectors (33% of cases) '<https://ct.catapult.org.uk/>'.

However, more effort is still needed to improve gene therapy approaches' efficiency and overcome the present challenges for ALS and other diseases. These challenges generally are concerned with the identification of the key target genes that are critical for modifying the disease pathology and the delivery of the therapeutic gene to the target tissue at an efficacious dose. There's also the issue of immune responses; the body might recognize and attack the viral vectors commonly used in gene therapies, reducing the effectiveness of the treatment or even causing harm to the patient. Durability of response remains a concern; in some cases, the therapeutic effects may be temporary, requiring repeated administrations that could further amplify the risk of immune reactions. The cost of gene therapy products, and potential triggering of the immune system represent additional common issues which need to be addressed (Shahryari et al., 2021). Recent advances in gene therapy for ALS are based primarily on RNA-mediated therapy which consist of RNA interference (RNAi), ASO and virus-based therapies (Wang et al., 2023).

1.9.4. Therapeutic Strategies for *C9ORF72*-ALS/FTD

Repeat expansion within intron 1 of the *C9orf72* gene is the most common genetic causes of both ALS and FTD. (DeJesus-Hernandez et al., 2011; Renton et al., 2011). The identification of the *C9orf72* gene mutation has highlighted the need for further research to fully understand the underlying pathogenic mechanisms of the disease and also has triggered greater research in developing potential therapies to treat the ALS and FTD. So far, researchers have primarily concentrated on discovering strategies to slow down or stop disease progression through alleviating toxic gain of function mechanisms. Therefore, strategies include targeting and reducing *C9orf72* repeat transcripts expression, inhibiting expression of toxic DPR protein through RAN translation, reducing the expression of expansion transcripts, and regulating the pathways impacted downstream such as nucleocytoplasmic transport and stress granules (Amado and Davidson, 2021; Hautbergue et al., 2021).

1.9.4.1. Targeting Mutant *C9orf72* Repeat RNA or DNA

1.9.4.1.1. ASO

Targeting the *C9orf72* repeat expansion with ASOs mitigates pathological features in several *in vitro* and *in vivo* models of the disease. ASOs targeting different *C9orf72* mRNA regions rescued repeat RNA-associated neurotoxicity in induced pluripotent stem cell (iPSC)-derived neurons from *C9ORF72*-ALS patients (Lagier-Tourenne et al., 2013; Sareen et al., 2013). In another study using ASOs, the number of GGGGCC-RNA foci in cells was reduced, and dysregulated gene expression was suppressed with biomarker genes in iPSC cells (Donnelly et al., 2013). The therapeutic efficacy of ASOs has been tested in *in vivo* models of *C9ORF72*-ALS/FTD. For example, an ASO targeting repeat G-quadruplexes in *Drosophila* model of *C9ORF72*-ALS prevented neurodegeneration by rescuing nucleocytoplasmic transport deficiency (Zhang et al., 2015a). In another study, a single dose intraventricular injection of an ASO targeting repeat RNA decreased selectively the formation of sense RNA foci and DPR expression in BAC transgenic mice through degradation of the target RNA via the RNase H1 pathway (Jiang et al., 2016). The promising outcomes from studies using ASOs sparked increased interest and accelerated use of ASOs in clinical trials for the treatment of ALS/FTD. For example, there is an ongoing phase I trial of *C9orf72*-ASO at the University of Sheffield since September 2019 (NCT03626012). The ASO aims to degrade *C9orf72* repeat transcripts via RNase H mediated degradation.

Another clinical trial for ASO therapy in *C9ORF72*-ALS/FTD called WWE-004 was developed by Wave Life Sciences. WWE-004, which is a stereopure ASO, aims to deplete the level of transcript variants that are responsible for the production of DPRs and RNA foci by targeting the splice site 1b (SS1b) sequence which can be found in all *C9orf72* transcripts. This sequence can specifically knockdown *C9orf72* transcripts containing the repeat expansion. It is known that the *C9orf72* gene encodes three main transcripts namely variant 1, 2, and 3 (DeJesus-Hernandez et al., 2011; Renton et al., 2011). When the (G₄C₂) repeats are located within the first intron of variant 1 and 3, this in turn leads to toxic DPR and RNA foci formation. Therefore, WWE-004 was designed to selectively reduce the level of variant 1 and 3 but not variant 2 in order to maintain expression of the normal *C9orf72* protein. The functional efficiency of WWE-004 was validated in patient derived neurons and C9-500 BAC mice. WWE-004 led to a dose-dependent reduction in the level of variant 3 and production of poly (GP) and RNA foci in spinal cord and cortex of C9-500 BAC mice while protecting *C9orf72* protein expression (Liu et al., 2022). WWE-004 is currently in Phase I/IIa stage (NCT04931862) where safety, tolerability, pharmacokinetic, and pharmacodynamic parameters are being evaluated within patients.

A similar approach was undertaken in another study utilising an ASO named afinersen (or ASO5-2), which was generated by introducing 2'-*O*-methoxyethyl (2'-MOE) at the 2'-position of the ribose sugar. This ASO selectively targets repeat-containing *C9orf72* variants 1 and 3, not functional variant 2. Pre-clinical investigations have shown that intracerebroventricular injection of afinersen to C9-500 BAC is able to effectively reduce the levels of poly (GP) in the cortex of transgenic mice in a dose-dependent manner. Afinersen treatment in mice showed that the therapy was safe and well-tolerated, with some evidence of potential pre-clinical benefit. Therefore, it was tested on *C9ORF72*-ALS patients harbouring 2,400 repeats, showing disease phenotypes, and having poly (GP) DPR in CSF ranging 0.01-0.03 ng/ml. Patients were first injected with ASO5 at 0.5, 1 and 1.5 mg/kg via intrathecal administration followed by repeat injections of ASO5 at 2 mg/kg. Repeated injection at 2 mg/kg led to an 80% reduction in the level of poly (GP). They also reported that consistent levels of the ASO in the cerebrospinal fluid and sustained reduction of poly GP were achieved through the administration of a 2 mg/kg dose every three months (Tran et al., 2022).

1.9.4.1.2. RNA Interference-mediated Degradation of *C9ORF72*-Repeat Transcripts

RNA interference (RNAi) is a biological process that plays a role in regulating gene expression. It is a mechanism by which small double stranded RNA molecules including short interfering RNAs (siRNAs), shRNAs, and microRNAs degrade specific mRNA molecules thereby preventing their translation into protein in the cytoplasm. This makes RNAi a potential therapeutic strategy for *C9ORF72*-ALS/FTD by targeting the *C9orf72* repeat expansion in the nucleus and reducing the expression of the toxic RAN protein (Castanotto and Rossi, 2009). One example study using single stranded siRNA to inhibit *C9orf72* sense or antisense repeat transcripts within the nucleus of patient-derived fibroblasts found that siRNA transfection inhibited the formation of nuclear RNA foci (Hu et al., 2017). In another example, AAV-miRNA targeting *C9orf72* was engineered to investigate its silencing efficacy in patient-derived neurons and the BAC transgenic (*C9orf72_3*) line 112 mouse model of ALS. They found that AAV-miRNA was able to efficiently reduce *C9orf72* mRNA levels up to 60% and 25% in patient-derived neurons and astrocytes from FTD patients, respectively. This study provides further evidence that *in vivo* treatment of AAV-miRNA can lead to a significant reduction in *C9orf72* mRNA and the sense intronic transcripts in the injected area as well as the formation of RNA foci in the nucleus or cytoplasm (Martier et al., 2019).

1.9.4.1.3. CRISPR

The CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9) gene-editing system is another powerful tool that has been explored as a potential therapy for *C9ORF72*-ALS/FTD. The goal of using CRISPR in this context is to reduce *C9orf72* induced toxicity (Heidenreich and Zhang, 2016). Researchers have used CRISPR to cut repeat expansions from the specific sequence of the *C9orf72* gene without affecting exons to produce *C9orf72* proteins and then subsequently repairing the gene (Jiang and Ravits, 2019). Proof-of-concept studies in cells and animal models have shown that CRISPR/Cas9 technology can reduce pathological hallmarks of the disease such as reduction in poly DPR and RNA foci formation. For example, a study published in 2022 used AAV9-mediated CRISPR/Cas9 to excise *C9orf72* repeat expansions using guide RNAs. They designed four gRNAs were designed flanking the HRE and exon 2, with two gRNAs targeting the 3' DNA sequence downstream of the HRE and two targeting the upstream, 5' sequence (gRNA1, gRNA2, gRNA3 and gRNA4).

To evaluate gRNA efficiencies, HEK293T cells were co-transfected with a plasmid expressing *S. pyogenes* Cas9 (SpCas9) and another plasmid expressing two gRNA combinations flanking the repeat in groups of gRNAs 1 and 3, gRNAs 1 and 4, gRNAs 2 and 3, or gRNAs 2 and 4, respectively. They found that gRNAs 2 and 3 and guide RNAs 2 and 4 in primary mouse cortical neurons and animal models, which led to a 38% and 29% reduction of nuclear RNA foci, respectively. Both gRNAs also reduced poly (GP) expression by 50% in mouse primary neurons (Meijboom et al., 2022). Same study also evaluated their therapeutic strategy in an animal model of *C9ORF72*-ALS/FTD. CRISPR-mediated targeting of the repeat expansion in C9-500 BAC and BAC111/Cas9 mice significantly decreased poly (GP) and (GR) expression and reduced nuclear foci (Meijboom et al., 2022). In another recent study, CRISPR was used to delete the specific promotor sequence of the *C9orf72* gene in patient derived neuron cells. Removing of 134 nucleotides from the transcription initiation site of sense repeat expansions resulted in elimination of poly (GA), poly (GP) and poly (GR) DPR production. Additionally, deletion of the promotor region alleviated axonal degeneration of motor neurons (Krishnan et al., 2020). Overall, the findings so far are promising and suggest that CRISPR-mediated therapy may be a promising new approach for treating these debilitating neurodegenerative diseases. However, despite its many advantages, there are several potential disadvantages and concerns associated with its use. One of the main concerns with CRISPR-Cas9 is the possibility of off-target effects. These happen when the Cas9 enzyme binds to and cleaves DNA sequences that are not the intended target, and such occurrences have been noticed with a frequency of 50% or higher (Zhang et al., 2015b). Another limitation is the absence of a protospacer adjacent motif sequence near the desired gene loci (Uddin et al., 2020). Additional concerns that have been raised include issues such as immunogenicity and DNA damage toxicity (Yang et al., 2021b).

1.9.4.1.4. Small Molecule Compounds

Small molecules represent another approach that targets *C9orf72* repeat transcripts by binding to their secondary structures, such as G-quadruplexes, R-loops, or hairpins. This binding results in inhibition of RAN translation and a reduction in RBP sequestration. For example, TMPyP4 (5,10,15,20-tetra(N-methyl-4-pyridyl) porphyrin), which is cationic porphyrin, has been shown to interact with G₄C₂ repeat transcripts to inhibit sequestration of RBPs on repeat transcripts.

In another study, TMPyP4 targeting G-quadruplexes was neuroprotective in *C9ORF72*-ALS/FTD *Drosophila* by correcting defects in the nucleocytoplasmic transport pathway (Zhang et al., 2015a) as well as decreasing the production of RNA foci and DPRs in iPSC-derived neurons from *C9ORF72* patients (Simone et al., 2018). In another study, a small molecule that can bind hairpins to selectively target (G₄C₂) repeat expansions was able to prevent RAN translation and the formation of RNA foci thus alleviating DPR-associated toxicity in *in vitro* models of *C9ORF72*-ALS/FTD (Wang et al., 2019).

1.9.4.2. Targeting DPRs

1.9.4.2.1. Immunotherapy Strategies

Immunotherapy or an antibody based therapeutic approach is another treatment strategy for *C9ORF72*-ALS/FTD. This approach utilises antibodies to slow down or halt the production of toxic DPRs (Hautbergue et al., 2021). One form of immunotherapy for ALS being researched is the use of α -GA antibodies in patient fibroblasts and primary neurons. The authors found that α -GA slowed down toxic poly (GA) production and prevented the spreading of GA accumulation in the cerebellum (Zhou et al., 2017). Administration of the α -GA vaccine into *C9ORF72*-ALS/FTD mice resulted in a reduction in poly(GA) expression, TDP-43 mislocalisation, and motor neuron degeneration (Zhou et al., 2020b). Another study found that treatment with α -GA antibodies reduced poly (GA) production, prevented motor neuron degeneration, and extended the life span of C9-500 BAC mice (Nguyen et al., 2020). Anti-GA based immunotherapy has been investigated as a method of inhibiting cell to cell transmission of poly (GA) DPRs. Neurons treated with 5F2-anti GA antibody showed a reduction in TDP-43 proteinopathy and poly (GA) DPR uptake in patient cells compared to controls (Khosravi et al., 2020). This suggests that α -GA antibodies may be a promising therapeutic approach for individuals with the *C9orf72* mutation and potentially for other forms of ALS. For example, in a recent study, researchers developed a new monoclonal antibody aimed at targeting misfolded SOD1. The findings indicated that administering this antibody intrathecally, starting 2 to 4 weeks prior to the onset of symptoms, could prolong life expectancy and postpone motor deficits in a SOD1 mouse model of ALS (Minamiyama et al., 2023).

1.9.4.2.2. Clearance of DPRs

Clearance of DPRs using the small heat shock protein HSPB8 is another therapeutic intervention that has been investigated. It has been shown that overexpression of the HSPB8 protein reduces the production of poly (GA), poly (GP), poly (GR), poly (PR) and poly (PA) DPRs in an NSC34 cell line via promoting the autophagy pathway (Cristofani et al., 2018). More recently, it has been discovered that an increase in the activity of the protein kinase A inhibitor (PKA) can promote production of the toxic DPRs. Therefore, researchers have proposed using H89 which is a specific inhibitor of PKA. Treatment with H89 in patient derived iPSC motor neurons and a *Drosophila* model of *C9ORF72*-ALS/FTD led to reduction of poly (GP) DPRs and conferred neuroprotective effect (Licata et al., 2022). In another study, it was reported that activation of PAR1 (activate protease-activated receptor) via 3K3A-APC, which is an anticoagulation-deficient form of activated protein C, reduces DPRs, improves survival of motor neurons from *C9ORF72*-ALS and sporadic ALS cases as well as reversing the depletion of TDP-43 from the nucleus (Shi et al., 2019).

1.9.4.3. Targeting RAN Translation

RAN translation is considered one of the main drivers of pathogenic mechanisms in *C9ORF72*-ALS/FTD and other microsatellite expansion disorders such as spinocerebellar ataxia type 8 and myotonic dystrophy (Becker et al., 2017). Therefore, developing therapeutic strategies targeting this mechanism would help in developing potential clinical candidates for the treatment of neurological diseases associated RAN translation. A recent study revealed that RNA-dependent protein K (PKR) is involved in regulating RAN translation and activation of PKR leads to elevated DPR production. Therefore, modulating PKR activity with potential therapies holds promise to ameliorate RAN associated toxicity. In this respect, the therapeutic efficiency of AAV- PKR-K296R and metformin were evaluated in *in vitro* and *in vivo* models of *C9ORF72*-ALS/FTD. Both approaches show that they can inhibit PKR activity with subsequent reduction in RAN protein levels. Indeed, they improved disease related phenotypes and pathological features of transgenic C9-500 BAC mice line (Zu et al., 2020).

1.9.4.4. Targeting Downstream Mechanisms

There is strong evidence showing that targeting the downstream mechanisms such as nucleocytoplasmic transport and stress granules in *C9ORF72*-ALS/FTD constitutes a potential therapeutic approach (Mayl et al., 2021). For example, partial knockdown of SRSF1, which acts as a nuclear export adaptor, using an RNAi strategy has been shown to inhibit the nuclear export of pathogenic sense and antisense repeat transcripts to the cytoplasm. This in turn resulted in inhibition of RAN translation and subsequently a decrease of sense and antisense DPR production in the cytoplasm and improved survival of iPSC derived neurons and rescue motor function of a *C9ORF72 Drosophila* model harbouring a (G₄C₂)₃₆ repeat expansion (Hautbergue et al., 2017). A similar neuroprotective effect was also seen in another study that aimed to either deplete the expression of exportin, which plays a role in large protein transport, or overexpressing importin α using RNAi in a *C9ORF72 Drosophila* model expressing a (G₄C₂)₃₀ repeat expansion (Zhang et al., 2015a).

In 2018, it was reported that ataxin-2 impairs the nucleocytoplasmic transport pathway in *C9ORF72*-ALS/FTD through leading to aggregation of stress granule proteins which includes RNA and RBP possessing domains with low complexity (Zhang et al., 2018a). Therefore, different approaches were taken to target ataxin-2 for inhibiting stress granule formation. For example, ASOs targeting ataxin-2 to reduce its expression led to alleviation of the defects in nucleocytoplasmic transport, TDP-43 proteinopathy and neurodegeneration in *C9ORF72*-ALS patient derived neuron as well as the rescue of the locomotor deficit in a *C9ORF72 Drosophila* model expressing (G₄C₂)₃₀ repeats (Zhang et al., 2018a). A similar ASO approach was also undertaken in another study which reported that reducing of the level of ataxin-2 expression ameliorated neurodegeneration caused by TDP-43 aggregation and produced extended survival in a mouse model of *C9ORF72*-ALS (Becker et al., 2017).

1.10. Limitation of Therapeutic Approaches

Although, there is a large body of research conducted in regard to discovering new treatments, pathogenic mechanism, mutations and phenotypes in the literature, we have still been unable to find therapeutic agents that significantly stop the progression of ALS and FTD. There are several challenges that can be attributed as the reasons for the therapeutic failure to date (Rosenfeld and Strong, 2015). Firstly, ALS and FTD are complex diseases and the exact pathophysiological mechanism are still not fully understood. Both ALS and FTD can be difficult to diagnose in their early stages thereby leading to delays in treatment and potentially reducing the effectiveness of therapies. Additionally, ALS and FTD affect multiple systems in the body and exhibit different clinical features as well as progressing at varying rates. This heterogeneity affects the ability to develop effective therapies (Yang et al., 2021a).

Furthermore, biomarkers in the CSF and blood for these diseases are critical for early diagnosis, monitoring of progression, and evaluation of the effectiveness of treatments. Therefore detecting new biomarkers or well-established single biomarkers that accurately reflect ALS/FTD is urgently needed (Witzel et al., 2022). So far, detection of neurofilament and poly (GP) proteins in CSF constitute promising biomarkers to track disease onset and progression (Jiang and Ravits, 2019). There is a constant effort from researchers to find new and more effective therapies that can fundamentally improve the course of these diseases (Hautbergue et al., 2021).

1.11. Cell-Penetrating Peptides

Cell-penetrating peptides (CPPs) are defined as short peptide families that are 5 to 30 amino acids in length (Raucher and Ryu, 2015; Vivès et al., 1997). CPPs are considered as potential cargo delivery macromolecules for DNA, viral vectors and drugs to penetrate inside cells due to their higher penetration ability across the cell membrane (Heitz et al., 2009). The first reported CPP was TAT peptide (trans-activator of transcription) which was produced from the HIV-1 protein in 1988 (Frankel and Pabo, 1988). Soon after, the penetratin peptide derived from the homedomain of antennapedia was discovered in 1991 (Joliot et al., 1991). It was shown that TAT and penetratin peptides successfully entered HELA (Frankel and Pabo, 1988) and nerve cells (mesencephalic cell, spinal cord neurons and striatal cells) (Derossi et al., 1996). From that time, CPPs have been extensively used and over 1850 peptides have been designed. Among them, the functionalities and structures of around 1700 CPPs were characterized (Agrawal et al., 2015).

1.11.1. Classification of CPPs

CPPs are classified into three groups based on their physical and chemical properties. These are amphipathic, cationic and hydrophobic peptides. Most CPPs are cationic such as penetratin and TAT due to being rich in arginine and lysine residues (Guidotti et al., 2017). They bear a positive charge at neutral pH. Cationic peptides show a higher ability to interact with the lipid bilayer of biological membranes and cross the membrane without consuming energy and also, they can escape degradation during endocytosis. The reason for displaying these properties lies in the two cationic residues (Wender et al., 2000). To investigate the effects of the arginine and lysine residues separately, solid state NMR analysis was performed using penetratin comprising 3 arginine and 4 lysine residues. It was found that deletion of arginines from penetratin showed less intracellular entry relative to removing lysines. The result was attributed to the presence of guanidinium head groups which form two hydrogen bonds with membrane components which triggers induction of transport. Since lysine residues do not have this group, the membrane penetrability of lysine is lower than arginine (Su et al., 2009). The second group is amphipathic CPPs which harbour both polar and non-polar amino acids. The amphipathic CPPs are broken down into three subsets which are primary, secondary and proline rich CPPs according to their sequence length and amphipathicity properties. Hydrophobic residues such as alanine and glycine are highly represented in amphipathic CPPs (Milletti, 2012; Zaro and Shen, 2015).

The primary class of amphipathicity CPPs consists of conjugating various domains, components, therapeutics, or other peptides through the formation of covalent bonds. For this reason, these CPPs are called fused or chimeric peptides (Fadzen et al., 2019). The secondary amphipathic CPPs are divided into α -helical peptides, which are composed of polar and non-polar groups on both sides of the helix, and β -sheet peptides which display polar and non-polar residues on each opposite side of a β -sheet. Transportan and VT5 are examples of this peptide class (Milletti, 2012). Another subset of peptide within the amphipathic CPP group is proline rich CPPs which display high intracellular performance. Bac-7 which belongs to the bactanecin family is one of the example of proline rich CPPs (Pujals and Giralt, 2008). The last group is hydrophobic CPPs which are composed of non-polar residues, thereby providing an advantage during cellular translocation of itself and causing low net charge. There are a few structurally and functionally defined CPPs in this group such as PFVYLI and pep-7 CPPs (Marks et al., 2011). Examples of structurally well-defined CPPs are given **Table 1.3**.

Table 1.3. Examples of various CPPs with their detailed information. The table adapted is from (Guidotti et al., 2017; Singh et al., 2018).

Name of CPPs	Sequence	Origin	Class	Reference
HIV-1 TAT protein, TAT ₄₈₋₆₀	GRKKRRQRRPPQ	(HIV-1 TAT protein)	Cationic	(Frankel and Pabo, 1988)
HIV-1 TAT protein, TAT ₄₉₋₅₇	RKKRRQRRR	(HIV-1 TAT protein)	Cationic	(Park et al., 2002)
HIV-1 TAT protein, TAT ₄₇₋₅₇	YGRKKRRQRRR	(HIV-1 TAT protein)	Cationic	(Popiel et al., 2007)
Antennapedia protein transduction domain (called also Penetratin), pAntp(43_58)	RQIKIWFQNRRMKWKK	Antennapedia Drosophila melanogaster	Cationic	(Futaki et al., 2001)
Polyarginines	R _n	Chemically synthesized	Cationic	(Futaki et al., 2001)
ARF(1-22)	MVRRFLVTLRIRACGPPR VRV	Natural	Amphipathic	(Milletti, 2012)
pVEC	LLIILRRRIRKQAHASK	Vascular endothelial cadherin	Amphipathic	(Elmqvist et al., 2006)
Transportan	GWTLSAGYLLGKINLKA LAALAKKIL	Chimeric galanin–mastoparan	Amphipathic	(Pooga et al., 1998)
C105Y	CSIPPEVKFNKPFVYLI	α 1-Antitrypsin	Hydrophobic	(Rhee and Davis, 2006)
PFVYLI	PFVYLI	Derived from synthetic C105Y	Hydrophobic	(Rhee and Davis, 2006)
Pep-7	SDLWEMMMVSLACQY	CHL8 peptide phage clone	Hydrophobic	(Gao et al., 2002)

1.11.2. Intracellular Uptake Mechanisms of CPPs

Over many years, many studies have been conducted to investigate the exact pathways of delivery of CPPs into the cell membrane but the mechanism of peptide translocation still is not elucidated. It was found that CPPs follow different routes depending on their structure, as well as the cell type and cargo type (Palm-Apergi et al., 2012). In this respect, the route of CPPs through the cell membrane is approached via two proposed mechanisms; direct penetration and endocytosis.

1.11.2.1. Direct Penetration

Direct penetration is one of the ways used by CPPs to enter the cell membrane via an energy independent pathway which basically consists of electrostatic interaction between positively charged CPPs with the negatively charged phospholipid layer and proteoglycans acting as a membrane receptor. This interaction results in changes in the membrane structure such as folding of the peptide on the lipid bilayer, thereby affecting the membrane stabilization (Guidotti et al., 2017; Madani et al., 2011). The uptake of CPPs inside the membrane is achieved by several potential mechanisms: ‘carpet-like (Matsuzaki et al., 1996), pore formation (Pouny et al., 1992) and reversed micelle (Derossi et al., 1996)’. Direct penetration occurs at high CPP concentrations. The carpet-like model is derived from formation of a carpet-like structure by coating of the peptide on the membrane surface following the binding of CPPs onto the membrane surface via electrostatic interaction (Madani et al., 2011; Matsuzaki et al., 1996).

The second model is ‘pore formation’ which comprises an explanation of how pores are formed via two subset models; barrel stave and toroidal models. In barrel stave, the hydrophobic side of α -helical amphipathic CPPs interact with the lipid core, which forms the pore’s interior and the hydrophilic side of the CPP in the internal side of the pore creates a channel or continuous layer to enable the CPP to be transported into the cell (Madani et al., 2011; Pouny et al., 1992). In the toroidal model, the CPP penetrates perpendicularly, thereby forming the pores (Guidotti et al., 2017; Kumar et al., 2018). The last proposed model for direct delivery of CPPs is proposed by the formation of reversed micelles in the lipid bilayer which results from an interaction of the hydrophobic side of the CPP with lipid heads on the membrane. The CPP is then enclosed with reversed micelles and enter the hydrophobic side of membrane. The micelles containing the CPP move the micelles through the membrane until the release of the CPP into the liquid part of cytoplasm (Alves et al., 2008; Derossi et al., 1996).

1.11.2.2. Endocytosis

Endocytosis is another mechanism used for the translocation of CPPs. This is an energy required process. Endocytosis is broken down into four main classes; pinocytosis, phagocytosis, clathrin receptor-mediated endocytosis and clathrin-dependent endocytosis. Pinocytosis and phagocytosis are responsible for the uptake of liquid and larger molecules inside the cell by creating small pits on the cell membrane (Koren and Torchilin, 2012). Receptor-mediated endocytosis is well characterized and occurs in all mammalian cells. Basically, this process is initiated by binding of ligand molecules to a specific membrane receptor. This binding leads to the formation of a pit with the help of invagination of the membrane. CPP containing pits are then covered with the clathrin protein. Finally, dynamin, a GTPase enzyme, is attached to the neck of the clathrin coated pit to disassociate it from the membrane (Sawant and Torchilin, 2010). The protein coating is separated as soon as the pit is fissured from the membrane. Additionally, receptors are recycled back to the membrane for the next endocytosis following its disassembly within the cytosolic endosome (Mayor and Pagano, 2007).

Clathrin receptor-dependent endocytosis shares similar steps with clathrin mediated endocytosis such as the formation of pits and fission from membrane (LeCher Julia et al., 2017). Endocytosis is considered as the main mechanism of translocation of CPPs but there is a significant hurdle which is endosomal entrapment affecting the internalization efficiency of CPPs and CPPs with cargo. To avoid enzymatic digestion from the lysosome, CPPs must be recovered during internalization (van den Berg and Dowdy, 2011). In this respect, new strategies have been developed to avoid this problem. For example, transportan 10 (TP-10), which is an amphipathic peptide, was exposed to N-terminal stearylation modification. The modified TP-10 peptide showed higher endosomal escape and internalization efficiency in comparison to unmodified CPPs (Mäe et al., 2009).

1.11.3. Pre-clinical and Clinical Studies

Due to having specific properties, CPPs have been intensively used in various applications. Amongst the important features are transporting various cargo molecules and crossing the cell membrane and it has also been shown that several CPPs achieved crossing of the blood-brain barrier, thereby leading to the evaluation of CPPs as powerful therapeutic molecular tools (Guidotti et al., 2017). Testing has already been initiated via in vivo and in vitro studies.

The promising results obtained from these studies has received great attention by drug companies and this resulted in the evaluation of over 25 CPP therapeutics in clinical trials for several disease such as cancer (Jia et al., 2011) and inflammation (Gurney et al., 2018). The details of pre-clinical and clinical studies are given **Table 1.4 and Table 1.5**. In addition, they are described in the discussion section.

Table 1.4. Pre-clinical studies of CPPs in animal models of various diseases. The table is adapted from (Guidotti et al., 2017).

CPPs-Cargo	Animal model	Application	Reference
PTD-FNK	Cerebral ischemia in gerbils	Cerebral ischaemia	(Asoh et al., 2002)
PTD-HA-Bcl-XL	MCAO in mice	Cerebral ischaemia	(Cao et al., 2002)
TAT-NBD	Infection-sensitized HI brain injury in neonatal rats	Perinatal infection in HI brain injury	(Yang et al., 2013)
TAT ₄₈₋₅₇ -BH4	hSOD1G93A mice	ALS	(Martorana et al., 2011)
TAT-JBD20 (D-JNKI-1)	TgCNRD8 mice	AD	(Sclip et al., 2014)
Antp-NBD	mdx mice	DMD	(Peterson et al., 2011)
TAT ₄₈₋₆₀ -BH4	Myocardial IR injury in mice	Myocardial IR injury	(Boisguerin et al., 2011)
RI-TAT-p53C	Terminal peritoneal carcinomatosis/lymphoma in mice	Cancer	(Snyder et al., 2004)
MPG-8/siRNA	Xenografted tumor in mice	Cancer	(Crombez et al., 2009)

Table 1.5. Clinical studies of CPP based therapies in various diseases. The table is adapted from (Guidotti et al., 2017).

Company	Compound	CPP-Cargo	Application	Status	ClinicalTrials.gov ID
Auris Medical	AM-111	TAT-JBD20 (D-JNKI-1)	Hearing loss	Phase II completed 2014	NCT00802425
CellGate, Inc	PsorBan	R7-cyclosporin A	Psoriasis	Phase IIb discontinued 2003	N/A
Capstone Therapeutics	AZX100	PTD4-HSP20 phosphopeptide	Scar prevention/reduction	Phase IIa completed 2012	NCT00451256 NCT00892723 NCT00811577
CDG Therapeutics, Inc.	p28	p28	Cancer	Phase I completed 2014	NCT00914914
KAI Pharmaceuticals	KAI-1678	TAT-ePKC inhibitor	Pain: postherpetic neuralgia, spinal cord injury, postoperative	Phase II completed 2011	NCT01106716 NCT01135108 NCT01015235
Revence Therapeutics, Inc.	RT002	TransMTS1-botulinum toxin A	Glabellar lines	Phase I/II completed 2016	NCT02303002
Sarepta Therapeutics	AVI-4658	N/A	Duchenne muscular dystrophy	Phase I/II completed 2010 and 2015	NCT00159250 NCT00844597
Sarepta Therapeutics				Phase III recruiting	NCT02255552
Xigen SA	XG-102	TAT-JBD20 (D-JNKI-1)	Inflammation	Phase I completed 2012	NCT01570205
Paediatric Brain Tumor Consortium	P28	P28	Central Nervous System Tumors	Phase I completed 2013	NCT01975116
Capstone Therapeutics	AZX-100	PTD4- HSP20 phosphopeptide	Wound healing	Phase II completed 2012	NCT00825916

1.12. Hypothesis and Aims

I hypothesise that intracellular delivery or expression of SRSF1-based CPPs that comprise the NXF1-binding site (Tintaru et al. 2007) would inhibit the interaction of endogenous SRSF1 with NXF1 preventing the nuclear export of pathological *C9orf72* repeat transcripts and the production of toxic DPRs as therapeutic strategy in cell and animal models.

The overall aim of my study is to design and evaluate the potential efficiency of a novel ALS-therapeutic concept based on the delivery of SRSF1-based cell permeable peptides (CPPs) and viral mediated gene therapy using a self-complementary adeno-associated virus serotype 9 (scAAV9) expressing SRSF1-CPPs that inhibit the SRSF1:NXF1 dependent nuclear RNA export pathway. Cell and mouse models of *C9ORF72*-ALS/FTD will be used to assess the efficacy of the SRSF1-CPP as a potential novel therapeutic approach.

- In the first results chapter, I will design, express and purify sufficient amount of recombinant SRSF1-linker CPP by performing several optimization experiments such as testing different types of constructs, bacterial culture medium and chromatography media.
- In the second results chapter, I will validate the functionality of recombinant SRSF1-linker CPP in HEK293T cells model of *C9ORF72*-ALS/FTD by investigating the potential nuclear export inhibition of *C9orf72*-repeat transcripts and inhibition of the production of sense and antisense DPRs as well as potential rescue of the DPR-associated cytotoxicity by measuring cell proliferation. Further, I will move into a short *in vivo* study to evaluate whether injection of the SRSF1-linker CPP inhibits DPR production in the CNS of *C9ORF72*-ALS/FTD mice.
- In the third result chapter, I will first engineer scAAV9 viral vectors expressing SRSF1-CPPs under RNA polymerase II or III promoters. The functionality of the generated viral plasmids will be then validated in the HEK293T cells model of *C9ORF72*-ALS/FTD. I will next perform a pilot proof-of-concept efficacy study to evaluate the potential DPR inhibitory effect of scAAV9 viruses in the CNS of *C9ORF72*-ALS/FTD mice.

2. Chapter 2 - Materials and Method

2.1. Materials

2.1.1. Strains and Cell Lines

2.1.1.1. Bacterial Cell Lines

Escherichia coli (E. coli) DH5-alpha

Genotype: F⁻ ϕ 80lacZ Δ M15 Δ (lacZYA-argF)U169 recA1 endA1 hsdR17(rK⁻, mK⁺) phoA supE44 λ ⁻ thi-1 gyrA96 relA1.

DH5-alpha (ThermoFisher Scientific) cells were used for cloning experiments based on their high transformation efficiency and the presence of a recA1 mutation which reduces unwanted recombination events on plasmids of interest.

Escherichia coli (E. coli) BL21 (Rosetta, RP and RIL)

Genotype: B F⁻ ompT gal dcm lon hsdSB(rB⁻mB⁻) λ (DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB⁺]K-12(λ S) pLysS[T7p20 orip15A](CmR).

E. coli BL21 competent cells (Novagen) derivatives were used hosting for plasmid. BL21 (DE3) cells harbour the T7 polymerase gene which can be induced by IPTG. This allows high efficiency recombinant gene expression under a T7 promoter. BL21 cells also either carry plasmids that confer resistance to chloramphenicol and allow expression of tRNAs recognising rare human codons for AGA, AGG, CCC (RP), AGA,AGG,AUA,CUA (RIL) and tRNAs for AGG, AGA, AUA, CUA, CCC, GGA (Rosetta).

2.1.1.2. Mammalian Cell Lines

HEK293T – Human Embryonic Kidney Cells

Human embryonic kidney cells 293T, a highly transfectable derivative of the 293 cells, were used for all transfection experiments.

2.1.2. Plasmids and Vectors

The details of all plasmids used in this study are given in **Table 2.1**.

Table 2.1. Plasmid source and description.

Plasmid Construct	Description	Source	Host Expression
pGEX-6P1-GST	- Ampicillin Resistant - Vector backbone using for GST-tagged control and SRSF1-linker CPPs	GE Healthcare	Bacterial
pGEX-6P1 SRSF1 V5 TAT PTD	Expressed the GST-SRSF1-linker CPP	Prof Guillaume Hautbergue	Bacterial
pGEX-6P1 V5 TAT	Expressed the GST-control CPP	Aytac Gul	Bacterial
pET24b -GB1-(His) ₆	- Kanamycin resistant 5.5 kbp - Vector backbone using for (His)₆ tagged SRSF1	Prof Guillaume Hautbergue	Bacterial
pET24b SRSF1 V5 TAT PTD	Expressed the (His)₆-SRSF1-linker CPP	Prof Guillaume Hautbergue	Bacterial
pET24b SRSF1 V5 W TAT PTD	Expressed the (His)₆-SRSF1-linker CPP	Prof Guillaume Hautbergue	Bacterial
pcDNA3.1 G ₄ C ₂ x38-V5	Expressed the sense dipeptide repeat proteins (DPRs)	Dr. Lydia Castelli (Castelli et al., 2023)	Mammalian
pcDNA5 3.1 C ₄ G ₂ x39-V5	Expressed the antisense dipeptide repeat proteins (DPRs)	Dr. Lydia Castelli (Castelli et al., 2023)	Mammalian

scAAV9_H1- SRSF1 linker CPP_EF1 α -eGFP	Expressed SRSF1-linker CPP under H1 promotor	Aytac Gul	Mammalian
scAAV9_CBH- SRSF1 Linker CPP_EF1 α -eGFP	Expressed SRSF1-linker CPP under CBH promotor	Aytac Gul	Mammalian
scAAV9_H1-SRSF1- RRM2 α 1 CPP_EF1 α -eGFP	Expressed SRSF1-RRM2α1 CPP under H1 promotor	Aytac Gul	Mammalian
scAAV9_CBH-SRSF1- RRM2 α 1 CPP_EF1 α -eGFP	Expressed SRSF1-RRM2α1 CPP under CBH promotor	Aytac Gul	Mammalian
pGEX-6P1-NXF1	Express GST-NXF1 protein	Prof Guillaume Hautbergue	Bacterial
pGEX-6P1-NXF2	Express GST-NXF2 protein	Prof Stuart Wilson	Bacterial
pGEX-6P1-NXF3	Express GST-NXF3 protein	Prof Stuart Wilson	Bacterial
P3X-Flag NXF1	Express Flag-NXF1 protein	Prof Guillaume Hautbergue	Bacterial
pET9a-p15	Contains p15 used for co-expression with TAP/NXF1	Prof Guillaume Hautbergue	Bacterial

2.1.3. Primers

Polymerase Chain Reaction (PCR) Primers

All primers used for molecular cloning in this study were ordered from Sigma-Aldrich. The name of each primer, primer direction, primer sequence and primer origins are provided in **Table 2.2**. Colour of text labels correspond to the DNA oligonucleotide elements highlighted in the sequences.

Table 2.2. PCR primer sequence for molecular cloning.

Primer	Primer Direction	Sequence	Supplier
SRSF1-89-CPP BamHI 5 scAAV Linker CPP_AAV H1	Forward	GGC GGATCC GAAGCCACC ATG CCGCGCAGCGGCCGCGGCACCG GCCGCGGTGGGGGCGGCGGTGGAGGTGGCGGAGCCCCGAGAGGCC GCTATGGACCGCCCAGCCGCCGGAGCGAA GGC	Sigma-Aldrich
SRSF1-120-V5-TAT-PTD-stp HindIII scAAV Linker CPP_AAV H1	Reverse	CCCGCC AAGCTT AAAAAATTA CCTCCGGCGCTGTCTGCGCTTTT GCGACCATA TCCGCC GGTGCTATCGAGGCCAGCAGCGGGTTGG GAATCGGTTTCCC ACCGCC TTCGCTCCGGCG	Sigma-Aldrich
SRSF1-89-CPP- EcoRI scAAV Linker CPP_AAV CBH	Forward	GGCGGG GAATTC GAAGCCACC ATG CCGCGCAGCGGCCGCGGCAC CGGCCGCGGTGGGGGCGGCGGTGGAGGTGGCGGAGCCCCGAGAGG CCGCTATGGACCGCCCAGCCGCCGGAGCGAA	Sigma-Aldrich
SRSF1-120-V5-TAT-PTD-stp XbaI 3 scAAV Linker CPP_AAV CBH	Reverse	CCCGCC TCTAGA TTA CCTCCGGCGCTGTCTGCGCTTTT GCGACC TA TCCGCC GGTGCTATCGAGGCCAGCAGCGGGTTGGAATCGGT TTCCC ACCGCC TTCGCTCCGGCGGCTGGG	Sigma-Aldrich

Table 2.2. PCR primer sequence for molecular cloning.

Primer	Primer Direction	Sequence	Supplier
SRSF1-132-V5-TAT-PTD - BamHI5 scAAV RRM2α1 CPP_AAV H1	Forward	GGCGGG GGATCC GTTTAGTGAACCGTCAGAAGCCACC ATG GGC AGCTGGCAGGATCTGAAAGATCATATGCGCGAAGCC GGCGGT G GAAACCGATTCCCAAC	Sigma-Aldrich
SRSF1-144-V5-TAT-PTD-stp HindIII scAAV RRM2α1 CPP_AAV H1	Reverse	CCCGCC AAGCTT AAAAAATTA CCTCCGGCGCTGTCTGCGCTTTTT GCGACCATA TCCGCC GGTGCTATCGAGGCCAGCAGCGGGTTGG GAATCGGTTTCCC ACCGCC GGCTTCGCGC	Sigma-Aldrich
SRSF1-132-V5-TAT-PTD - EcoRI scAAV RRM2α1 CPP_AAV CBH	Forward	GGCGGG GAATTC GAAGCCACC ATG GGCAGCTGGCAGGATCT GAAAGATCATATGCGCGAAGCC GGCGGT GGGAAACCGATTCCC AAC	Sigma-Aldrich
SRSF1-144-V5-TAT-PTD-stp XbaI3 scAAV RRM2α1 CPP_AAV CBH	Reverse	CCCGCC TCTAGA TTA CCTCCGGCGCTGTCTGCGCTTTTTGCGACC ATA TCCGCC GGTGCTATCGAGGCCAGCAGCGGGTTGGGAATCG GTTTCCC ACCGCC GGCTTCGCGC	Sigma-Aldrich
X6P-CPP_Delta_BamHI5	Forward	GGCGGGGGATCCGGCGGTGGGAAACCGATTC	Sigma-Aldrich
X6P-CPP_Delta_BamHI3	Reverse	CCCGCCGGATCCCAGGGGCCCCCTG	Sigma-Aldrich
Sex 1 sry	Forward	TTGTCTAGAGAGCATGGAGGGCCATGTCAA	Sigma-Aldrich
Sex 1 sry	Reverse	CCACTCCTCTGTGACACTTTAGCCCT CGA	Sigma-Aldrich
C9-GT-F	Forward	5'-AGT TGG GTC CAT GCT CAA CAA-3'	Sigma-Aldrich
C9-GT-R	Reverse	5'-ACT GTT CTA GGT ACC GGG CT-3'	Sigma-Aldrich
Vgll4-F	Forward	5'-TTG GAT GGA GAA GGA TGG AG-3'	Sigma-Aldrich
Vgll4-R	Reverse	5'-GTC TCC ACA AGC CCA TGA GT-3'	Sigma-Aldrich

Quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR) Primers :

All qRT-PCR primers used in this study are outlined in **Table 2.3** with their name, direction, restriction sequence, R² value and origins.

Table 2.3. Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) primers.

Primer	Primer Direction	Sequence	R2 value	Supplier
Human <i>U1 snRNA</i>	Forward	5-CCATGATCACGAAGGTGGTT-3'	0.999	Sigma-Aldrich
Human <i>U1 snRNA</i>	Reverse	5'-ATGCAGTCGAGTTTCCCACA-3'		Sigma-Aldrich
C9-RAN reporter	Forward	5'-GGGCCCTTCGAACCCCCGTC-3'	0.987	Sigma-Aldrich
C9-RAN reporter	Reverse	5'GGGAGGGGCAAACAACAGAT-3'		Sigma-Aldrich
Human <i>NXF1</i>	Forward	5'-CGTTGTCCTGAATCGCAGAAGC-3'	0.997	Sigma-Aldrich
Human <i>NXF1</i>	Reverse	5'-GTTGGGTGCCTTCTGAACAATGC-3'		Sigma-Aldrich
Human <i>NXF2</i>	Forward	5'-CTGCTTCTCCTTGGCTATTCCC-3'	0.994	Sigma-Aldrich
Human <i>NXF2</i>	Reverse	5'-CGTTTTGTGCGCCTCAGCAGTT-3'		Sigma-Aldrich
Human <i>NXF3</i>	Forward	5'-CCTTCTCGGATACCTCCAGCAA-3'	0.995	Sigma-Aldrich
Human <i>NXF3</i>	Reverse	5'-GGTTGGTAACCTCTTGTGGGCT-3'		Sigma-Aldrich

Table 2.3. Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) primers.

Primer	Primer Direction	Sequence	R2 value	Supplier
Mouse <i>NXF1</i>	Forward	5'-CCGAGAACCTAAAGAGTCTGGTC-3'	0.998	Sigma-Aldrich
Mouse <i>NXF1</i>	Reverse	5'-TCGGACAGGATTCTGAGGGTTG-3'		Sigma-Aldrich
Mouse <i>NXF2</i>	Forward	5'-CCAGAAAGCTCTGGACCTCGAA-3'	0.989	Sigma-Aldrich
Mouse <i>NXF2</i>	Reverse	5'-CTGTAGTGTGGCAACCATGCAG-3'		Sigma-Aldrich
Mouse <i>NXF3</i>	Forward	5'-ACCAACTCTCTGTGGTCCTGAAG-3'	0.991	Sigma-Aldrich
Mouse <i>NXF3</i>	Reverse	5'-CTGCTTCAGGAAGTGTAGGACC-3'		Sigma-Aldrich
H1cas2	Forward	5'- TAAGTTGGGTAACGCCAGGG-3'	0.995	Sigma-Aldrich
H1cas2	Reverse	5'- CGGATCCGAGTGGTCTCATAC-3'		Sigma-Aldrich
eGFP	Forward	5'- AGCTGAAGGGCATCGACTTC-3'	0.997	Sigma-Aldrich
eGFP	Reverse	5'- TGCTCAGGTAGTGGTTGTCGG-3'		Sigma-Aldrich
BGH	Forward	5'- TGTGCCTTCTAGTTGCCAGC-3'	0.992	Sigma-Aldrich
BGH	Reverse	5'- TAGAAGGCACAGTCGAGG-3'		Sigma-Aldrich

2.1.4. Media and Antibiotics for Bacterial and Mammalian Cell Lines

2.1.4.1. *E.coli*

LB (Luria-Bertani broth)

- 10 g/l. Tryptone, 5 g/l. yeast extract, 10 g/l. NaCl, (pH 7.0)
- Used for growing bacteria in liquid media.

LB Agar, Miller (Fisher Scientific)

- 10 g/l. Tryptone, 10 g/l. NaCl, 5 g/l. Yeast extract, 15 g/l. of agar
- Used for growing transformed bacteria on solid media.

TB (Terrific Broth)

- 12 g/l. Tryptone, 24 g/l. Yeast extract, 2.31 g/l. KH_2PO_4 , 12.54 g/l. K_2HPO_4 , 4 ml Glycerol
- Used for recombinant protein expression.

SB (Super Broth)

- 35 g/l. Tryptone
- 20 g/l. Yeast extract
- 5 g/l. NaCl
- 5 ml NaOH (1M)
- Used for recombinant protein expression.

Antibiotics used in this study are shown in **Table 2.4**. They were diluted in LB or TB media or agar plates at 1/1000 dilution except for LB + Kanamycin agar plates in which the final concentration is 30 mg/ml.

Table 2.4. Antibiotics used for bacterial cell culture.

Antibiotic	Stock Concentration (mg/ml)	Dissolved in
Ampicillin	100	H ₂ O
Kanamycin	50	H ₂ O
Chloramphenicol	34	100 % Ethanol

2.1.4.2. Mammalian Cell Lines

Dulbecco's Modified Eagle's Medium (DMEM) – high glucose (Sigma)

- Supplemented with 10 % v/v (Tetracycline-Free) Fetal Bovine Serum (FBS) (Sigma) and 1 % v/v Penicillin/Streptomycin (Penstrep) (Lonza).
- Used for culturing HEK293T cell lines.

Phosphate Buffered Saline (PBS)

- Used for washing cells

Opti-MEM

- Used for plasmid transfection

2.1.5. Restriction and DNA Modifying Enzymes

Restriction and DNA modifying enzymes used in this study are provided in **Table 2.5**. All enzymes were purchased from ThermoFisher Scientific and utilised according to the manufacturer's instructions.

Table 2.5. Restriction and DNA modifying enzymes.

Restriction enzyme	DNA modifying enzyme
BamHI	T4 Ligase
XbaI	Herculase DNA polymerase
XhoI	
HindIII	
EcoRI	

2.1.6. Antibodies

All primary and secondary antibodies used in this study and their details (antibody specificity, molecular weight, source, species, type, dilution, application) are shown in **Tables 2.6 and 2.7**.

Table 2.6. Primary antibodies used in western Blotting and immunofluorescence microscopy.

Antibody	Molecular weight (kDa)	Host Species	Dilution and Application	Source
α -Tubulin	52	Mouse, Monoclonal	WB: 1/5000	Santa Cruz sc-32293
β - III Tubulin (Tuj 1)	55	Chicken, Polyclonal	IF: 1/1000	Abcam, ab41489
α -V5	N/A	Mouse, Monoclonal	WB: 1/5000 IF: 1/1000	Invitrogen R96025
α -Flag clone M2	13	Mouse, Monoclonal	WB: 1/2000 IF: 1/800	ThermoFisher MA1-980
α -SRSF1	27	Rabbit, Monoclonal	WB: 1/1000	Abcam, ab38017
α -GFP	N/A	Chicken, Polyclonal	WB: 1/2000	Abcam, ab13970
Tap/NXF1	65	Mouse, Monoclonal	WB: 1/2000	Abcam, ab50609
NXF2	72	Rabbit, Polyclonal	WB: 1/2000 IF: 1/500	Proteintech 13275-1-AP
NXF3	60	Rabbit, Monoclonal	WB: 1/2000	Proteintech 12185-1-AP
SSRP1	81	Mouse, Monoclonal	WB: 1/500	Abcam, ab26212
Vimentin	N/A	Chicken, polyclonal	IF: 1/5000	Sigma-Aldrich AB5733

Table 2.7. Secondary antibodies used in western blotting and immunofluorescence microscopy.

Antibody	Host Species	Dilution	Source
Goat-Anti Chicken Alexa Fluor 647	Chicken	IF: 1/1000	Abcam, ab150171
Goat-Anti Mouse Alexa Fluor 488	Mouse	IF: 1/1000	Abcam, ab150113/ab150077
Anti-Mouse-IgG HRP conjugated	Goat	WB: 1/5000	Promega, W4021
Anti-Rabbit- IgG HRP conjugated	Goat	WB: 1/5000	Promega, W4011
Anti-Chicken-IgG HRP conjugated	Rabbit	WB: 1/5000	Promega, W1351

2.1.7. Molecular Biology Kits

The name of each molecular biology assay kits and their source details that were used in this are shown in **Table 2.8**.

Table 2.8. Molecular biology assay kit details.

Kit	Source
QIAprep Plasmid Spinprep Kit	QIAGEN, 27106
Plasmid Miniprep Kit	QIAGEN, 12123
Plasmid Plus Midiprep Kit	QIAGEN 12143
Plasmid Plus Megaprep	QIAGEN, 12118
QIAquick Gel Extraction Kit	QIAGEN, 28704
GenElute™ Mammalian Genomic DNA Miniprep Kit	Sigma-Aldrich, G1N350
Pierce™ Quantitative Colorimetric Peptide Assay	ThermoFisher Scientific, 23290
DNeasy Blood & Tissue Kit	QIAGEN, 59504

2.1.8. Equipment

- Probe sonicator (SONICS)
- High speed centrifuge Avanti J-E series (Beckman Coulter)
- AKTA chromatography system including fraction collector, UV monitor and detector (GE HealthCare)
- Ion-exchange chromatography columns: Sulphopropyl (SP) and Mono-S (GE HealthCare)
- Affinity chromatography columns: Cobalt (Talon, Hyclone) and Glutathione S-transferase (GST) resin
- Heparin adsorption column (GE HealthCare)
- Superdex 30 PG size exclusion chromatography column (GE HealthCare)
- UV/VIS spectrophotometer (190-800nm) (Biochrom)
- Amicon stirred cells for sample concentration (Merck)
- Whatman filtration system (GE HealthCare)
- Incubator shaker for large scale bacteria culture (New Brunswick™ Innova)
- G-box for visualizing western blot and agarose gel (SYNGENE)
- RT-PCR and RT-qPCR (Bio-Rad)
- Protein Quantification Equipment; electrophoresis chambers, power supplies, SDS-PAGE gel cassettes (Bio-Rad)
- Semi-dry western blot transfer machine (Analytikjena)
- Agarose gel electrophoresis (Bio-Rad); electrophoresis chambers, power supplies
- NanoDrop™ 1000 Spectrophotometer (ThermoFisher Scientific)
- High Content Screening System (PerkinElmer Opera Phenix™)
- The PHERAstar® FSX Microplate reader (BMG LABTECH)
- MESO QuickPlex SQ 120 MSD-Elisa (MESO SCALE DISCOVERY)
- Cryostat-Microtome (Leica BIOSYSTEM)
- Precellys® Evolution - Tissue Homogenizer (Bertin Instruments)
- AxioPlan 2 fluorescence microscope (Zeiss)
- InCell Analyzer 2500 HS (GE HealthCare)

2.2. Methods

2.2.1. Maintenance of Plasmids

2.2.1.1. Competent E.coli Cells Preparation

Induced competent cells, which have two types of chemical and electro - competence, have been largely preferred due to their ability to easily take foreign DNA into the cell by creating temporary pores in cell membrane. The competent cells were produced chemically. The experimental steps for generation of chemically competent cells are explained below:

Day 1

50 ml LB was inoculated with bacterial strains (DH5 alpha, BL21 Rosetta, RIL or RP) and incubated at 37°C overnight in a shaking incubator at 180 rpm.

Day 2

400 ml LB broth was inoculated with bacteria culture grown overnight to OD= 0.05 (optical density) then incubated at 37°C in a shaking incubator (180 rpm) until optical density had reached OD₆₀₀ ~ 0.48. Sterile TfbI and TfbII buffers were prepared (see **Table 2.9**) and kept on ice. Once the optical density had been reached, the bacterial cells were centrifuged for 10 minutes at 6000 g, 4°C. The pellet was then resuspended in 50 ml TfbI (pH 5.8) buffer by inverting gently and left for 10 minutes on ice followed by additional centrifugation for 10 minutes at 6000 g, 4°C. The supernatant was discarded, and the pellet was reconstituted in 20 ml TfbII (pH 6.5) buffer and left for 15 minutes on ice. The cells were immediately aliquoted into pre-chilled tubes and snap-frozen in liquid nitrogen.

Table 2.9. The compositions of TfbI and TfbII buffers.

TfbI	TfbII
30 mM Kac	10 mM MOPS
100 mM RbCl ₂	75 mM CaCl ₂
10 mM CaCl ₂	100 mM RbCl ₂
50 mM MnCl ₂	15% v/v Glycerol
15% v/v Glycerol	-

2.2.1.2. Transformation

10 µl ligated DNA was mixed with 90 µl of competent cells and left for 10 minutes on ice. The cells were then exposed to heat shock at 37°C for 5 minutes. 900 µl LB broth was added, followed by incubation at 37°C for 45 minutes-1 hour to grow bacteria. During this incubation time, LB agar plates containing appropriate antibiotics were placed into an incubator at 37°C for drying. The tube also was occasionally inverted to provide oxygen in order to improve bacterial growth. The cells were then pelleted by centrifugation at 10,000 rpm for 1 minute at room temperature (RT). 900 µl of supernatant was discarded and the pellet was resuspended in the remaining supernatant. E.coli cells were spread onto LB agar plates containing appropriate antibiotic (s) and incubated at 37°C overnight.

2.2.1.3. DNA Preps

2.2.1.3.1. Rough Minipreps

Miniprep is a technique used to isolate plasmid DNA from transformed bacteria to determine successful integration of DNA into the host plasmid. Minipreps were performed using a QIAGEN Plasmid Mini Kit according to the manufacturer's instructions. Briefly, bacterial colonies were picked and inoculated in 2 ml LB medium containing appropriate antibiotics at 37°C. The following day, 1.5 ml of overnight bacterial culture was pelleted by centrifugation at 13,000 rpm for 1 minute at RT and the supernatant was discarded. Bacterial pellets were resuspended in 200 µl of resuspension buffer (P1). 200 µl lysis buffer (P2) was added to the sample and mixed gently until the solution turned blue, before being incubated for 5 minutes at RT. 300 µl neutralization buffer (S3) was added, and the sample was inverted until the mixture turned colourless. The lysate was centrifuged at 13,000 rpm for 10 minutes at RT. The supernatant was transferred to a clean tube and 700 µl of isopropanol was added followed by a 10 minutes incubation at RT to precipitate the DNA. Plasmid DNA was pelleted at 13,000 rpm for 10 minutes at RT and the supernatant was carefully discarded. A quick pulse spin was applied to remove the remaining isopropanol in the tubes. The pellet was then air-dried for 10-20 minutes at RT and then reconstituted in 25 µl dH₂O. The final plasmid DNA then was digested with appropriate restriction enzymes and resolved by agarose gel electrophoresis at 80 V for 40-50 minutes. The agarose gel was screened using G-BOX (SYGENE). Positive colonies were subjected to midi-preps for large scale plasmid preparation.

2.2.1.3.2. Spinpreps

Spinprep is a quick and efficient technique to generate high purity DNA by removing salts, DNA polymerase and primers. Spinpreps were performed using a QIAGEN Spinprep Kit according to the manufacture's protocol. Initially, bacterial cultures were grown overnight in LB medium containing appropriate antibiotics. The following day, cultures were pelleted by centrifugation at 13,300 rpm for 1 minute at RT. The supernatant was discarded, and cell pellets were resuspended in 250 µl P1 resuspension buffer. 250 µl P2 lysis buffer was added and gently mixed until the solution turned blue. 350 µl P3 neutralization buffer was applied to the mixture and inverted until it became colourless. The sample was centrifuged at 13,300 rpm for 10 minutes at RT. The supernatant was then applied onto a DNA binding column and passed through by centrifugation at 6,000 rpm for 1 minute at RT. Columns were initially washed with 500 µl PB buffer followed by an additional wash with 750 µl PE buffer by centrifugation at 13,000 rpm for 1 minute at RT. To elute the DNA, 50 µl elution buffer was added to the column and incubated for 5 minutes to improve DNA yield. Columns were centrifuged at 13,000 rpm for 1 minute at RT. The integrity of the eluted DNA was determined by Sanger sequencing.

2.2.1.3.3. Midipreps

Colonies were inoculated in 100 ml LB containing the appropriate antibiotic and incubated at 37°C overnight (16-18 hours). The following day, 100 ml bacterial culture was harvested by centrifugation at 6,000 g for 15 minutes. The QIAGEN midiprep kit was utilized according to the manufacture's guidelines. The bacterial pellet was resuspended in 4 ml P1 resuspension buffer. 4 ml of P2 was then added and mixed by gentle inversion until the solution turned blue. To neutralise the lysis reaction, P3 buffer was added and mixed thoroughly by inverting until white and viscous. The solution was incubated for 10 minutes at RT. The lysate was transferred to QIAfilter cartridge and incubated for 10 minutes prior to clearing the lysate by filtering into a fresh collection tube. 2 ml binding buffer (BB) was added to the lysate and mixed by inversion. The lysate was then applied onto a QIAGEN DNA binding column and drawn through via centrifugation at 6000 rpm for 1 minute at RT. Columns were washed with 700 µl ETR followed by an additional wash with 750 µl PE buffer by centrifugation at 13,000 rpm for 1 minute at RT. To remove excess ethanol, the column was centrifuged at 13,000 rpm for 5 minutes and incubated for 5 minutes at 37°C.

To elute the DNA, 200 µl elution buffer was added to centre of the column and incubated for 5 minutes then centrifuged at 13,000 rpm for 1 minute at RT. To improve the DNA yield, the elution buffer was placed into 37°C prior to eluting. The concentration of the DNA was quantified using a Nanodrop 1000 spectrophotometer (ThermoFisher Scientific) and the eluted sample was diluted to a final concentration of 500 ng/µl.

2.2.1.3.4. Megapreps

Mega-prep is a purification process to obtain larger amounts of DNA with purity. 1.5-3 mg of DNA can be obtained if a high copy number plasmid is used. In our experiment, high copy number pcDNA5 3.1 (G₄C₂)₄₅-V5 (sense) and pcDNA5 3.1 (C₄G₂)₄₃-V5 (antisense) plasmids were purified to be used in cell transfection experiments for expressing sense or antisense DPR proteins. Plasmid purification was performed with QIAGEN megaprep kit according to the manufacture's protocol. Pre-culture containing colonies having the correct DNA insert of interest confirmed by miniprep (see **section 2.2.1.3.1**) or midiprep (see **section 2.2.1.3.3**) were inoculated in 400 ml LB containing appropriate antibiotics and incubated at 37°C overnight. The following day, bacterial cultures were harvested in 400 ml autoclaved bottles at 6000 g for 15 minutes at 4°C. The pellet was resuspended in 50 ml P1 resuspension buffer by vortexing thoroughly until completely dissolved. 50 ml P2 lysis buffer was added to the mixture and inverted gently to release the DNA from cells followed by a 5-minute incubation at RT. 50 ml prechilled buffer P3 was added, and the solution was gently inverted 5-6 times and incubated on ice for 30 minutes. The solution was transferred to fresh tubes and centrifuged at 20,000 g for 45 minutes at 4°C. During the centrifugation time, QIAGEN-columns were equilibrated with 35 ml QBT buffer. The supernatant was applied onto the column and drawn through via gravity flow. The column was washed using 200 ml QC buffer. DNA was eluted with 35 ml QF buffer and precipitated by adding 24.5 ml isopropanol. The DNA was pelleted by centrifugation at 15,000 g for 30 minutes at 4°C. The DNA pellet was washed with 7 ml 70 % ethanol and further centrifuged at 15,000 g for 30 minutes at 4°C. The pellet was air-dried for 20-30 minutes at RT before being reconstituted in 500 µl TE buffer. The concentration of the DNA was quantified using a Nanodrop 1000 spectrophotometer (ThermoFisher).

2.2.1.4. Restriction Check

2.2.1.4.1. Restriction Digest

Plasmids extracted from transformed bacterial colonies were digested with appropriate restriction enzymes and components at 37°C overnight to check for the successful integration of the DNA insert. The restriction digest reaction is outlined in **Table 2.10**.

Table 2.10. Restriction digest reaction for PCR product.

2 µl Eluted DNA
6 µl dH ₂ O
1 µl 10X Fast Digest Buffer
0.5 µl Restriction Enzyme 1
0.5 µl Restriction Enzyme 2
Total volume: 10 µl

2.2.1.4.2. Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to visualise DNA fragments produced by PCR or restriction digest according to their molecular weight. The percentage of the agarose gel (w/v) was determined depending on the size of DNA fragments. In brief, a required amount of agarose powder was dissolved in 1X TAE (Tris-acetate-EDTA) Buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) by heating in a microwave. The agarose solution was cooled with tap water prior to adding 0.5 µg/µl of ethidium bromide. 100 mL agarose solution was then poured into a clean tray with combs and left to set at RT for 30 minutes. Samples were loaded onto the gel alongside a DNA ladder (ThermoFisher Scientific). Electrophoresis was carried out at 80 V for 45 or 50 minutes. Agarose gels were visualized using a G:BOX, or a UV-transilluminator.

2.2.1.5. DNA Sequencing

Plasmids were sequenced at the core genomic facility of Source Bioscience using an Applied Biosystems' 3730 DNA Analyser.

2.2.2. Molecular Cloning

2.2.2.1. Insert Preparation

2.2.2.1.1. Polymerase Chain Reaction (PCR)

I have performed two different types of PCR in the work described in this thesis depending on the size of fragments. These are Herculanase II fusion DNA polymerase and Taq DNA polymerase respectively. Herculanase PCR reaction was assembled according to **Table 2.11** and Herculanase PCR programme parameters were outlined respectively in **Table 2.12**.

Table 2.11. Inverse PCR reaction assembly.

DNA template	1µl (100 ng)
Forward Primer	1.25 µl
Reverse Primer	1.25 µl
10X dNTP	8 µl
5X Herculanase II reaction buffer	10 µl
DMSO (Dimethyl sulfoxide)	1 µl
Herculanase II fusion DNA polymerase	1 µl
dH ₂ O	8 µl

Table 2.12. Herculanase II DNA fusion polymerase PCR programme.

PCR program for Herculanase II DNA fusion polymerase	
Heated Lid	110°C
Temperature Step	95°C 1 minute
Start Cycle	26x
Denaturing temperature	95°C 30 second
Annealing temperature	58°C 1 minute
Elongation temperature	72°C 10 minute
End Cycle	
Temperature Step	72°C 10 minute
Hold	10°C Infinite

Taq DNA polymerase PCR reaction was assembled according to **Table 2.13** and Taq PCR programme parameters are outlined respectively in **Table 2.14**.

Table 2.13. Taq PCR reaction assembly.

Forward Primer (10 µM)	2 µl
Reverse Primer (10 µM)	2 µl
10X dNTP	10 µl
10X Taq reaction buffer	10 µl
MgCl ₂ (50 mM)	2 µl
Taq DNA polymerase	2 µl
dH ₂ O	72 µl

Table 2.14. Taq DNA polymerase PCR programme.

PCR program for Taq DNA polymerase	
Heated Lid	110°C
Temperature Step	94°C 3 minutes
Start Cycle	25x
Denaturing temperature	94°C 30 seconds
Annealing temperature	55°C 1 minute
Elongation temperature	72°C 30 seconds
End Cyle	
Temperature Step	72°C 10 minute
Hold	10°C Infinite

2.2.2.1.2. Gel Extraction

Gel extraction was performed using a Qiagen Gel Extraction kit according to the manufacturer's protocol. Initially, excised gel fragments were weighed and dissolved in an appropriate volume of QG buffer. 200 µl isopropanol was added to the sample and mixed well. The mixture was applied onto a QIAquick column and centrifuged at 4,000 rpm for 1 minute at RT. The column was then washed with 500 µl QG wash buffer by centrifugation at 13,300 rpm for 1 minute. To remove salt residues, 750 µl of buffer PE was applied to the column and centrifuged for 1 minute at 13,300 rpm followed by additional centrifugation at 13,300 rpm for 1 minute to remove excess PE buffer. The column was transferred into a clean eppendorf tube and incubated at 37°C for 5 minutes to evaporate residual ethanol from the column prior to performing elution in 50 µl elution buffer via centrifugation at 13,300 rpm 1 minute at RT.

2.2.2.1.3. DNA Precipitation with Phenol-chloroform

100 µl of phenol-chloroform was added and mixed by vortexing. Samples were centrifuged at 13,300 rpm for 10 minutes and approximately 90 µl of supernatant was transferred into fresh tubes. 10 µl of 3 M sodium acetate (NaOAc) was added and quickly vortexed. 300 µl 100 % ethanol was added and the sample was inverted 5 times. The sample was then incubated at -20°C for 1 hour. After incubation, the sample was centrifuged at 13,300 rpm for 20 minutes at RT. The supernatant was discarded and the pellet was washed with 70 % ethanol then air-dried at RT. The pellet was re-constituted in 41 µl of dH₂O.

2.2.2.1.4. Restriction Digest

The resulting DNA (from **section 2.2.2.1.3**) was digested as follows:

- 41 µl of DNA
- 5 µl 10X restriction buffer (ThermoFisher)
- 2 µl of Restriction enzyme 1 (10 units / µg DNA)
- 2 µl of Restriction enzyme 2 (10 units / µg DNA)

2.2.2.2. Vector Preparation

Vectors were digested with the same restriction enzymes as highlighted in the above section. The digest reaction contained: 5 µg vector DNA (10 µl), 5 µl 10X restriction buffer (ThermoFisher), 2 µl of restriction enzyme 1 and 2 µl of restriction enzyme 2 (10 units / µg DNA) and the volume made up to 50 µl total with ddH₂O. The samples were then incubated at 37°C for 2 hours. Digested vectors were dephosphorylated by adding 10 U/µl of calf intestinal alkaline phosphatase (CIAP; Biolabs) and incubated for 15 minutes at 37°C and then 5 minutes at 90°C. Digested vectors were fractionated on a 0.8 % agarose gel at 80 V for 1.5 hour. DNA fragments were excised and the DNA was extracted using QIAquick Gel Extraction Kit (QIAGEN) following the manufacturer's protocol as previously described (see **section 2.2.2.1.2**).

2.2.2.3. Ligation

Ligation reactions were assembled by mixing 15 µl of DNA fragment with 2 µl cut vector, 2 µl 10X T4 DNA ligation buffer and 2 µl T4 DNA ligase. The ligation reaction then was left for overnight incubation at 16°C in a thermocycler.

2.2.2.4. Plasmid Transformation

After ligation, transformation of plasmids was performed. The detail of the transformation protocol is given section **2.2.1.2**.

2.2.3. Recombinant Protein Expression and Purification

2.2.3.1. Growth of Bacteria and Induction

Bacterial growth consists of two steps: (1) preparation of overnight pre-culture (50 ml), and (2) inoculating 750 ml fresh TB medium with the pre-culture to logarithmically increase bacterial cell biomass. Pre-cultures were generally prepared in TB medium except for some optimization experiments that required the use of SB medium. To prepare the pre-culture, the plasmid of interest was transformed into *E.coli* BL21 cells (RP/RIL/Rosetta) and inoculated in 50 ml TB supplemented with appropriate antibiotics. Cultures were grown overnight at 37°C in a shaking incubator. The following day, the pre-culture was inoculated into 750 ml TB supplemented with appropriate antibiotics at 37°C until an optical density of ~ 0.7 was reached (OD₆₀₀), indicating that the bacteria had reached mid-log phase which is the optimum time for protein expression. The bacterial cells were then induced with 1 mM IPTG at 37°C and harvested at different time points, ranging from 1 hour to overnight.

2.2.3.2. Preparation of Protein Extract for Purification

Induced cell pellets were harvested by centrifugation at 6,000 g for 15 min at 4°C. 3 - 3.5 g bacterial pellets were resuspended in 30 ml lysis buffer in accordance with the type of tags fused to the CPP or chromatography columns. The pellets were then lysed by sonication 10 times for 30 seconds at 100 % amplitude (with 30 second incubations on ice between each sonication) in lysis buffer, supplemented with 300 µl of 1 M phenylmethylsulphonyl fluoride (PMSF) and 150 µl of EDTA-free protease inhibitor cocktail (PIC) (Roche) (20 µl/ml). The lysate was subsequently centrifuged for 20 minutes at 20,000 g, 4°C. The supernatant containing soluble protein was then applied onto a column. The column was washed several times with suitable washing buffer. CPP was eluted in appropriate elution buffer. The composition of lysis, loading, washing and elution buffers that were used for purification of His₆-tagged and GST-tagged SRSF1-linker and Control (Ctrl) CPPs is outlined in **Tables 2.15 and 2.16**.

Table 2.15. The loading, lysis, washing, and elution buffers used in affinity and ion exchange chromatography systems for purification of His₆-tagged SRSF1-linker CPP.

His ₆ -Tagged SRSF1/Ctrl CPP Purification		
Co²⁺ (Cobalt) Affinity Chromatography		
<u>Lysis/Loading Buffer</u>	<u>Washing buffer</u>	<u>Elution buffer</u>
• 20 mM Na-P • 1 M NaCl • 0.5 % v/v Triton X-100	• 20 mM Na-P • 250 mM NaCl • 5 mM Imidazole	• 20 mM Na-P • 250 mM NaCl • 200 mM Imidazole
Mono-S/ Sepharose Fast Flow (SP) Ion Exchange Chromatography		
<u>Lysis/ Loading Buffer</u>	<u>Washing buffer</u>	<u>Elution buffer</u>
• 20 mM Na-P • 200 mM NaCl • 0.5 % v/v Triton X-100	• 20 mM Na-P	• 20 mM Na-P • 2 M NaCl

Table 2.16. The loading, lysis, washing, and elution buffers used in affinity and ion exchange chromatography systems for purification of GST-tagged SRSF1-linker and Ctrl CPPs.

GST-Tagged SRSF1/ Ctrl CPP Purification		
Glutathione-Sepharose (GST) Affinity Chromatography		
<u>Lysis/Wash Buffer</u>	<u>Digestion Buffer</u>	<u>Elution Buffer</u>
• 50 mM Tris-HCl • 150 mM NaCl • 1 mM EDTA • 0.5 % Triton X-100	• 50 mM Tris-HCl, • 150 mM NaCl, • 1 mM EDTA, • 1 mM DDT	• 50 mM Tris-HCl • 40 mM reduced Glutathione, pH: 8.0 • 100 mM NaCl
Sepharose Fast Flow (SP) Ion Exchange Chromatography		
<u>Lysis Buffer</u>	<u>Washing buffer</u>	<u>Elution buffer</u>
• 50 mM Tris-HCl • 150 mM NaCl • 1 mM EDTA • 0.5 % Triton X-100 • 1 mM DDT	• 50 mM Tris-HCl • 150 mM NaCl • 1 mM EDTA • 1 mM DDT	• 1 M NaCl/PBS • 2 M NaCl/PBS
Carboxymethyl (CM) Ion Exchange Chromatography		
<u>Loading Buffer</u>	<u>Washing buffer</u>	<u>Elution Buffer</u>
• 200 - 500 mM NaCl/PBS containing 6 M Urea	• 200 - 500 mM NaCl/PBS	• 1 M NaCl/PBS
Heparin ion exchange chromatography		
<u>Loading Buffer</u>	<u>Washing Buffer</u>	<u>Elution Buffer</u>
• 1X PBS	• 1X PBS	• 1 M NaCl/PBS • 2 M NaCl/PBS
Size Exclusion Chromatography		
<u>Loading Buffer</u>	<u>Washing buffer</u>	<u>Elution buffer</u>
• 1 M NaCl/PBS	• 1 M NaCl/PBS	• 500 mM NaCl/PBS

2.2.3.3. Chromatography Systems

The following sections on chromatography have been subdivided into the purification of (His)₆-tagged and GST-tagged CPPs.

2.2.3.3.1. Purification of (His)₆-Tagged-CPP

Cobalt Affinity Chromatography

The protein extract was loaded onto a talon cobalt column (Cytiva) which was previously equilibrated in cobalt lysis buffer. The column was washed 3 times using cobalt wash buffer and elution was performed with cobalt elution buffer. The elution fractions were analysed by western blotting (see **section 2.2.5.3**) and Coomassie staining (see **section 2.2.5.2**) to determine the integrity and abundance of the peptide within each fraction.

Mono-S/ SP Ion Exchange Chromatography

Initially, a mono-s column was equilibrated in mono-S loading buffer. The protein extract was loaded onto the equilibrated mono-s column and connected to a fast protein liquid chromatography system (AKTA FPLC, GE Healthcare). The column was washed with Mono-S wash buffer and elution was performed with a continuous imidazole gradient between (0-200 mM). Eluted samples were collected using a fraction collector. The eluted fractions were analysed by western blotting (see **section 2.2.5.3**) and Coomassie staining (see **section 2.2.5.2**) to determine the integrity and abundance of peptide within each fraction.

2.2.3.3.2. Purification of GST-tagged -CPP

The GST column was equilibrated by washing with first dH₂O and then GST lysis buffer. The column was controlled manually instead of connecting AKTA FPLC. The protein extract was loaded onto the pre-equilibrated GST column and passed through the column by gravity flow. The column was then washed 3 times with lysis buffer followed by a final wash with digestion buffer to remove Triton X-100. The digestion buffer was mixed with preScission protease enzyme (PSP) and applied onto the column to cleave the GST tag and release the CPP into solution. The column was then incubated at 4°C overnight on a rotating wheel.

Heparin/ SP/ CM Ion Exchange Chromatography

The cleaved CPP was collected following overnight incubation of GST column. The GST column was washed with digestion buffer to collect the remaining CPP on the column. The two PSP-cleaved CPP fractions (10 ml in each fraction) were combined followed by dissolving 6 M urea into the peptide solution. In this step, the peptide solution was loaded on three different columns at different times (SP (sulfopropyl), CM (carboxymethyl) and heparin) to investigate which column would give a higher amount peptide and a higher degree of purity. The column was washed 3 times with column-specific wash buffer. This was carried out by elution of CPP with column-specific elution buffer and collecting the fractions by fraction collector. The eluted fractions were analysed by western blot (see **section 2.2.5.3**) and Coomassie staining (see **section 2.2.5.2**) to determine which fractions contained an abundant amount of high purity peptide.

Size Exclusion Chromatography

Size exclusion chromatography was applied after heparin chromatography to remove any remaining impurities from the cell permeable peptide. The eluted peptide from heparin chromatography was loaded onto a Superdex 30pg size exclusion column connected to a fast protein liquid chromatography system (FPLC) (AKTA, GE Healthcare). Initially, the column was washed with dH₂O and then equilibrated with 1M PBS/NaCl which was used to elute the peptide from heparin chromatography. Peptide fractions were loaded on to pre-equilibrated size exclusion columns. The peptide was eluted using 500 mM PBS/NaCl and all fractions were collected by a fraction collector. The eluted fractions based on the chromatogram were analysed by Coomassie staining (see **section 2.2.5.2**) to determine which fractions contained an abundant amount of high purity peptide.

2.2.3.4. Optimised Protocol

Cell pellets (3-3.5 g) were resuspended in 30 ml GST lysis buffer supplemented with 1 M PMSF and PIC (20 µl/ml). Cell pellets were exposed to 10 cycles of 30 seconds sonication at 100 % amplitude (with 30 second incubations on ice between each sonication). The lysate was then centrifuged at 20,000 g, at 4°C for 20 minutes. During this time, a 5 ml GST column was equilibrated by first washing in 15 ml dH₂O and subsequently 15 ml GST lysis buffer.

The supernatant was loaded onto a GST column (5 ml) in RT and washed with 3 times GST wash buffer (15 ml), followed by a final wash using digestion buffer (15 ml). Next, the digestion buffer (10 ml) was mixed with 1 mg/ml preScission protease enzyme (600 µl) for cleavage of the GST-tag and loaded onto a GST affinity column in RT. The column was then incubated overnight at 4°C on rotating wheel. The following day, first CPP fraction was collected (CPP-1). The GST column was washed with digestion buffer (10 ml) and the second CPP fraction was collected (CPP-2). 6 M urea was dissolved in RT by slowly adding it into the two combined CPP fractions. The heparin column (5 ml) was equilibrated with PBS (15 ml) prior to applying the peptide solution onto the heparin column. The column was washed 3 times with washing buffer (15 ml) and elution of CPP was performed with 5 ml elution buffer (1 M NaCl/ PBS). Once elution from the heparin column was completed, the GST column was also eluted with GSH elution buffer (10 ml) in RT. Since the purity of the eluted peptide from heparin chromatography was not sufficient, the eluted peptide fractions were further applied to size exclusion chromatography. The elution fractions were analysed by western blot (see **section 2.2.5.3**) and Coomassie staining (see **section 2.2.5.2**) to determine which fractions contained the desired and abundant amount of peptide.

2.2.3.5. Quantification of Cell Permeable Peptide Concentration

The recombinant SRSF1-linker and Ctrl CPP concentrations were determined using a Pierce Quantitative Fluorometric Peptide Assay (ThermoFisher Scientific), according to the manufacturer's instructions. In brief, the peptide digest assay standard, SRSF1-linker and Ctrl CPPs were serially diluted two-fold in PBS (150 µl for each point) starting from 1:2 to 1:128. 10 µl of assay standard and peptide samples were applied in duplicate on a 96-well plate. 70 µl of fluorometric peptide assay buffer was then applied to each well of 96-well plates followed by the addition of 20 µl of fluorometric peptide reagent per well. The plate was incubated for 5 minutes at RT. The plate was then placed in the PHERAstar plate reader (BMG Labtech) and absorbance at 390 nm/ 475 nm was measured. The standard curve was generated with 8 points and the slope of the curve was used for calculating the peptide concentration using the following formula:

$$\text{Concentration } (\mu\text{g/mL}) = \text{Peptide molecular weight (Da)} \times \text{Concentration } (\mu\text{M})/1000$$

2.2.4. Mammalian Protein Expression

2.2.4.1. Tissue Culture

Human embryonic kidney 293 cells (HEK293T) were cultured in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich) containing 10 % Foetal Calf Serum (FCS; Sigma-Aldrich) and 1 % Penicillin Streptomycin (Penstrep; Lonza). The cells were subcultured every 3-4 days with fresh supplemented DMEM. Cells were plated on different plates depending on the nature of the experiment:

- Western blot: 50,000 cells in 500 µl media per well in a 24-well plate
- Nuclear export assay: 200,000 cells in 2 ml media per well in a 6-well plate
- MTT assay: 25,000 cells in 500 µl media per well in a 24-well plate
- Immunofluorescence: 5000 cells in 100 µl media per well in a 96-well plate
- Immunofluorescence: 25,000 cells in 500 µl media per well in a 24-well plate
- Co-immunoprecipitation assay: 1×10^6 cells in 15 ml media per plate

2.2.4.2. Cell Transfection

HEK293T cells were transfected 18-20 hours after plating. For transfection, plasmid DNA was mixed with polyethyleneimine (PEI, Sigma) to mediate endosomal escape. Opti-MEM media (reduced serum media, Sigma) was then added to dilute the DNA/PEI complex and improve transfection efficiency by keeping cells in reduced serum conditions. The transfection mixture was mixed well by vortexing 16 times and incubated for 10 minutes at RT. The transfection complex was added dropwise to each well of cells in culture and gently mixed prior to incubation at 37°C, 5 % CO₂. The transfection conditions according to experiment and plate size are outlined in **Table 2.17**.

Table 2.17. Transfection condition for HEK293T cell line.

Reagents	Plate Size				
	10 cm	6-well plate	12-well plate	24-well plate	24-well plate MTT
DNA (µg)	15 µg	2 µg	1 µg	700 ng	500 ng
PEI (µg)	52 µg	7 µg	3.5 µg	1.75 µg	1.75 µg
Opti-MEM (µl)	500 µl	200 µl	100 µl	50 µl	50 µl

2.2.4.3. Protein Extraction from Mammalian Cells

Cells were lysed approximately 72-hours post transfection to extract protein. Cell culture media was discarded and cells were washed with cold sterile PBS (137 mM NaCl, 2.7 mM KCL, 10 mM Na₂HPO₄, pH 7.4). Cold IP lysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, 0.5 % Triton X-100, 10 % Glycerol, 1 mM EDTA) supplemented with PIC and PMSF was added to per well of each plate (50 µl for 24-well plate, 350 µl for 6-well plate and 500 µl for 10 cm plate). Following this, the cells were scraped into IP lysis buffer using a P-200 pipette tip for 24-well plate or scraper for 6 well and 10 cm plates. The lysed cells were collected into 1.5 ml tubes and passed through a 25G gauge needle 10 times and incubated on ice for 5 minutes. The lysate was cleared by centrifugation at 13,000 rpm at 4°C for 5 minutes. The supernatant containing soluble protein extract was transferred into a new tube. The protein concentration was determined using the Bradford assay (2.2.4.4).

2.2.4.4. Bradford Assay

Bradford assays were used to quantify protein concentration within cell lysates. In brief, 5X Bradford reagent was diluted to 1 X with dH₂O and 2 µl of cell lysate was added to 1 ml 1X Bradford reagent and mixed by inverting 5-10 times. The total protein concentration was quantified with the Beer–Lambert equation based on absorbance values at 595 nm taken using a UV-Vis spectrophotometer.

Beer-Lambert equation: $A = \epsilon cL$ (A: absorbance, ϵ : 2/15 (molar extinction coefficient) c: concentration, L : 1 cm (light length)). The equation was simplified, and protein concentration was measured based on the equation below:

$$C = A \times 7.5 \times 1$$

2.2.5. Analysis of Protein by SDS-PAGE Gel Electrophoresis

2.2.5.1. SDS- PAGE Gel

Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a common biochemistry technique used for the separation of proteins or nucleic acids based on their molecular weight. SDS polyacrylamide gels were prepared depending on the molecular weight of the protein of interest. The compositions of gels of different resolving percentages were made with Bio-Rad Mini-PROTEAN® Tetra Cell cassettes, glass plates and 1 mm 15 well combs (see **Table 2.18**). Proteins were mixed with 4 % SDS page loading buffer (4 % SDS, 20 % glycerol, 10 % 2-mercaptoethanol, 0.004 % bromophenol blue and 0.125 M Tris HCl) and placed into 95°C for 5 minutes to denature the proteins. A 30-40 µg protein sample was loaded onto the SDS gel alongside a protein marker for molecular weight reference. The gels were then run at 80 V for 30 minutes followed by 150 V for approximately 1 hour until the bromophenol blue had reached the bottom of the gel.

Table 2.18. Composition of SDS polyacrylamide gels.

Reagents	5 % Stacking	10 % Resolving	12 % Resolving	15 % Resolving	20 % Resolving
dH ₂ O	5.8 ml	4.2 ml	3.5 ml	2.5 ml	830 µl
30 % w/v bis-acrylamide	1.7 ml	3.3 ml	4 ml	5 ml	6.6 ml
Resolving Buffer	N/A	2.5 ml	2.5 ml	2.5 ml	2.5 ml
Stacking Buffer	2.5 ml	N/A	N/A	N/A	N/A
10 % Ammonium persulfate (APS)	50 µl	50 µl	50 µl	50 µl	50 µl
N,N,N', N'-Tetra methyl ethylenediamide	20 µl	20 µl	20 µl	20 µl	10 µl

2.2.5.2. Staining of SDS PAGE Gel

Coomassie Blue Staining

Coomassie blue staining is a quick and simple method for detecting proteins on gels. The detection limit of Coomassie blue is approximately 0.1– 0.5 μg protein. Gels were transferred into plastic trays and submerged in Coomassie blue staining solution for 1 hour at RT. The gels were then transferred into de-stain buffer for 3-4 hours at RT. The composition of Coomassie blue stain and de-stain solutions is provided in **Tables 2.19 and 2.20**, respectively.

Table 2.19. Composition of Coomassie blue staining solution.

Component	Final concentration	Volume
Methanol	40 % (v/v)	400 ml
Acetic acid	10 % (v/v)	100 ml
Coomassie Brilliant Blue R-250	0.1 % (w/v)	1 g

Table 2.20. Composition of de-staining solution.

Component	Final concentration	Volume
Methanol	40 % (v/v)	400 ml
Acetic acid	10 % (v/v)	100 ml

Ponceau Staining

Ponceau staining is used to detect protein on PVDF or nitrocellulose membranes following Western blotting (see **section 2.2.5.3**). Membranes were treated with ponceau staining solution for 1 minute at RT followed by washing with 1X Tris-buffered saline with 0.2 % Tween[®] 20 (TBST). The composition of the ponceau staining solution is outlined in **Table 2.21**.

Table 2.21. Composition of Ponceau staining solution.

Component	Final concentration	Volume
Ponceau S	0.1 % (w/v)	1 g
Acetic acid	5 % (w/v)	50 ml
dH ₂ O	95 % (w/v)	950 ml

2.2.5.3. Western Blotting

To transfer proteins from SDS gels to nitrocellulose membranes (Amersham), gels were placed on top of nitrocellulose membranes between two layers of transfer buffer-soaked filter paper. Semi-dry electro transfer was performed at 0.15 A per gel for 1 hour. Membranes were blocked with 5 % milk in 1X TBST for 1 hour to minimize nonspecific antibody binding. Membranes were then incubated with primary antibody in 5 % milk + 1X TBST overnight at 4°C. The following day, membranes were washed 3 times with 1X TBST (5 minutes per wash) followed by incubation in a horseradish peroxidase (HRP) conjugated secondary antibody in 5 % milk + 1X TBST for 1 hour. The membrane was washed 5 times 1X TBST (5 minutes per wash). The membrane was then incubated with ECL1 (1 M Tris-HCl pH: 8.5, 250 mM luminal and 90 mM p-coumaric acid in 10 ml dH₂O) and ECL2 (1 M Tris-HCl pH: 8.5, H₂O₂ in 10 ml dH₂O) solution for 1 minute and then visualized using a G: BOX and protein band intensity was quantified using Genetools software (Syngene).

2.2.6. Immunocytochemistry

2.2.6.1. Immunofluorescence Assay for SRSF1-linker and Control (Ctrl) CPPs Uptake

HEK293T cells were seeded into 24-well plates (2.5×10^4 cells per well). The following day, cell culture media was changed and new media containing Ctrl CPP at 5 μm or SRSF1-linker CPP at 0.5, 1 and 5 μm was added. 48 hours post peptide treatment, cell media was removed, and cells were washed three times with PBS. Cells were fixed and permeabilised in 4 % paraformaldehyde (PFA) containing 0.2 % Triton X-100 for 10 minutes at RT. Cells were then washed with PBS before blocking with 4 % goat serum in PBS for 1 hour at RT. Cells were then incubated in primary antibody solution prepared in 4 % goat serum in PBS for 1 hour at RT. Cells were then washed three times with PBS containing 0.2 % tween-20 and then incubated with secondary antibodies (Alexa Fluor 488 nm, Alexa Fluor 647 nm) for 1 hour at RT in the dark. 2 % Hoechst bizBenzimide prepared in 4 % goat serum was applied to cells and incubated for 45 minutes at RT. The cells were washed three times with 0.2 % tween-20 in PBS. After the final wash, 100 μl of PBS was added without tween-20 and plates were wrapped in aluminium foil prior to imaging. Cells were imaged using the PerkinElmer Opera Phenix™ High Content Imaging System.

2.2.6.2. Immunofluorescence Imaging for SRSF1-linker and Control CPPs

HEK293T cells were seeded into 24-well plates containing glass coverslips, which were coated with appropriate volume of 0.1 mg/ml polylysine (2.5×10^4 cells per well), and the plate was left overnight at 37°C, 5 % CO₂. The following day, media was replaced with new media containing Ctrl CPP at 1, 5 and 10 μm or SRSF1-linker CPP at 0.5, 1, 5 and 10 μm . 48 hours post peptide treatment, media was removed and cells were washed with 500 μl PBS. Cells were fixed with 500 μl of 4 % PFA containing 0.2 % Triton X-100 for 20 minutes at RT. The cells were washed with PBS and blocked with 4 % goat serum in PBS and incubated for 1 hour at RT. Primary antibody solutions (50 μl /coverslip) were diluted in blocking solution. The glass coverslips were removed from the plate and placed on parafilm. Cells were incubated in primary antibody solution at 4°C overnight. The next day, coverslips were put back into the wells and washed with three times with PBS.

2.2.7. GST-NXF1, NXF2 and NXF3 p15 Pull-Down Assays

Pull-down assays are an important and useful technique to identify and confirm physical protein-protein interaction. The principle of this method involves immobilization of tagged proteins (bait) in affinity resin to provide a capture platform for interaction with other proteins (prey). When the prey proteins are incubated with tagged proteins, bait and prey protein interacting complexes are eluted with relevant elution buffer depending on the type of affinity resin. In our pull-down experiment, recombinant GST and GST-NXF1, GST-NXF2, GST-NXF3 were expressed in *E. coli* BL21-RP cells induced with 1 mM IPTG at 37°C overnight. (Details of protein expression step are provided in **section 2.2.3.1**). Induced cell pellets were obtained by centrifugation at 6,000 g for 15 minutes at 4°C. The GST, GST-NXF1, GST-NXF2 and GST-NXF3 bacterial pellets (0.1 g) were then lysed with 1 ml PBS + 0.5 % Triton X-100 buffer supplemented with 1 M PMSF and PIC (20 µl/ml) and sonicated 10 times for 30 seconds at 100 % amplitude (with 30 second incubations on ice between each sonication). The lysate was centrifuged at 17000 g, 4°C for 5 minutes and the soluble protein containing supernatant was then loaded onto 30 µl GST resin (which had been blocked with 1 % BSA in lysis buffer overnight) and incubated at 4°C for 1 hour on a rotating wheel. The GST beads were then washed 3 times with lysis buffer (1 ml) prior to incubation of GST beads with SRSF1-linker and Ctrl CPPs. Ctrl and SRSF1-linker CPPs were added onto GST beads at different concentrations in the presence of RNase A and incubated at 4°C for 1 hour on a rotating wheel. To elute the interacting complex, beads were washed 5 times with lysis buffer (1 ml). After the final wash, all liquid was removed and 50 µl GST elution buffer was added and left on ice for 10 minutes. The samples were centrifuged at 17,000 g at 4°C for 1 minute. 15 µl eluted proteins were mixed with 4X SDS gel loading dye and analysed with Coomassie blue staining (see **section 2.2.5.2**) and western blotting (see **section 2.2.5.3**).

2.2.8. Co-Immunoprecipitation (CoIP) Assay

HEK293T cells were split into 10 cm cell culture plates (1×10^6 cells per plate). The following day, the cells were transfected with 15 μ g p3x FLAG or p3x FLAG-NXF1 plasmids (see **section 2.2.4.2**). 6 hours post transfection, cell culture media was changed and fresh media containing Ctrl CPP at 5 μ m or SRSF1-linker CPP at 0.5, 1 and 5 μ m was added. 48 hours post transfection, cells were washed with ice-cold PBS and lysed with IP lysis buffer supplemented with PIC (20 μ l/ml) and 1 M PMSF. The lysed cells were passed through a 25G needle 10 times and left on ice for 10 minutes prior to clearing the cell lysate by centrifugation at 13,300 rpm at 4°C for 5 minutes. Protein concentration was determined using a Bradford assay (see **section 2.2.4.4**) and diluted to 2 mg/ml. 40 μ l anti-FLAG M2 affinity gel beads (Sigma), which had been blocked with 1 % BSA in IP lysis buffer overnight, were washed with IP lysis buffer and incubated with protein lysate at 4°C for 2 hours on a rotating wheel. The beads were washed 5 times with IP lysis buffer. To elute the protein, all the liquid was drained from the beads after the final wash, and 60 μ l IP lysis buffer containing 100 μ g FLAG peptide was added and incubated for 30 minutes at 4°C on a rotating wheel. To obtain eluates, samples were centrifuged at 13,300 rpm at 4°C for 1 minute. 15 μ l eluted proteins and 30 μ g total input protein samples were mixed with 4X SDS gel loading dye for Western blotting (see **section 2.2.5.3**).

2.2.9. Total and Cytoplasmic and Nuclear RNA Extraction

All solutions utilised within this section were formed using 0.1 % diethyl pyrocarbonate (DEPC)-treated water to inactivate the RNase enzyme, followed by autoclaving to inactivate DEPC. For total, nuclear, and cytoplasmic RNA extraction, HEK293T cells were split into 6x 6-well plates. The following day, cells were transfected with 2 μ g sense ($(G_4C_2)_{45}$ -3xV5) or antisense ($(C_4G_2)_{43}$ -3xV5) plasmids. 6-hours post transfection, media was changed and replaced with fresh media containing Ctrl or SRSF1-linker CPP (0 and 1 μ m). Cells were lysed, and RNA was extracted 72 hours post transfection (see **section 2.2.9.1**).

2.2.9.1. Total RNA Extraction

HEK293T cells were washed with ice-cold DEPC-PBS (1 ml per well of a 6-well plate). 350 µl 1X Reporter lysis buffer (Promega) supplemented with PIC-DEPC (20 µl/ml), 2 mM PMSF and 0.16 U/µl RNase inhibitors was added to each well of a 6-well plate, and cells were collected using a scraper. Cells were lysed for 10 minutes on ice prior to centrifugation at 13,300 rpm at 4°C for 5 minutes. 250 µl supernatant from the cell lysate was taken for RNA extraction. The remaining supernatant was used for western blot analysis (see **section 2.2.9.4**).

2.2.9.2. Cytoplasmic RNA Extraction

HEK293T cells were washed with ice-cold DEPC-PBS (1 ml per well of a 6-well plate) then collected by pipetting and transferred to pre-cooled Eppendorf tubes. Cells were centrifuged at 400 g for 4 minutes at 4°C. The supernatant was discarded, and cell pellet was washed with 350 µl Hypotonic lysis buffer (10 mM HEPES pH:7.9, 1.5 mM MgCl₂, 10 mM KCl and 0.5 mM DTT) to remove excess DEPC-PBS without disturbing the cell pellet. Cell pellets were lysed with 350 µl hypotonic lysis buffer containing PIC-DEPC (20 µl/ml), 2 mM PMSF and 0.16 U/µl RNase inhibitors. The cells were fully resuspended using 3 different P1000 tip size (big>medium>small) to avoid shearing of the cells. Once cell pellets had been fully resuspended, the lysate was incubated on ice for 10 minutes prior to undergoing serial centrifugation steps. Initially, samples were centrifuged at 1,500 g for 3 minutes at 4°C to separate the nuclear and cytoplasmic fractions. The pellet was kept for nuclear fractionation and the supernatant was transferred into a fresh tube and re-centrifuged at 3,500 g for 8 minutes at 4°C. The supernatant was transferred into a new tube and centrifuged at 17,000 g for 1 minute at 4°C. The final 250 µl of supernatant containing the cytoplasmic fraction was taken for RNA extraction (see **section 2.2.9.5**) and 50 µl was transferred into another tube for anti-SSRP1 western blotting to confirm the absence of nuclear fractionation.

2.2.9.3. Nuclear RNA Extraction

The nuclear fraction collected as a pellet in the previous (see **section 2.2.9.2**) was washed four times with hypotonic buffer with centrifugation at 1,500 g for 3 minutes at 4°C between each wash. The pellet was then resuspended in 350 µl 1X reporter lysis buffer containing PIC-DEPC (20 µl/ml), 2 mM PMSF and 0.16 U/µl RNase inhibitors followed by a 5 minutes incubation on ice. Samples were passed through a 25G needle 10 times before centrifugation for 5 minutes at 17,000 g, 4°C. The supernatant was transferred into a new tube and 250 µl of supernatant was stored for RNA extraction. The remaining volume was subjected to western blotting for Tuj1.

2.2.9.4. Western Blot

Total, cytoplasmic and nuclear fractionated samples (15 µl extracts + 5 µl 4X SDS loading dye) were loaded onto 2X 12 % SDS acrylamide gels followed by western blotting (see **section 2.2.5.3**). One set of western blots was probed with anti-Tuj1 (cytoplasmic marker) to confirm the absence of cytoplasmic leakage, and another set was probed for anti-SSRP1 (nuclear marker) to confirm the absence of the nuclear fraction.

2.2.9.5. PureZOL™ RNA Extraction

750 µl (three times volume of fractionated sample) PureZOL (Bio-Rad) was added directly to each of fractionated total, cytoplasmic and nuclear samples to extract RNA. Extracts were incubated for 10 minutes at RT followed by addition of one-fifth of the volume of chloroform (generally 200 µl) to each tube. The tubes were shaken vigorously by hand for 15 seconds and incubated for 10 minutes at RT. The mixture was cleared by centrifugation for 10 min at 17,000 g, 4°C. the 450 µl upper clear layer was transferred into a new tube and 1 µl of 5 µg/µl glycogen (Ambion), 50 µl of 3 M NaOAc, and 500 µl 100 % isopropanol were added. To precipitate the RNA, tubes were stored overnight at -20°C. The following day, samples were centrifuged for 10 minutes at 17,000 g, 4°C. The supernatant was discarded without disturbing the pellet. RNA pellets were washed with 500 µl 70 % DEPC-ethanol and then centrifuged at 17,000 g, 4°C for 10 minutes. The RNA pellets were air-dried for 10 minutes at RT.

2.2.9.6. DNaseI Treatment

After ensuring that all the ethanol had evaporated, RNA pellets were resuspended in 25 µl DNaseI solution (22 µl DEPC water + 2.5 µl 10X DNase buffer + 0.5 µl DNaseI (Roche)) followed by incubation at 37°C for 30 minutes followed by 70°C for 10 minutes to inactivate the DNaseI.

2.2.9.7. Reverse Transcription for cDNA Synthesis

RNA concentration was measured with a Nanodrop 1000 spectrophotometer (ThermoFisher Scientific). 2 µg RNA was converted for cDNA synthesis using a reverse transcription reaction. Initially, 11 µl RNA + 1 µl 40 µM dN₆ + 1 µl 10 mM dNTPs were mixed to anneal random hexamer primer (random priming). Samples were incubated at 70°C for 5 minutes in a thermocycler and immediately placed on ice. The RT-PCR reaction mixture was assembled accordingly; 13 µl random priming premix, 4 µl 5X first strand buffer, 2 µl 0.1 M DTT and 1 µl M-MLV reverse transcriptase (RT; 200 U). cDNA synthesis was completed by incubating samples at 25°C for 10 minutes followed by at 42°C for 60 minutes and 10 minutes at 85°C.

2.2.9.8. Quantitative Reverse Transcription PCR (qRT-PCR)

Synthesised cDNA was diluted (1:3) by adding 40 µl DEPC-water prior to qRT-PCR. qRT-PCR reactions were performed in duplicate and prepared by mixing 1 µl synthesized cDNA, 5 µl 2X Brilliant III Ultra - Fast SYBR Green QPCR master mix (Agilent), and 1 µl primer mix (5 µM Human shRNA U1 and C9RNA qPCR primers). The samples were loaded into each well of 96-well qPCR plate (Bio-Rad) and plastic caps were placed on the plate prior to centrifugation at 2000 rpm for 1 minute at RT. Two step qRT-PCR was performed using a CFX96TM Real-Time qPCR system (Bio-Rad) as follows: initial denaturation at 95°C for 10 minutes followed by 45 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 1 minute. qRT-PCR data were normalized to Human shRNA U1 and analysed with CFX qPCR software (Bio-Rad).

2.2.10. MTT Cell Proliferation Assay

HEK293T cells were seeded into 24-well plates at a density of 2.5×10^4 cells per well. The following day, cells were transfected with 500 ng sense (G_4C_2)₄₅-3xV5) or antisense (C_4G_2)₄₃-3xV5) plasmids (see **section 2.2.4.2**). 6-hours post transfection, media was replaced with fresh media containing Ctrl at 0.5 and 2 μ M or SRSF1-linker CPP either synthetic or recombinant at 0.0625, 0.125, 0.25, 0.5, 1 μ M. 72 hours post transfection, 50 μ l of MTT (thiazolyl blue tetrazolium bromide, 0.5 mg/ml) dye was added to each well of the 24-well plate and incubated for 1 hour at 37°C, 5 % CO₂. After incubation, the cells were lysed with 550 μ l of MTT lysis buffer (20 % w/v SDS, 50 % v/v dimethylformamide) followed by another incubation for 1 hour at RT on an orbital shaker in the dark. The calorimetric signal was measured at 590 nm using a PHERAstar plate reader (BMG Labtech).

2.2.11. Animal Work

2.2.11.1. Ethics Statement

The section is adapted from the work published as part of my PhD (Castelli et al., 2023). Ethical approval for animal experiments was acquired from the University of Sheffield Animal Welfare and Ethical Review Body and the UK Animal Procedure Committee (Scientific Procedures). All *in vivo* experiments were carried out under a UK Home Office project licence (P608B5AD3), in accordance with the Act 1986. The daily care of mice was carried out in accordance with the Code of Practice of the Ministry of the Interior for the Housing and Care of Animals Used in Scientific Procedures. ARRIVE 2.0 guidelines were followed in the conduct of all *in vivo* experiments.

2.2.11.2. Mouse Stock and Husbandry

The section is adapted from the work published as part of my PhD (Castelli et al., 2023). BAC transgenic *C9ORF72*-ALS/FTD mice expressing approximately 500 G_4C_2 repeats within the human *C9orf72* gene (FVB/NJ- Tg(*C9orf72*)500Lpwr/J) were created by Professor Laura Ranum's group and obtained from the Jackson Laboratory (Stock No: 029099, also known as C9-500). Mice were housed in groups of 2 to 5 with access to food and water ad libitum (Envigo, standard rodent diet 2018). The temperature was maintained at 21°C, with 12-hours light/dark cycles. A plastic house was provided in each cage with sawdust (Datesand) to cover the floor of the cages and nesting material provided (Datesand paper wool).

2.2.11.3. Mouse Sex Determination

Tail clippings between 1-2 mm length were collected from new-born *C9ORF72*-ALS/FTD mice and submerged in 20 µl Quick Extract buffer. Samples were placed in a thermocycler and subjected to the following protocol to extract genomic DNA: Heated lid (110°C); 65°C for 15 min; 98°C for 2 min; 15°C hold. To determine the sex of the mice, genomic DNA was amplified by PCR using Sex1 and Sex 2 primers as shown in **Table 2.1**. The PCR reaction was assembled accordingly: 18 µl water, 5 µl FirePol 5X master mix solution, 1 µl of 3 µM primer working solution and 1 µl DNA template (50-200 ng). PCR mixtures were held at 95°C for 10 minutes (initial denaturation) followed by 30 cycles at 95°C for 30 seconds, 65°C for 30 seconds, 72°C for 30 seconds, and a final extension at 72°C for 10 minutes. 15 µl PCR samples were then resolved on a 2 % agarose gel alongside a DNA ladder (ThermoFisher Scientific) at 80 V for 30 minutes. A single band was expected in males at 200 bp and no band was expected in females.

2.2.11.4. Mouse Genotyping Analysis

DNA was extracted as previously described (see **section 2.11.3**). A PCR reaction for determining the genotype was assembled as follows: 10 µl water, 2X 12.5 µl Quick-Load Taq master mix (New England Biolabs) solution, 0.5 µl C9-GT-F (10 µM), 0.5 µl C9-GT-R (10 µM), 0.25 µl VgII4-F (10 µM), 0.25 µl VgII4-R (10 µM) and 1 µl extracted DNA template. PCR reactions were run in a thermocycler as follows: 96°C for 3 minutes (initial denaturation) followed by 32 cycles at 94°C for 45 seconds, 55°C for 45 seconds, 72°C for 1 minute, and a final extension at 72°C for 6 minutes. 15 µl PCR samples were then resolved on a 2 % agarose gel alongside a DNA ladder (ThermoFisher) at 80 V for 1 hour. A single band was expected in transgenic mice at 200 bp and no band was expected in non-transgenic mice.

2.2.11.5. Southern Blot for *C9orf72* Repeat Expansion Measurement

DNA was extracted using either a GenElute Mammalian Genomic DNA Miniprep Kit (Sigma Aldrich) or a DNeasy Blood & Tissue Kit (Qiagen) from either tail or brain, according to the manufacturer's instructions. Approximately 1-2 cm of tail or 0.5 x 0.3 cm of brain tissue were lysed using a Precellys evolution homogenizer at 7,200 rpm for 90 seconds followed by an overnight incubation at 56°C.

Day 2- Restriction Digest

The following day, restriction digest reactions were assembled as follows:

- 11 µL DNA (~10ug)
- 5 µL Cutsmart buffer (10X)
- 2 µL AluI restriction enzyme
- 2 µL DdeI restriction enzyme
- 30 µL Milli-Q H₂O

Reactions were incubated in a thermocycler overnight at 37°C.

Day 3 - Denaturation and Transfer

Digested DNA was mixed with 10 µl of 6X DNA loading buffer and then resolved on a 1 % agarose gel with 5 µl of DIG-labelled marker on either side of the gel at 80 V for 2 hours. The agarose gel was taken from the tank and excess wells which did not contain samples were excised. The gel was then flipped over with using two clean glass plates and transferred to a tray containing 1 L of denaturation buffer (1.5 M NaCl, 0.5 M NaOH) then left gently rocking for 45 minutes at RT. The gel was transferred to another tray containing 1 L of neutralisation solution (3 M NaCl, 0.5M Tris, pH: 7.5) then incubated on a rocking platform for 45 minutes at RT.

During the denaturing/neutralising of the gel, a transfer apparatus was prepared by dipping two long sheets of Whatman paper into the bottom of the tank and cutting to 4 gel-sized sheets of Whatman paper, 8 gel-sized sheets of blotting paper, a positively charged membrane (Amersham Hybond-N+), a plastic mask with an opening cut to the size of the gel, cling film, and 2 half-folded stacks of tissue paper.

A tray was filled with 20X SSC buffer (0.3 M sodium citrate, 3 M NaCl, pH 7.0). Two long sheets of Whatman paper were placed on glass, which is wide enough to comfortably fit the gel on the edges of the tray and dipped into the 20X SSC buffer until the paper was completely saturated in the solution. The gel with the well facing upwards was then transferred on the wetted Whatman paper which was sat atop the glass sheet. Next, the mask was placed on top of the gel and edges of tray were wrapped with cling film to improve transfer efficiency. The membrane was soaked in 5X SSC buffer and placed on the surface of the gel. Two squares of wet Whatman paper were applied on the membrane and the rolled out to remove bubbles. Two further dry sheets of Whatman paper and 8 sheets of blotting paper were placed on top of the wetted Whatman paper. This was followed by adding 2 rolls of tissue. A glass sheet was placed evenly on top of the tissue paper and then weight was placed evenly on the glass sheet followed by leaving to transfer overnight at room temperature. A representative image of the transfer steps is shown in **Figure 2.1**.

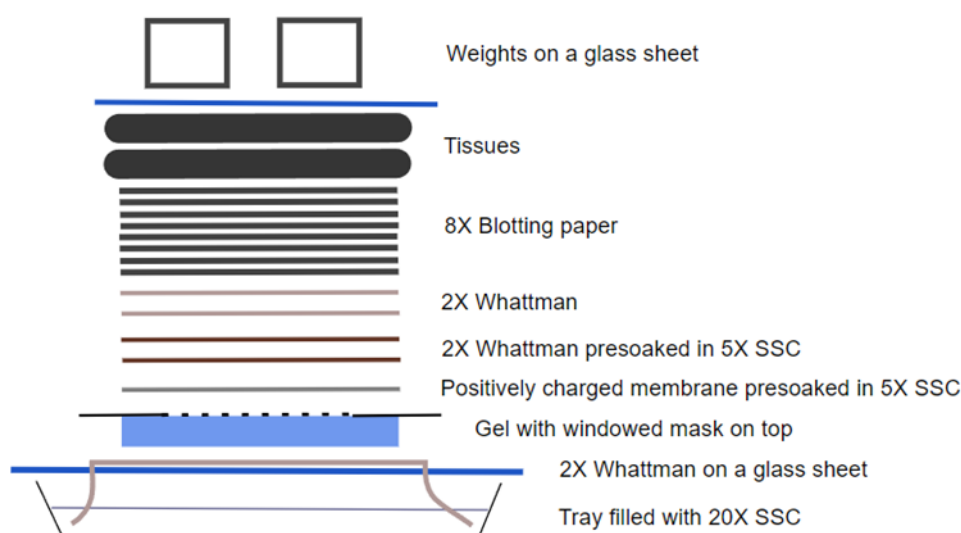


Figure 2.1. The representative image of southern blot transfer steps.

Day 4 - Probe Hybridisation

The following day, the upper layers of transfer materials were removed, leaving the 5X soaked Whatman paper for support. The membrane and gel were flipped over using a glass sheet and the wells of the gel were marked through on the membrane surface using a pencil. The membrane was moved to a plastic tray and washed on the rocking platform in 2X SSC for 5 minutes at RT. The membrane was then placed on a dry piece of Whatman paper in the transilluminator and cross-linked at 210,000 uJ. The membrane was washed again in 2X SSC for 5 minutes on a rocker followed by placing the membrane on top of a nylon mesh and transferring to 30 ml pre-hybridization buffer in a glass hybridization tube with the DNA-coated side facing inwards. The membrane was then rotated in an oven for 4 hours at 48°C. Hybridisation buffer supplemented with a Dig labelled probe at a concentration of 100 ng/ml (3 µl of probe to 30 ml buffer) was placed in an oven at 48°C to pre-warm. After incubation, the pre-hybridization buffer was replaced with hybridisation buffer and left to rotate overnight at 48°C.

Day 5 - Probe Detection and Imaging

The following day, the hybridisation buffer was poured out and the membrane was washed with pre-warmed high stringency washes as listed below:

- 50 ml 2X SSC 0.1 % SDS for 10 min while the temperature was ramping from 48°C to 65°C
- 50 ml 2X SSC 0.1 % SDS for 10 min at 65°C
- 100 ml 0.5X SSC 0.1 % SDS for 15 min 65°C
- 100 ml 0.2X SSC 0.1 % SDS for 15 min 65°C

The membrane was transferred to the plastic tray and washed with a DIG wash and block buffer according to the manufacturer's protocol (Roche, 11585762001). The name of the reagents, volume and times are provided in **Table 2.22**.

Table 2.22. The name of reagents, volumes and times for probe detection and imaging.

Reagent	Volume (ml)	Time (min)
Wash Buffer	200	5
Blocking Solution	200	60
Anti-DIG-AP Fab fragment 35 mU/ml (1:20,000 dilution of stock) in block	50	30
Wash Buffer	100	10
Wash Buffer	100	10
Wash Buffer	100	10
Detection Buffer	50	5

The first wash was performed with vigorous rocking (~ 40 rpm), but this was lowered to gentle rocking (~ 25 rpm) for blocking and incubation with Anti-DIG (11093274910, Sigma Aldrich). The final three washes were performed with vigorous rocking (~ 40 rpm). The membrane was placed on plastic sheet and 50 drops of CDP were added followed by 5 minutes incubation prior to pressing out excess CDP onto a tissue and placing the membrane at 37°C in the hybridisation oven for 5 minutes. The membrane was then exposed to film for at least 3 hours in a dark room followed by placing the membrane in the developer solution and rocking until seeing a signal appear. The membrane was rinsed with tap water and placed in fixative solution. The membrane was rinsed with tap water for a further 5 minutes.

2.2.11.6. Cisterna Magna Injection of the scAAV9 Viral Vector into Neonatal Mice

Neonatal transgenic pups first were tattooed after their birth to allow identification during administration of scAAV9_H1-SRSF1 linker CPP_EF1 α -eGFP and scAAV9_H1-SRSF1-RRM2 α 1 CPP_EF1 α -eGFP directly into the cerebrospinal fluid (CSF). Before starting the injection, a stereotaxic frame and syringe pump system were installed and connected to an anaesthetic machine (Burtons Veterinary) and filling isoflurane (Iso Vet) unit. 5 μ l of 5x10¹⁰ vg of scAAV9_H1-SRSF1 linker CPP_EF1 α -eGFP and scAAV9_H1-SRSF1-RRM2 α 1 CPP_EF1 α -eGFP in PBS was loaded into a 33-gauge Hamilton syringe, ensuring no air bubbles were present in the solution.

Pups were anaesthetized by placing them in an induction chamber and administering oxygen at a flow rate of 4 L/min with 5 % isoflurane. The pups were monitored until their breathing had slowed. Oxygen flow rate and the percentage of isoflurane were reduced to 0.5 L/min and 2 %, respectively. The mouse pups were then placed on Wee sight with their nose in anaesthetic tubing, ensuring that the nose was totally covered. The needle was slowly advanced around a quarter turn to pierce the atlanto-occipital membrane and virus was administered at 1 μ l per minute. After the injection was complete, the injected substance was left for one minute to allow dispersion from the injection site. The needle then was removed slowly from the injection site and the pup was returned to its home cage.

2.2.11.7. Cisterna Magna Injection of SRSF1-linker CPP into P30 Mice

BAC transgenic mice were anaesthetized by placing them in an induction chamber and administering oxygen at a flow rate of 4 L/min with 5 % isoflurane. Mice were monitored until breathing had slowed to approximately 30 breaths/ minute. Fur was then shaved from the back of the mouse's head and neck. The oxygen flow rate and isoflurane were reduced to 0.5 L/min and 2 %, respectively. The nose of the mouse was placed securely on a frame by adjusting and fixing with ear bars. All loose fur from the shaved areas was cleaned by wiping with 100 % ethanol. 10 μ l lignol was injected subcutaneously between the mouse's ears prior to surgery. 5 μ l of synthetic or recombinant SRSF1-CPP or Ctrl-CPP at 1 mM in PBS was loaded into a 33-gauge Hamilton syringe, ensuring that no air bubbles were present in the solution. A vertical incision was made in the skin between the ears. Connective tissue that brings together muscles at the back of the neck was carefully cut and the muscle tissue was separated slowly.

The next layer of muscle behind the neck was separated using forceps. The cistern magna was identified as a dark inverted triangle with the cerebellum above and the brainstem below. The needle was slowly advanced until it pierced the atlanto-occipital membrane, and the peptide solution was administered at 1 μ l per minute. The injected substance was left for one minute to disperse from the injection site and the needle was slowly removed. The cut muscle layer was sutured using an absorbable three-knotted suture. The cut skin then was closed using a non-absorbable suture. The mouse was then placed into an incubator at 37°C and its recovery was monitored. Once the mouse was ambulatory, it was returned to its home cage.

2.2.11.8. Tissue Collection

P34 mice (for the SRSF1-linker CPP study) and p30 mice (for the scAAV9_H1-SRSF1 linker and RRM2 α 1 CPP study) were first perfused under terminal anaesthesia with pentobarbital (stock 100 mg/ml) and then perfused with sterile 25 ml of PBS. Brain, spinal cord, heart, liver, spleen, lung, kidney, CSF, blood, and muscle tissues were collected and snap frozen for analysis. The brains and cervical spinal cords of one of the animals in each treatment group were fixed in 4 % PFA for histological analysis (see **section 2.2.12**). The tissues then underwent two different staining processes, one for V5 staining with frozen tissue (see **section 2.2.11.11.1**), and another for GFP staining with paraffin embedded tissue (see **section 2.2.11.11.2**). Three female mice per group were used as the minimum to enable statistical testing. Mice were not randomized to groups, instead litters were distributed evenly across groups.

2.2.11.9. DNA Extraction

Genomic DNA was extracted from brain, liver, lung, kidney, spinal cord, muscle, spleen and heart samples using a GenElute Mammalian DNA Miniprep Kit (Sigma) according to the manufacturer's protocol. In brief, tissue of up to 25 mg was excised using a sterile blade and transferred to a tube then incubated on ice. 180 μ l lysis buffer T containing 20 μ l Proteinase K was added to the tissue and incubated for 4 hours at 55°C with vortexing occasionally until the tissue completely dissolved. 200 μ l lysis buffer C was added to the sample and vortexed then incubated at 70°C for 10 minutes. Columns were equilibrated by adding 500 μ l of column preparation solution followed by centrifugation at 13,330 rpm for 1 minute. 200 μ l of 100 % ethanol was added to mixture and mixed by vortexing.

The entire solution was applied to the pre-equilibrated column and centrifuged at 7000 rpm for 1 minute. The column was washed twice with 500 µl wash buffer. To eliminate any residual wash buffer, the column was centrifuged 13,300 rpm for 5 minutes and incubated at 37°C for 5 minutes to evaporate residual ethanol. To elute the DNA, 200 µl of pre-warmed (50-60°C) elution buffer was loaded onto the centre of the column, incubated for 5 minutes at RT and centrifuged at 13,300 rpm for 5 minutes. The DNA concentration was quantified using a NanoDrop 1000 spectrophotometer (ThermoFisher).

2.2.11.10. Protein Extraction

Reporter lysis buffer (Promega) was diluted from 5X to 1X in dH₂O and supplemented with 1 M PMSF and PIC (20 µl/ml). 400 µl lysis buffer was added to ~ 0.1 g tissue, followed by adding 8-10 (1.4 mm) zirconium oxide beads into the tube. The sample was then lysed using a Precellys evolution homogenizer at 5500 rpm for 2 x 30 seconds followed by a 10 minutes incubation on ice; this process was repeated. The tissue underwent centrifugation at 17,000 g for 10 minutes at 4°C. The supernatant was transferred into a fresh tube and protein extract concentrations were determined using a Bradford assay (Bio-Rad) (see **section 2.2.4.4**).

2.2.11.11. Histopathological Analysis

2.2.11.11.1. Tissue Preparation for Frozen Tissue and V5 Immunofluorescence Staining

The post-fixed brain tissue in 4 % PFA was first incubated in 20 % sucrose at 4°C for 24 hours. The following day, tissue was incubated in 30 % sucrose at 4°C for 12 hours. The resulting tissues were mounted in optimal cutting temperature (OCT) compound by freezing in liquid nitrogen and iso-pentane. The brain tissue was sectioned into 8 µm thickness sections using a cryostat at -21°C and directly placed on the charged slides and left to dry for 1 hour before storage at -20°C. The slides were thawed for 30 minutes at RT and rehydrated with PBS for 5 minutes at RT. The sections were then blocked for 1.5 hours at RT with blocking buffer (0.5 % Triton X-100, 5 % BSA, 2% heat-inactivated goat serum, 1 % FBS, PBS). Tissue sections were incubated overnight with a primary V5 antibody (1:500 dilution in PBS with 0.1 % Triton X-100 and 1 % heat-inactivated goat serum). The following day, the sections were washed 3 times (10 minutes per wash) with buffer (PBS with 0.1 % Triton X-100 and 1 % heat-inactivated goat serum).

Tissue sections were then incubated for 1 hour at RT with a secondary fluorescence antibody (Alexa Fluor 488) (1:1000 dilution in PBS with 0.1 % Triton X-100 and 1 % heat-inactivated goat serum) prior to washing 3 times (10 minutes per wash) with PBS. The sections were dried by wiping around the edges and one drop of Hardset Vectashield medium with DAPI was applied. Coverslips were slowly lowered on top of the sections to ensure even spreading of the Vectashield mounting medium over the surface.

2.2.11.11.2. Tissue Preparation for Paraffin Embedding and GFP Immunostaining

Day 1

Brain tissue fixed in 4 % PFA was washed with cold PBS. The tissues were then embedded in paraffin before being sectioned into 10 μ m slices using a cryostat at -21°C and directly placed on the charged slides. The tissues were rehydrated to remove wax through placing in xylene twice for 5 minutes followed by graded ethanol incubation (100 % for 5 minutes, 95 % for 5 minutes, 70 % for 5 minutes) and dH₂O for 5 minutes. The slides were placed in antigen retrieval solution and heated in a pressure cooker for 30 minutes at 20 psi to break methylene bridges formed during fixation, thereby allowing access to antigen binding sites. Slides were then gently washed with dH₂O followed by a second wash with PBS. The back of the slides was dried carefully without touching them. The covering section of each slide was then blocked with 5 % BSA + 0.25 % Triton X-100 and incubated for 20 mins at RT. The blocking solution was removed and then primary antibodies diluted (1:1000) in PBS containing 1 % BSA and 0.25 % Triton X-100 and 75 μ l were added per section in a cold room and incubated overnight at 4°C.

Day 2

The slides were washed three times with PBS (8 minutes per wash) on a shaker. Sections were incubated with secondary antibody (500 μ l per section) (1:1000 dilution in 1 % BSA + PBS) for 1.5 hours at RT in the dark. Sections were then washed four times with PBS (8 minutes per wash). Sections were dried by wiping around the edges and one drop of Hardset Vectashield medium with DAPI was added to the surface. Coverslips were slowly lowered over the sections to ensure that Vectashield mounting medium was spread evenly over the surface of the sections. The sections were left to dry at RT in the dark and stored at 4°C.

2.2.11.12. Biodistribution of Viral Copies in Mouse Tissue

Biodistribution of scAAV9_H1-SRSF1-Linker CPP-GFP and scAAV9 SRSF1 scAAV9_H1-SRSF1-RRM2 α 1 CPP-GFP across animal tissues was established by qPCR. Prior to performing qPCR, the stock plasmids packed into scAAV9_H1-SRSF1-Linker CPP-GFP and scAAV9_H1-SRSF1-RRM2 α 1 CPP-GFP were digested with HindIII enzyme. This was followed by performing gel extraction (see **section 2.2.2.1.2**). The linearized DNA was then diluted from 10^{-2} to 10^{-9} to create a standard curve with 8 points. Then, the extracted DNA from all collected tissues of each animal was diluted to 50 ng/ μ l. The qPCR reactions were then assembled per well as follows:

- 1 μ l of DNA
- 5 μ l 2X Brilliant III Ultra-Fast SYBR Green qPCR master mix (Agilent)
- 2 μ l of DEPC-H₂O
- 0.5 μ l of 10 μ M Forward primer (H1cas2, GFP and BGH qPCR primers)
- 0.5 μ l of 10 μ M Reverse primer (H1cas2, GFP and BGH qPCR primers)

The samples and master mix were loaded in duplicate into a 96-well qPCR plate (Bio-Rad) in accordance with depicted plate design in **Figure 2.2** and plastic caps were placed on the plate prior to centrifugation at 2,000 rpm for 1 minute at RT.

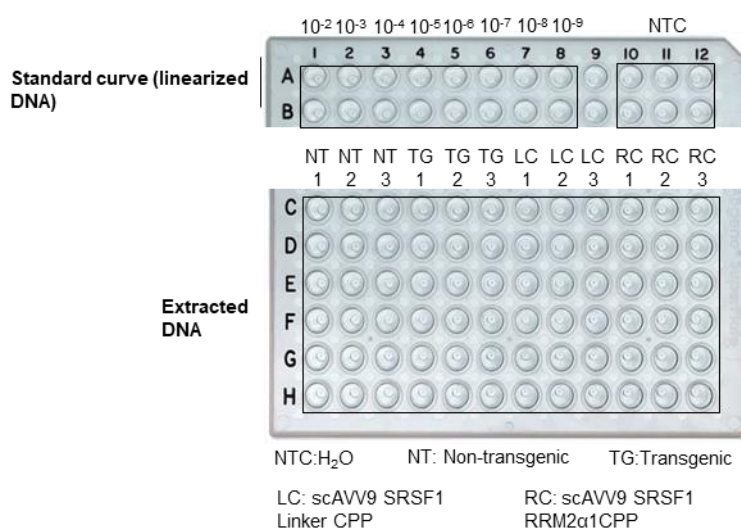


Figure 2.2. qPCR plate design for viral biodistribution across several organs.

Having completed sample loading on the plate, the qPCR was performed using a CFX96™ Real-Time qPCR system (Bio-Rad) as follows: initial denaturation at 95°C for 5 minutes followed by 39 cycles of 95°C for 10 seconds, 60°C for 30 seconds and plate read then followed by 60 cycles of 65°C for 31 seconds, 65°C for 5 minutes + 0.5°C/cycle, Ramp 0.5°C/s and then plate read. Once the qPCR had been completed, the standard curve was plotted with log (viral copies) and Ct values of each dilution point. Log of viral copies was calculated based on the equation given below.

$$\text{Viral copies} = \frac{\text{Amount of DNA (ng)} \times \text{Avogadro's Constant}}{\text{Plasmid length (bp)} \times 650 \times 10^9}$$

Ct values from unknown samples entered into the software (GraphPad) which allowed us to interpolate logarithmic viral copies of each samples. Viral copies, corresponding to the reverse log of these interpolated values, were normalised per total ug of DNA in each sample.

2.2.11.13. Poly-GP MSD ELISA Assay

Day 1

Un-biotinylated capture antibody was diluted in 1X TBS followed by vortexing for 3 to 5 seconds. 30 µl of poly GP capture antibodies (custom synthesis, Eurogentec) was added to each well of multi-array 96-well plate. The plate was then sealed with an adhesive sealing sheet. The plate then underwent differential shaking (15 seconds at 1250 rpm then 15 minutes at 600 rpm), ensuring that the bottom of each well was coated sufficiently. The plate was incubated overnight at 4°C.

Day 2

The plate was washed three times with 150 µl TBST per well. 150 µl blocking buffer (3 g milk powder/ 100 ml TBST) was added to each well and the plate was shaken for 2 hours at 700 rpm. During this incubation, brain tissue was lysed (as explained in **section 2.11.10**), and extracted protein was normalized to 4 mg/ml using reporter lysis buffer. Samples were further diluted to 2 µg/µl in EC buffer (5 mM NaH₂PO₄, 20 mM Na₂HPO₄, 400 mM NaCl, 2.5 mM EDTA, 0.05 % (w/v) CHAPS, 0.2 % (w/v) BSA, 0.4 % (w/v) Block ACE (Bio-Rad) (40 µl EC buffer + 40 µl protein samples at 4 mg/ml) and 75 µl were aliquoted in a 96-well plate.

To create a standard curve, standard samples were prepared by performing serial dilutions with poly-GPx7 peptide CGPGPGPGPGPGPGP in duplicate. After incubation, the plate was washed three times with 150 μ l TBST. The samples and standards were added to a pre-capture antibody coated plate followed by shaking for 2 hours at 700 rpm at RT. The required amount of detector antibody solution was prepared to be 25 μ l/per well by adding detector antibody and sulfotag streptavidin to the blocking buffer, and then vortexed for 5 seconds. The plate was washed again three times with 150 μ l TBST after incubation. 25 μ l of detector antibody solution was added to each well and then the plate was placed on shaker for 2 hours at 700 rpm at RT. The detector antibody solution was removed and the plate was washed three times with 150 μ l TBST. After the final wash, 150 μ l 1X reading buffer (1:1 ratio with dH₂O) was added to each well and the plate was immediately analysed using an MSD plate reader, and expression levels of poly-GP DPRs were calculated.

2.2.12. Statistical Analysis of Data

Gene Tools (Syngene) was used for the quantification of band intensity in western blot images. GraphPad Prism was used for the statistical analysis of data. All data were expected to be normally distributed in the experiments presented in this thesis and one-way or two-way tests were carried out by using recommended multiple comparison tests. Significance values were as follows: NS = Non-significant, $p \geq 0.05$; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; **** = $p < 0.0001$.

3. Chapter 3 - Expression and Purification of Cell Permeable Peptides

3.1. Introduction

My supervisor designed SRSF1-CPP constitute binding site of SRSF1 to NXF1 driven by the fact that amino acids 89-120 of the SRSF1 protein have a direct interaction with NXF1 (Tintaru et al., 2007). Furthermore, they integrated amino acids 47-57 from the HIV-1 trans-activator of transcription protein transduction domain (TAT PTD) into the SRSF1 peptide sequence, with the aim of enhancing intracellular delivery (Frankel and Pabo, 1988). To enable detection, a V5 tag was introduced, facilitating visualization of the peptide through Western blotting or immunohistochemistry techniques. This peptide was custom-chemically synthesized and it was tested in reporter HEK293T cells, patient-derived neurons and *Drosophila* model of *C9ORF72*-ALS/FTD (Castelli et al., 2023). The next stage of this project involved testing the peptide in the *C9ORF72*-ALS/FTD animal model. However, in order to demonstrate *in vivo* proof of principle for the SRSF1-CPP, a substantial amount of peptide (10 mg/ml) was required. This situation led to issues with both cost and time. The synthesis of the synthetic peptide took approximately 2 months and was not cost-effective (2000 £/ml). Therefore, in this chapter, I developed a robust methodology to produce recombinant peptide, aiming to eliminate both time and cost constraints. With the optimized production protocol, the recombinant peptide was generated within a week, and the average cost of the optimized recombinant protocol was significantly lower than that of the synthetic peptide. The average cost for the peptide expression step was calculated as 40.78 £ by dividing the quantities of chemicals used in the TB buffer by the cost of the purchased stock chemicals. The purification cost was calculated as 76.8 £ by dividing the volume of GST and heparin beads (5 ml each) used in the optimized protocol by the cost of the purchased stock. Developing a robust recombinant technique for peptide production presents a significant strategic edge. This approach not only refines the production phase but also accelerates the rapid screening of a broad spectrum of peptides, allowing for expedited evaluations across different cellular models. Moreover, recombinant SRSF1-CPP can be used as molecular tool to investigate the nuclear export pathway in any cells. For instance, patient-derived neurons are not amenable to plasmid transfection. If one aims to manipulate the nuclear export in neurons, it necessitates the creation of a lentivirus, a process that requires a time to constructing both the plasmid and the virus. However, with my method, I can simply introduce the peptide directly to the cell, significantly simplifying and speeding up the process. Also, beyond its utility as nuclear export tool, this peptide also holds promise as a potential therapeutic strategy for neuroprotection.

3.2. Expression and Purification of Recombinant Peptides

3.2.1. Recombinant Peptide Expression

Peptide production in general can be performed using three different methods: chemically, recombinantly, and enzymatically (Chao et al., 2020). Peptide production via the chemical method, based on solution-based synthesis and solid phase peptide synthesis, is the most prevalent technique due to providing flexibility in scale, sequence, and modification. Indeed, this method is efficient to produce a peptide shorter than 30 amino acids (Chandrudu et al., 2013). However, production of a peptide larger than 30 amino acids is challenging and expensive because of sequence related and high purification cost (Gaglione et al., 2019). Although chemical synthesis is an efficient method, it is not a favourable approach for large-scale peptide production due to the high cost of raw materials and the relatively low yield (Aguilar-Montes de Oca et al., 2022; Zhao et al., 2016). These limitations have led to the use of recombinant production platforms, which are capable of producing large peptides economically and on a large scale, while generating less waste for the environment (Wegmuller and Schmid, 2014). Several eukaryotic and prokaryotic expression systems such as insects and bacteria have been used for recombinant peptide production. In this thesis, I utilized a bacterial expression system based on *Escherichia coli* (*E.coli*) which is the most commonly used bacterial system and is relatively simple to manipulate, inexpensive to culture, with a short amount of time required to generate a heterologous recombinant peptide (Palomares et al., 2004). Recombinant peptide production in *E. coli* is a simple process containing serial steps which include: identifying the protein of interest; cloning its gene into a vector; transforming it into a suitable host station and inducing protein expression in an appropriate medium. After that, cells are harvested and lysed to perform downstream processes depending on the features of the target peptide such as the isoelectric point and the molecular mass of the protein of interest and affinity tag of the plasmid. The final step of the production process is the functional and structural characterisation of the peptide. Following the approval of recombinant insulin in the 1980s, there has been a significant increase in the production of recombinant peptides being used for therapeutic approaches. At present, there are more than 130 peptides which are used in clinical settings, and numerous others are in various stages of development, indicating that peptide-based drug development has become a highly popular and significant area (Wang et al., 2022).

Researchers have been focused on recombinant technology to express and purify various type of therapeutic peptide due to cost effective unlimited production, the ability to develop novel peptides and the ease of generating functional and structural modifications (Borghouts and Weiss, 2009). However, the production of recombinant peptides still faces common challenges such as poor expression, poor solubility and loss of functionality (Sarah and Sergio, 2014). To overcome these problems, several parameters of expression and purification are usually optimised. For example, the choice of host cells is important to achieve a high yield of peptide/protein expression. Host systems include mammalian cells (Bandaranayake and Almo, 2014), insect cells, plants, yeasts, and bacteria (Gecchele et al., 2015). Among them, *E.coli* BL21 (DE3) competent cells is the preferred method of producing recombinant peptides/proteins based on their high growth rate. Most functional therapeutic peptides have been obtained using *E. coli* (Rosano and Ceccarelli, 2014).

Culture medium composition, incubation temperature and induction time are considered to improve the expression yield of peptides in prokaryotic systems (Volontè et al., 2010). For example, TB medium has been shown to produce 117 mg / L wet biomass of an antimicrobial peptide, Pa-MAP 2, which is nearly twice as high than in LB culture (Sousa et al., 2016; Volontè et al., 2010). Fusion tags such as histidine (Kreisig et al., 2014; Svensson et al., 2006), NT11 (Nguyen et al., 2019) and glutathione-S-transferase (GST) (Harper and Speicher, 2001; Zhou et al., 2020a) have also been used as an alternative way to increase expression yield and also overcome solubility and purification challenges. For instance, thioredoxin (TrxA), a small molecular weight (12 kDa) soluble fusion tag was fused with viscotoxins to increase solubility. The result revealed that viscotoxins fused with TrxA gave a higher yield of soluble viscotoxins compared to processing without a tag (Bogomolovas et al., 2009).

The propensity of peptides to be degraded by proteases is a big challenge due to the small size of peptides and their usually high cationic charges. This proteolytic degradation leads to loss of function, limiting the peptide's stability and membrane permeability. Therefore, many peptides fail in *in vitro* assays as well as in pre-clinical studies and clinical trials even if they are successfully expressed and purified (Haney et al., 2019; Lee et al., 2019). For example, antimicrobial peptides (AMPs), which is a component of the innate defence system in most multicellular organisms, exhibit significant activity against bacteria, but the majority of AMPs are not able to maintain their *in vitro* activity under physiological salt and serum conditions thereby becoming less functional. (Lei et al., 2019).

In another example, only four out of ten novel antimicrobial peptides demonstrated good antimicrobial activity against *E.coli* strains (Wu et al., 2014). It is therefore important to screen for potential functionality to ensure that the peptide will persist and remain active. Aggregation and precipitation are also common issues in the peptide development process, which may result from effects of sequence, concentration, pH, net charge, excipients, chemical degradation and modification, surfaces and interfaces, and impurities (Zapadka et al., 2017). Indeed, aggregation problems reduce the physical stability of the peptide, leading to a loss of activity as well as toxicity and immunogenicity (Moussa et al., 2016; Zapadka et al., 2017).

Taking into consideration the frequent problems with peptides, many parameters are usually investigated to optimize their expression and purification such as the fusion tags, media composition, and *E.coli* BL21 host strains harboring additional tRNAs that recognize rare human codons. Different chromatography techniques such as affinity, ion exchange and size exclusion can also be used in combination to improve the purity of the peptide preparations. In the light of these suggestions, I tested multiple parameters for optimizing the expression and purification of the SRSF1-linker CPP.

3.2.2. Methodologies Used for Purification of Peptides

Different chromatography techniques are utilised for the purification of peptides based on their biochemical properties such as size, solubility, and charge. Some of the most commonly used methods are ion exchange chromatography, affinity chromatography, and size exclusion chromatography (Liu et al., 2020).

3.2.2.1. Ion Exchange Chromatography

Ion exchange chromatography (IEX) is a widely used purification technique, which separates molecules based on their net surface charge (Fritz, 2004). It can be used at different stages such as capture, intermediate purification, and polishing. It is able to separate similar proteins/peptides differing by just one charged group and also provide high resolution, high sample loading capacity, low cost, and rapid separation (Cummins et al., 2017). The purification principle of IEX is dependent on a reversible interaction which means that charged molecules and an oppositely charged IEX matrix interact with each other electrostatically to allow binding and elution of target molecules (Karlsson and Hirsh, 2011).

Since proteins possess zwitterionic and amphoteric properties, their net surface charge varies with pH; therefore, two forms of IEX chromatography media are employed as cation and anion exchangers. If the isoelectric point (pI) of the protein is above the pH of the solution, the protein is positively charged and binds to the cation exchanger (negatively charged), however when the pH is above the pI, the protein is negatively charged and binds to the anion exchanger (positively charged) (Cummins et al., 2017). The functional chemical groups providing the surface charges of anionic and cationic exchangers are shown in **Table 3.1**. These functional groups separate ion exchangers into two types known as strong and weak exchangers. The meaning of the weak and strong ion exchanger indicates change in the ionization degree of functional groups with pH because when the pH of the solution changes, the state of ionization changes (Mikšík, 2000). For strong exchangers, the changes in pH do not alter the ion exchange capacity, therefore protons are not lost. This situation is reversed for weak ion exchangers because they are prone to lose protons with pH changes and hence their ion exchange capacity shows alteration with pH. The meaning of strong and weak does not come from the strength of the functional groups that bind the target molecules (Cummins et al., 2011).

Table 3.1. Ion-exchange chromatography resins and their functional groups.

Anion exchangers	Weak or Strong	Functional group	Functional pH Range
Quaternary ammonium (Q)	strong	$-\text{CH}_2-\text{N}^+(\text{CH}_3)_3$	pH 1-14
Diethylaminoethyl (DEAE)	weak	$-\text{CH}_2-\text{CH}_2-\text{N}^+(\text{CH}_2\text{CH}_3)_2$	pH 2-9
Diethylaminopropyl (ANX)	weak	$-\text{CH}_2-\text{CHOH}-\text{CH}_2-\text{N}^+(\text{CH}_2\text{CH}_3)_2$	pH 2-9
Cation exchangers	Weak or Strong	Functional group	Functional pH Range
Sulfopropyl (SP)	strong	$-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{SO}_3^-$	pH 2-14
Methyl sulfonate (S)	strong	$-\text{CH}_2-\text{SO}_3^-$	pH 2-12
Carboxymethyl (CM)	weak	$-\text{CH}_2-\text{COO}^-$	pH 5-10

IEX purification consists of several steps including: exchange separation, equilibration, sample loading and washing, elution, and regeneration (**Figure 3.1**). IEX separation begins with equilibration of the stationary phase. Generally, 5 or 10 columns volume of equilibration buffer is applied over the stationary phase. The pH and ionic strength of the equilibration buffer should be compatible with the charged biomolecules of interest, otherwise their binding on the column will be compromised and denaturation can occur (Özlem Bahadır, 2013). The next step is application of the sample or mobile phase onto the column. To increase the binding yield of sample to column, loading and equilibration buffers should share the same conditions such as pH and ionic strength or be prepared at same composition. Once the molecules (peptide or protein) immobilize oppositely charged groups, the stationary phase is washed to remove molecules that bind non-specifically or weakly (Duong-Ly and Gabelli, 2014). After the washing step, the protein/peptide of interest attached on the column is eluted with either increasing salt concentration/ ionic strength or by changing the pH. At high salt concentrations, salt ions compete with immobilized biomolecules of interest, disrupting the electrostatic interaction of bound molecules with the stationary phase. The target proteins move down from the column based on the degree of their net charge. The lowest charged molecules are eluted first and the highest are eluted last due to weak and strong electrostatic interaction with column. In the final recovery stage, the column is washed with high salt buffer to remove any impurities. The above-mentioned purification stages are generally applied for all types of chromatography techniques (Cummins et al., 2017; Fritz, 2004).

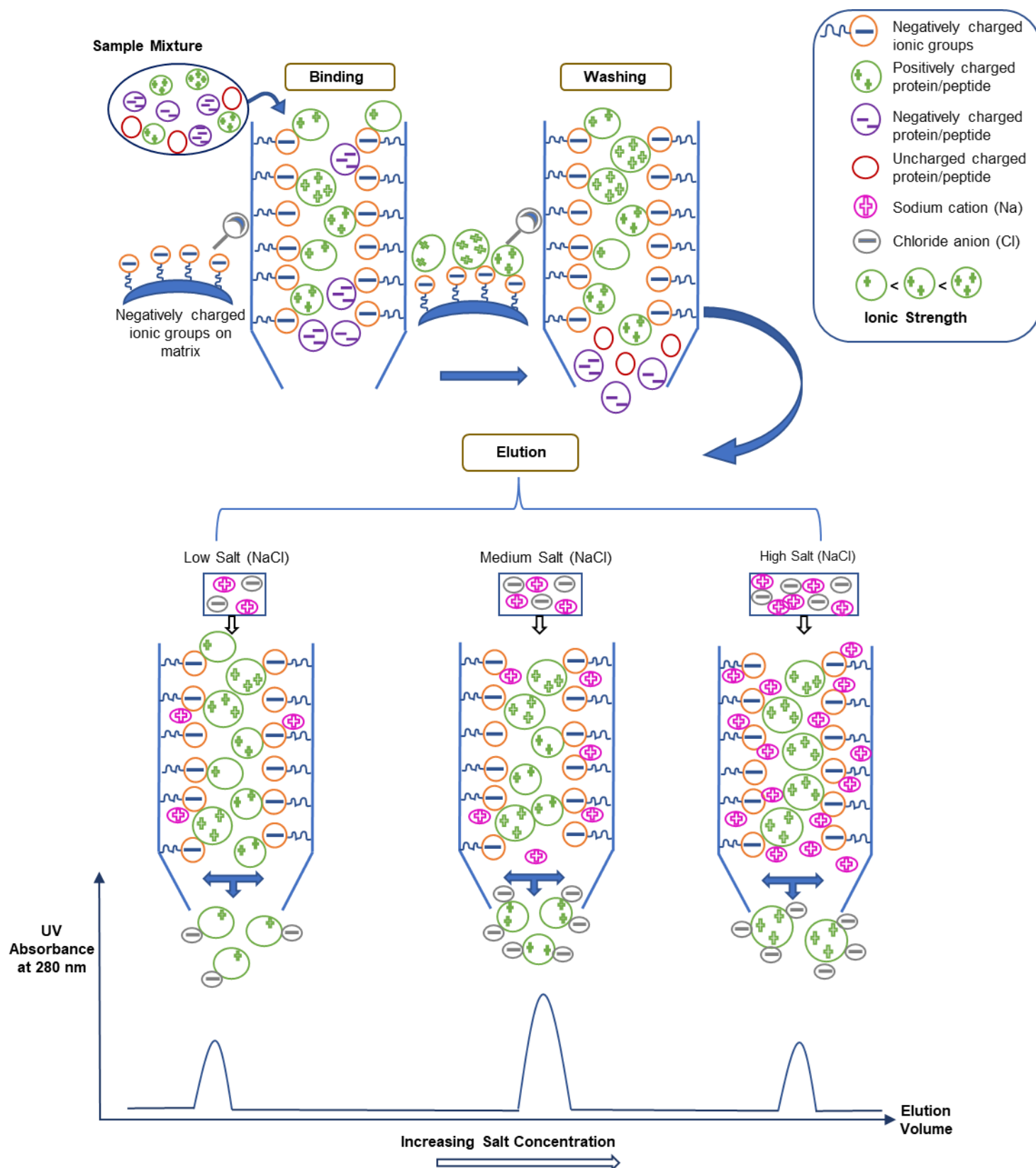


Figure 3.1. Schematic representation of typical anion exchange chromatography separation steps.

3.2.2.2. Affinity Chromatography

Affinity chromatography (AC) is one of the most efficient purification methods of biomolecules based on specific reversible interaction between the target molecule and the specific ligand which is covalently attached to the matrix (Hage, 1999). The ligand has to interact strongly with the target molecules and keep its functionality or activity during the washing steps (Lowe, 2001). AC purification is generally employed for capture purposes and intermediate purification due to offering good resolution, high purification efficiency and selectivity. A diagram representing a typical affinity purification is shown in **Figure 3.2**. Separation begins with loading the sample containing mixture of target and non-target molecules to the stationary phase and then biological interaction between the target molecule and ligand occurs. These interactions can be enzyme: substrate; antibody: antigen; lectin: polysaccharide; ligand: cell surface receptor; glutathione: glutathione-S-transferase (GST); or GST fusion proteins and metal ions: histidine-tagged proteins. These intermolecular interactions can be driven by electrostatic, hydrophobic interactions, van der Waals forces and hydrogen bonding (Özlem Bahadır et al., 2013; Urh et al., 2009). After sample loading, the column is washed to remove unbound molecules and impurities. To recover the immobilized target molecules from the matrix, the interaction between the target and ligand is reversed either by specifically using a competitive ligand or non-specifically by changing conditions such as pH and ionic strength. Finally, the column is regenerated by washing to remove any residues for subsequent purification. Immobilized metal ion and glutathione affinity chromatography are the most commonly utilised form of AC technique to respectively purify hexa-histidine-tagged proteins and GST-tagged proteins (Hage and Cazes, 2005).

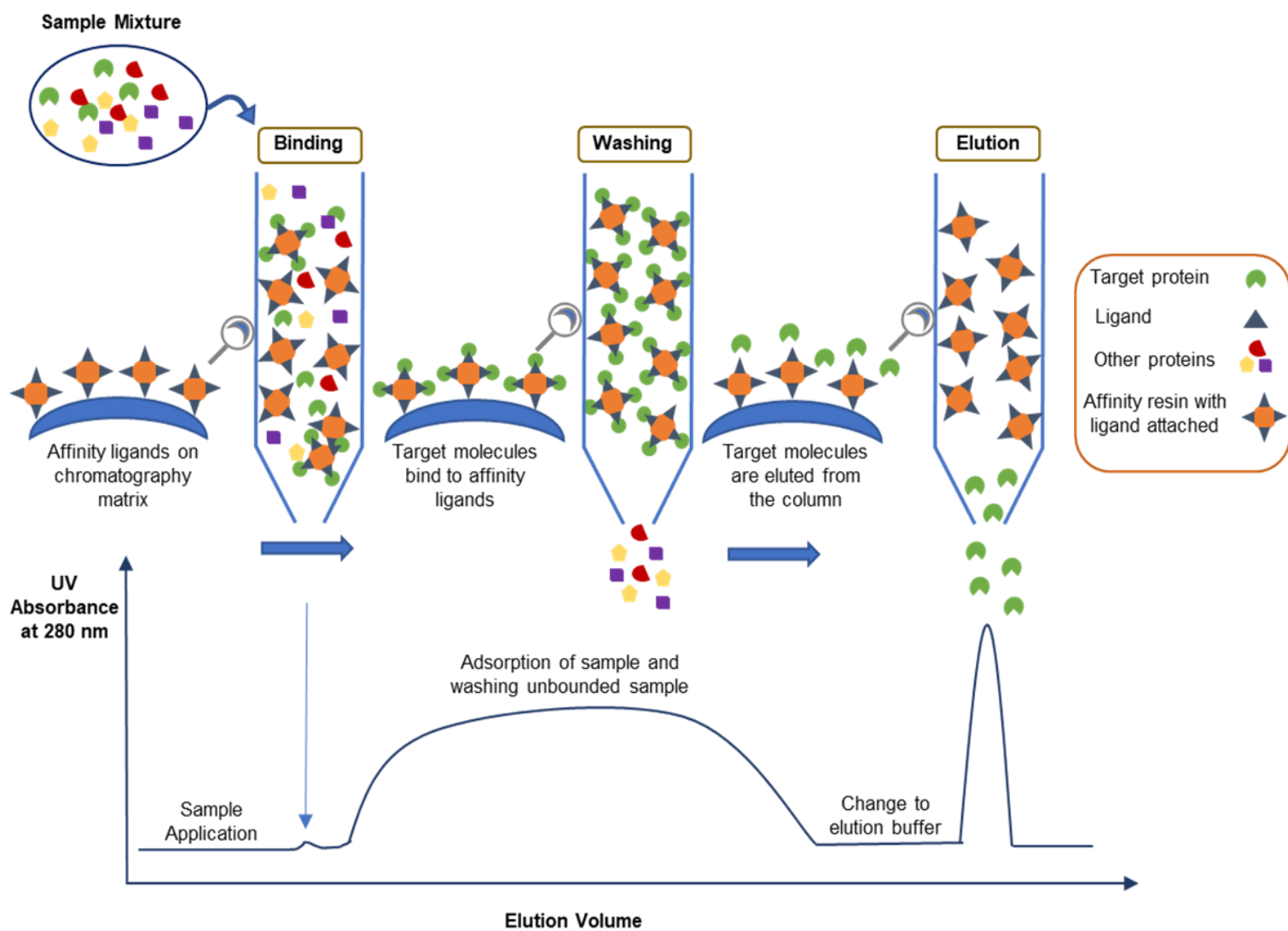


Figure 3.2. Schematic representation of general affinity chromatography purification steps.

3.2.2.2.1. Immobilized Metal Ion Affinity Chromatography

Immobilized metal ion affinity chromatography (IMAC) is the most ideal purification technique for biological molecules which have an affinity for divalent metal ions such Zn^{2+} , Ni^{2+} and Co^{2+} or other transition ions which attach to the surface of the chromatography matrix by chelation, whereby molecules with affinity bind selectively and others do not bind or weakly bind (Fanou-Ayi and Vijayalakshmi, 1983; Ueda et al., 2003). Among the molecules with affinity, 6 to 14 histidine (His) tag fused molecules have the highest affinity for Ni^{2+} ions, making this the most preferred affinity tag (Zhao et al., 2013). The purification steps of IMAC are similar to other chromatography techniques beginning with equilibration of the column, followed by sample loading and washing. Elution of histidine tagged molecules is performed using buffer containing imidazole which competes with histidine fused molecules to bind immobilized Ni^{2+} ions. The reason for use of imidazole is that molecules with histidine tags bind to metal ions with the help of a histidine imidazole ring therefore imidazole eases the recovery of bound molecules from the column (Ueda et al., 2003).

3.2.2.2.2. GST-Fusion Protein Purification

Glutathione S-transferase is a 26-Kda protein which is widely used as an affinity tag due to having strong specificity for reduced glutathione (GSH), which is composed of three peptides: glutamic acid (Glu), cysteine (Cys), glycine (Gly). These peptides provide a binding platform for a GST fused protein and hence separating them from other proteins occurs selectively (Schäfer et al., 2015). GST is ideal for enhancing the expression and solubility of fusion proteins in prokaryotic cells and protecting recombinant proteins from cleavage by proteases. GST is a cytosolic protein which is not present in prokaryotic cells (*E.coli*) and is highly soluble. The purification steps of GST fusion proteins are similar to other affinity tag purifications (Harper and Speicher, 2011). Unlike IEX chromatography, AC may require specific buffer types and conditions. For example, Tris buffer at pH 7.5 is recommended to increase the binding capacity of the GST-tagged protein to the column. The target protein is recovered from the column with the help of a buffer containing reduced glutathione that competes with immobilized glutathione bound GST on the column, enabling dissociation of the GST fusion protein. After GSH purification, GST may be subjected to enzymatic cleavage if not required. For enzymatic cleavage, Tris buffer is preferred as it protects the structure of the fusion protein and enhances the activity of the enzyme (PreScission protease) (Harper and Speicher, 2011).

3.2.2.3. Size Exclusion Chromatography

Size exclusion chromatography (SEC) is another commonly used purification method that separates molecules such as proteins, polymers and macromolecules based on differences in molecular size according to their Stokes radius as they pass through the stationary phase (Kilz and H.Pasch, 2000). SEC is ideal for polishing or intermediate purification to remove retained impurities. It is rarely preferred for first step purification or capture, but it is powerful method for group separations due to simplicity of use and scale up, versatility, reliability (Hong et al., 2012). The purification stage of SEC is similar to AC and IEX (**Figure 3.3**). In SEC purification, the column is equilibrated prior to sample application. Then, the sample mixture containing variably sized molecules is applied to the column. Molecules larger than the matrix pores directly pass through the column. The other small and intermediate sized molecules which are able to diffuse across the pores of the matrix are eluted isocratically in order of decreasing molecular size. Other molecules smaller than the pores such as salts pass through the column and are eluted last and collected in the total liquid volume (Striegel et al., 2009). SEC chromatography provides several advantages. One of the main advantages is that the buffer composition does not have an effect on the degree of purification because molecules do not immobilize to the stationary phase. Since there is no binding step occurring in the column, this reduces the potential risk for molecules that are prone to precipitation or aggregation (Kilz and H.Pasch, 2000). Furthermore, SEC purification can be performed under mild conditions which do not affect the stability and activity of target molecules (Held and Kilz, 2021). However, it is important to note that when the volume of sample exceeds the capacity of column, this results in band broadening which adversely affects the column separation efficiency and performance and also leads to poor resolution. In addition to this, a large volume of elution causing sample dilution and low resolution are another common issue during size exclusion chromatography purification. Overall, SEC is still very common purification technique that is preferred in a variety of applications (O'Fágáin et al., 2011).

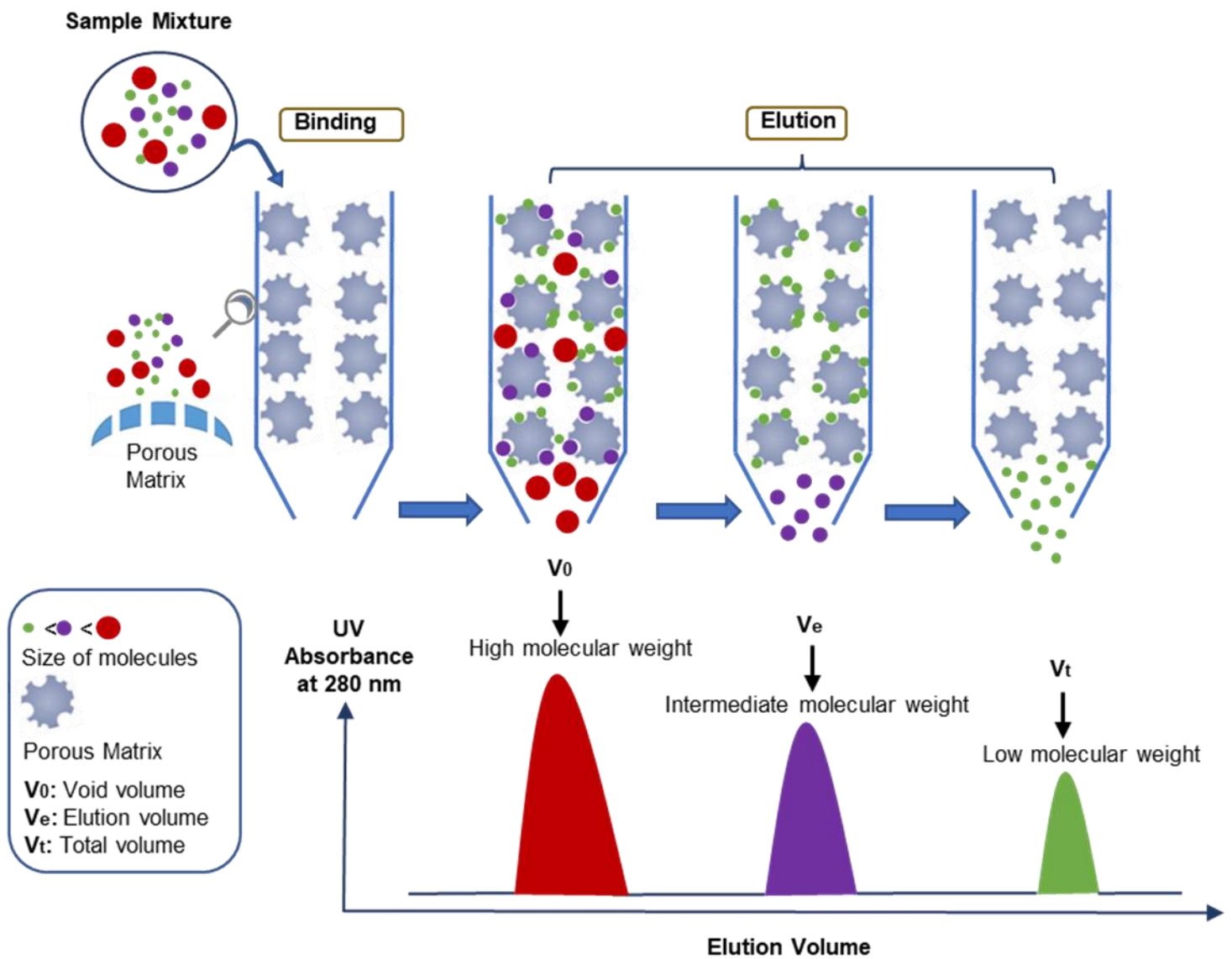


Figure 3.3. Flow diagram illustrating steps for size exclusion chromatography.

3.3. Aims and Objectives

- To design an SRSF1-derived CPP that competitively inhibits the interaction between NXF1 and SRSF1, therefore disrupting the nuclear export of the pathological repeat transcripts and downstream DPR-mediated cytotoxicity in initially cell models.
- To evaluate various strategies of expression and purification, including harvesting times of bacteria, bacterial culture medium, lysis buffers compositions, different chromatography media to optimize a protocol to produce recombinant SRSF1-linker and control peptides with a high degree of purity and abundance to be used in *in vitro* and *in vivo* C9ORF72-ALS/FTD models.

3.4. Results

3.4.1. Designing SRSF1(89-120)-V5-TAT PTD CPPs

Obtaining a high yield and efficient purification of peptides from host cells is challenging; therefore, peptides are usually chemically synthesised by companies. However, since the overall goal of my PhD is to investigate the biodistribution and neuroprotective potential of the SRSF1 cell-permeable peptide in mice, large amounts of the peptide are required. In terms of cost and time considerations, the first aim of my project was to develop a robust method for the production of recombinant CPPs from *E. coli*.

In order to explore methodologies used to express and purify peptides, three plasmids consisting of SRSF1, V5 and TAT PTD sequences expressing either a carboxyl-terminal hexahistidine or an amino-terminal cleavable GST fusion of SRSF1 89-120 were engineered by Prof Hautbergue. These plasmids were built into the pGEX-6P1-GST (Glutathione-S-transferase) and pET24b-GB1 (immunoglobulin-binding domain B1 of streptococcal protein G)-His₆ vector backbones to tag the peptides for affinity chromatography and potentially to enhance the expression of the recombinant peptides with the highly-expressed and soluble GST/GB1 tags. In addition, a tryptophan sequence was inserted or not between the V5 and TAT PTD domains of the His₆-tagged SRSF1 peptide to increase UV absorbance, making it easier to follow the peptides eluted from fast protein liquid chromatography (FPLC). The generated plasmid sequences are provided below:

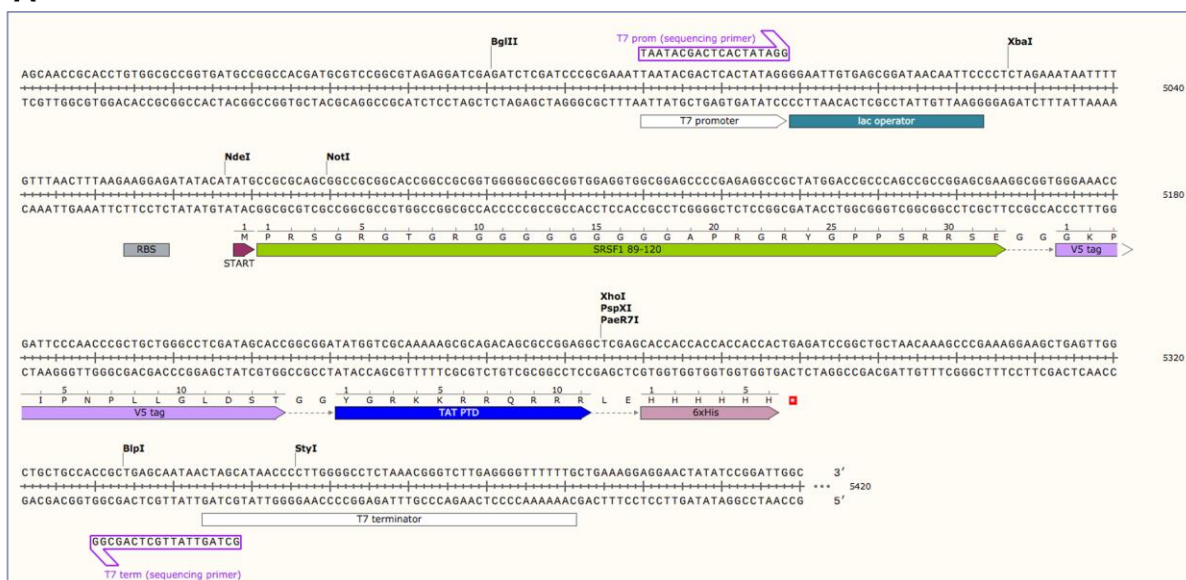
Strategies 1 and 2 (with or without presence of tryptophan):

SRSF1(89-120)-V5-TAT PTD- (His)₆

PRSGRGTGRGGGGGGGGGAPRGRYGPSSRRSEGGKPIPNNLLGLDSTGGYGRKKRR
QRRRHHHHHH

The full sequence maps of plasmids expressing SRSF1(89-120)-V5-TAT PTD-His₆ (**Figure 3.4.A**) and SRSF189-120-V5-W-TAT PTD-His₆ peptides (**Figure 3.4.B**) are provided below. The peptide expression starts from methionine to six histidine tag residue (His₆-tag) which possesses strong interaction with metal ions. Several studies suggest that short histidine tags eliminate the possibility of impaired protein function compared to the use of long histidine tags (Mohanty and Wiener, 2004). His-tags are easily incorporated into the protein of interest and directly allow expression of the SRSF1 cell-permeable peptide for subsequent purification using Co²⁺ immobilized metal affinity chromatography (IMAC) (Zhao et al., 2013). However, this method might not be sufficient for expressing the peptide at a high level /or expressing insoluble peptide. The low expression level of the peptide may lead to non-specific binding of untagged protein to the IMAC column, thus causing contamination in the protein of interest (Stevens, 2000).

A



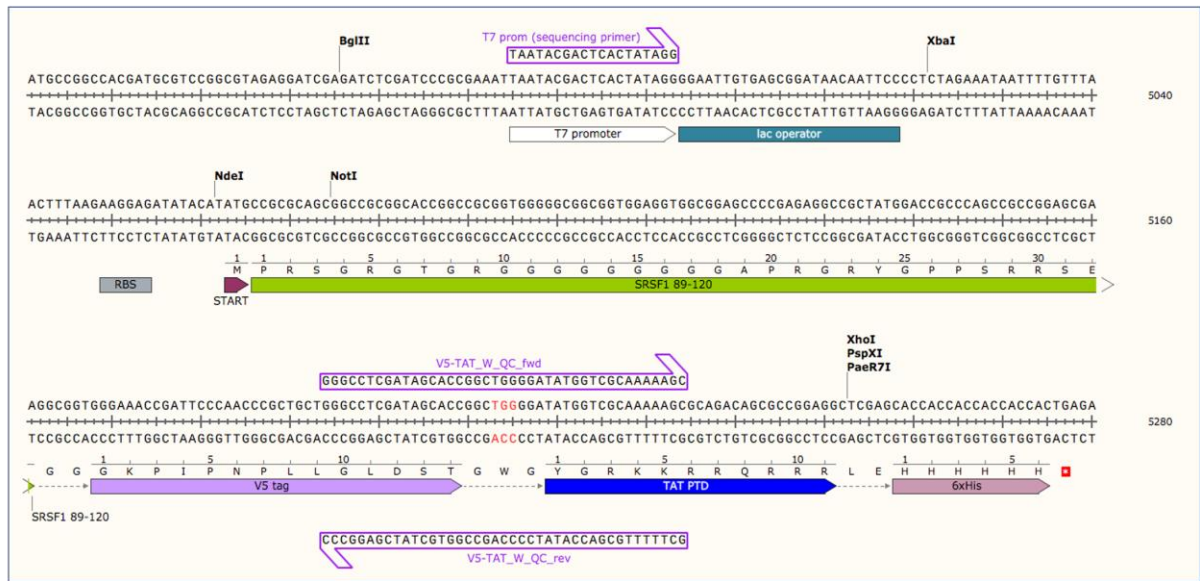
B

Figure 3.4. Plasmid maps of engineered SRSF1(89-120)-V5-TAT PTD-His6 and SRSF1 (89-120)-V5-W-TAT PTD-His6. (A) pET24b-GB1 plasmid map showing integration of His₆-tag into SRSF1(89-120)-V5-TAT PTD. **(B)** pET24b-GB1 plasmid map showing integration of His₆-tag into SRSF1(89-120)-V5-W-TAT PTD. Both DNA sequences and encoded amino-acids in the one-letter code are highlighted on the map.

Strategy 3:

GST-SRSF1(89-120)-V5-TAT PTD

GST-PreScission protease site-

PRSGRGTGRGGGGGGGGGAPRGYGPSSRRSEGGKPIPNNLLGLDSTGGYGRKKRR
QRRR

Another strategy is to fuse the pGEX-6P1 vector backbone with a GST affinity tag to enhance the expression of the SRSF1 peptide and improve its solubility. The GST-tag also acts as a chaperone to promote protein folding. GST fusion proteins are purified easily with glutathione affinity chromatography because glutathione is a specific substrate for GST which contains Glu-Cys-Gly peptides (Brown et al., 2008; Esposito and Chatterjee, 2006). However, a PreScission protease cleavage step is required to release the protein of interest. Also, the protein/peptide might be contaminated with uncleaved or cleaved GST products, therefore additional purification step(s) may be required (Deceglie et al., 2012). The sequence map of the plasmid expressing GST- SRSF1-V5-TAT PTD is shown in **Figure 3.4.C**. Since GST is a large tag, only a small part of its sequence is shown.

3.4.2. Production of Recombinant SRSF1(89-120)-V5-W-TAT (His)₆ CPP

3.4.2.1. Expression of Recombinant SRSF1(89-120)-V5-W-TAT (His)₆ CPP

E.coli BL21-RIL, *E.coli* BL21-RP and *E.coli* BL21 -Rosetta competent cells used for the expression of SRSF1(89-120)-V5-W(tryptophan)-TAT PTD-(His)₆ were grown in medium to investigate which bacterial strains would express the highest amount of His-tagged SRSF1-CPP containing tryptophan. The harvesting times of cells between un-induced (0 h) and IPTG induced cultures at different induction times (1 hour to overnight (O/N)) were tested to determine the optimal expression time of the SRSF1 peptide. Harvested total protein extracts were assessed by SDS-PAGE gel followed by Coomassie staining and western blot (**Figure 3.5**).

Coomassie blue staining of the SDS-PAGE did not detect specific expression of the peptide around 6.5 kda (approximately expected in 6500 g/mol) over increasing induction time for all bacterial strains. However, the SRSF1 peptide expression was detected by western blot using a V5 monoclonal antibody, indicating overall that the expression of the His₆-tagged SRSF1 peptide is very low. Western blotting indicated that the total amount of expressed His₆-tagged SRSF1 peptide in bacterial lysate initially increased following IPTG induction then subsequently decreased for all bacteria strains. The maximum peptide expression time was attained 5 hours post IPTG induction for RP cells (**Figure 3.5.A**), 4 hours for Rosetta cells (**Figure 3.5.B**), and 3 hours for RIL cells (**Figure 3.5.C**). RP cells at 5 hours IPTG post induction time has shown the maximum peptide expression because it gave more cells weight in total compared to just 3 hours induction time. Therefore, subsequent expression experiments were carried out using *E.coli* BL21-RP at 5 hours induction time.

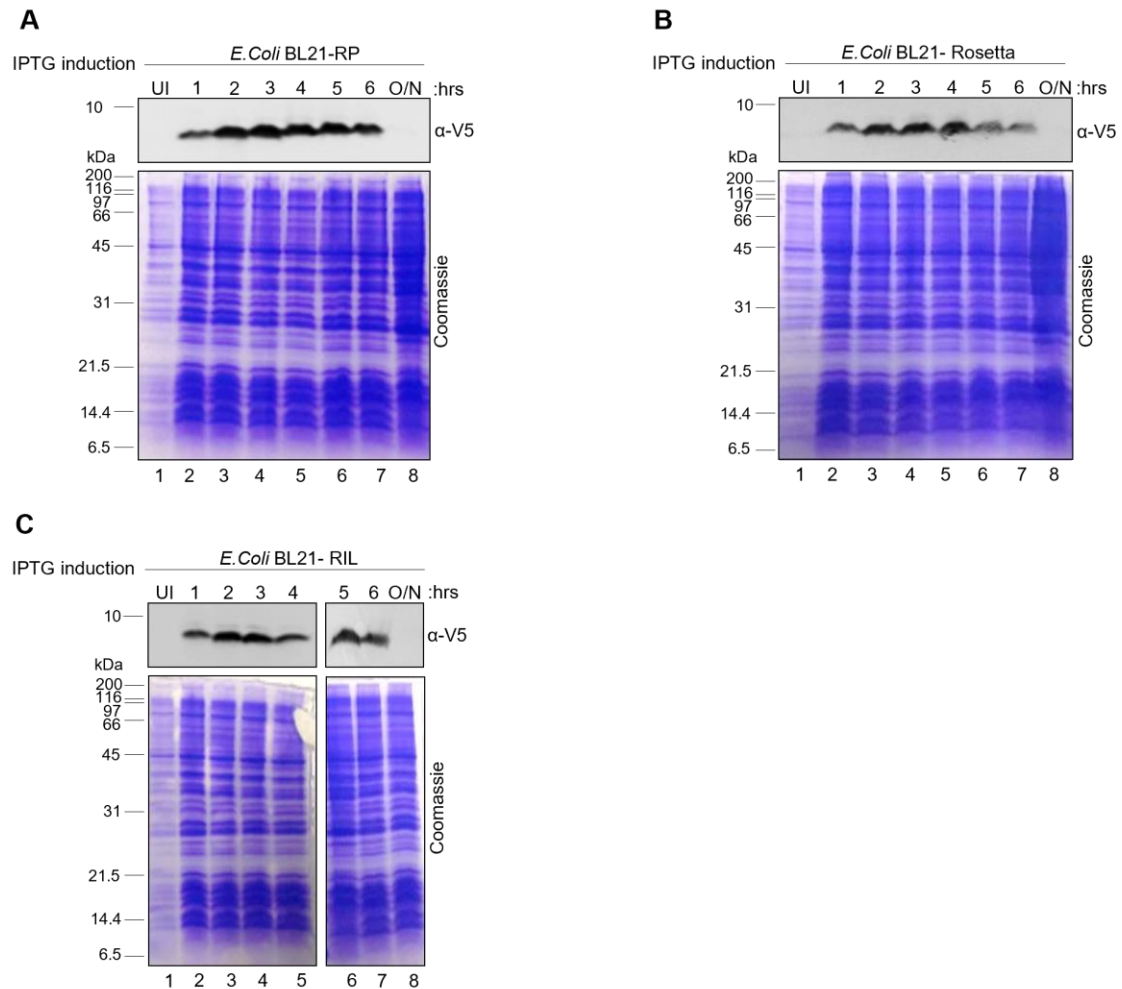


Figure 3.5. Induction time course of SRSF1(89120)-V5-W-TAT PTD-(His)6 in BL21-RP, BL-21 Rosetta and BL21-RIL *E. coli* strains. (A) Western blot and Coomassie blue stained 15 % resolving acrylamide gel to analyse SRSF1-linker CPP expression of: (A) BL21-RP cells, (B) BL21-Rosetta cells, and (C) BL21-RIL cells. **Lane 1:** Un-induced bacteria culture. **Lane 2-8:** IPTG induced culture of 1, 2, 3, 4, 5, 6 hours and overnight (O/N). Western blots were probed with V5 antibody. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

3.4.2.2. Purification of Recombinant SRSF1(89-120)-V5-W-TAT (His)₆ CPP with Mono-S Ion Exchange Chromatography

Purification of recombinant SRSF1(89-120)-V5-W(tryptophan)-TAT PTD-(His)₆ peptide expressed in BL21-RP cells was performed by two-step purifications. Histidine tagged SRSF1 peptide was first purified by batch cobalt ion metal affinity chromatography (IMAC, 4 ml), which was used to isolate the peptides bound on Co²⁺ coated beads, depending on coordinated interaction between histidine and Co²⁺ metal ions. Histidine-tagged SRSF1 peptides have a high affinity in IMAC because of the multiple hexa-histidine ligands and are usually the strongest binder among other cellular proteins in a crude sample extract (Hochuli et al., 1987). To facilitate cobalt affinity purification, cell extract was lysed in cobalt lysis buffer containing 1 M NaCl. The cobalt column beads were then equilibrated by washing with cobalt wash buffer (250 mM NaCl) and the cell extract was applied to the column. Peptide was eluted with 200 mM imidazole which competes with His₆ tagged peptides to bind to the IMAC column. The eluted fraction, flow through (FT), and total cell lysate (Input) were resolved on 15 % SDS-PAGE gel followed by anti-V5 western blotting to confirm its identity as SRSF1-CPP and Coomassie blue staining to confirm its purity (**Figure 3.6.A**).

The results showed that there is sufficient expression of SRSF1 peptide at approximately 6.5 kDa in total extract as seen in cobalt input lane (**Figure 3.6.A**). The SRSF1-CPP binds well to the cobalt resin as there is a limited amount of peptide detected in FT. The bound complex successfully eluted with imidazole, but it was insufficiently pure based on the presence of contaminants above the peptide band. Therefore, the second step of purification was performed with a mono-S (methyl sulfonate) strong cationic ion exchange chromatography column. This chromatography method is suitable because the TAT protein transduction domain inserted into the peptide sequence is rich in arginine and lysine conferring positive charges at the pH used here. The eluted peptide fraction was loaded (input) at a flow rate of 1 ml/min over the mono-S column by diluting in loading buffer containing 200 mM NaCl. The stationary phase was then washed with mono-S washing buffer containing 200 mM NaCl and flow through was collected and subjected to western blot and Coomassie blue staining (**Figure 3.6.A**).

The elution of SRSF1 peptide began after 28 ml elution buffer, which contains ~ 1.9 M high salt, passed through the column. The eluates were selected using a chromatogram of mono-S purification because the SRSF1 peptide possesses a tryptophan residue which is expected to increase UV absorbance at 280 nm (**Figure 3.6.B**). The selected fractions were analysed by V5 western blotting (**Figure 3.6.C**). Results show that SRSF1-CPP also binds efficiently to the mono-S column as no peptide was detected in the FT. However, the peptide was not eluted from the stationary phase in a sharp peak as shown by anti-V5 western blotting (**Figure 3.6.C**). The peptide elution commences from fractions 11 to 32, with a high salt concentration towards the gradient's conclusion (2 M NaCl). This suggests that the peptide wasn't successfully eluted but remained adhered to the stationary phase. Consequently, the peptide's presence extended across the entirety of the eluted fractions, resulting in the manifestation of multiple peaks rather than a one peak. Incorporation of the tryptophan sequence into the SRSF1 peptide does not appear to be a good method of detecting the peptide through UV absorbance because the peptide-containing fractions assessed with V5 western blot (**Figure 3.6.C**) did not match the fractions in the chromatogram that were assumed to contain peptides. Moreover, the peptide could not be detected in selected fractions at high intensity using the mono-S chromatogram. A weaker exchange column was therefore tested in section 3.3.2.3.

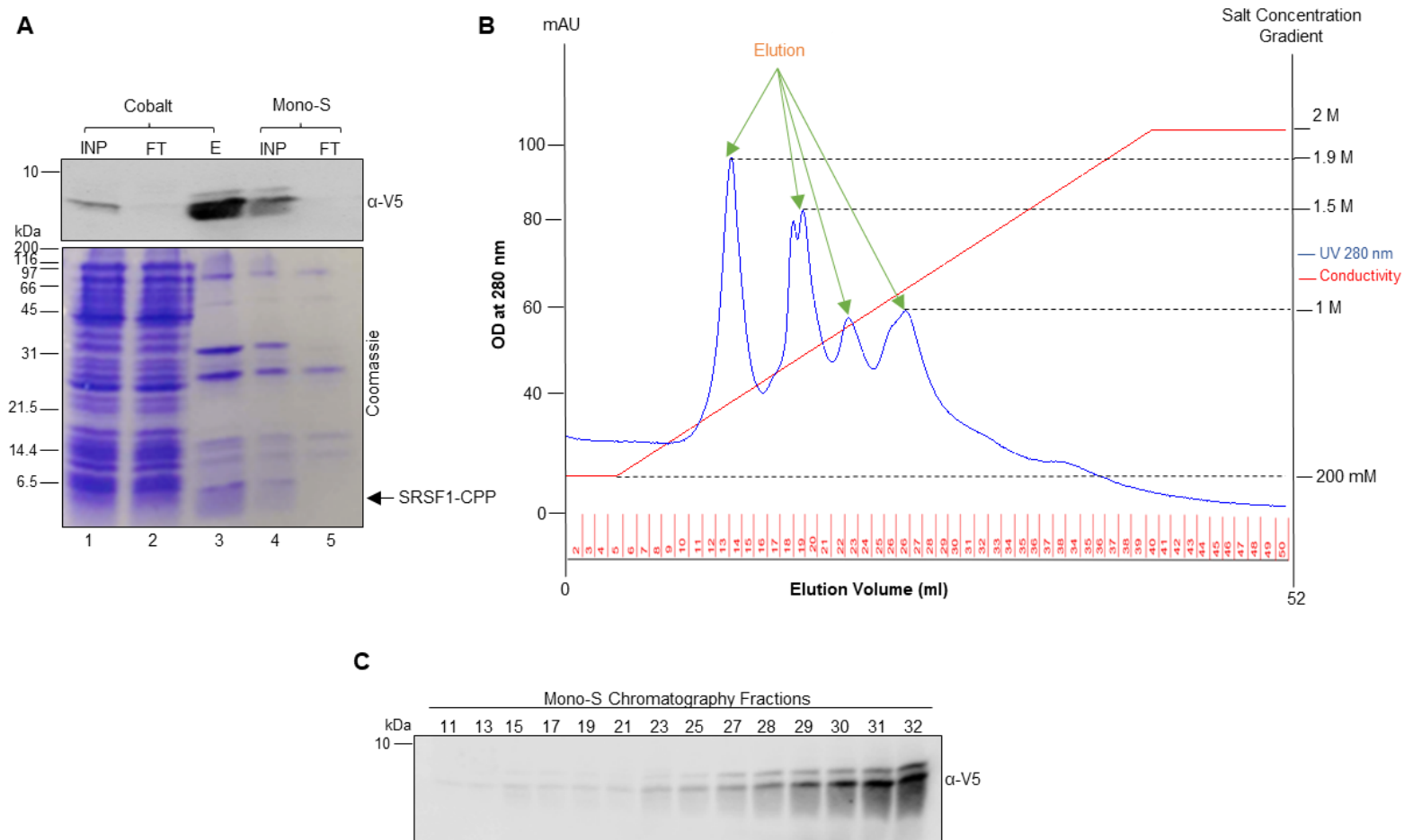


Figure 3.6. Two-step purification of SRSF1(89-120)-V5-W-TAT PTD-(His)₆ CPP with Cobalt affinity and Mono-S strong cationic ion exchange chromatography. **(A)** Western blot and Coomassie blue stained 15 % resolving acrylamide gel to analyse to input (INP), flow through (FT) and eluates (E) from cobalt and mono-S chromatography. **(B)** AKTA FPLC elution profile/chromatogram of eluted fractions (Fr) from Mono-S strong cationic ion exchange chromatography. **(C)** Western blot analysis of selected fractions with high intensity based on the chromatogram. Arrows indicates high salt elution of SRSF1(89-120)-V5-W-TAT-(His)₆ CPP. Western blots were probed with V5 antibody. **Lane 1:** Input, **Lane 2:** Flow through, Lane 3: Elution from Cobalt chromatography. **Lane 4:** Input for mono-S column, **Lane 5:** Flow through from mono-S chromatography. Western blots were probed with V5 antibody. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

3.4.2.3. Purification of Recombinant SRSF1(89-120)-V5-W-TAT (His)₆ CPP with SP Ion Exchange Chromatography

To overcome the difficulties in eluting the peptide from mono-S strong ion exchange chromatography, a weaker cation exchange chromatography column called sulfopropyl (SP) was utilized to perform large scale purification of SRSF1(89-120)-V5-W-TAT PTD-(His)₆ CPP. As before, the recombinant SRSF1(89-120)-V5-W-TAT PTD-(His)₆ CPP was expressed in BL21-RP cells. The cell pellet was lysed in appropriate lysis buffer containing 200 mM NaCl, then applied at a flow rate of 2 ml/min on pre-equilibrated 10 ml SP ion exchange chromatography columns controlled by an AKTA FPLC system. The column was washed with SP wash buffer to eliminate contamination by collecting unbound proteins in the eluates. The peptide was eluted with a stepwise salt gradient (2 M NaCl) to ensure that the proteins with the lowest net charge would be eluted from the column first and the proteins with the highest charge will be eluted last as the salt ions compete with the bound protein. Eluted fractions were collected in 5 ml elution volume after 50 ml of elution buffer passed through the column.

The fractions with high intensity peptide based on the chromatogram (**Figure 3.7.A**) were selected and analysed by resolving on 15 % SDS-PAGE gel followed by anti-V5 western blotting and Coomassie blue staining (**Figure 3.7.B and 3.7.C**). Coomassie blue staining and western blotting indicate that the peptide elution starts at fraction 16 and continues until fraction 25. However, the majority of the peptide was eluted in fractions 19 and 20, when the salt gradient reaches 400 mM. This observation suggests that the peptide is eluted within a single distinct peak. Purification on an SP column also revealed two other bands which may represent contaminants with higher molecular weight. Indeed, SP chromatography showed improved purity compared to the first purification step on cobalt chromatography. Also, the peptide yield from the SP column was higher than with the mono-S column. Therefore, the SP column was utilised for further experiments.

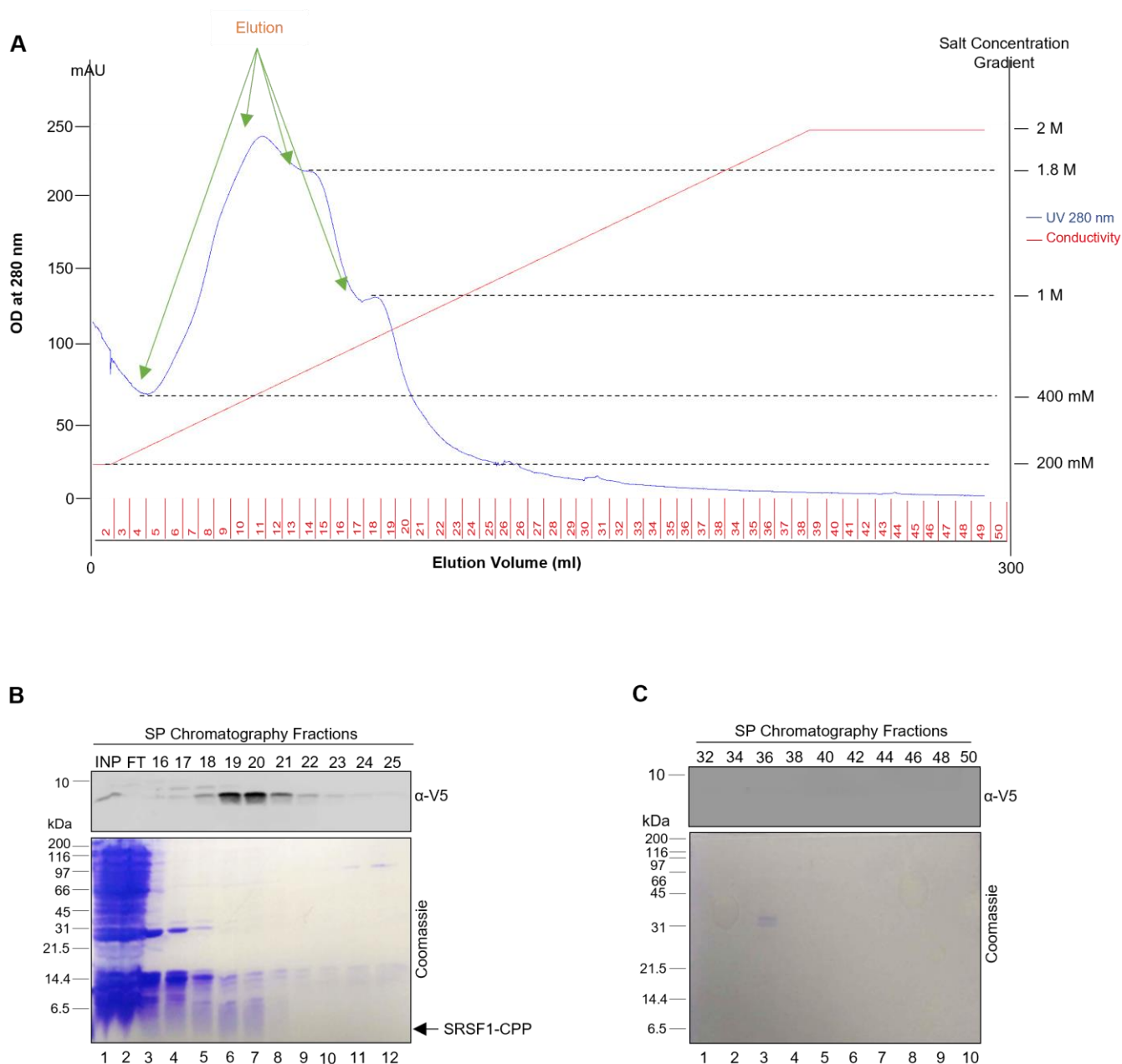


Figure 3.7. Two-step purification of SRSF1(89-120)-V5-W-TAT PTD-(His)₆ CPP with Cobalt affinity and SP weak cationic ion exchange chromatography. (A) AKTA FPLC elution profile/chromatogram of eluted fractions (Fr) from Mono-S strong cationic ion exchange chromatography. **(B)** Western blot and Coomassie blue stained 15 % resolving acrylamide gel to analyse to input (INP), flow through (FT) and fractions 16-25. **Lane 1:** Input, **Lane 2:** Flow through, **Lane 3-12:** Fractions between 16-25. **(C)** Western blot and Coomassie blue stained 15 % resolving acrylamide gel to analyse to input, flow through and fractions between 32-50. **Lane 1-12:** Fractions between 32-50. Arrows indicate high salt elution of SRSF1(89-120)-V5-W-TAT-(His)₆ CPP. Western blots were probed with V5 antibody. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

The fractions that yielded the highest amount of peptide (fractions 19 and 20) were combined, resulting in a total volume of 10 ml. This pooled solution was subsequently introduced to a stationary phase employing cobalt affinity chromatography, aiming to concentrate the peptide. In this process, 1 ml of cobalt affinity slurry was utilized. The concentrated peptide solution was analysed by SDS-PAGE followed by anti V5 western blotting and Coomassie blue staining (**Figure 3.8.A**) to visualize final purity and quantity of peptide. The Coomassie blue and western blot results showed that histidine tagged SRSF1 peptide bound very well to the stationary phase of cobalt chromatography which represented an efficient means of concentrating the peptide. However, the final concentrated peptide pool exhibited a higher level of contaminants/impurities or degradations products. Also, the quantity of peptide was estimated to be around 100 µg when comparing with a reference SDS-PAGE gel (**Figure 3.8.B**) upon which increasing known amounts of synthetic SRSF1-CPP were loaded. This indicated that the yield of the recombinant peptide was not purified in high enough concentration (approximately 10 mg) which would allow future use in *in vitro* and *in vivo* studies. Therefore, new expression strategies were developed to obtain a larger amount of peptide with enhanced purity.

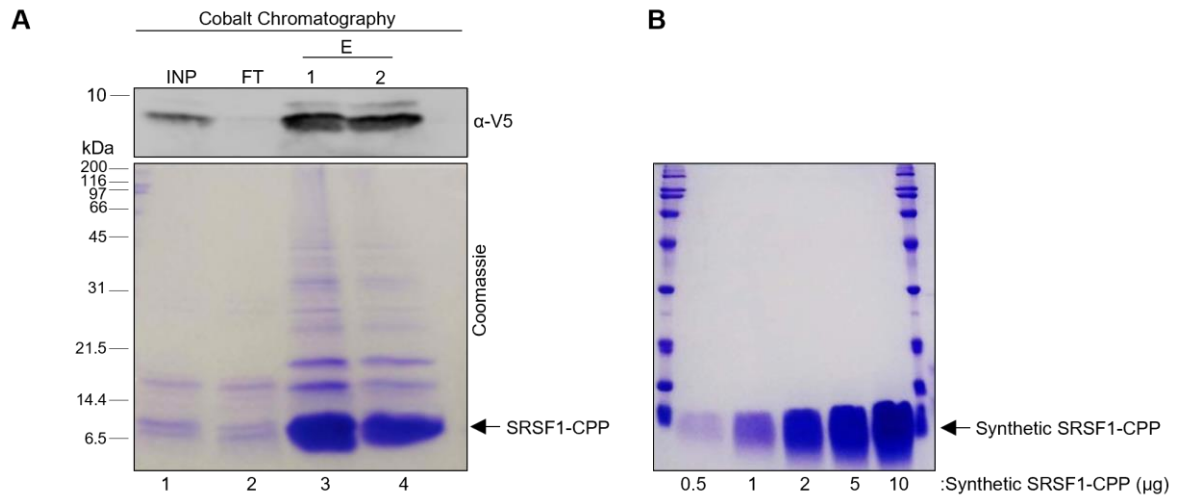


Figure 3.8. Concentrating SRSF1(89-120)-V5-W-TAT PTD-(His)₆ CPP fractions with Cobalt affinity chromatography. (A) Coomassie stained 15 % acrylamide gel analysis of input, flow through (FT) and elution 1 and 2 after loading SRSF1(89-120)-V5-W-TAT PTD-(His)₆ to the stationary phase of Cobalt affinity chromatography. **(B)** Coomassie stained 15 % SDS-PAGE gel analysis of increasing concentration of synthetic SRSF1-CPP. **Lane 1:** Input (INP), **Lane 2:** Flow through (FT) **Lane 3:** Eluate-1 (E1) from SP chromatography, **Lane 4:** Eluate-2 (E2) from SP chromatography. Arrow depicts concentrated SRSF1(89-120)-V5-W-TAT PTD CPP. Western blot was probed with a V5 antibody. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

3.4.3. Optimisation of SRSF1(89-120)-V5-W-TAT PTD-(His)₆ CPP Expression (with the presence of tryptophan)

Since the desired amount of SRSF1(89-120)-V5-W-TAT PTD-(His)₆ peptide was not initially obtained for future experiments, several strategies and parameters were investigated for enhancing the expression yield of the histidine tagged SRSF1 peptide with tryptophan. The first approach was to test different growth media to explore optimal growing conditions of BL21-RP cells for SRSF1 peptide expression.

BL21-RP *E.coli* cells expressing histidine tagged SRSF1 peptide were grown in SB medium and the samples corresponding to different IPTG induction time points were collected to compare the amount of peptide expressed with the peptide expressed in TB medium. The total expressed amount of peptide in SB medium was determined by resolving the samples on 15 % SDS-PAGE gels followed by Coomassie blue staining and anti V5 western blotting (**Figure 3.9**). The results show that in SB medium, the expressed peptide was not visible following Coomassie blue staining, but the western blot showed that peptide expression could be detected 1 hour post IPTG induction, and that the peptide intensity starts decreasing gradually over time. SB medium did not make any positive contribution to peptide expression in comparison to TB culture (**Figure 3.5**) because the intensity of the band corresponding to peptide at 1 hour induction, which is the highest intensity in the SB culture, is considerably lower than the band density at the 5th hour attained from TB.

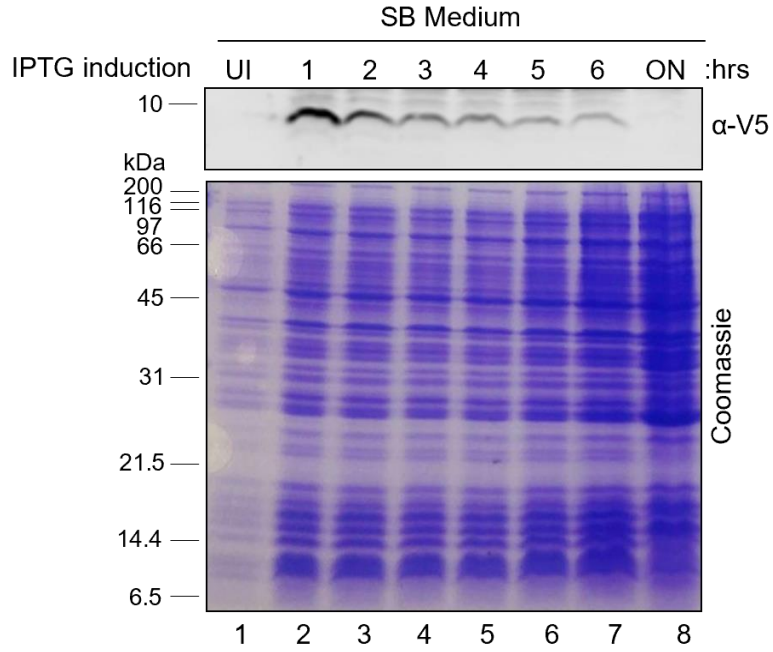


Figure 3.9. Induction time course of SRSF1(89120)-V5-W-TAT PTD-(His)6 CPP expressed in BL21-RP *E. coli* strains. Western blot and Coomassie blue stained 15 % resolving SDS-PAGE gel to analyse to expression level of SRSF1-CPP in SB growth medium. **Lane 1:** Un-induced bacteria culture. **Lane 2-8:** IPTG induced culture of 1, 2, 3, 4, 5, 6 hours and O/N. Western blot was probed with an anti-V5 antibody. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

3.4.4. Optimisation of SRSF1(89-120)-V5-TAT PTD-(His)₆ CPP Expression (without the presence of tryptophan)

A second approach was to use the histidine tagged SRSF1-CPP construct without tryptophan, which is encoded by single codon and is the rarest amino acid in *E.coli*, to increase SRSF1 peptide expression in TB and SB medium. As previously described, SB medium did not result in enhanced peptide expression. However, SB medium was tested alongside TB medium as it was used for a different plasmid expressing SRSF1 peptide without tryptophan. BL21-RP *E.coli* cells expressing SRSF1 peptide were grown in SB and TB medium.

The culture samples from both growth media were collected at different time points and peptide expression was assessed by Coomassie blue staining and anti V5 western blotting (**Figure 3.10.A and 3.10.B**). The SRSF1 peptide was not detectable following Coomassie blue staining (**Figure 3.10.A and 3.10.B**) due to the quantity of expressed peptide being insufficient. The western blots (**Figure 3.10.A and 3.10.B**) show that absence of the tryptophan codon in histidine tagged SRSF1-CPP did not result in increased peptide expression in both TB and SB medium, as the intensity of expressed peptide band was still relatively low compared to SRSF1-CPP with tryptophan codons. Therefore, it was decided that the SB medium approach and histidine tagged SRSF1 peptide would not be used for further peptide expression experiments.

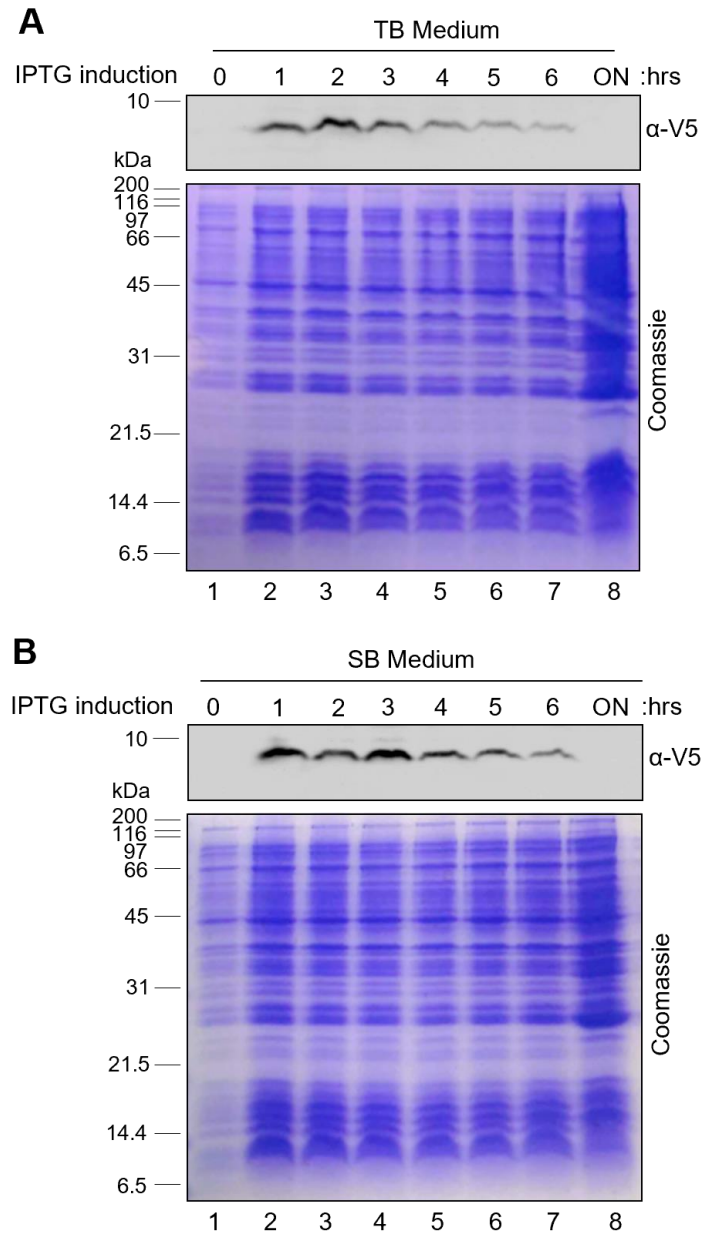


Figure 3.10. Induction time course of SRSF1(89-120)-V5-TAT PTD-(His)₆ CPP expressed in BL21-RP *E. coli* strains. (A) Western blot and Coomassie blue stained 15 % acrylamide gel for analysing the expression of SRSF1(89-120)-V5-TAT PTD-(His)₆ CPP in TB medium. **(B)** Western blot and Coomassie blue 15 % acrylamide gel for analysing the expression of SRSF1(89-120)-V5-TAT PTD-(His)₆ CPP in SB medium. **Lane 1:** Un-induced bacteria culture. **Lane 2-8:** IPTG induced culture of 1, 2, 3, 4, 5, 6 hours and O/N. Western blot was probed with an anti-V5 antibody. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

3.4.5. Production of Recombinant GST-SRSF1(89-120)-V5-TAT PTD CPP

3.4.5.1. Expression of Recombinant GST-SRSF1(89-120)-V5-TAT PTD CPP

Due to previous unsuccessful attempts to generate a sufficient purity and yield of histidine tagged SRSF1-CPP (approximately 100 µg), a new strategy was considered which involved the expression of a GST fusion of the SRSF1 peptide that would further require cleavage by PreScission protease. BL21-RP and BL21-Rosetta *E.coli* expression strains were previously transformed with the plasmid construct to express GST-SRSF1-V5-TAT PTD. In order to maximize GST fusion protein expression, it was important to determine the optimal expression time points. Un-induced (0 h) and IPTG induced cultures at different induction times (1 hour to overnight (O/N)) were collected and analysed by SDS-PAGE followed by Coomassie blue staining of 15 % resolving gels and anti V5 western blotting.

The results show that the GST fusion protein was expressed at a higher quantity than the histidine tagged peptide (**Figure 3.11.A and 3.11.B**). Also, a major single band at the expected molecular weight was detectable by Coomassie staining when comparing non-induced and induced cells. The electrophoretic migration of protein marker between 31 and 45 kDa confirmed the predicted molecular weight of the GST fusion protein which includes GST (~27 kDa) and the SRSF1-V5-TAT peptide of interest (~6.5 kDa). Further to this, there was a difference in peptide expression between BL21-RP and BL21-Rosetta cells. The maximum GST fusion protein expression was attained in BL-21 Rosetta cells at 4, 5 and 6 hours post induction time. However, it was difficult to determine exactly which time point was optimal due to excess protein within lanes that resulted in overflowing of the sample. Therefore, determination of optimum expression time was achieved by re-analysing the time course samples at 4, 5 and 6 hours of BL21-RP and BL21-Rosetta cells by SDS-PAGE followed by Coomassie staining and western blotting (**Figure 3.11.C**). 5 hours post induction time was considered optimal for CST fusion protein expression in BL21- Rosetta cells. Therefore, all subsequent experiments were carried out with this optimised protocol.

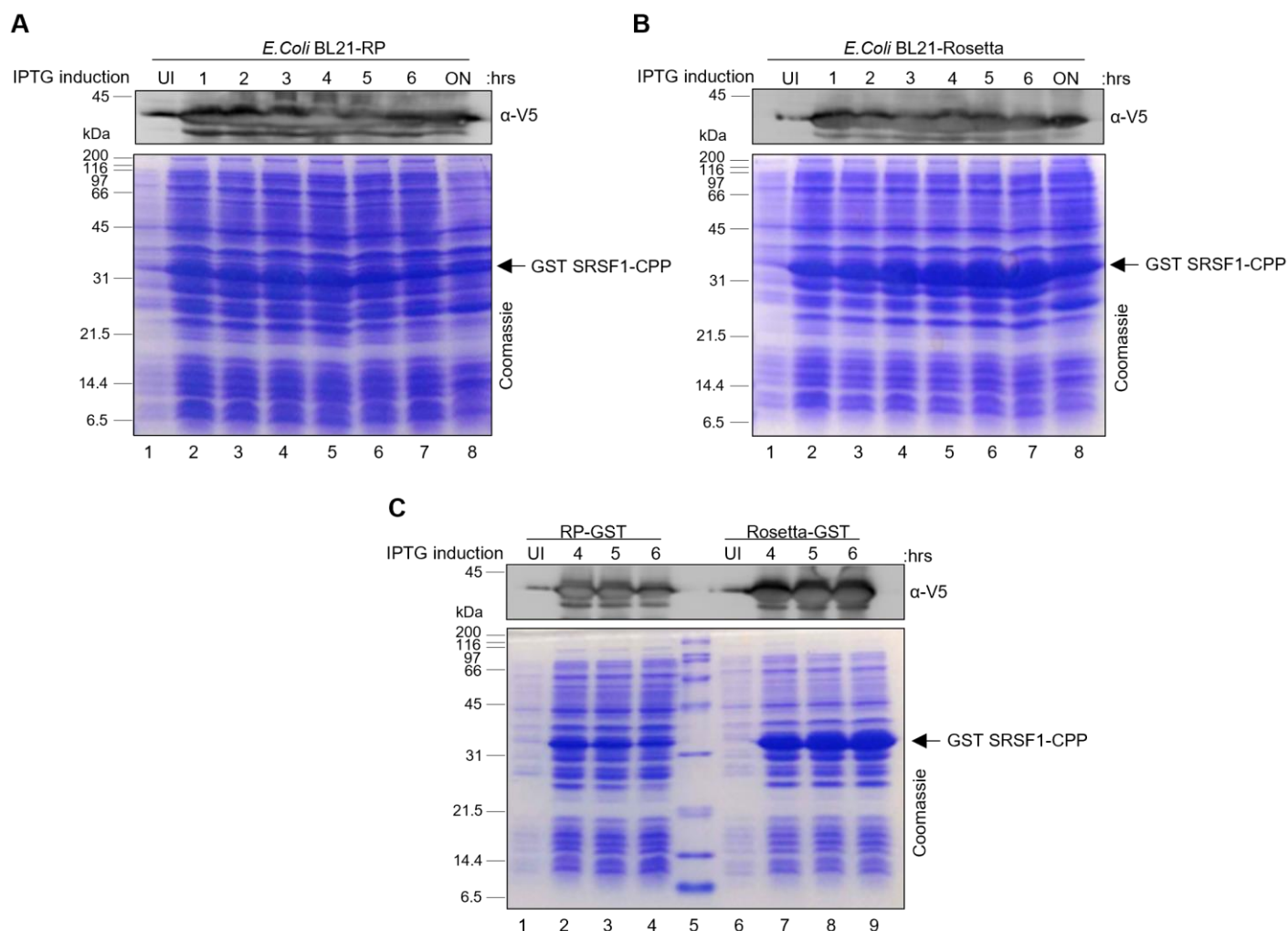


Figure 3.11. Induction time course of GST-SRSF1(89120)-V5-TAT PTD in BL21-RP, BL-21 Rosetta *E. coli* strains. Western blot and Coomassie stained 15 % acrylamide gel to analyse GST-SRF1 fusion protein expression from (A) BL21-RP *E. coli* cells and (B) BL21- Rosetta *E. coli* cells. (C) Western blot and Coomassie stained 15 % acrylamide gel analysis to compare RP and Rosetta *E. coli* expression strains at selected time points (4, 5 and 6 hours) to determine the best induction time. **Lane 1:** Un-induced bacteria culture. **Lane 2-8:** IPTG induced culture of 1, 2, 3, 4, 5, 6 hours and O/N. Western blot was probed with a V5 antibody. Arrows depict expressed GST-SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the western blots and Coomassie gel images presented within this Figure.

3.4.5.2. Small Scale Purification of GST-SRSF1(89-120)-V5-TAT CPP

Once the GST-SRSF1 fusion protein expression trials had obtained a high amount of total fusion protein expression in the BL21- Rosetta expression strain, the GST fusion protein was purified using Glutathione-Sepharose (GSH) affinity chromatography. Enzymatic cleavage of SRSF1-V5-TAT from the GST moiety using PreScission protease was used to release the SRSF1-CPP of interest between glutamine and glycine amino acids upon recognition of an octapeptide sequence (Leu-Glu-Val-Leu-Phe-Gln-Gly) (Waugh, 2011). To assess several conditions required to optimize protease cleavage of the SRSF1 peptide, a small-scale purification was performed which is recommended prior to large scale purification (Shukla et al., 2009). In general, Tris buffer is recommended for the removal of GST by enzymatic cleavage due to its ability to maintain a stable pH during cell lysis, promote solubility, and create an optimal pH (7-9) for PreScission protease function. All previous peptide purifications were performed in PBS buffer; therefore, we first tested PBS lysis buffer (PBS strategy 1) to optimize the working condition of PreScission protease activity. Also, the effect of the reducing agent, dithiothreitol (DTT), on GST binding activity to glutathione in the presence of PreScission protease was examined. Therefore, Tris buffer without DTT (strategy 2) and Tris buffer in the presence of PreScission protease with DTT (strategy 3), were tested. The conditions of these strategies are shown in **Table 3.2**. Subsequent analysis was performed via by Coomassie staining (**Figure 3.12.A, 3.12.B and 3.12.C**).

Table 3.2. The conditions used in Strategy 1, Strategy 2, and Strategy 3.

	Lysis Buffer	Digestion Buffer	Elution Buffer
Strategy 1	1X PBS + 0.5 % Triton X-100	50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.4	50 mM Tris-HCl, 100 mM NaCl, 40 mM reduced glutathione, pH 7.5-8.0
Strategy 2	50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.5 % Triton X-100	50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 7.4	50 mM Tris-HCl, 100 mM NaCl, 40 mM reduced glutathione, pH 7.5-8.0
Strategy 3	50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.5 % Triton X-100	50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.4	50 mM Tris-HCl, 100 mM NaCl, 40 mM reduced glutathione, pH 7.5-8.0

The expressed GST-SRSF1-CPP (**section 3.4.5.1**) was applied on GST beads (30 μ l slurry) and then protease enzyme was added prior to incubating overnight. Next day, the beads were washed and eluted. The eluted peptide under three different condition was analysed by Coomassie staining.

The Coomassie staining shows that there was strong band in ‘GST-fusion’ lane at ~34 kDa corresponding to GST-SRSF1-CPP, suggesting that the majority of GST fusion protein bound to the GST resin (**Figures 3.12.A, 3.12.B and 3.12.C**). Adjacent to the ‘GST-fusion’ lane, the PreScission protease enzyme (PSC) was loaded to detect possible contamination of the cleaved SRSF1 peptide with the protease. Retention of the protease on the GST beads is expected due to the GST tag. Also, it was found that the enzymatic cleavage reaction resulted in a very high yield for all three conditions, and only a faint band corresponding to GST fusion protein was detected in the elution.

The SRSF1 peptide minus the GST tag and the GST remaining on the beads were eluted and shown respectively in elution (E) and GST lanes. The quantity of the resulting SRSF1 peptide varied between the 3 different strategies, although was successfully expressed and cleaved. The peptide was not detectable in the gel in **Figure 3.12.A**. The cleaved peptide was detected on the gel using strategy 2 (**Figure 3.12.B**); however, the intensity of peptide band in strategy 2 was not strong indicating that the amount of cleaved peptide was relatively less. Strategy 3 shows that the SRSF1 peptide was almost completely eluted (**Figure 3.12.C**). Further to this, the PreScission protease enzyme was not detected in the eluate from strategy 3, whereas very faint bands were detected in the eluates from strategies 1 and 2. Consequently, strategy 3 was used in subsequent GST purification experiments.

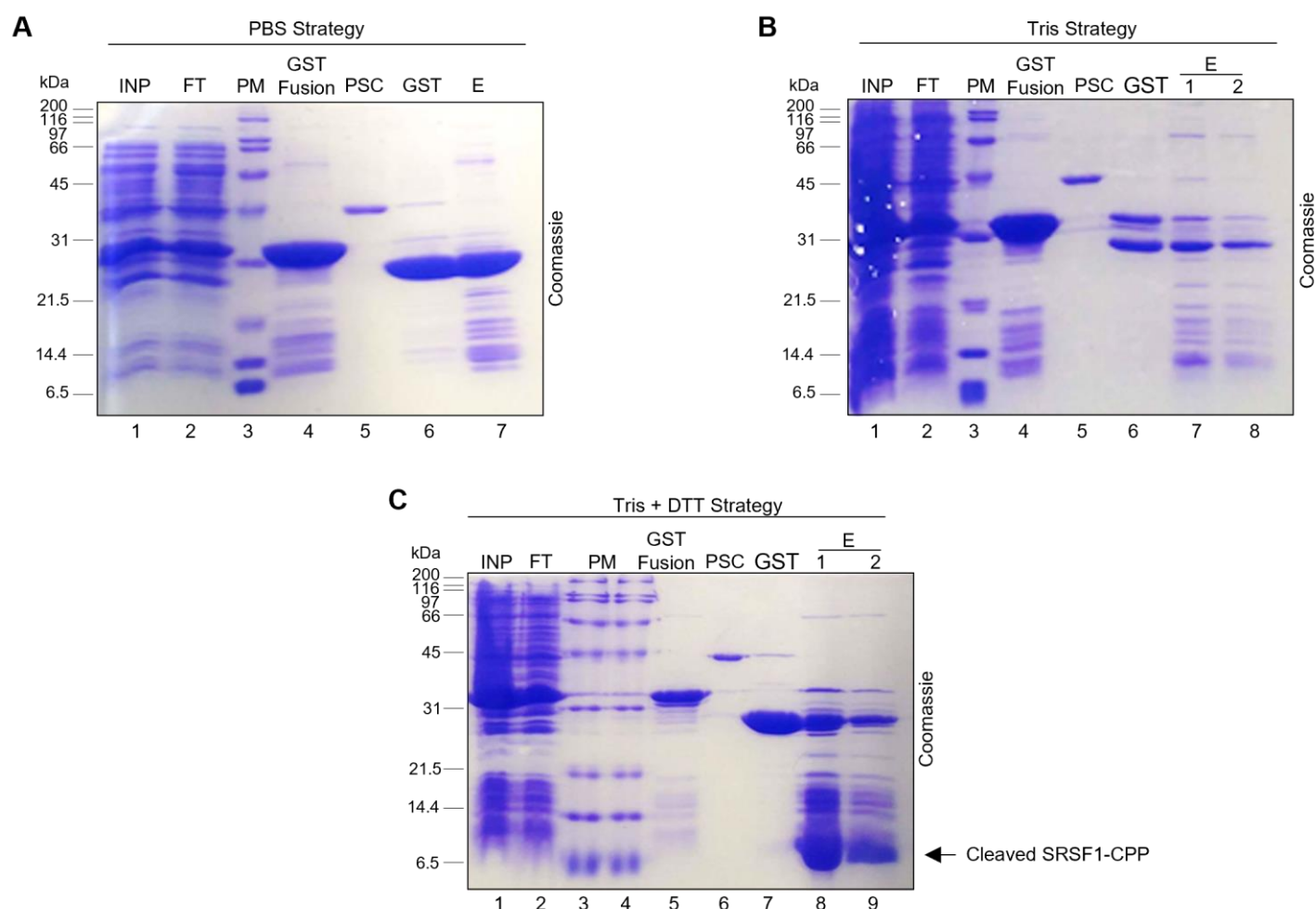


Figure 3.12. Small scale purification of GST-SRSF1-V5-TAT PTD CPP with GSH affinity chromatography. Coomassie stained 15 % acrylamide gel was used to analyse efficiency of three different strategies upon cleavage of GST. Coomassie stained 15 % acrylamide gel analysis of GST cleavage with using conditions in **(A)** Strategy 1 ‘PBS’, **(B)** Strategy 2 ‘Tris’, and **(C)** Strategy 3 ‘Tris + DTT’. **Lane 1:** Input (INP). **Lane 2:** Flow Through (FT). **Lane 3 and 4:** Protein Marker (PM). **Lane 5:** GST Fusion (GST- SRSF1-V5-TAT protein). **Lane 6:** PreScission protease enzyme (PSC). **Lane 7:** GST Residual Elution (GST). **Lane 8:** PSC-Cleaved SRSF1-CPP eluate-1 (E1). **Lane 9:** PSC-Cleaved SRSF1-CPP eluate-2 (E2). Arrow depicts cleaved SRSF1-V5-TAT PTD CPP. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.5.3. Large Scale Purification of GST-SRSF1(89-120)-V5-TAT CPP

The cleavage conditions of SRSF1-V5-TAT separating from the GST fusion complex were previously optimised by small scale purification (see **section 3.3.5.2**). However, large scale purification of GST-SRSF1-CPP fusion protein was required to enhance the yield. Two-step purification was applied to the GST fusion protein. The initial step of purification was performed by applying the sample to the stationary phase of glutathione sepharose affinity chromatography (5 ml) to purify the GST-fusion protein and then cleave by PreScission protease to release the peptide from the GST moiety bound on the beads. An additional purification step using SP ion exchange chromatography was used to remove contaminants. Input (INP), flow through (FT) and both eluates from GSH affinity and SP ion exchange chromatography (pre-packed column connected to AKTA FPLC) were assessed by Coomassie blue staining (**Figure 3.13**). Total bacterial extract (Input) was loaded twice to maintain the binding of the GST fusion protein at maximum level by saturating the GST stationary phase. FT-1 was obtained just after the first loading of analyte over the stationary phase. FT-1 was applied again onto the stationary phase of chromatography and FT-2 was collected.

FT-1 and FT-2 displayed similar patterns on the gel, indicating that highly concentrated GST fusion protein was present in both lanes after binding to the GSH column, resulting in smearing of lanes (**Figure 3.13**). Therefore, it was difficult to predict the relative amount of expressed GST fusion protein at this point by assessing the relative band intensities. However, GST eluate lanes (GST-1 and GST-2) showed that there was a high amount of fusion protein expression. Once FT-2 was loaded, the GST stationary phase was incubated with protease enzyme for cleavage at 4 °C overnight on a rotating wheel. The following day, the cleaved peptide was eluted twice from the GSH column resulting in a high amount of peptide in eluates 1 (PSC Cleaved-1) and 2 (PSC Cleaved-2) whilst also showing evidence of contamination. The eluates were applied onto SP in the stationary phase at a flow rate of 2 ml/min. FT-3 was collected, and the majority of peptide was bound to the SP stationary phase. The stationary phase connected to AKTA was exposed to high salt gradient elution.

The 5 ml elution fractions (E1, E2 and E3) were collected and the concentration of the eluates was estimated to be ~1 mg/15 ml. The estimated peptide concentration was calculated by loading several dilutions of synthetic SRSF1 peptide of known concentration. The approximate amount of purified recombinant peptide produced was estimated by matching the intensity of the band of synthetic peptide with a band of similar intensity of the recombinant SRSF1 peptide. Although that amount of peptide was not enough to perform all *in vitro* and *in vivo* characterisation studies, it was excellent progress in comparison to histidine tagged peptide purification. It is important to note that precipitation of peptide occurred soon after the cleavage step. Therefore, 6 M urea was slowly dissolved into the eluates to allow re-solubilisation of SRSF1-CPP by concomitantly providing the chemical energy and denaturing protein to decompose aggregates thermodynamically. The urea strategy worked well; however, it was important to test the binding of peptide onto SP chromatography in the presence of urea. To summarize, the peptide yield was considerably higher than that of his-tagged purification. Two-step purification resulted in achieving a higher peptide yield than that of his-tagged purification. However, the second step of purification still required some optimization.

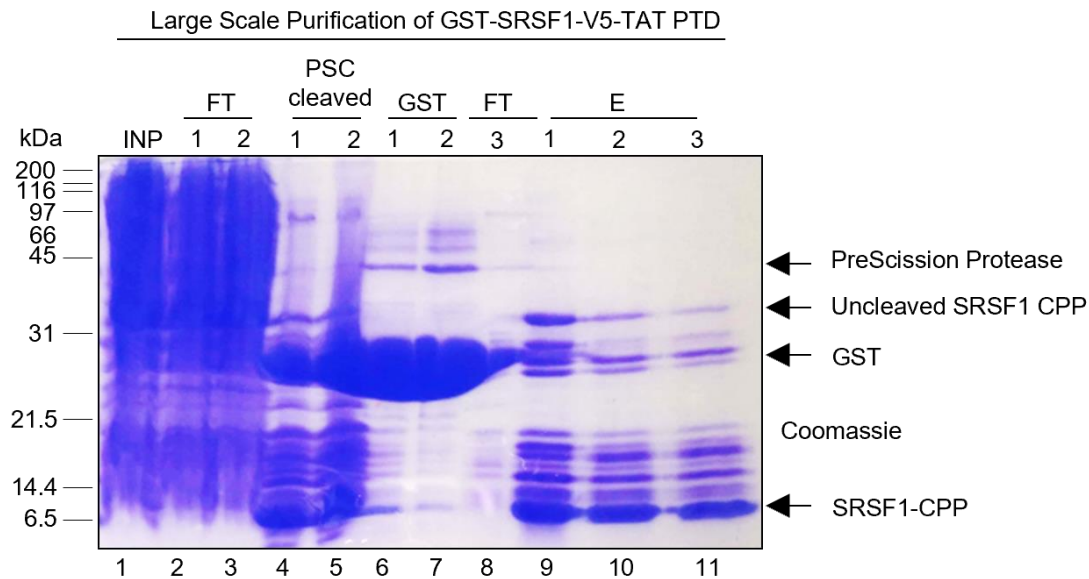


Figure 3.13. Large scale two-step purification of GST-SRSF1-V5-TAT PTD CPP including GSH affinity and SP ion exchange chromatography. Coomassie stained 15 % acrylamide gel was utilized to analyse the purification of GST tagged SRSF1-CPP. **Lane 1:** Input (INP). **Lane 2:** Flow Through-1 (FT1). **Lane 3:** Flow Through-2 (FT2). **Lane 4:** PSC-Cleaved SRSF1-CPP eluate-1. **Lane 5:** PSC-Cleaved SRSF1-CPP eluate-2. **Lane 6:** GST Residual Elution - 1 (GST) **Lane 7:** GST Residual Elution - 2 (GST). **Lane 8:** Flow Through-3 (FT3). **Lane 9:** Eluate 1 (E1). **Lane 10:** Eluate-2 (E2). **Lane 11:** Eluate 3 (E3). Arrows depict respectively from top to bottom of the gel; PreScission Protease enzyme, Un-cleaved GST-SRFS1 CPP, GST and Cleaved SRSF1-CPP. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.5.4. Testing The Binding of Cleaved SRSF1(89-120)-V5-TAT PTD CPP in The Presence of Urea and Elution onto Various Chromatography Media

3.4.5.4.1. SP Weak Ion Exchange Chromatography

The purified SRSF1 peptide is now in presence of 6 M urea and it was therefore important to investigate the binding efficiency of the peptide on chromatography media in the presence of urea. In addition to this, a range of salt (NaCl) concentrations, which is known to have a major role in binding and desorption of protein or peptide (Wang et al., 2021), from 200 mM to 500 mM in loading and washing buffer for SP chromatography was characterised to optimize the conditions. A further strategy was to investigate the amount of SRSF1 peptide remaining in the stationary phase and thus prevent possible peptide loss to maximise yield. Binding of the peptide to the stationary phase was performed in the presence of 6 M urea while washing buffer did not contain the urea to subsequently remove it when the SRSF1 peptide binds to the stationary phase. The SRSF1 peptide was diluted in different salt concentrations. Prior to performing a binding efficiency test of SRSF1-CPP on the different chromatography medium, it should be noted that the same amount of peptide was loaded onto each different chromatography media. An equal volume of beads (30 μ l slurry) was also present in the reaction and elution was performed with the same volumes. Finally, the same volume of eluates in the presence of different salt concentrations was analysed by Coomassie blue staining (**Figure 3.14**). In this way, it was possible to compare the efficiency of binding and elution in the presence of the different chromatography media.

SRSF1-CPP binds well to binding site SP resin between 200 and 400 mM NaCl in a Na-Phosphate buffer, as seen in the elution lanes (**Figure 3.14**). The most efficient binding was obtained at 250 mM salt concentration compared to 500 mM NaCl which inhibits binding of the peptide. This may account for decreasing peptide solubility. However, over half the quantity of the SRSF1-CPP was not eluted from the stationary phase even at 2 M NaCl elution. Therefore, it was decided to change the pH of buffers to test whether this could alter the charge of the peptide, thus modifying its SP-binding properties.

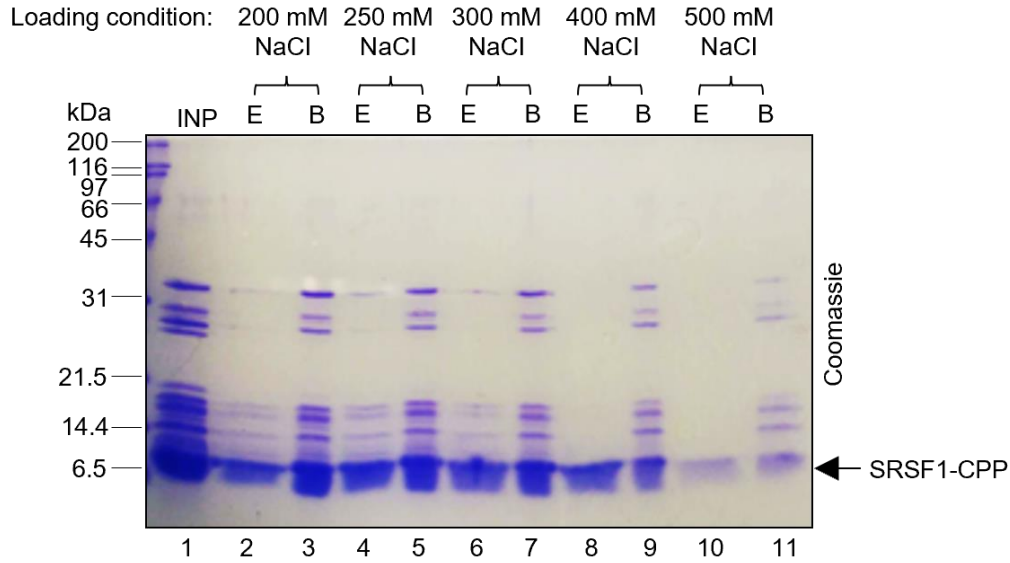


Figure 3.14. Effect of mobile phase with a range of salt concentrations (200 mM – 500 mM) in the presence of urea on the binding ability of cleaved SRSF1-CPP to SP chromatography resin. Lane 1: Input (Cleaved SRSF1-CPP). **Lane 2:** Elution (E) at 2 M NaCl (peptide at 200 mM NaCl in Mobile phase). **Lane 3:** Beads (B) (The amount of peptide at 200 mM NaCl remained attached on beads after 2 M salt elution) **Lane 4:** Elution (E) at 2 M NaCl (peptide at 250 mM NaCl in Mobile phase). **Lane 5:** Beads (B) (The amount of peptide at 250 mM NaCl remained attached on beads after 2 M salt elution) **Lane 6:** Elution (E) at 2 M NaCl (peptide at 300 mM NaCl in Mobile phase). **Lane 7:** Beads (B) (The amount of peptide at 300 mM NaCl remained attached on beads after 2 M salt elution) **Lane 8:** Elution (E) at 2 M NaCl (peptide at 400 mM NaCl in Mobile phase). **Lane 9:** Beads (B) (The amount of peptide at 400 mM NaCl remained attached on beads after 2 M salt elution) **Lane 10:** Elution (E) at 2 M NaCl (peptide at 500 mM NaCl in Mobile phase). **Lane 11:** Beads (B) (The amount of peptide at 500 mM NaCl remained attached on beads after 2 M salt elution). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

Optimization studies were carried out to find a way to eliminate the retention of peptide on the chromatography medium. Tris and Hepes buffers at 250 mM NaCl were used in the presence of different pH values in the mobile phase. In parallel, the effect of high salt concentration in elution buffer on releasing of peptide from chromatography matrix was also investigated (1 M and 2 M NaCl). The conditions of mobile phase and elution buffers are listed in **Table 3.3**. The purified SRSF1-CPP in each condition (obtained from large scale GST-SRSF1 purification, **Figure 3.13**) was separately loaded onto SP beads (30 µl slurry); these buffers were used in the washing steps. The effect of loading buffer/elution buffer at different salt concentrations and along with loading SRSF1-peptide as an input on the gel were monitored by SDS-PAGE followed by Coomassie staining (**Figure 3.15**).

Table 3.3. The composition of tested loading and elution buffers conditions.

	Loading	Salt Concentration in Elution Buffer
Condition 1	50 mM Tris, 250 mM NaCl, pH : 8	1 M NaCl
	50 mM Tris, 250 mM NaCl, pH : 8	2 M NaCl
Condition 2	50 mM Hepes and 250 mM NaCl pH: 8.5	1 M NaCl
	50 mM Hepes and 250 mM NaCl pH: 8.5	2 M NaCl
Condition 3	50 mM Tris, 250 mM NaCl, pH : 10	1 M NaCl
	50 mM Tris, 250 mM NaCl, pH : 10	2 M NaCl

Coomassie staining shows that the binding of SRSF1-CPP to the SP stationary phase in conditions 1, 2 and 3 resulted in a high intensity peptide band from CPP desorption/elution lanes compared to the input. However, the majority of CPP remained adsorbed in the stationary phase, resulting in significant loss of the peptide. The only significant difference relating to the amount of CPP remaining bound to the chromatography medium was observed in the presence of 1 M and 2 M salt in the elution buffer. The peptide elution performed with elution buffer containing 2 M salt resulted in less dissociation of peptide from the column in comparison to 1 M salt in the elution buffer (**Figure 3.15**), which should have resulted in obtaining a higher peptide yield than elution with 1 M salt. Therefore, it was decided to perform elution at 1 M salt concentration for further experiments in the presence of a strong ion exchanger.

In conclusion, the use of Tris and Hepes buffer in the presence of different pH values as a mobile phase, as well as eluates with different salt concentrations, were not found to have a positive effect on releasing the peptide from the chromatography resin. Therefore, it was decided to test different chromatography media, with a weaker CM, carboxymethyl cationic exchange, to try reducing the remaining amount of SRSF1 peptide on the beads.

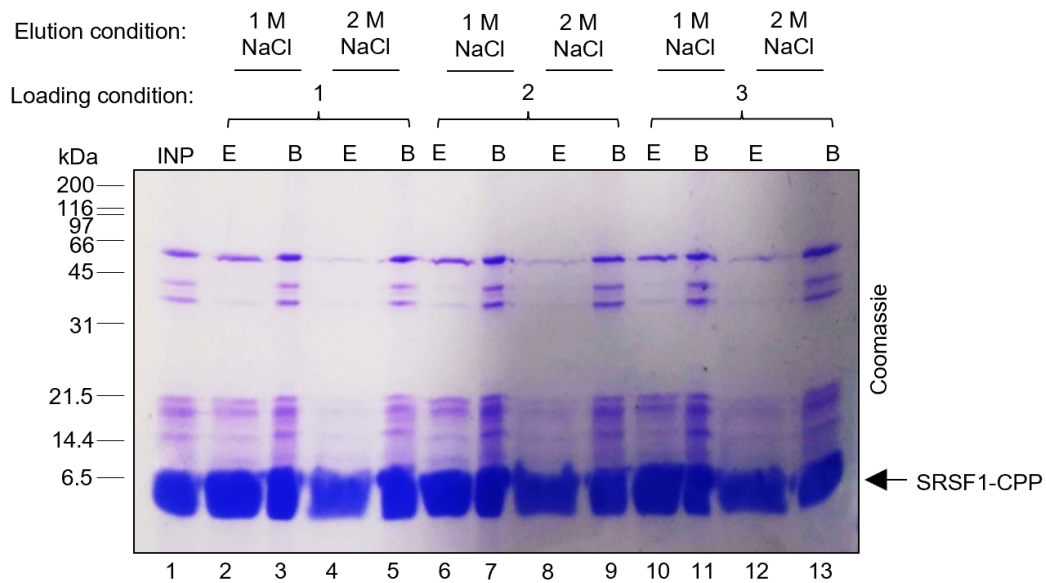


Figure 3.15. Effect of three different buffer compositions on the binding ability of cleaved SRSF1-CPP and salt concentrations of elution buffer to SP chromatography resin. Lane 1: Input (INP). Lane 2: Elution (E) at 1 M NaCl (Mobile phase **condition 1). Lane 3: Beads (B) (The amount of peptide remained attached after 1 M salt elution). Lane 4: Elution (E) at 2 M NaCl (Mobile phase **condition 1**). Lane 5: Beads (B) (The amount of peptide remained attached after 2 M salt elution). Lane 6: Elution (E) at 1 M NaCl (Mobile phase **condition 2**). Lane 7: Beads (B) (The amount of peptide remained attached after 1 M salt elution). Lane 8: Elution (E) at 2 M NaCl (Mobile phase **condition 2**). Lane 9: Beads (B) (The amount of peptide remained attached after 2 M salt elution). Lane 10: Elution (E) at 1 M NaCl (Mobile phase **condition 3**). Lane 11: Beads (B) (The amount of peptide remained attached after 1 M salt elution). Lane 12: Elution (E) at 2 M NaCl (Mobile phase **condition 3**). Lane 13: Beads (B) (The amount of peptide remained attached after 2 M salt elution). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.**

3.4.5.4.2. CM Ion Exchange Chromatography

A weaker CM (carboxymethyl) cationic ion exchanger was used to overcome the retention issue of peptide within the stationary phase. The optimization studies in CM ion exchange chromatography were similar to those made with SP chromatography. The peptide was loaded onto the CM stationary phase in the presence Na-Phosphate buffers which have different salt concentrations to determine the binding efficiency of CPP on CM beads. All loading buffers have 6 M urea to avoid precipitation, whereas washing buffers do not include urea.

The SRSF1-CPP with a range of salt concentrations from 200 mM to 500 mM NaCl was loaded directly onto CM weak anion-exchange chromatography resins (30 μ l slurry). The beads were washed prior to desorption of SRSF1-CPP with 2 M NaCl. Coomassie blue staining revealed that a significant fraction of the cleaved SRSF1-CPP did not attach onto the surface of CM beads at 200 mM salt concentrations. Only a small fraction of peptide remained attached on resin in comparison to the amount loaded onto the beads (input lane). Further to this, the use of the CM column at increasing salt concentrations did not allow SRSF1-CPP for adsorption on resin and this resulted in the loss of almost all of the peptide during elution (**Figure 3.16**).

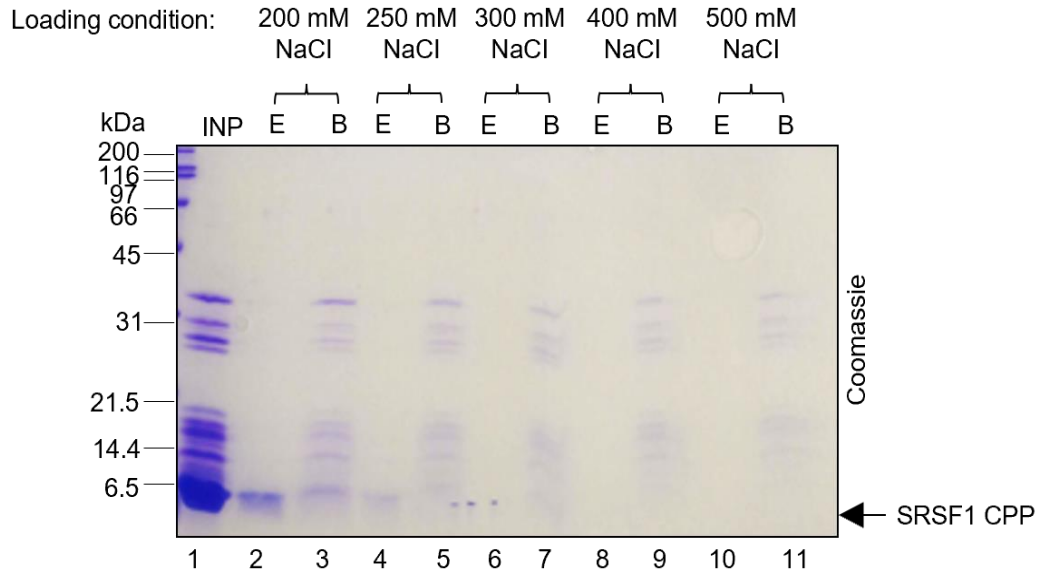


Figure 3.16. Investigating binding ability of cleaved SRSF1-CPP onto CM chromatography resin at range of salt concentrations (200 mM – 500 mM) and in presence of urea in mobile phase. Lane 1: Input (INP). Lane 2: Elution (E) at 2 M NaCl (peptide at 200 mM NaCl in Mobile phase). Lane 3: Beads (B) (The amount of peptide at 200 mM NaCl remained attached on beads after elution with 2 M NaCl) Lane 4: Elution (E) at 2 M NaCl (peptide at 250 mM NaCl in Mobile phase). Lane 5: Beads (B) (The amount of peptide at 250 mM NaCl remained attached on beads after elution with 2 M NaCl) Lane 6: Elution (E) at 2 M NaCl (peptide at 300 mM NaCl in Mobile phase). Lane 7: Beads (B) (The amount of peptide at 300 mM NaCl remained attached on beads after elution with 2 M NaCl) Lane 8: Elution (E) at 2 M NaCl (peptide at 400 mM NaCl in Mobile phase). Lane 9: Beads (B) (The amount of peptide at 400 mM NaCl remained attached on beads after elution with 2 M NaCl) Lane 10: Elution (E) at 2 M NaCl (peptide at 500 mM NaCl in Mobile phase). Lane 11: Beads (B) (The amount of peptide at 500 mM NaCl remained attached on beads after elution with 2 M NaCl). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

The CM ion exchange column containing a range of salt concentrations between 200-500 mM in the mobile phase accounted for loss of peptide. This suggests that the peptide might be released earlier than elution step from the stationary phase as an increase in salt concentration of mobile phase can interfere with the binding of peptide to CM resins, therefore significant amount of peptide was not seen in elution lanes. In an attempt to avoid losing the peptide in the chromatography medium, the peptide was loaded onto a CM column at lower salt concentrations in Na-Phosphate buffer at pH 7.4 (50, 100, and 150 mM NaCl). Flow through (FT) samples were also collected to determine whether the peptide disassociated from the column during the washing step (**Figure 3.16**). The conditions were analysed with SDS-PAGE followed by Coomassie blue staining (**Figure 3.17**).

CPP binding at low salt concentrations improved the binding of CPP on medium compared to 200 mM salt concentration. In addition, peptide was seen in the FT at 150 mM salt indicating the point at which high salt concentrations interfere with the interaction of the peptide with the CM matrix. However, the amount of the peptide eluted at the lowest binding salt concentration is only a small fraction of the amount of peptide loaded onto the beads (input lane), indicating that the majority of the peptide remained attached even after the 2 M NaCl elution. A similar result was also observed at 100 mM and 150 mM binding salt concentration. Therefore, additional changes to the chromatography media were required.

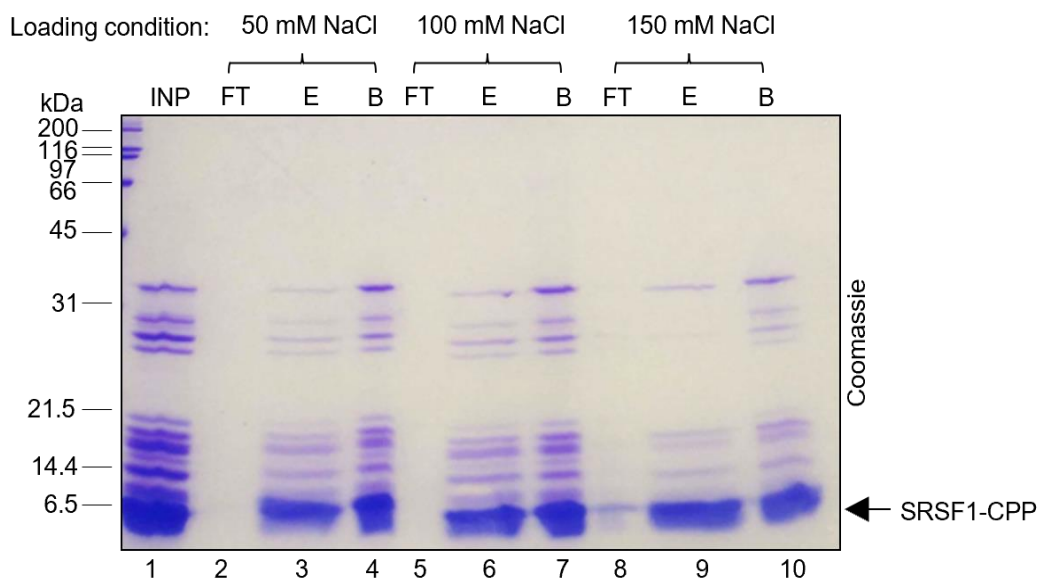


Figure 3.17. Investigating binding ability of cleaved SRSF1-CPP onto CM chromatography resin at low range of salt concentrations (50 mM – 150 mM) and in the presence of urea in mobile phase. **Lane 1:** Input (INP). **Lane 2:** Flow Through (FT) at 50 mM NaCl in Mobile phase. **Lane 3:** Elution (E) at 2 M NaCl (peptide at 50 mM NaCl in Mobile phase). **Lane 4:** Beads (B) (The amount of peptide at 50 mM NaCl remained attached on beads after 2 M salt elution). **Lane 5:** Flow Through (FT) at 100 mM NaCl in Mobile phase. **Lane 6:** Elution (E) at 2 M NaCl (peptide at 100 mM NaCl in Mobile phase). **Lane 7:** Beads (B) (The amount of peptide at 100 mM NaCl remained attached on beads after 2 M salt elution). **Lane 8:** Flow Through (FT) at 150 mM NaCl in Mobile phase. **Lane 9:** Elution (E) at 2 M NaCl (peptide at 150 mM NaCl in Mobile phase). **Lane 10:** Beads (B) (The amount of peptide at 150 mM NaCl remained attached on beads after 2 M salt elution). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.5.4.3. Heparin Adsorption Chromatography

Heparin adsorption chromatography, which involves both ion exchange and affinity chromatographies, is a commonly used purification method to fractionate different classes of protein specifically and reversibly from total extract with higher level of selectivity (Bolten et al., 2018). Therefore, heparin chromatography was utilised as the final strategy. Similar to previous strategies, the effect of salt concentration on CPP adsorption in the mobile phase was optimised. SRSF1-CPP with salt concentrations ranging from 50 mM to 300 mM NaCl was applied directly onto Heparin resins (30 µl slurry). SRSF1-CPP underwent desorption with 2 M NaCl containing elution buffer, and the product was analysed by Coomassie blue staining (**Figure 3.18. and 3.19**). The Coomassie gel shows efficient adsorption of SRSF1-CPP on the Heparin stationary phase at lower salt concentrations based on a similar band intensity to the input (**Figure 3.18**).

A very small fraction of peptide remained bound to the chromatography media, suggesting that the peptide was successfully released from the chromatography media. Furthermore, the peptide was not detected in the FT at 50 mM and 100 mM salt concentrations indicating that the peptide was not eluted earlier during the washing step. A small amount of peptide was detected in the FT when loading was performed at 150 mM NaCl. The Coomassie gel shows that the peptide does not adsorb well to the stationary phase at higher salt concentrations particularly at 250 mM and 300 mM (**Figure 3.19**). This suggests that the peptide was lost during the washing step and appeared in the FT, similar to what was seen in CM chromatography (see section 3.4.5.4.2). Since the heparin column solves the issue of CPP binding and the inability to release a large amount of peptide from the stationary phase, it was decided to use the heparin column in subsequent experiments. However, the optimum salt concentration was not yet fully determined as similar band intensities were observed for loading between 100 mM and 150 mM NaCl.

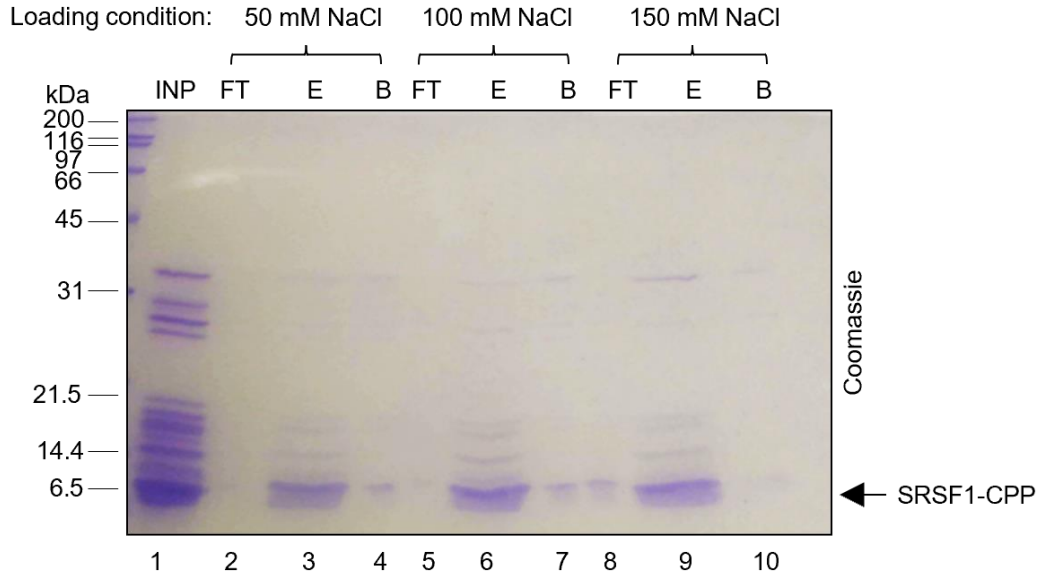


Figure 3.18. Investigating binding ability of cleaved SRSF1-CPP onto Heparin adsorption chromatography resin at low range of salt concentrations (50 mM – 150 mM) and in the presence of urea in mobile phase. Lane 1: Input (INP) (Cleaved SRSF1-CPP). Lane 2: Flow Through (FT) at 50 mM NaCl in Mobile phase. Lane 3: Elution (E) at 2 M NaCl (peptide at 50 mM NaCl in Mobile phase). Lane 4: Beads (B) (The amount of peptide at 50 mM NaCl remained attached on beads after 2 M salt elution). Lane 5: Flow Through (FT) at 100 mM NaCl in Mobile phase. Lane 6: Elution (E) at 2 M NaCl (peptide at 100 mM NaCl in Mobile phase). Lane 7: Beads (B) (The amount of peptide at 100 mM NaCl remained attached on beads after 2 M salt elution). Lane 8: Flow Through (FT) at 150 mM NaCl in Mobile phase. Lane 9: Elution (E) at 2 M NaCl (peptide at 150 mM NaCl in Mobile phase). Lane 10: Beads (B) (The amount of peptide at 150 mM NaCl remained attached on beads after 2 M salt elution). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

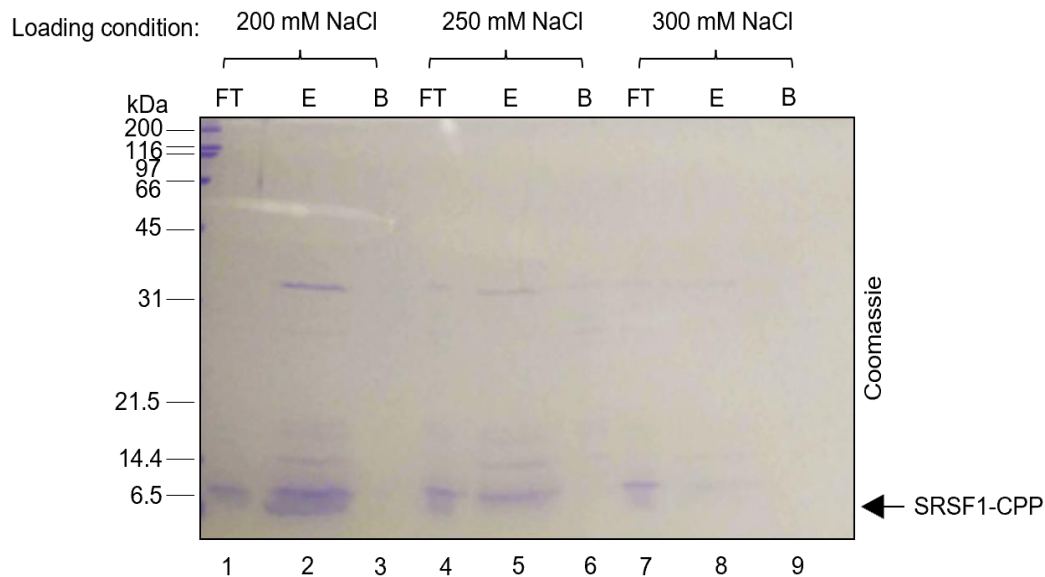


Figure 3.19. Investigating the binding ability of cleaved SRSF1-CPP onto Heparin adsorption chromatography resin at higher range of salt concentrations (200 mM, 250 mM and 300 mM) and in the presence of urea in mobile phase. Lane 1: Flow Through (FT) at 200 mM NaCl in Mobile phase. **Lane 2:** Elution (E) at 2 M NaCl (peptide at 200 mM NaCl in Mobile phase). **Lane 3:** Beads (B) (The amount of peptide at 200 mM NaCl remained attached on beads after 2 M salt elution). **Lane 4:** Flow Through (FT) at 250 mM NaCl in Mobile phase. **Lane 5:** Elution (E) at 2 M NaCl (peptide at 250 mM NaCl in Mobile phase). **Lane 6:** Beads (B) (The amount of peptide at 250 mM NaCl remained attached on beads after 2 M salt elution). **Lane 7:** Flow Through (FT) at 300 mM NaCl in Mobile phase. **Lane 8:** Elution (E) at 2 M NaCl (peptide at 300 mM NaCl in Mobile phase). **Lane 9:** Beads (FT) (The amount of peptide at 300 mM NaCl remained attached on beads after 2 M salt elution). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

The mobile phase at 100 mM and 150 mM salt concentrations were tested to determine the optimal NaCl concentration in terms of CPP binding capacity to the Heparin column. Elution of the CPP was also carried out at two different total salt concentrations (1 M and 2 M) to investigate effect of salt concentration in the elution buffer. Coomassie staining shows that when the CPP was applied on heparin beads in the presence of 100 mM salt concentration, a significant fraction of the peptide attached to the stationary phase of heparin because the intensity of eluate bands at either 1 M or 2 M total salt concentrations was higher than that of eluates obtained from the mobile phase at 150 mM and 200 mM (**Figures 3.20 and 3.21**). Additionally, both gels show that the CPP desorption commenced at 1 M salt elution indicating that the majority of the peptide had been pulled down from the beads. At 2 M elution, the same trend was seen by the mobile phase at 100 mM, 150 mM and 200 mM NaCl concentrations; therefore, elution buffer comprising 1 M salt was used for subsequent purification. Further to this, the peptide eluted from the heparin resins exhibited higher purity than input which is purified SRSF1-CPP obtained from SP chromatography.

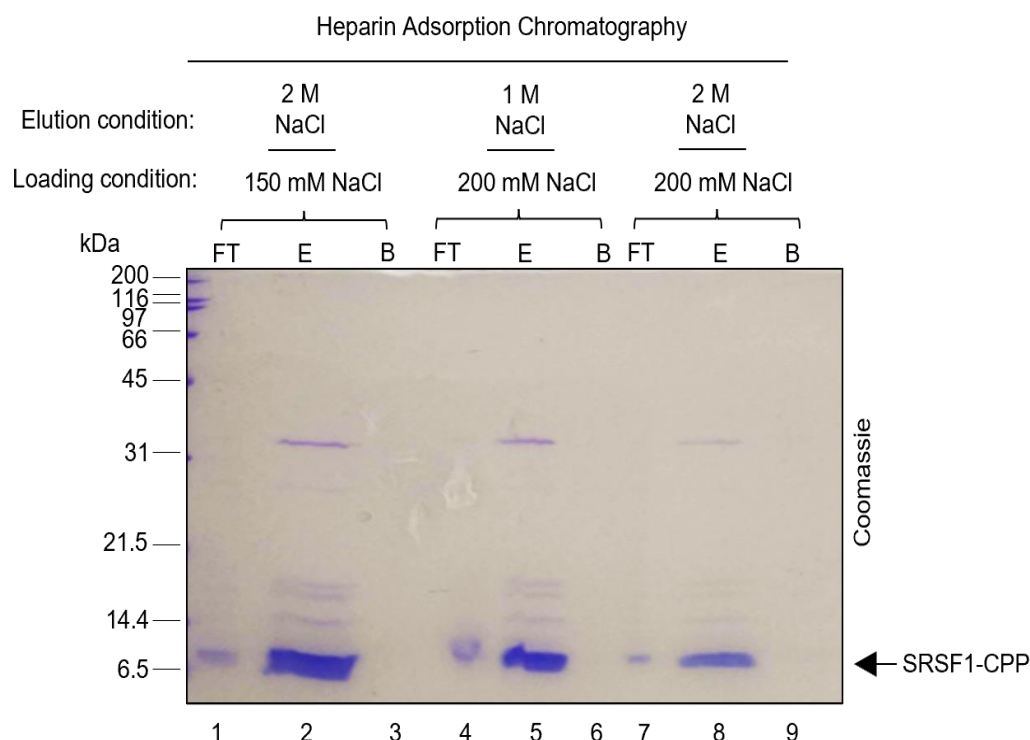


Figure 3.21. Investigating total salt concentrations in the Elution buffer for desorption of SRSF1-CPP at 150 mM and 200 mM NaCl from Heparin adsorption chromatography. The mobile phase of SRSF1-CPP at 150 and 200 mM was applied onto the Heparin matrix. Coomassie Blue stained gel analysed the yield elution that performed at either 1 M or 2 M NaCl in elution buffer from Heparin Adsorption Chromatography. **Lane 1:** Flow Through (FT) of 150 mM NaCl in Mobile phase. **Lane 2:** Elution (E) at 2 M NaCl (Peptide at 150 mM). **Lane 3:** Beads (B) (The amount of peptide remained attached after 2 M salt elution). **Lane 4:** Flow Through (FT) of 200 mM NaCl in Mobile phase. **Lane 5:** Elution (E) at 1 M NaCl (Peptide at 200 mM). **Lane 6:** Beads (B) (The amount of peptide remained attached after 1 M salt elution). **Lane 7:** Flow Through (FT) of 200 mM NaCl in Mobile phase. **Lane 8:** Elution (E) at 2 M NaCl (Peptide at 200 mM). **Lane 9:** Beads (B) (The amount of peptide remained attached after 2 M salt elution). Arrow depicts cleaved SRSF1-V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

To determine the total amount of peptide purified using the Heparin strategy, large-scale expression of the GST-SRSF1-CPP fusion protein in BL-21 Rosetta strains from six culture flasks of 750 ml TB medium followed by two-step purification steps involving GSH affinity (5 ml slurry) and Heparin adsorption chromatography (5 ml slurry) were performed with optimised conditions. The resulting eluates were analysed by Coomassie staining. The ‘input’ lane showed that GST fusion peptide was expressed at a high level (**Figure 3.22**). The intensity of the FT band was considerably lower compared to the input lane indicating strong binding of the fusion protein to the GSH column. Following enzymatic cleavage of GST fusion protein, the PSC-cleaved peptide fractions were collected from the GST column and applied onto the stationary phase of the heparin column. FT-2 showed that a very faint band of peptide was observed in FT-2 lane indicating that the maximum quantity of peptide was bound to the heparin column. Desorption of peptide was then achieved at 1 M NaCl, and eluates (E1, E2, E3 and E4) were collected in 5 ml of elution buffer. The purified fractions exhibited a high concentration and purity of SRSF1-V5-TAT peptide at the expected size of 6.5 kDa. However, the purity of SRSF1-CPP was not still sufficient; therefore, the eluted peptide from Heparin chromatography was subjected to further purification by size exclusion chromatography (see section 3.4.5.4.4) in order to remove contaminants.

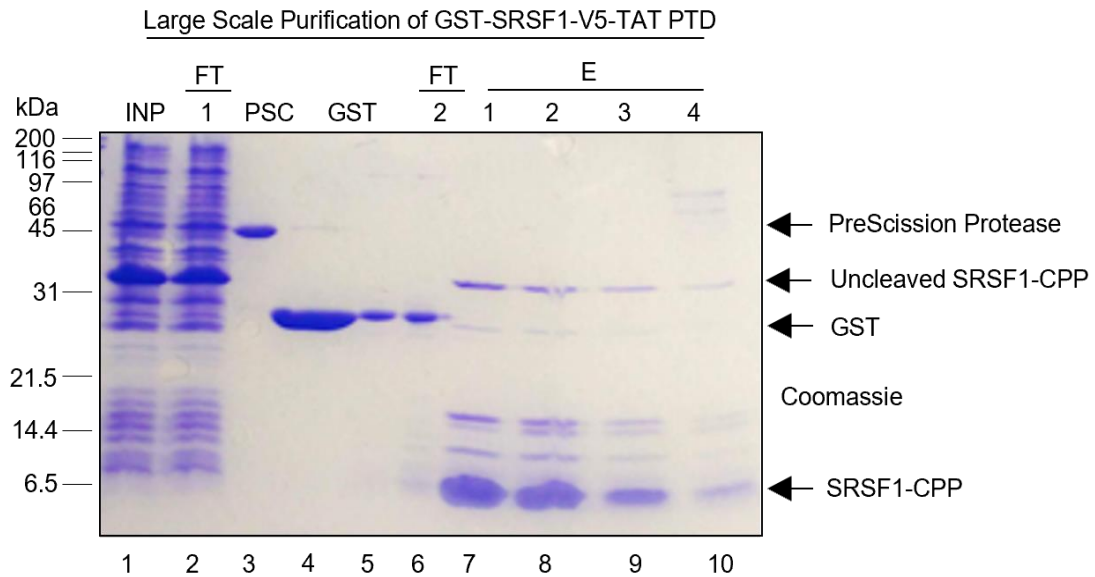


Figure 3.22. Large scale two-steps purification of GST-SRSF1-V5-TAT PTD CPP with GSH affinity and Heparin adsorption chromatography. Coomassie stained 15 % acrylamide gel for analysing the two-step purification process of GST tagged SRSF1-CPP. **Lane 1:** Input (INP) (total cell extract). **Lane 2:** Flow Through-1 (FT-1). **Lane 3:** PreScission Protease enzyme (PSC). **Lane 4:** GST Residual Elution - 1 (GST) **Lane 5:** GST Residual Elution - 2 (GST). **Lane 6:** Flow Through-2 (FT-2). **Lane 7:** Eluate 1 (E1). **Lane 8:** Eluate-2 (E2). **Lane 9:** Eluate 3 (E3). **Lane 10:** Eluate 4 (E4). Arrows depict respectively from top to bottom of the gel; PreScission Protease enzyme, Uncleaved GST-SRFS1 CPP, GST and Cleaved SRSF1-CPP. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.5.4.4. Size Exclusion Chromatography

Size exclusion chromatography, as part of optimization trials, was conducted after Heparin adsorption chromatography to enhance the purity of SRSF1-CPP. The chromatogram of size exclusion chromatography (pre-packed column) is provided (**Figure 3.23.A**). It shows that molecules with a higher molecular weight than SRSF1-CPP were seen in an initial sharp peak when approximately 45 ml elution buffer had passed through the column, indicating that the size exclusion column was starting to fractionate contaminants from the peptide. The second small wide peak at ~ 60 ml elution volume corresponds to the fractions containing the peptide (lanes 20-32). The fractions selected based on the chromatogram were assessed by Coomassie blue staining, which revealed that fractions 26-33 contained fewer contaminants and high purity peptide (**Figure 3.23.B**). It was concluded that contaminants were removed by size exclusion chromatography and therefore no additional purification steps were required, indicating that a well-defined and robust chromatography model had been established for the purification of SRSF1-CPP.

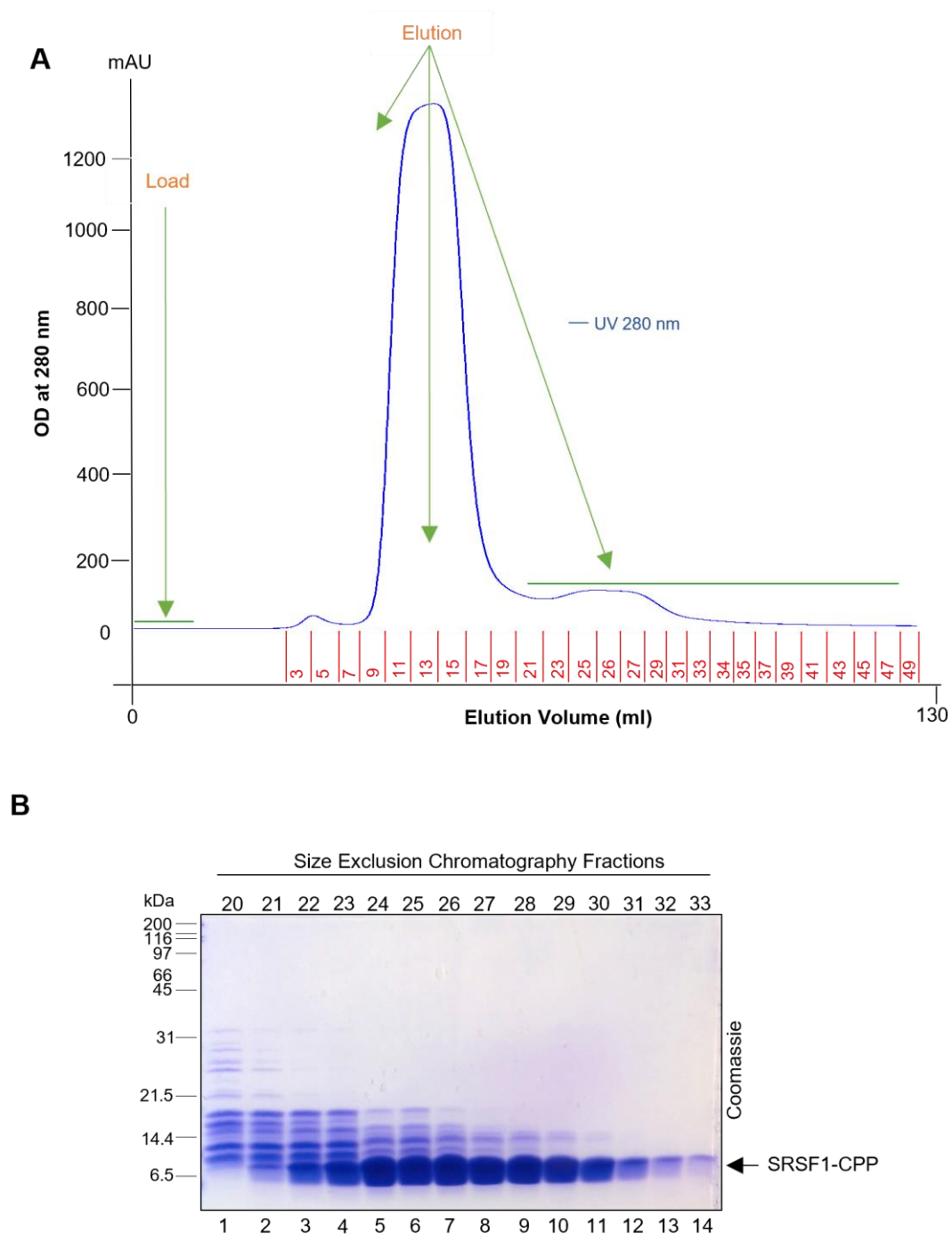


Figure 3.23. Size exclusion chromatography of SRSF1-CPP. SRSF1-CPP eluates from the Heparin column were run on a Superdex 30 PG size exclusion column. 1 ml fractions were collected. **(A)** Elution profile of Size exclusion of chromatography. **(B)** Coomassie stained 15 % acrylamide gel analysis of 15 μ l of each fraction. Fraction numbers are given above the SDS-PAGE gel. **Lane 1 to 14** indicates fractions number starting from 20 to 33. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.5.5. Quantification and Concentration of SRSF1-CPP

Fractions 26-33 containing high purity SRSF1-CPP were pooled and diluted in phosphate buffer to make the salt concentration equal to the salt concentration in PBS. This ensures that the composition of the resulting peptide will mimic the condition of human body fluid mitigating cell death or unwanted side effects. The peptide fractions were concentrated using an Amicon Stirred Cells device with a 6 kDa membrane pore size which allows any molecule larger than 6 kDa to remain in the peptide solution, whereas smaller molecules pass through the membrane. The resulting peptide concentration was quantified with a Pierce Quantitative Fluorometric Peptide Assay kit. The fluorescence intensity of standard peptide is provided in **Figure 3.24.A**. The purity stock of concentrated recombinant SRSF1 peptide was compared with a stock of synthetic peptide and analysed by Coomassie staining (**Figure 3.24.B**). The Coomassie gel shows that 1 ml highly pure and highly concentrated recombinant peptide at 1 mM was produced.

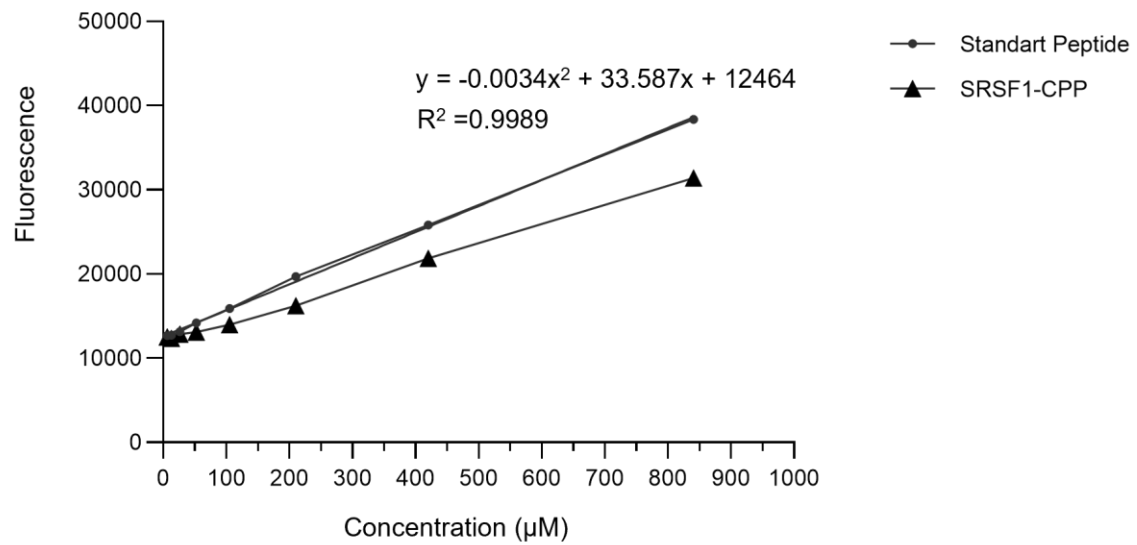
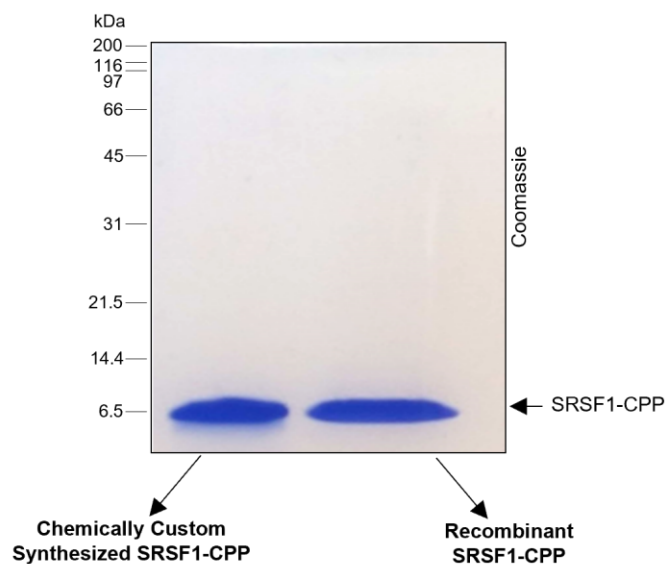
A**B**

Figure 3.24. Quantification and concentration of SRSF1-CPP. The eluate containing SRSF1-CPP was concentrated using an Amicon Stirred Cells device with 6 kDa molecular weight membrane. **(A)** Fluorescence intensity standard curve. **(B)** Coomassie stained 15 % acrylamide gel analysis of 1 µl of synthetic and recombinant SRSF1-CPP. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.6. Expression and Purification of GST-control V5-TAT PTD CPP

A robust protocol for expression and purification to the milligram range of SRSF1-CPP had been optimised. The next step of my project was to produce a control peptide lacking the SRSF1 aa 89-120 sequence. Therefore, a GST -control V5-TAT PTD expressing plasmid was generated.

3.4.6.1. Generation of Plasmid for GST-control V5-TAT PTD CPP

The plasmid to be used to produce the control (ctrl) CPP was obtained using an inverse PCR strategy (**Figures 3.25.A and 3.25.B**). Initially, the pGEX6P-1-SRSF1-V5-TAT plasmid was amplified using the two BamHI primers which are annealing in opposite directions to remove the SRSF1 sequence.

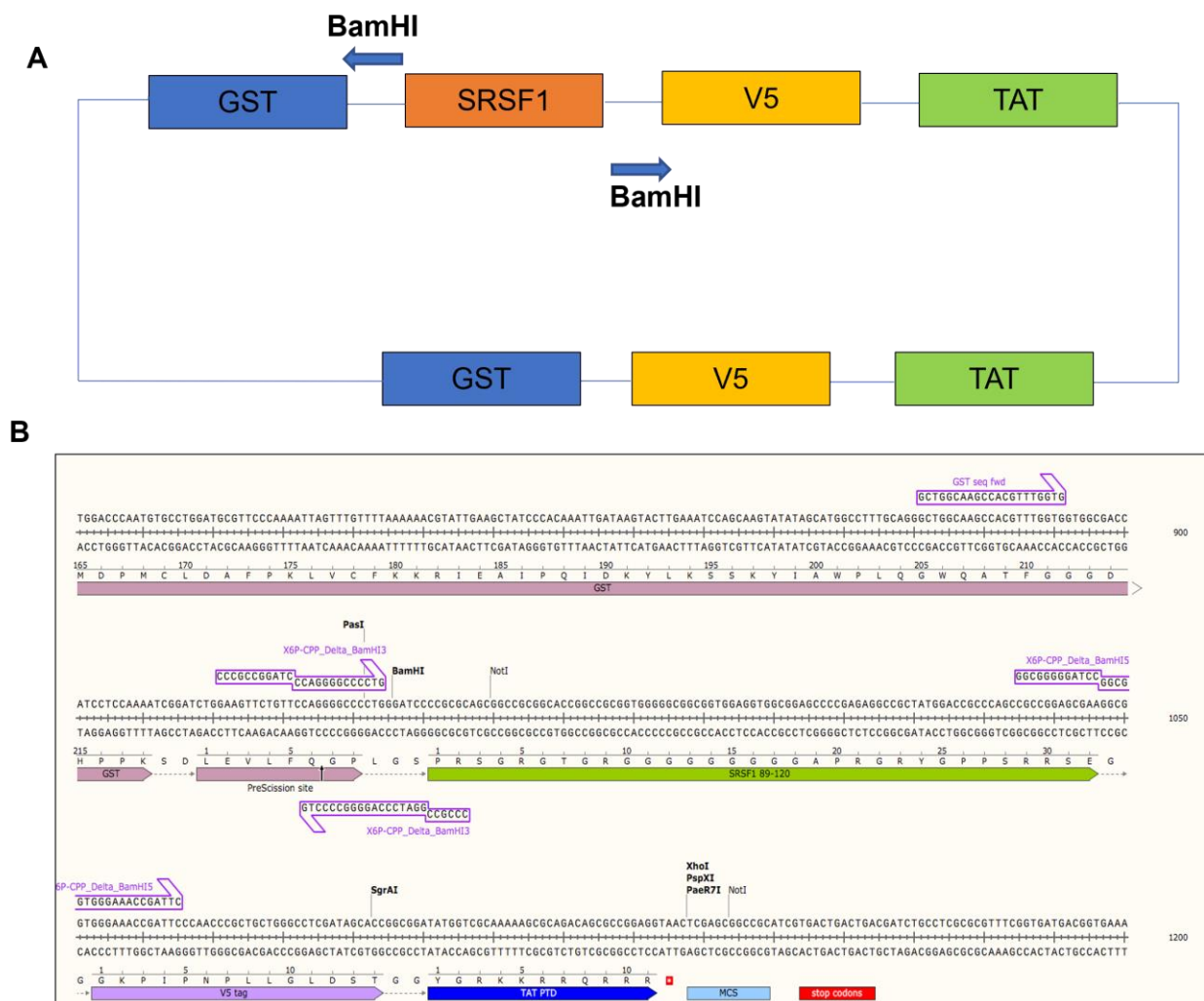
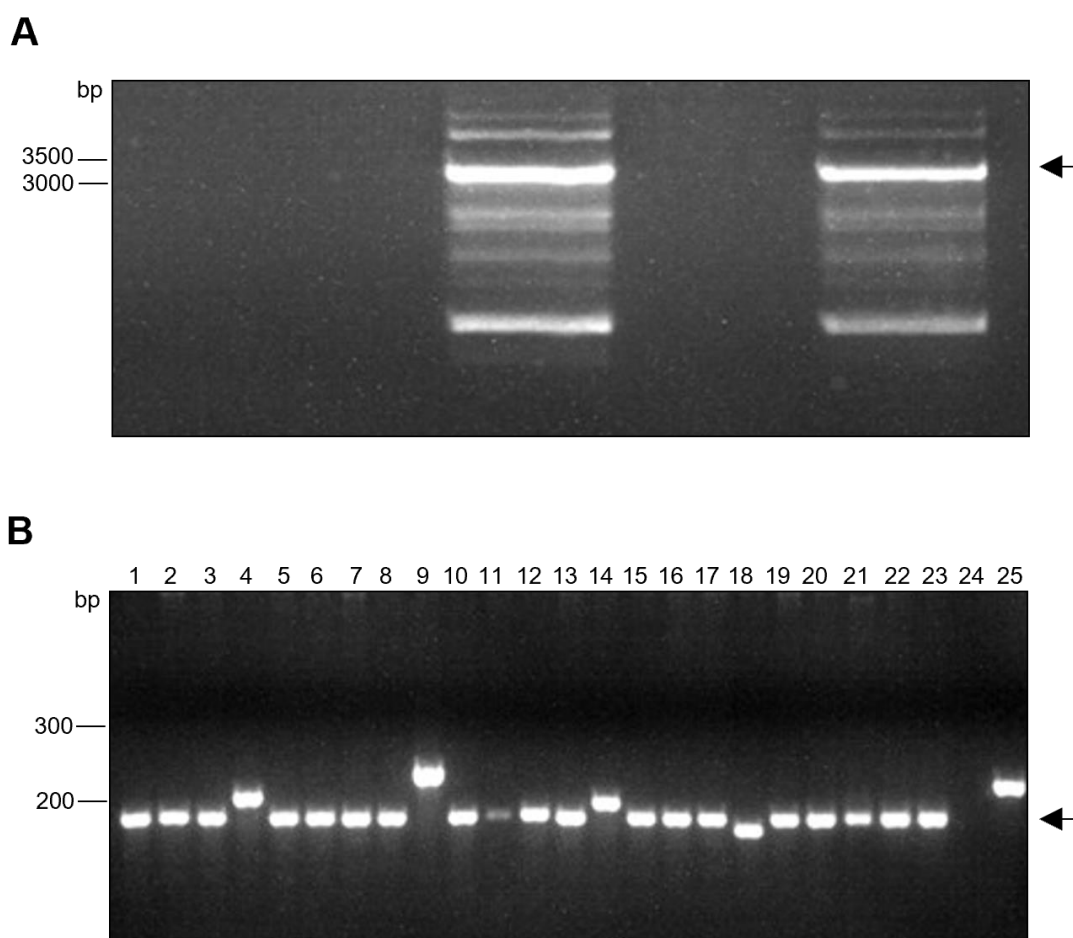


Figure 3.25. Cloning strategy of Control-V5-TAT PTD CPP. (A) Schematic representation of inverse PCR strategy used for generating the GST Control-V5-TAT PTD plasmid. **(B)** The sequence map of the plasmid expressing GST- SRSF1(89-120)-V5-TAT PTD. Both DNA sequences and encoded amino-acids in the one-letter code are highlighted on the map.

Multiple PCR products were detected on a 1 % agarose gel (**Figure 3.26.A**); however, the PCR product of interest with BamHI on either side minus the SRSF1 sequence had migrated between 3000 and 3500 bp. The product was purified using gel extraction and digested with BamHI. Restricted DNA fragments were then ligated into a GST expression vector prior to transformation into BL21-Rosetta *E.coli* cells. Colonies were screened and confirmed by BamHI restriction digest, producing a 247 bp fragment with correct integration after removing SRSF1 from the plasmid which has 96 base pairs (**Figure 3.26.B**). It was decided that clone #3 and #5 of each of the control peptide plasmids would be used for large-scale production of the plasmid. The integrity of the plasmid was confirmed using Sanger sequencing. The full sequence of GST-control-V5-TAT PTD is shown in **Figure 3.26.C**.



C

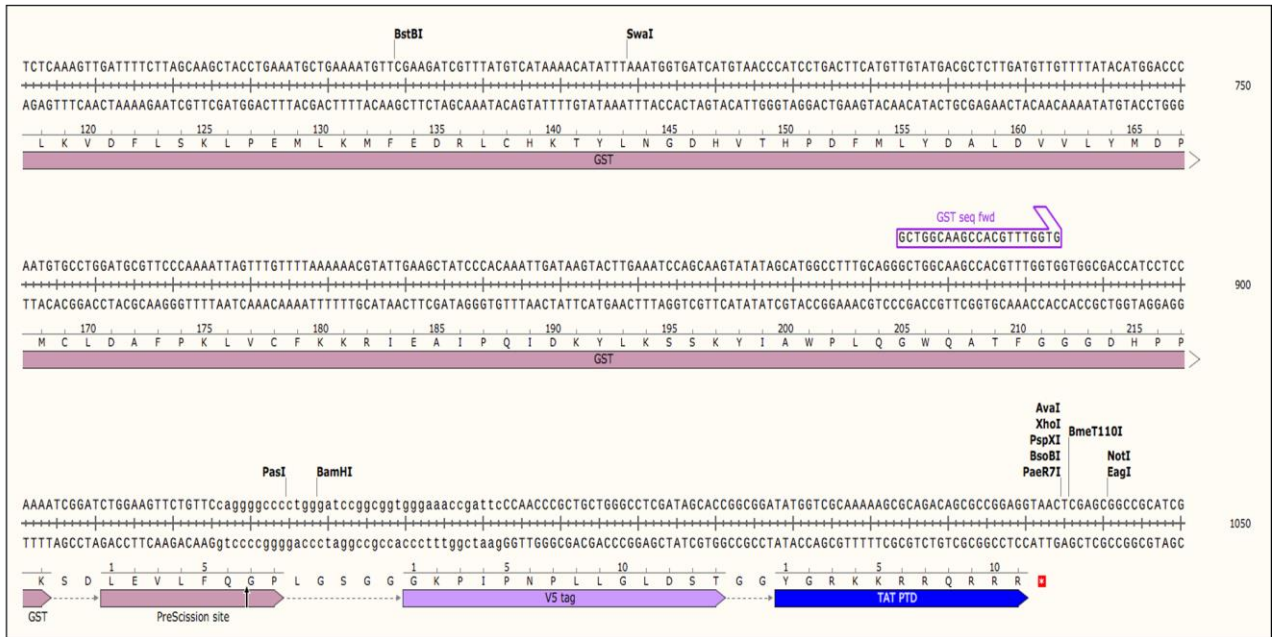


Figure 3.26. Generation of GST-Control-V5-TAT PTD CPP plasmid. PCR amplification of GST-Control-V5-TAT PTD CPP. **(A)** Arrow depicts the band that was extracted on a 1 % agarose gel. **(B)** Bacterial colony screening after digesting with BamHI and run a 3 % agarose gel. Arrow depicts successful clones. One biological repeat was performed of the Agarose gel images presented within this Figure. **(C)** The sequence map of GST-Control-V5-TAT PTD. Both DNA sequences and encoded amino-acids in the one-letter code are highlighted on the map. Molecular weight marker bands are indicated in bp.

3.4.6.2. Expression of BL21-Rosetta-GST- Control V5-TAT PTD CPP

BL21-Rosetta *E. coli* expression strains previously transformed using the plasmid construct to express GST-V5-TAT PTD were grown in TB culture. The optimal expression time point was determined by collecting sample from each post-induction. Un-induced (0 h) and IPTG induced cultures at different induction times (1 hour to overnight (O/N)) were collected and analysed using SDS-PAGE followed by Coomassie blue staining of 15 % resolving gel and western blotting. There was a very high expression of GST fusion protein, which increased gradually from 1 hour to 5 hours, at the expected molecular weight (about 32 kDa), when comparing un-induced and induced cells (**Figure 3.27**). The maximum GST fusion protein expression was attained in BL-21 Rosetta cells at 5 hours post induction time. It was decided that 5 hours of post induction time in BL21- Rosetta cells was optimal for the highest expression of this protein. Therefore, all subsequent experiments concerning the control CPP were carried out with this optimised expression protocol.

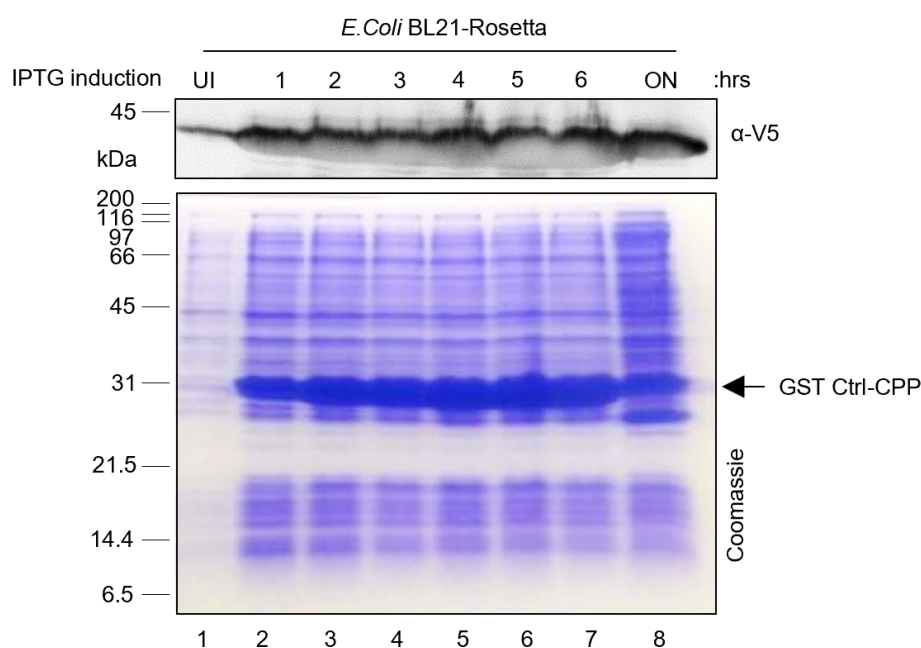


Figure 3.27. Induction time course of GST-V5-TAT PTD in BL-21 Rosetta *E. coli* strains. Western blot and Coomassie stained 15 % acrylamide gel to analyse GST-Ctrl fusion protein expression from BL21-Rosetta *E. coli* cells to determine the best induction time. **Lane 1:** Un-induced bacteria culture. **Lane 2-8:** IPTG induced culture of 1, 2, 3, 4, 5, 6 hours and O/N. Western blot is probed with a V5 antibody. Arrow depicts expressed GST- V5-TAT PTD. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.6.3. Purification of GST-control V5-TAT PTD CPP

Large scale purification of GST-Ctrl-CPP fusion protein was performed by initially applying total crude extract of GST fusion protein to the stationary phase of GSH affinity chromatography (5 ml slurry) to enable maximal binding of GST-fusion protein followed by PreScission protease cleavage to release the peptide from the GST moiety. The further purification step was carried out by loading cleaved Ctrl peptide to Heparin adsorption chromatography (5 ml slurry) to remove contaminants followed by a washing step and then performing elution in a 5 ml fraction with 1 M PBS/NaCl elution buffer. The diluted input and flow throughs (FTs), GST residues left on GSH column and eluates (E1, E2, E3 and E4) were assessed by Coomassie blue stained SDS-PAGE gel (**Figure 3.28**). The result reveals that the majority of the GST fusion protein in input binds well on GSH column as the intensity of the FT band was considerably lowered compared to the input lane. The eluted fractions also exhibited a strong band at around 5.8-6 kDa, which corresponds to the expected size of the control V5-TAT PTD peptide. However, the purity of the control peptide was not at the desired level.

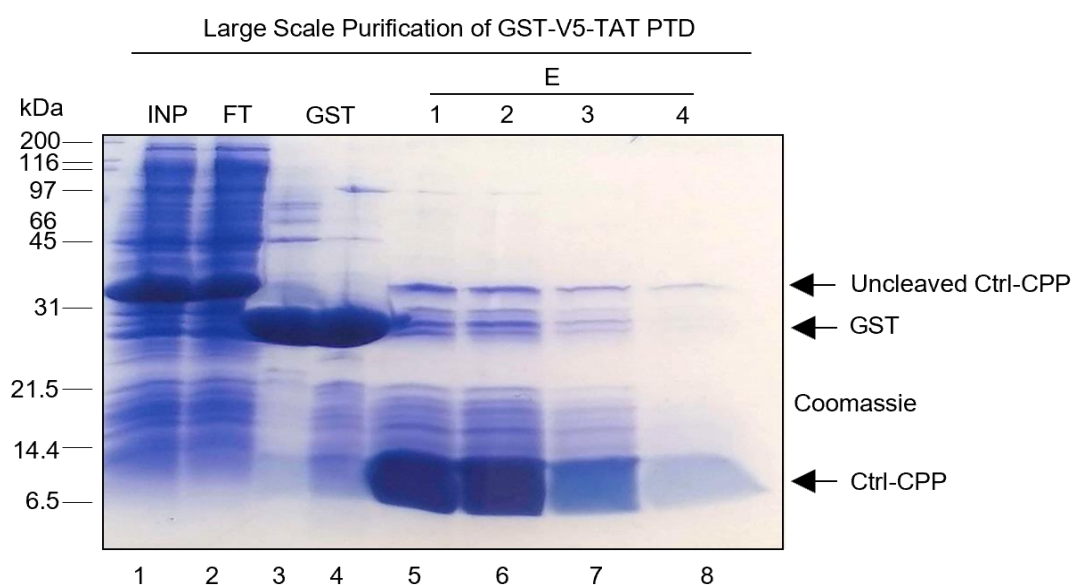


Figure 3.28. Large scale purification of GST-Ctrl-V5-TAT PTD CPP with GSH affinity and Heparin adsorption chromatography. Coomassie stained 15 % acrylamide gel analysed the two-step purification process of GST tagged SRSF1-CPP. **Lane 1:** Input (INP). **Lane 2:** Flow Through (FT). **Lane 3:** GST Residual Elution - 1 (GST) **Lane 4:** GST Residual Elution - 2 (GST). **Lane 5:** Eluate 1. **Lane 6:** Eluate-2. **Lane 7:** Eluate 3. **Lane 8:** Eluate 4. Arrows depict respectively from top to bottom of the gel; Un-cleaved GST-SRFS1 CPP, GST and Cleaved SRSF1-CPP. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

An additional purification step was required and performed using Size exclusion chromatography (pre-packed column). The eluted peptide fractions were subsequently applied to the column via a sample loop. The peptide was eluted with 0.5 mM PBS/NaCl and collected in 1 ml fraction after 40 ml elution buffer had passed through the column. The chromatogram of size exclusion chromatography is provided in **Figure 3.29.A**. It shows that the small wide peak after first sharp peak, which indicates high molecular weight contaminants at around 70 ml elution volume, corresponding to the fractions containing peptide (lanes 20-30). The fractions selected based on the chromatogram were monitored by Coomassie blue staining (**Figure 3.29.B**). The gel result displays that fractions 20-25 still contain relatively high levels of contaminants, but fractions 26-30 represent highly pure peptide with little contamination or degradation.

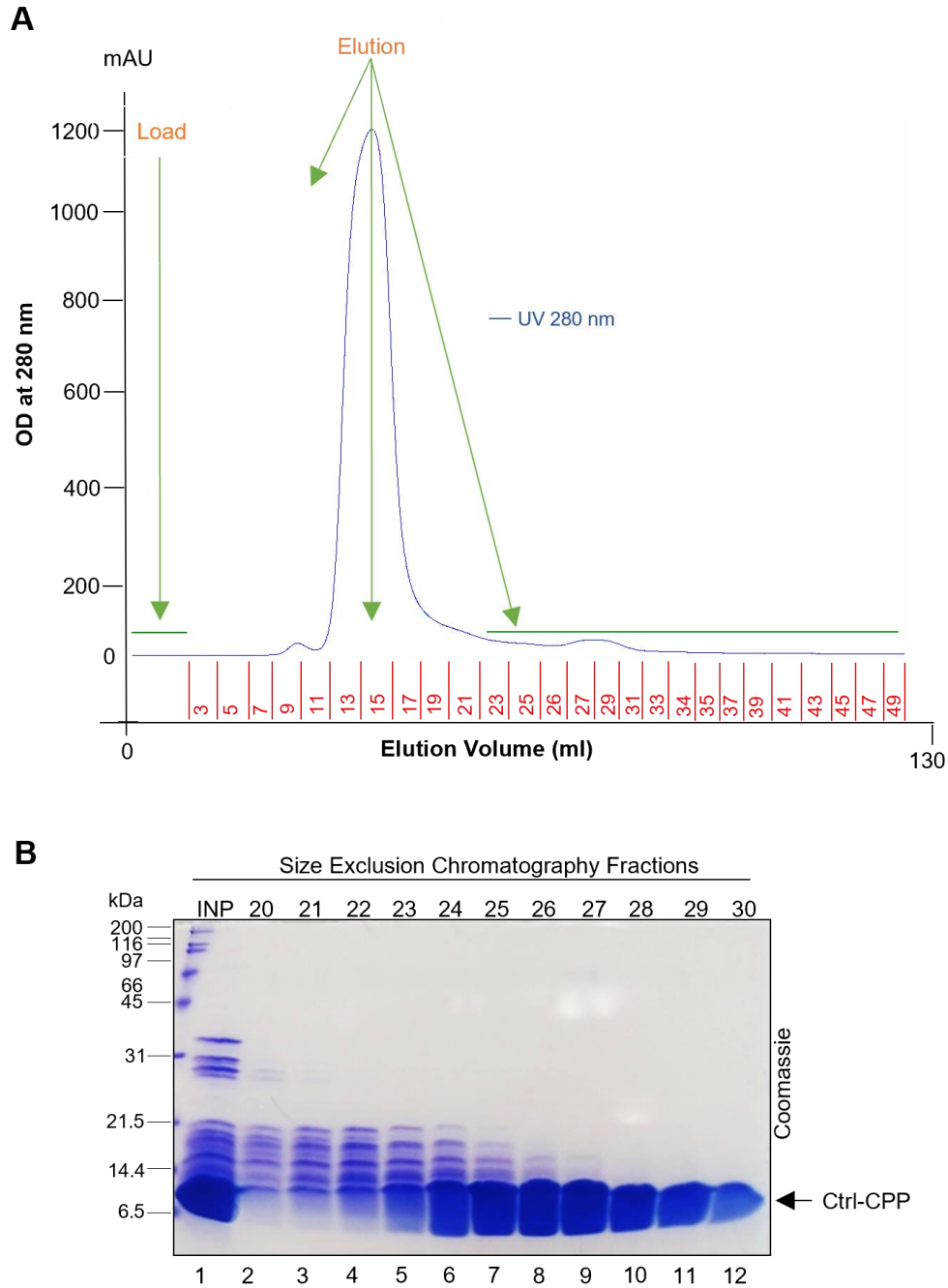
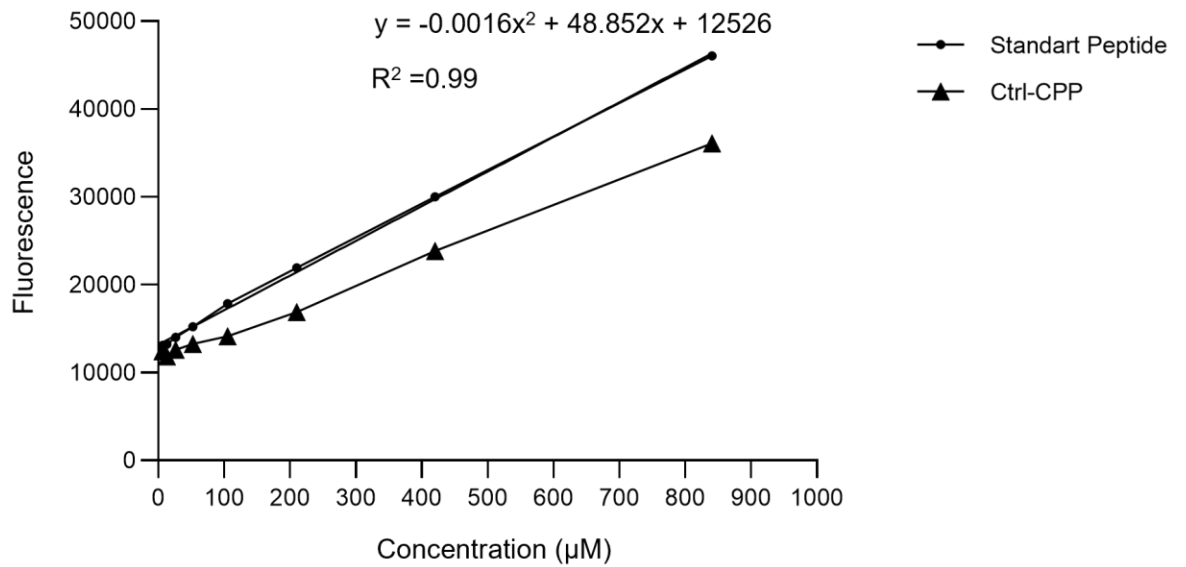


Figure 3.29. Size exclusion chromatography of Ctrl CPP. Ctrl CPP eluates from the Heparin column were run on a Superdex 30 PG size exclusion column. 1 ml fractions were collected. **(A)** Elution profile of Size exclusion of chromatography. **(B)** Coomassie stained 15 % acrylamide gel analysis of 15 μ l of each fraction. Fraction numbers are given above the SDS-PAGE gel. **Lane 2 to 12** indicates fractions number starting from 20 to 30. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.4.6.4. Quantification of Ctrl-CPP Concentration

Fractions 26-30 containing SRSF1-CPP with high purity were pooled and diluted in phosphate buffer. The peptide fractions were concentrated using an Amicon Stirred Cells device with 3 kDa membrane pore. The resulting peptide concentration was quantified with a Pierce Quantitative Fluorometric Peptide Assay kit. A standard curve was generated to standard peptide diluted to concentrations between 0 to 841 μ M CPP (**Figure 3.30.A**). The purity stock of concentrated recombinant Ctrl peptide was compared with a stock of Ctrl peptide and analysed by Coomassie staining (**Figure 3.30.B**). The Coomassie gel in Panel B shows that 1ml highly pure and concentrated recombinant control peptide at 1 mM was produced.

A



B

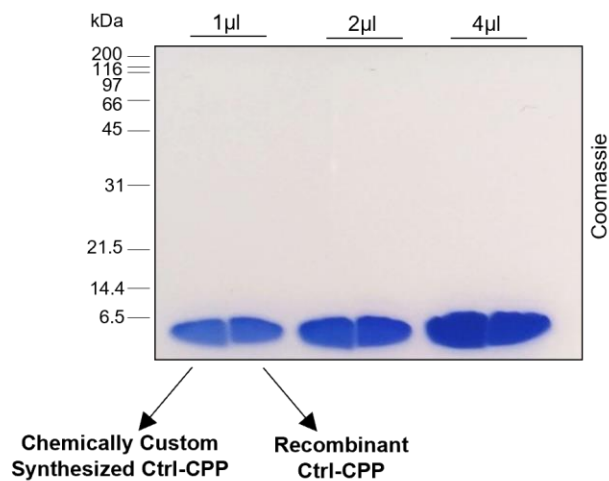


Figure 3.30. Quantification and concentration of Ctrl CPP. The eluates containing Ctrl CPP were concentrated using an Amicon Stirred Cells device with 6 kDa molecular weight membrane. **(A)** Fluorescence intensity standard curve. **(B)** Coomassie stained 15 % acrylamide gel analysis of 1 μl of synthetic and recombinant Ctrl-CPP. Molecular weight marker bands are indicated in kDa. One biological repeat was performed of the Coomassie gel images presented within this Figure.

3.5. Discussion

This chapter presents different optimization attempts concerning the expression and purification of SRSF1-V5-TAT PTD CPP for use in subsequent *in vitro* and *in vivo* studies of *C9ORF72*-mediated ALS/FTD to validate the functionality of the recombinant peptides. However, to perform these experiments *in vitro and in vivo*, a high degree of purity and a large quantity of recombinant peptide is needed.

Peptide production is a very challenging process, therefore there are several proposed methods for peptide production which are chemical (synthetic), recombinant and enzymatic methods. Peptide production via synthetic method, based on solution-based synthesis and solid phase peptide synthesis, is the most prevalent technique due to providing flexibility in scale, sequence, and modification. Indeed, this method is efficient to produce a peptide shorter than 30 amino acids (Chandrudu et al., 2013). However, production of peptide having more than 30 amino acids is challenging and expensive because of highly sequence related and requiring high purification cost (Gaglione et al., 2019). Additionally, this method is not ideal for creating peptides that are uniformly or selectively labelled with isotopes, which are necessary for studying the interactions between ligands and receptors using high-resolution NMR spectroscopy in liquid or solid form (Zorko and Jerala, 2010).

In our laboratory, I have developed a recombinant method for the production of peptides. This technique, although it took approximately 6 months to establish, offers significant advantages over synthetic peptide production approaches. Notably, it permits the purification of the desired peptide within a week, compared to the two-month duration required for the conventional synthetic methods. In terms of yield, the recombinant production system consistently provides 10 mg of peptide post-purification, a quantity that is commensurate with yields from chemical synthesis. Equally important is the purity level; our recombinant method achieves a degree of purity that parallels that of chemical synthesis. One of the other advantages of this recombinant approach is the turnaround time. Given the stability issues associated with peptides, which often limit their shelf-life to a few months, the ability to produce a fresh batch within a short period is invaluable for experimental continuity. In contrast, when relying on custom synthesis, we previously encountered numerous challenges concerning the functionality and stability of peptides, leading to extended waiting periods before resuming our experiments.

Given the evident benefits of this method, we are in the process of publishing our findings, specifically focusing on the purification of cell-permeable peptides. This publication will facilitate the adoption of our method by other laboratories, streamlining peptide production and research globally.

In the work described in this chapter, I have utilized a recombinant production method for obtaining SRSF1 and ctrl-CPPs. However, expression of peptide in *E.coli* and purification is difficult using a recombinant peptide production system. Therefore, multiple parameters were tested from review of the literature to enhance the production yield. For this goal, several strategies during designing, expression and purification of SRSF1-V5-TAT PTD were approached to generate a robust methodology to produce peptide with a high degree of purity and in a large quantity for biochemical characterisation in further studies.

The first parameter considered to achieve recombinant protein expression was to choose a strong promoter. The selected promoter determines whether the protein of interest is obtained through inducible or constitutive system. Expression of protein constitutively may result in plasmid instability because of metabolic drain or burden. Therefore, that system generally used in a case that growth rate of host is not affected significantly by the expression of protein (Palomares et al., 2004). In many instances, the best conditions for cell growth are different from those for recombinant protein production (Balbás, 2001). For our study, we used T7 for the pET expression system in BL21 strains with IPTG-inducible T7 RNA polymerase and the constitutive *tac* promoter for the pGEX6P-1 expression system, both of which have strong activity and common promoters enabling the generation of large amount of mRNA and concomitant recombinant protein expression after IPTG induction (Gräslund et al., 2008).

Another strategy for increasing peptide expression was to determine the optimal expression strains of *E.coli* which play an important role in acquiring a soluble and large amount of peptide. Up to now, many commercial *E.coli* strains have been developed to maximize protein expression. BL21(DE3) and its derivatives are one of the commonly used strains which lack Lon and OmpT proteases, thereby providing enhanced protein stability (Francis and Page, 2010).

In this study, we used BL21 codon plus (DE3) RP, RIL and Rosetta competent cells which are engineered to harbour additional copies of gene coding for rare tRNA, thereby enhancing expression of gene sequences that contain rarely used *E. coli* codons in the expression of heterologous protein in *E. coli* because the high throughput recombinant protein expression is often limited by the rarity, number, position and combination of some tRNAs, which are abundant in organisms from which the proteins are produced (Gustafsson et al., 2004; Rosano and Ceccarelli, 2009).

The use of fusion tags has been also considered as another strategy for enhancing the production of the desired protein. The appropriate fusion tag should be able to enhance expression, solubility and folding of the target protein (Jia and Jeon, 2016). The choice of fusion tag also plays a role in facilitating the downstream process. Therefore, it is expected that the fusion tag should enhance the purification and detection of the target protein from *E. coli* (Rosano et al., 2019). However, there is not one ideal tag that can provide all of these features, and they all offer different advantages over one other. Therefore, it was necessary to test several fusion tags to decide which to use in further studies (Rosano and Ceccarelli, 2014). In the present study, Histidine and GST fusion tags were evaluated.

I first started to express histidine₆-tagged SRSF1(89-120)-V5-W-TAT CPP in *E. coli* BL21 competent cells using different induction times and cells to determine the optimal conditions for peptide expression (**Figure 3.5**). The Histidine tag is the most popular and widely used fusion partner for expression and purification studies (Kimple et al., 2013). They are small fusion tags ranging from 2-10 histidine epitopes and their charge does not generally influence protein structure and function (Li, 2011). Histidine tags also provide rapid single affinity purification of the target protein using metal ions immobilized within the column such as nickel, cobalt and copper because of higher affinity of histidine residues for metal ions (Gaber-Porekar and Menart, 2001).

In the histidine tag fusion strategy, the length of the histidine residues constitutes a key factor during construction of the peptide or protein because it has effects on the expression and purification steps (Malhotra, 2009). In this manner, a study examined the length and position of histidine tags on expression and purification of membrane proteins. They expressed the protein with 6 and 10 histidine residues and showed that the protein yield with 6 histidine residues was four-fold higher than that of 10 histidine residues (Mohanty and Wiener, 2004). Therefore, in our study, six -histidine residues (His₆) were preferred which is the most popular and widely used strategy for the purification of the target protein. Indeed, His₆ is generally considered a good starting point during the design of a new protein (Mohanty and Wiener, 2004).

The SRSF1-CPP tagged with (His)₆ was successfully expressed in multiple *E. coli* strains. For purification, I initially utilized a cobalt column due to the known affinity between histidine and cobalt. Despite this approach, the eluted peptide from the column was riddled with contaminants (**Figure 3.6**). In response, a secondary purification was performed using diverse chromatography media, such as mono-S strong (**Figure 3.6**) and SP weak ion exchange (**Figure 3.7**). Altering chromatography methods did enhance the yield, reaching approximately 100 µg (**Figure 3.8**). Notably, this yield was sufficient to support subsequent in vitro and in vivo experiments. However, the overall yield was constrained, primarily due to the suboptimal expression levels of the (His)₆-tagged SRSF1-CPP. To address this, I am exploring the possibility of amplifying the peptide expression by leveraging alternative culture mediums. The composition of growth media is known to have a significant impact on the yield of protein expression in bacteria or cells because some components like glycerol and yeast extract, which are responsible for providing carbon and nitrogen to growth medium, can inhibit protein production when they are present in higher or lower concentrations in the growth medium; therefore, formulation of medium conditions is required to be in balance (Volontè et al., 2011). In this manner, in order to enhance cell density and histidine tagged SRSF1-CPP expression, I decided to use SB medium, which contains more tryptone and yeast extract compared to TB medium but no glycerol which is a carbon source for bacteria. (Kram and Finkel, 2015; Piubelli et al., 2013).

SB medium has previously been used to increase protein expression. For example, Piubelli *et al.* aimed to produce Trx-TB10.4 which is an immunodominant protein that is part of esat-6 family and recognized by T-cells from tuberculosis patients and therefore it is promising vaccine candidate for TB patients. The Trx-TB10.4 was expressed first in LB medium and subsequently purified by metal-chelating chromatography. However, the yield of final pure protein was insufficient for vaccine production, indicating that expression of the target protein was low. In order to improve expression of the protein, they evaluated four different media, which are LB, TB, SB and M9, and revealed that the highest protein expression was attained with SB medium (Piubelli *et al.*, 2013). It was decided to investigate the efficiency of SB medium on peptide expression in our study. However, we found that histidine tagged SRSF1 peptide expression was lower than when TB medium was used (**Figure 3.9**). A similar result was observed in a study that compared survival rates of *E.coli* K-12 in TB and SB medium. They found that >80 % of viable *E. coli* cells died when incubated in SB medium. However, approximately 10 % reduction in viable *E. coli* cells was observed in TB medium (Kram and Finkel, 2015). This may be due to the absence of glycerol in SB medium which results in acceleration of alkalization, thereby increasing pH more rapidly and subsequently causing cells to enter the death phase sooner or inhibiting growth and protein expression. Therefore, glycerol is recommended as a necessary component for bacteria to maintain growth and protein expression (Kram and Finkel, 2015) (Kopp *et al.*, 2018). In addition to SB medium, a plasmid harbouring no tryptophan codon, which is still able to express histidine tagged SRSF1 peptide, was tested (**Figure 3.10**). The result showed that the expression of SRSF1 peptide was lower in the plasmid without tryptophan when compared with plasmid in the presence of tryptophan codons. Overall, *e. coli* expression and purification of histidine tagged SRSF1-CPP was insufficient to generate the required amount of pure peptide because the amount of soluble expressed peptide was not enough in *E.coli*. Various structural studies found that up to 50 % of His-tagged proteins expressed within *E.coli* are insoluble (Edwards *et al.*, 2000) (Stevens, 2000).

To overcome this, the SRSF1-CPP construct was cloned downstream of a GST fusion tag known to express well and to be soluble. GST fusion is an increasingly popular method for difficult to express peptides and proteins as it enhances their solubility and expression (Riley *et al.*, 2013). It also provides rapid and single high-affinity purification with glutathione beads.

Additionally, GST fusion has been utilized to determine the structure of several small peptides and regulatory proteins such as gp41 from HIV (Lim et al., 1994) and the ankryn-binding domain of α -Na/K ATPase (Zhang et al., 1998). However, this strategy has some challenges. In order to maintain solubility and protect the structure and function of the target protein, fusion tags need to be removed using a PreScission protease enzyme (Bell et al., 2013). The SRSF1-CPP, fused with GST, was successfully expressed across multiple *E. coli* strains, as illustrated in **Figure 3.11**. Subsequently, I performed a small-scale GSH affinity purification, employing a variety of lysis buffers. The primary objective of assessing three distinct lysis buffers, depicted in **Figure 3.12**, was to optimize the enzymatic cleavage conditions, facilitating the detachment of the peptide from the GST moiety. Upon evaluation, the 'Tris + DTT' methodology emerged as the most efficacious. Notably, while each buffer solution allowed for the successful cleavage of the SRSF1 peptide, the inclusion of DTT seemingly enhanced binding to glutathione beads. This is likely attributed to DTT's role in averting the oxidation of disulphide bonds, which could either induce the formation of oxidized dimers of glutathione or stymie peptide aggregation.

The proper organization of disulphide bonds is related to obtain functional activity of recombinant protein and any errors that happen during the formation of disulphide bonds can result in misfolding and aggregation of protein. Several studies have aimed to enhance disulphide bond formation in protein by adding external additives to the medium (Alliegro, 2000; Messens and Collet, 2006). DTT is one of the components that can be added externally and was used in a study which aimed to investigate the effect of DTT on accumulation of ribonuclease inhibitor (RI) expression in cytoplasmic and periplasmic regions of bacterial cells. They first tested 36 constructs expressing RI in the periplasmic region and found that high RI production was attained by several constructs; however, the expressed RI was insoluble and functionally inactive. To overcome these issues, DTT was included externally in their growth medium. They found that DTT significantly enhanced the activity and expression of soluble RI in the strains. Indeed, SDS PAGE analysis of the expressed protein showed that the addition of DTT generated enhanced intensity of bands corresponding to RI (Šiurkus and Neubauer, 2011). DTT is also suggested to overcome the GST dimerization issue, which prevents elution of fusion proteins from the GST chromatography resin (Maru et al., 1996). A similar elution issue was observed when PBS and Tris strategies were used. The use of DTT in the last strategy prevented dimerization thereby eluting the cleaved SRSF1 peptide from GST resins.

After optimizing GST affinity chromatography with the Tris + DTT method, I conducted large-scale purification (**Figure 3.13**). This approach yielded more peptide than previous methods. However, further optimization was needed during the second purification steps due to CPP precipitation issue. This problem was overcome by introducing urea into the CPP solution. Urea, which is a hydrophilic chemical additive, has been widely used to improve the solubility of active protein. For example, Patra *et al* have investigated the effect of several buffers with and without the addition of urea to enhance the solubility of recombinant human growth hormone (r-hGH) which was expressed as inclusion bodies. They found that addition of 2 M urea enhanced 94 % r-hGH solubility from inclusion bodies without disturbing the structure of the protein (Patra et al., 2000). It should be noted that the ideal urea concentration varies from protein to protein because some proteins such as tubulin may require higher urea concentrations to obtain better solubilization (Rabilloud et al., 1997). Although peptide solubility was improved by the addition of urea, the binding ability of the mobile phase of the peptide on cationic exchangers had to be tested in the presence of urea because it was reported that urea containing elution buffer has a significant effect on protein retention during separation from ion exchange chromatography. Parente and Wetlaufer examined the effect of urea on the retention behaviour of hen egg white lysozymes from cationic exchange chromatography. They revealed that the presence of urea decreases retention of lysozymes (Parente and Wetlaufer, 1984).

In another study, the retention behaviour of eight proteins was investigated with the presence of urea in the mobile phase of hydrophobic interaction chromatography (HIC). They observed that urea containing mobile phase of protein caused changes in surface tension of the mobile phase indicating that urea has a role in the retention of proteins (Feng et al., 1995). Therefore, it was important to investigate the binding kinetics of the peptide in urea. In parallel to urea, several strategies including examining a range of salt concentrations in the loading and washing buffers, as well as investigating different salt concentrations in the elution buffer, were also tested in several ion exchange/adsorption chromatography media. The optimization of the second purification step began by evaluating various salt concentrations in the presence of urea on SP chromatography beads (**Figure 3.14**). Determination of the correct salt concentration is a highly important parameter for column chromatography because it directly affects the binding capacity of the column and the selectivity of the target protein, thus enabling non-specific suppression of protein binding (Tsumoto et al., 2004); (Ejima et al., 2005).

There are two effects that are commonly seen during the purification of any target molecules in certain salt concentrations; these are called ‘salting-out’ and ‘salting-in’. These effects are independent from any chromatography system and can occur in all types of chromatography system. ‘Salting-out’ and ‘salting-in’ effects are generally referred to respectively as a decrease and increase in protein solubility (Tsumoto et al., 2007). In ion exchange chromatography, the binding of target molecules to chromatography resins is mediated by electrostatic interaction. However, salts interfere with the binding efficiency of the protein to the surface of the resin (Tsumoto et al., 2007; Von Hippel and Wong, 1965). Several studies have shown that increased salt concentration in the mobile phase corresponds with decreased binding affinity of the protein. This is because higher salt concentrations decrease protein solubility leading to aggregation and excessive charge shielding, thus reducing the amount of protein bound to the resin (Arakawa and Timasheff, 1982; Melander and Horváth, 1977; Tsumoto et al., 2007). Based on these lines of evidence, it was observed that the highest peptide binding was attained at 250 mM NaCl, and further increasing the salt concentration inhibited the binding of the peptide to the column, which aligns with the literature (Tsumoto et al., 2007). However, the main problem was retention of the peptide in the chromatography resin as a significant fraction of the peptide was not released from the matrix in the presence of elution buffer containing a high salt concentration (2 M NaCl/PBS). Under normal elution conditions of ion exchange chromatography, use of higher salt concentration (1M NaCl or above) during elution needs to be carefully performed with molecules that are prone to aggregation (Kumar et al., 2000). However, performing harsh elution was required in our case because binding of the peptide to the column was strong and binding to the resin occurred at a higher concentration. Therefore, this led us to perform elution of the bound peptide at a high salt concentration.

The optimization experiment and large-scale purification were performed with sodium phosphate buffer which is commonly used based on its ability to maintain a stable pH in specific environments and provide efficient target protein recovery (Pavani et al., 2021). However, it also carries some disadvantages in biological systems. For example, it demonstrates inefficient buffering capacity when used over pH 7.5 and also prevents metal-ion dependent biochemical reactions (Ferreira et al., 2015). In our situation, the peptide in the mobile phase of sodium phosphate buffer with a range of salt concentration was applied to the SP resin, but the majority of the peptide was not dissociated from the SP resin.

In order to eliminate retention of peptide from resins, I tested Tris and Hepes buffers at different pH levels and elution buffers in different salt composition (**Figure 3.15**). The use or choice of the ideal buffer is an important step in the purification process because it affects several parameters which are directly related to the stability, binding kinetics, activity, and structural changes of the protein (Brudar and Hribar-Lee, 2021; Loreto et al., 2017). In this regard, Tris and Hepes buffer can work effectively in a wider pH range and are therefore widely preferred in purification studies (Loreto et al., 2017); (Liu et al., 2019). For example, it was reported that the use tris buffer pH 8.0 influenced the stability of high-performance liquid chromatography and the performance of the stationary phase where poly(methyloctylsiloxane) was immobilized on silica-PMOS-SiO₂. The authors also highlighted that tris buffer provides unique selectivity and stability for the mobile phase thereby making them promising candidates for pharmaceutical research (Ferreira et al., 2015).

I also investigated the effect of pH on peptide retention. It is known that pH plays a significant role in peptide/protein adsorption to and desorption from the column (Parkes et al., 2014). Borges and Collins evaluated the effect of different mobile phase pH (pH:6,7,8 and 11) with the same buffer on stationary phase retention and revealed that increased pH of the mobile phase reduces interactions between the basic solutes and the stationary phase thereby resulting in lower retention times (Borges and Collins, 2012) . In our case, the mobile phase of peptide with high pH was also chosen to decrease peptide interaction with the stationary phase by making them less protonated, thereby having less positive charge for the peptide as we were using SP cationic ion exchange chromatography. In theory, the peptide would not be able to facilitate strong binding to a negatively charged surface, which would accelerate protein desorption from the column. However, this was not observed in our experiment. Overall, from this optimization study, I found that using different pH values in the mobile phase and also using elution buffers with different salt concentration were not found to have a positive effect on disassociating the peptide from the chromatography resin.

Since the majority of the peptide was not fractionated from the SP stationary phase, a weak ion exchanger chromatography based on the carboxymethyl (CM) functional group was used (**Figure 3.16**). Weak ion exchange chromatography resins such as diethylaminoethyl (DEAE) (Easton et al., 2010) and CM are widely used when strong ion exchange does not work properly (Birney and O'Connor, 2001). Weak ion exchangers are preferred due to being more flexible in terms of selectivity and retaining their functionality over the pH 6–9 (Cummins et al., 2011). The optimization studies conducted with the CM matrix revealed that the peptide was not able to absorb onto the CM matrix in the presence of a high salt concentration in its mobile phase. The reason could be due to the high salt concentration increasing the hydrophobic interaction between the peptide and the CM column matrix, thus promoting precipitation of the peptide and interfering with the binding of the peptide to the stationary phase. Once the salt concentration was decreased in the mobile phase, the peptide absorbed well to the CM stationary phase (**Figure 3.17**). However, the majority of the peptide remained attached on the beads even after 2 M salt elution.

Further optimization was attempted with Heparin adsorption chromatography as a means of overcoming the issue of eluting the remaining peptide from the beads (**Figure 3.18 and 3.19**). Heparin confers a higher negative charge density than other molecules therefore it is able to bind a variety of positively charged molecules (around 400 different types of protein) (Bolten et al., 2018; Jayanthi et al., 2017) such as the SRSF1 peptide, which is rich in arginine and lysine, conferring a positive charge. In heparin purification, heparin is acting as both an affinity chromatography ligand and a cation exchanger, which also explains the reason why this purification method is widely preferred (Staby et al., 2005). For example, purification of extracellular vesicles (EVs) was conducted with a heparin column and highly pure EVs were collected in the eluted fractions. It was found that the recovery of EVs from the stationary phase of heparin was higher than with standard purification methods (Barnes et al., 2022). Other groups also attempted to purify lactoferrin (Yuan et al., 2021), DNA ligase, DNA and RNA polymerases, and fibroblast growth factor (Karlsson and Winge, 2004), and these purification attempts resulted in obtaining high levels purity of the target molecules. Considering the advantages of the heparin purification system, identical optimization experiments were performed with heparin chromatography to those performed with other chromatography media.

The results showed that the removal of bound peptide from the stationary phase, which was our main issue, was eliminated and a large amount of peptide was eluted (**Figure 3.20 and 3.21**). Indeed, heparin purification proved to be an adequate method to obtain a high purity of SRSF1-CPP with limited contamination (**Figure 3.22**). For the final stage of optimisation, the SRSF1 peptide eluted from the heparin column was subjected to purification size exclusion chromatography (SEC) (**Figure 3.23**). This enables separation of the peptide from contaminants based on their molecular weight. SEC has the ability to separate target molecules from contaminants based on their molecular weight under mild conditions, which protects the biological functionality of target molecules (Adawy and Groves, 2017). This purification method not only removes small contaminants, but it is also used to exchange buffer present in the protein or peptide and clean up reagents such as enzymes, inhibitors, and detergents. SEC is a widely used purification process in several research fields (Hong et al., 2012; Sun et al., 2004). For example, Guo et al established a robust dichotomic SEC method to purify EVs from serum samples (foetal bovine serum (FBS), human serum (HS) and FBS-free cell culture supernatants), which are a challenging sample type from which to purify EVs. With the optimised SEC purification method, they successfully obtained EVs with high purity and recovery rate from the column (Guo et al., 2021).

In another study, nicotinamide mononucleotide molecules (NMN), as an anti-aging product (Nadeeshani et al., 2022) in DNA repair and improvement of cell growth and survival (Hou et al., 2018), were purified with SEC from bacterial lysate. In the past, ion exchange chromatography was utilized for purification; however, they reported that this technique was complex, time consuming and not cost effective. Therefore, the authors developed an optimised purification method of NMN with SEC. Another reason for using SEC in this study was to reduce the salt concentration of the NMN sample for use in mass spectrophotometry (Marinescu et al., 2018). The desalting property of the SEC purification technique is why we performed the elution of SRSF1 and Ctrl peptide with 0.5 mM salt elution in our study. The SRSF1 peptide eluted from the Heparin column was applied onto size exclusion chromatography to filter out contaminants from the peptide mixture.

The SEC chromatography resulted in attaining very pure SRSF1 peptide fractions. This was followed by pooling the high purity peptide fractions together to concentrate the peptide. The resulting recombinant SRSF1 peptide was concentrated to 1 mM based upon the results from a fluorometric quantitation assay and exhibited the same purity level as the chemically synthesised peptide (**Figure 3.24**). The final step of this chapter was to produce a control peptide lacking the SRSF1 aa89-120 sequence. A GST-control-V5-TAT PTD plasmid (**Figure 3.25 and 3.26**) was generated and then the Ctrl CPP was highly expressed (**Figure 3.27**) and then subjected to purification with an optimised protocol comprising GSH affinity, Heparin adsorption (**Figure 3.28**), and Size Exclusion chromatography (**Figure 3.29**). At the end of the three-step purification and concentration steps, 1 mM pure stock recombinant Ctrl peptide was achieved based upon results from a fluorometric quantitation assay (**Figure 3.30**).

3.6. Experimental Conclusion

The work presented here describes the optimization of a protocol for the expression and purification of recombinant cell permeable peptides for use in subsequent *in vitro* (Chapter 4) and *in vivo* studies (Chapter 5) to validate the functionality of the purified peptides. The efficient, precise, and reproducible recombinant expression and purification method was required for obtaining a high degree of pure SRSF1 peptide in a feasible way. Therefore, several parameters and strategies were investigated systematically for SRSF1-V5-TAT PTD expression and purification. A summary of all the attempts used to optimize the expression and purification are detailed in **Figures 3.31 and 3.32**. Red boxes indicate that the parameter tested gave a positive result during the optimization studies.

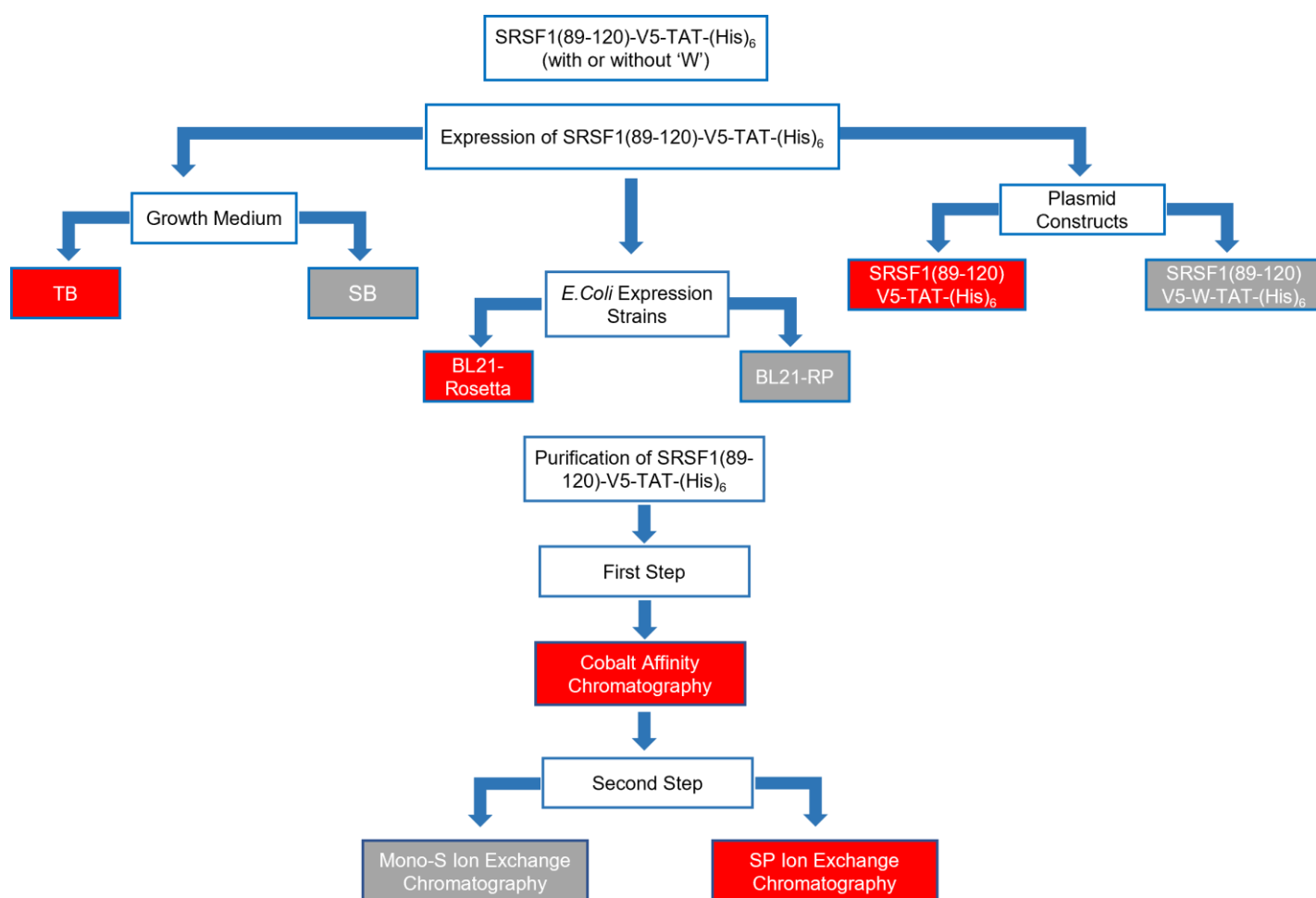


Figure 3.31. Schematic representation of the optimization strategies for expression and purification of SRSF1(89-120)-V5-TAT PTD-(His)₆ CPP (with or without tryptophan).

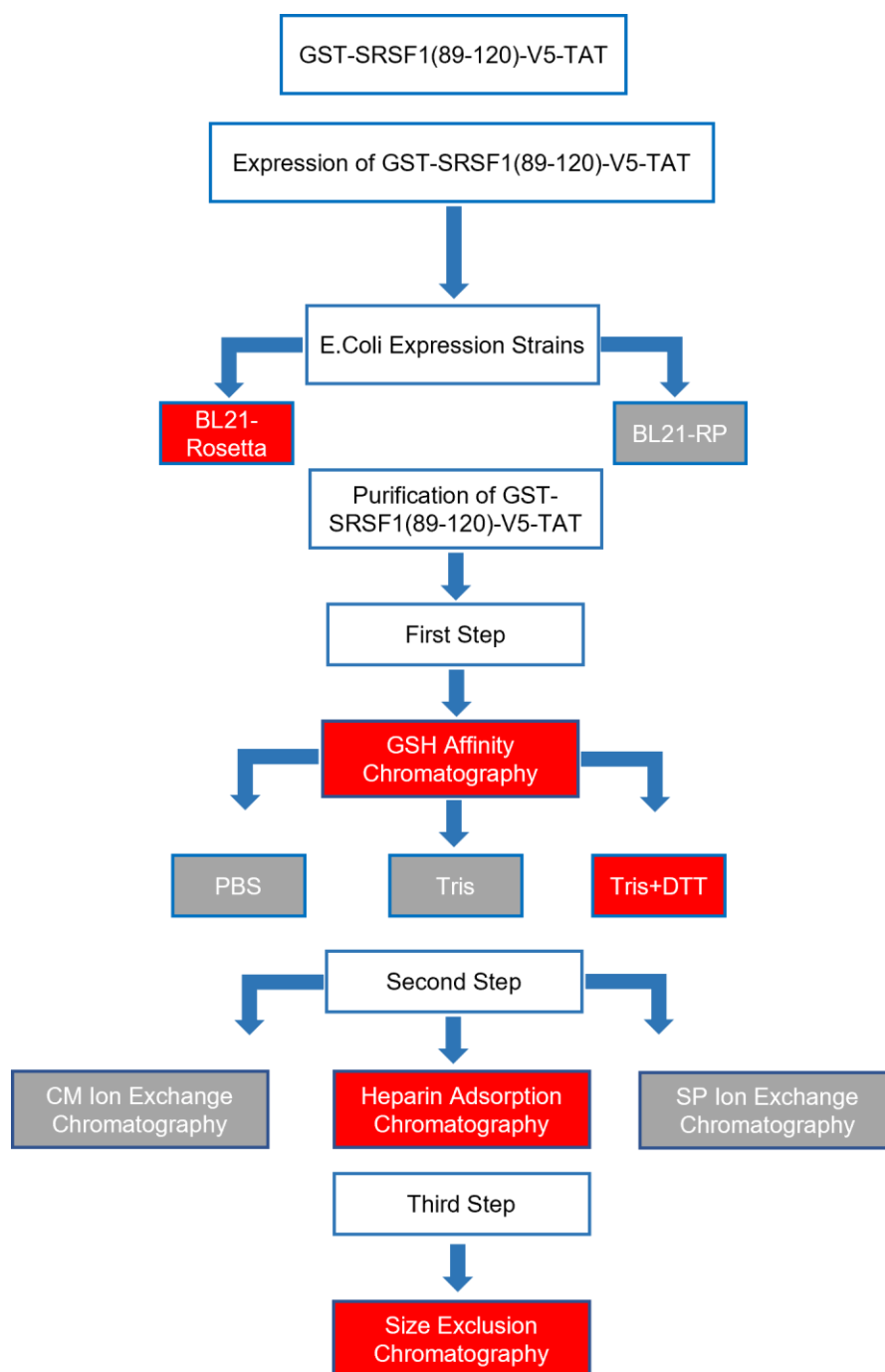


Figure 3.32. Schematic representation of the optimization strategies for expression and purification of GST SRSF1(89-120)-V5-TAT PTD.

In conclusion, we have developed an adequate and scalable protocol for the large-scale production of SRSF1 and Ctrl peptide using a GST fusion system in BL21- Rosetta *E. coli* cells. Further to this, an efficient and robust multi-step purification method, as shown in **Figure 3.33**, comprising GSH affinity, Heparin adsorption and Size Exclusion chromatography for the purpose of capture, intermediate and polishing was also developed to attain a very high degree of peptide purity and yield, homogenous and confirmed with an anti-V5 antibody. The purification step resulted in a yield of peptide in the milligram range (1 mM stock concentration of SRSF1 and Ctrl peptide). The method described in this chapter to produce SRSF1 and Ctrl peptide might be a valuable protocol to produce other proteins and peptides in *E. coli*.

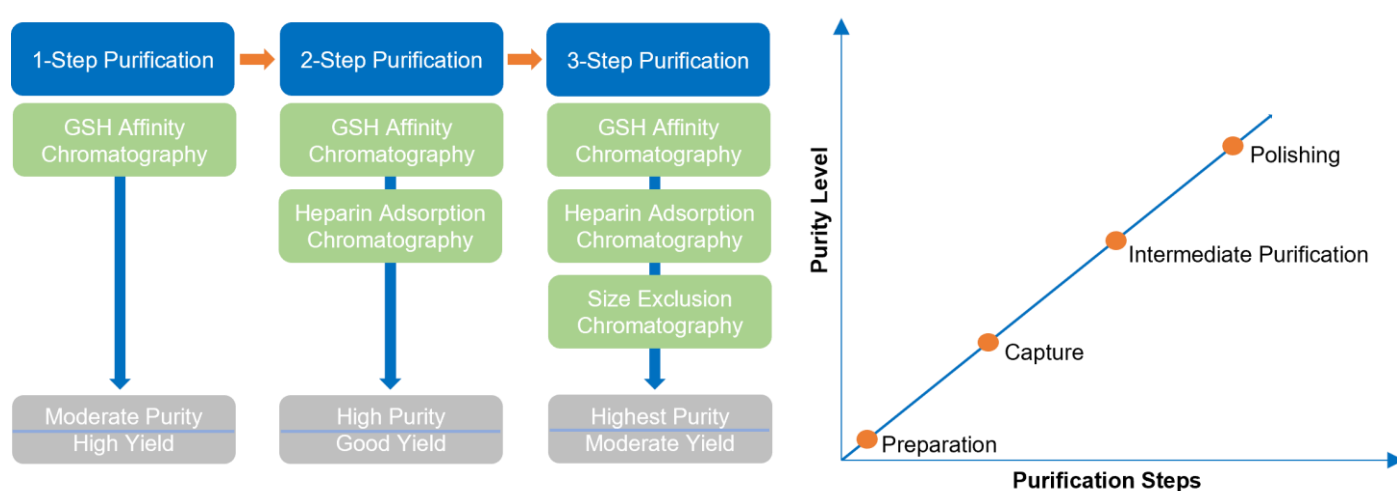


Figure 3.33. The optimised purification method for SRSF1 and Ctrl CPP. A three-step purification method involving GSH Affinity, Heparin Adsorption and Size Exclusion chromatography was developed.

4. Chapter 4 - Functional Characterization of Recombinant SRSF1-V5-TAT CPP in *In Vitro* and *In Vivo* models of C9ORF72-ALS/FTD

4.1. Introduction

The experiments presented in this chapter are an extended version of study that has been published during my PhD. Unless otherwise stated, I am the person who conducted all experiments. Any partners that contributed are acknowledged. The majority of this research was conducted using the synthetic SRSF1-CPP before I embarked on my PhD thesis. This peptide was custom-synthesized chemically and subsequently tested in reporter HEK293T cells, patient-derived neurons, and a *Drosophila* model of C9ORF72-ALS/FTD. However, During the revision stage of the paper, I performed additional experiments, notably the SRSF1:NXF1 pull-down assay and mouse study. These findings are elaborated upon in the second results chapter. As a result, data exclusively related to this synthetic peptide has been excluded from this thesis (Castelli et al., 2023).

In previous chapter, I developed a novel recombinant production method for SRSF1-CPP. One unique aspect of this protocol involves the use of urea, which denatures the peptide. Also, there is a notable difference between our recombinant peptides and their synthetic counterparts is the addition of a 5 amino acid sequence, GPLGS, to the amino terminal end of the recombinant peptide. This addition results from precision protease cleavage during the production process. Given this variation, as well as the denaturing conditions under which the recombinant peptide is produced, it is imperative to assess the biological activity of the recombinant variant. Thus, I have conducted experiments comparing the activity of the recombinant peptide against that of the synthetic version, which is produced under non-denaturing conditions. The goal is to determine if the recombinant peptide, despite its production differences, exhibits biological activity equivalent to the synthetic peptide. Another aim is also to examine whether this recombinant peptide can be employed as a tool to inhibit the nuclear export pathway in a HEK cells model (Hautbergue et al., 2017).

4.2. Aim and Objectives

The work described in this chapter investigates the functionality of recombinant SRSF1-linker CPP as a drug-like therapeutic agent in an *in vitro* and *in vivo* model of *C9ORF72*-ALS/FTD. Its functionality will be compared with a synthetic (chemically custom synthesized) peptide (Castelli et al., 2023). The fundamental aims of this chapter are as follows:

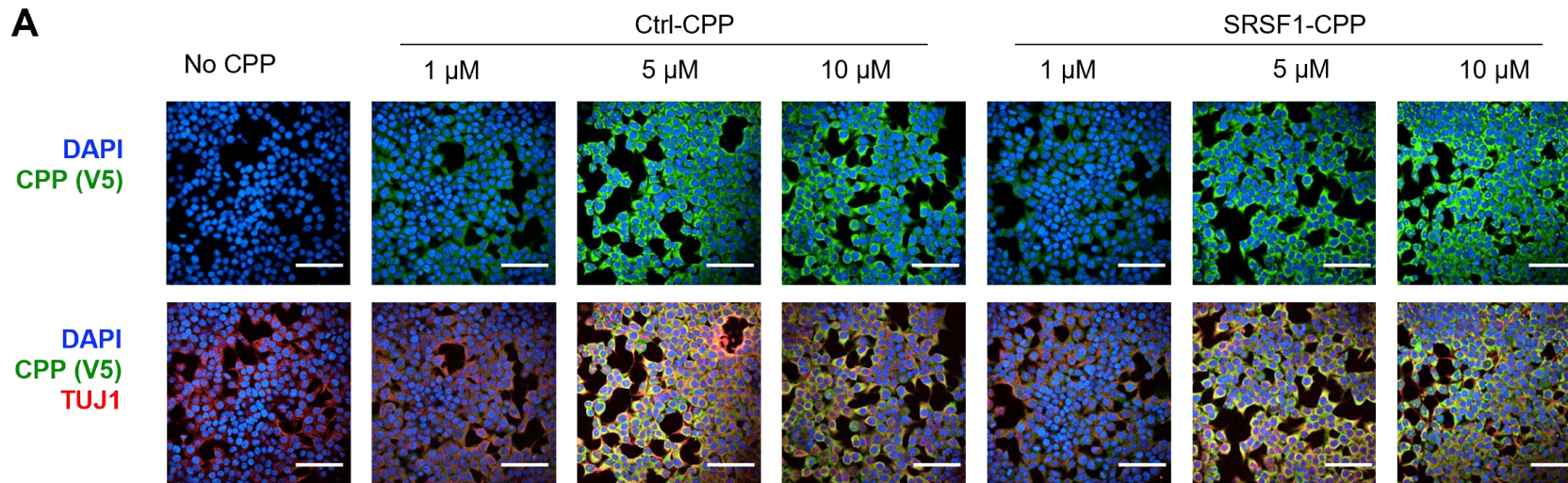
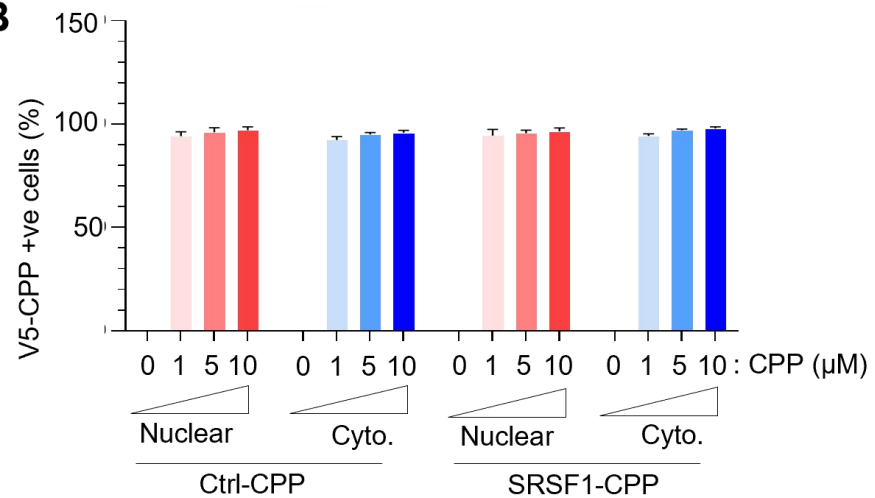
- To evaluate the ability of recombinant SRSF1-linker CPP to enter the nuclear and cytoplasmic compartments of HEK 293T cells.
- To determine whether recombinant SRSF1-linker CPP is able to bind directly to NXF1.
- To assess whether recombinant SRSF1-linker CPP inhibits the nuclear export of pathological sense and antisense repeat transcripts and concomitantly reduces sense and antisense DPR expression and their associated toxicity in HEK 293T cells.
- To test whether oral administration of recombinant SRSF1-linker CPP confers neuroprotection *in vivo* in a *C9ORF72*-ALS/FTD *Drosophila* model in collaboration with Dr Alexander Whitworth's group (University of Cambridge).
- To test whether a single cisterna magna injection of recombinant SRSF1-linker CPP leads to reduced DPR expression *in vivo* in a *C9ORF72*-ALS/FTD mouse model.

4.3. Results

4.3.1. Recombinant CPPs are delivered into the nuclear and cytoplasmic compartments

CPPs are short peptides fused to the protein transduction domain (PTD) of the HIV1 TAT protein which confers high membrane permeability and the capability for intracellular transportation of cargo. As such, CPPs have attracted significant attention in many research fields (Guidotti et al., 2017). This raises the question of whether the recombinant SRSF1 and Ctrl peptides can cross the plasma membrane. To confirm this, immunofluorescence studies were performed to investigate dose dependent cytoplasmic and nuclear delivery of peptides into HEK cells. HEK cells were treated with a range of SRSF1 and Ctrl peptide concentrations at 1 μ M, 5 μ M and 10 μ M prior to visualisation 72 hours post-treatment. The visualisation of cellular uptake of CPPs at lower magnification was undertaken by using a high content imaging system (PerkinElmer Opera PhenixTM) to maximize the number of cells within a field (**Figure 4.1.A**) and the acquired fluorescence signal at each concentration was quantified (**Figure 4.1.B**). The higher magnification images in separate channels were taken to obtain detailed images of CPPs within the nucleus and cytoplasm of individual HEK cells. (**Figure 4.1.C**). Panels A and C show the intracellular localization of SRSF1 and Ctrl-CPPs across the nucleus and cytoplasm of cells in comparison to untreated cells (0 μ M CPPs). Both peptides were successfully delivered into the nucleus and cytoplasm at each concentration, but the nuclear and cytoplasmic localization of peptide at 10 μ M was more pronounced than that observed with 1 μ M peptide. Overall, immunofluorescence images indicated that a TAT sequence is necessary for internalization. **Figure 4.1.B** shows statistical quantification of the delivery percentage for each CPP at different concentrations performed with three replications in HEK cells.

Untreated cells show no background signal from both SRSF1 and control peptides. Also, SRSF1 and Ctrl peptides remain predominantly in the nucleus and cytoplasm 72-hours post treatment; however, the accumulation of SRSF1 peptide in the nucleus was slightly higher than in the cytoplasm. Additionally, increasing concentrations of the SRSF1 and control peptide resulted in increasing signals as expected but quantification of SRSF1 and control peptide levels at 5 μ M and 10 μ M remained unchanged, which is consistent with the immunofluorescence images. Lastly, the SRSF1 and Ctrl peptides were detected and remained highly expressed in the nucleus and cytoplasm suggesting more than 90 % uptake efficiency after 72 hours of treatment, indicating that peptide is still functionally active.

A**B**

C

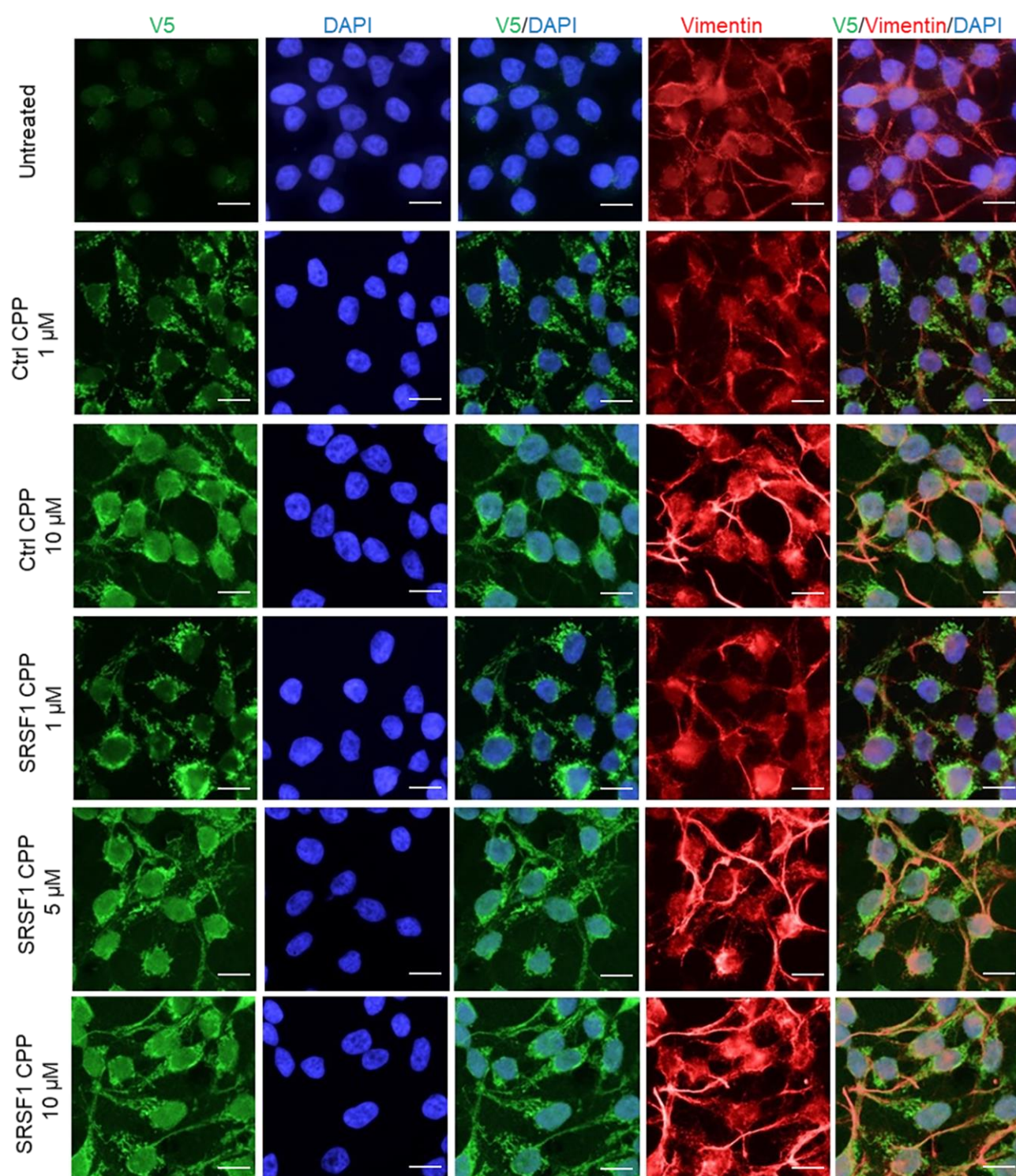


Figure 4.1. The recombinant CPPs are detected in the nucleus and cytoplasm of > 90 % of HEK cells. (A) Low magnification immunofluorescence microscopy of HEK cells cultured in media containing recombinant control (ctrl) or SRSF1-CPP for 72 hours. CPPs were detected in the green channel using anti-V5 and anti-mouse ALEXA488 antibodies. DAPI was used to delineate the nucleus, while anti-TUJ1 and anti-chicken ALEXA647 (far red) were used to delineate the cytoplasm. Scale bars 20 μ m. (B) Statistical assessment of the HEK cell counts visualised in (A) was performed from approximately 3000 cells cultured with each CPP per biological replicate (n=3, Bar charts are means of cell numbers $\pm$ SD or SEM; one-way ANOVA with Tukey post hoc's correction for multiple comparisons). (C) High magnification immunofluorescence microscopy of HEK cells cultured in media containing control or SRSF1-CPP for 72 hours. CPPs were detected in the green channel using anti-V5 and anti-mouse ALEXA488 antibodies. DAPI was used to delineate the nucleus, while Vimentin and anti-chicken ALEXA647 (far red) were used to delineate the cytoplasm. Scale bars 20 μ m.

4.3.2. Evaluation of Intracellular Distribution of Synthetic CPPs

HEK cells were treated with either 1 μ M or 10 μ M synthetic SRSF1 peptide alongside a Ctrl peptide to compare the efficiency of dose-dependent cytoplasmic and nuclear delivery. At 72 hours post peptide treatment, the cellular uptake of synthetic peptides was visualized (**Figure 4.2.A**) and acquired fluorescence signal of both peptides at each concentration was quantified using a high content imaging system (PerkinElmer Opera PhenixTM) (**Figure 4.2.B**). **Panel A** shows immunofluorescence images of the untreated cells (0 μ M), SRSF1 and Ctrl-CPPs at 10 μ M which indicates that SRSF1 and control peptide are delivered successfully and maintained in the nucleus and cytoplasm 72 hours post-treatment, as for the purified peptide. **Panel B** indicates that the fluorescence signal of the SRSF1 peptide in the nucleus appeared slightly higher than in the cytoplasm, but more than 90 % of both synthetic SRSF1 and control peptide were taken up by the cells. Increasing concentrations of the SRSF1 and control peptide resulted in increasing signals as expected, but quantification of SRSF1 and control peptide levels at 5 μ M and 10 μ M are close each other. Overall, these immunofluorescence results indicated that recombinant SRSF1 and Ctrl peptide exhibited similar levels of nuclear and cytoplasmic delivery compared to synthetic CPPs.

***These data were produced by Dr Lydia Castelli in the lab** and used to compare the uptake functionality of my recombinant cell permeable peptide purified from *E. coli*.

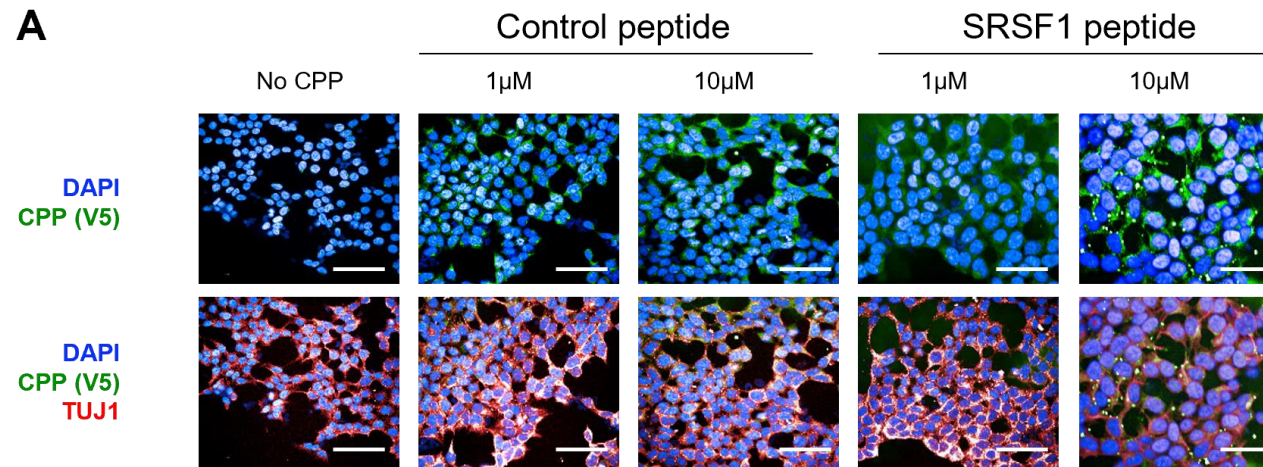
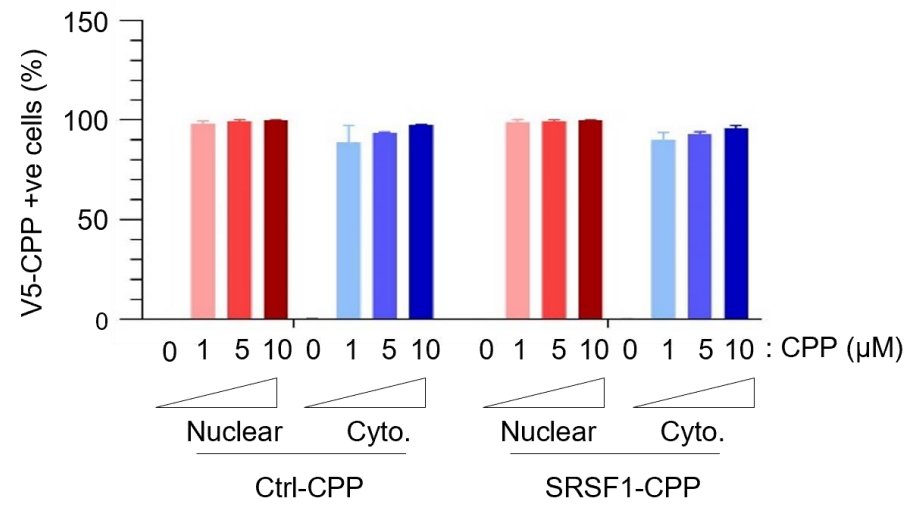
A**B**

Figure 4.2. The synthetic CPPs are taken up by > 90 % cells in concentrations which were detected in both the cytoplasm and nucleus of HEK cells. (A) Immunofluorescence microscopy of HEK cells cultured in media containing synthetic control (ctrl), SRSF1-CPP for 72 hours. CPPs were detected in the green channel using anti-V5 and anti-mouse ALEXA488 antibodies. DAPI was used to delineate the nucleus, while anti-TUJ1 and anti-chicken ALEXA647 (far red) were used to delineate the cytoplasm. Scale bar 20 μ m. (B) Statistical assessment of the HEK cell counts visualised in (A) was performed from approximately 3000 cells cultured with each CPP per biological replicate (n=3, Bar charts are means of cell numbers $\pm$ SD or SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons).

4.3.3. Investigating The Potential In Vitro Interaction of Recombinant SRSF1-CPP with Tap/NXF1

In order to check potential interaction of recombinant SRSF1-CPP with NXF1, I transfected HEK cells with flag-tagged NXF1 and flag-tagged ctrl plasmids. 48-hours post transfection, co-immunoprecipitation was performed using protein lysate which was incubated with pre-blocked anti-flag beads for immobilization of flag-tagged NXF1 protein to determine whether addition of SRSF1-CPP at different concentrations affects the interaction between NXF1 and SRSF1. Input on **Figure 4.3.A** and eluted proteins on **Figure 4.3.B** were analysed by western blotting, (**Figure 4.3.A and 4.3.B**). The western blot result of the input revealed that flag-immunoprecipitation was observed in all the lanes which contain the expressed Flag-NXF1 protein, but not with Flag Ctrl Lane. The Flag-IP in **Figure 4.3.B** showed that Flag-NXF1 was immobilized onto the anti-flag beads very well as seen by increased signal in comparison to the input image. It was also observed in SRSF1 co-IP western blot that the lanes in which cells were treated with recombinant SRSF1-CPPs at 1 μ M and 5 μ M co-immunoprecipitated with Flag-NXF1 thereby leading to inhibition of endogenous SRSF1 interaction with NXF1. This resulted in decreased signal in comparison to the input and ctrl-CPP lanes. This experiment shows that recombinant SRSF1-CPP interacts with NXF1 in RNA independent manner which is also consistent with my hypothesis. Moreover, the V5 western blot in **Figures 4.3.A and 4.3.B** show that the SRSF1 and Ctrl peptide were detected in eluates with V5 antibody indicating that both CPPs are structurally stable and functionally active during co-immunoprecipitation.

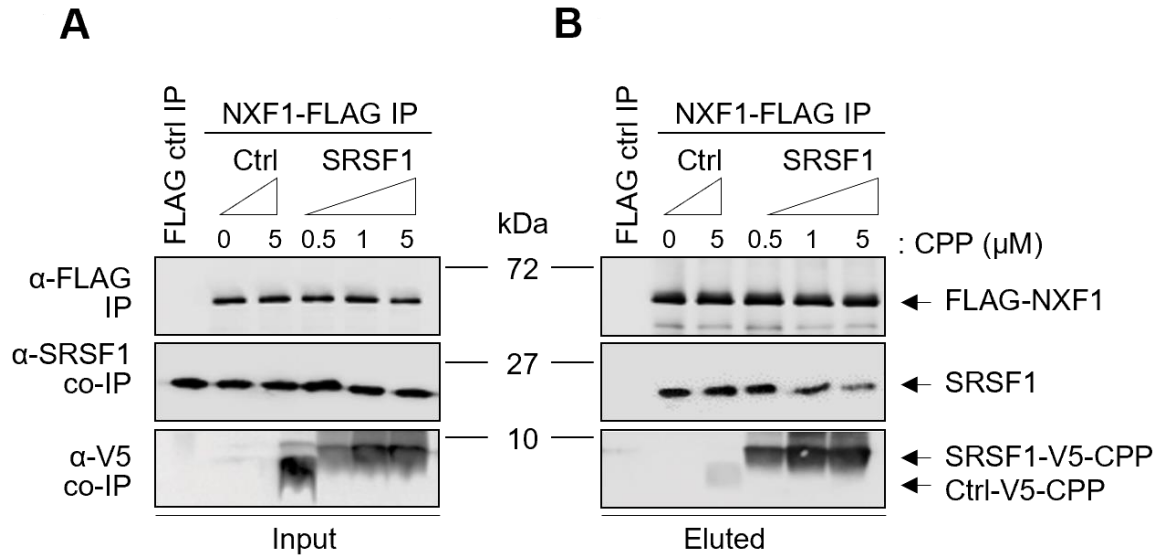


Figure 4.3. The recombinant SRSF1-CPP inhibits interaction of endogenous SRSF1 with NXF1 in human cells. HEK293T cells were transfected with either FLAG control or FLAG-NXF1 plasmids and treated with ctrl-CPP (5 μM) and SRSF1-CPPs (0.5, 1 and 5 μM) for 48 hours. Whole cell lysates from each condition were applied onto anti-FLAG beads and then the protein complex was eluted. Input **(A)** and eluates **(B)** were subjected to western blot by running on 12 % and 20 % SDS-PAGE gel. Blots were probed with Flag, SRSF1 and V5 antibodies to analyse level of NXF1, SRSF1 and ctrl/SRSF1-CPPs. The left panel shows the western blot of input. The right panel indicates western blot of eluted samples. Molecular weight marker bands are indicated in kDa. The western blot images presented in this Figure are from one biological repeat.

4.3.4. Investigating The Potential Direct Interactions of Recombinant SRSF1-CPP with TAP/NXF1, NXF2 and NXF3

I next performed a GST pull-down assay to assess direct interaction of recombinant SRSF1-CPP with Tap/NXF1 protein. GST tagged Tap/NXF1 recombinant protein expressed in *E. coli* was purified in high salt and incubated with GSH beads. Increasing concentrations of SRSF1 (0.125, 0.25, 0.5, 1, 2, 4, 8 μ M) and Ctrl peptide (8 μ M) were incubated with immobilized Tap/NXF1 protein on the beads prior to washing and elution. The reaction was treated with RNase to evaluate direct protein: protein interaction and eliminate false positive interaction caused by RNA molecules. The eluted protein samples were analysed via western blotting and Coomassie blue staining (**Figure 4.4.A**). Recombinant GST-NXF1 p15 protein was present in equal amounts in all lanes treated with either Ctrl or SRSF1 peptide (**Figure 4.4.A**). The Coomassie-stained gel reveals that GST (negative control) and GST-NXF1 p15 proteins immobilized on the GSH beads successfully (**Figure 4.4.A**). Immobilization of GST-NXF1 protein on the beads leads to specific pull down of the SRSF1 V5 tagged CPP showing that increasing concentration of SRSF1 peptide enhanced the affinity of interaction with GST-NXF1 not with GST. In addition, the pull-down assay of Ctrl-CPP proved the specificity of SRSF1-CPP and NXF1 interaction. Ctrl-CPP at 8 μ M was incubated with lysate expressing GST-NXF1 protein showing that the binding of Ctrl peptide at maximum concentration (8 μ M) with GST-NXF1 did not appear on both Coomassie stained gel and V5 western blot.

NXF1 belongs to a family of three cellular proteins, the others being NXF2 and NXF3. Therefore, I investigated the binding potential of SRSF1-CPP to NXF2 and NXF3, two proteins with tissue-restricted expression (testis mainly). To perform this, GST-NXF2 and GST-NXF3 proteins were expressed in *E. coli* and purified in high salt. Increasing concentrations of SRSF1 and Ctrl peptide were incubated with immobilized Tap/NXF2 and Tap/NXF3 proteins. A pull-down assay was performed using immobilized GST-NXF2 and GST-NXF3 proteins and the eluates were analysed by Coomassie blue staining and western blotting (**Figures 4.4.B and 4.4.C**). The results suggest that SRSF1-CPP does not interact with GST-NXF2 or GST-NXF3, indicating the NXF1 specificity of SRSF1-CPP.

Next, we tested the expression of NXF1, NXF2 and NXF3 in HEK cells, patient-derived neurons, HEK cells mouse brain and testis tissues. The NXF1 antibody specifically recognises NXF1, however the other commercial NXF2 and NXF3 antibodies appeared cross-reactive with the other purified recombinant NXF proteins (**Figure 4.4.D**). Therefore, we extracted RNA from NXF1, NXF2 and NXF3 protein lysates and used qRT-PCR to investigate NXF1, NXF2 and NXF3 mRNA expression levels (**Figure 4.4.E**). The results show expression of NXF2 and NXF3 specifically in mouse testis but not in mouse brain, HEK cells or patient-derived neurons. Overall, these data further validate that SRSF1-CPP specifically targets NXF1 in the human and mouse models used in this project.

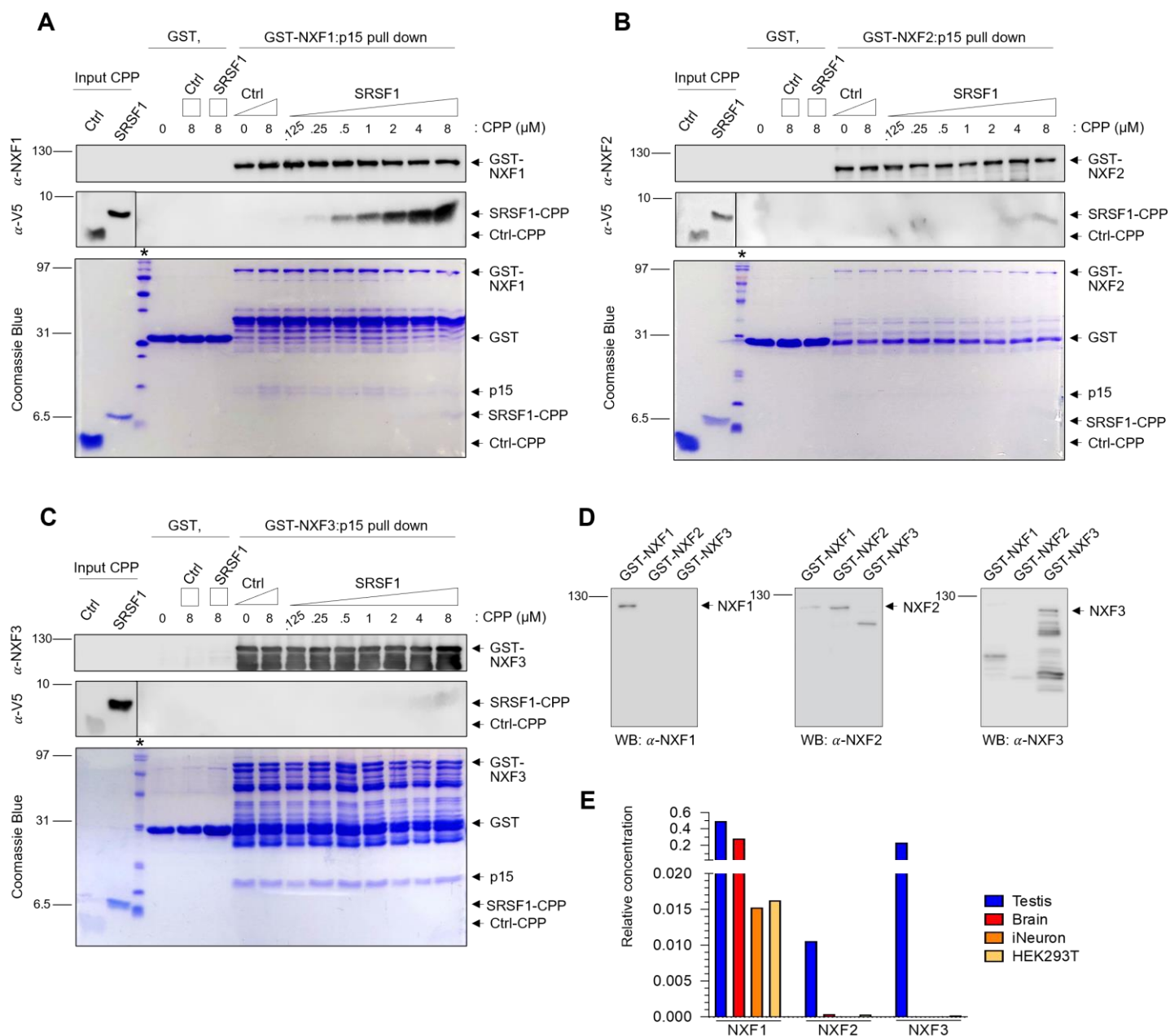


Figure 4.4. The recombinant SRSF1-CPP specifically interacts with TAP/NXF1: p15 in a dose dependent manner, but not with NXF2 and NXF3 which are expressed in only mouse testis. (A-C) GST, GST-TAP/NXF1: p15, GST-TAP/NXF2: p15, GST-TAP/NXF3: p15 were expressed in E. coli and immobilized onto GSH beads. Increasing concentration of Ctrl and SRSF1-CPPs were incubated with immobilized GST, GST-NXF1 **(A)**, NXF2 **(B)** and NXF3 **(C)**. GST was used as negative control. Binding reactions were analysed by Coomassie stained 15 % SDS-PAGE gel and western blot. Blots were incubated with anti-V5 and anti-NXF1, NXF2, NXF3 antibodies. **(D)** Purified GST-NXF1/NXF2/NXF3:p15 were analysed with anti-NXF1, NXF2, NXF3 antibodies. The expression of NXF1, NXF2 and NXF3 mRNAs was investigated in mouse testis (positive control), mouse cortex homogenate, human- derived neurons and HEK293T cells using qRT-PCR. The chart shows the relative concentrations of NXF1-3 mRNAs normalized to U1 snRNAs. Note that primers are different for mouse and human samples, and direct comparisons cannot be extrapolated. Molecular weight marker bands are indicated in kDa. The western blot and Coomassie gel images presented in this Figure are from one biological repeat. **This figure is published in the Journal of Science Translational Medicine where I am co-author in that publication (Castelli et al., 2023). I performed the experiments on Panel A, B, C and D.**

4.3.5. Recombinant SRSF1-CPP Inhibits SRSF1:NXF1 Dependent Nuclear Export of Sense and Antisense-*C9orf72*-Repeat RNA

Binding of the recombinant SRSF1-CPP to NXF1 is hypothesised to inhibit the nuclear export of sense and antisense repeat transcripts. Therefore, I investigated whether pathological repeat transcripts in the nucleus and cytoplasm were altered upon treatment with recombinant SRSF1-CPP. To determine the effect of the peptide on SRSF1:NXF1-dependent nuclear escape of pathological sense and antisense *C9orf72* repeat transcripts in human cell models of *C9ORF72*-ALS/FTD, HEK cells were transfected with either sense or antisense repeat transcripts (pcDNA3.1/(G4C2)₄₅-3xV5, pcDNA3.1/(G2C4)₄₃-3xV5) prior to treatment with SRSF1 and Ctrl-CPPs for 72 hours. Cellular fractionation was performed using hypotonic lysis buffer to separate the cytoplasm from the nuclei. The integrity of the nuclear and cytoplasmic fractions was analysed by western blot (**Figure 4.5.A and Figure 4.5.B**) to confirm that the fractionation has worked by using the nuclear modelling factor SSRP1 as a nuclear marker and Tuj 1 as a cytoplasmic marker. The results (**Figures 4.5.A and 4.5.B**) show that nuclear SSRP1 is not present or minimal in cytoplasmic fraction as expected and minimal cytoplasmic Tuj 1 is present in the nuclear fraction indicating that successful fractionation has been performed with absence of nuclear contamination. After confirming the integrity of the nuclear and cytoplasmic fractions, RNA was extracted from the fractions in independent triplicates and converted to cDNA. The levels of pathological *C9orf72* repeat transcripts were quantified by qRT-PCR. The Ct (cycle threshold) values for each condition were normalized against U1 snRNA and the relative percentage expression was calculated for sense and antisense repeat transcripts (**Figures 4.5.C and 4.5.D**). No difference was observed between total, nuclear and cytoplasmic fractions that had been treated with ctrl peptide. Indeed, the total level remained the same after the addition of either Ctrl or SRSF1 peptide. However, upon addition of SRSF1 peptide at 1 μ M, a nuclear retention of pathological *C9orf72* sense (**Figure 4.5.C**) and antisense (**Figure 4.5.D**) repeat transcripts was observed. This corresponded with a significant reduction in the cytoplasmic content of the pathological transcript which fits the model that SRSF1 cell permeable peptide is interacting with NXF1 which prevents pathological repeat bound SRSF1 from interacting and being exported in the cytoplasm and which is therefore retained in nucleus. This indicates the potential that the SRSF1-CPP may confer a cryoprotective effect in the *C9ORF72*-ALS/FTD cell model.

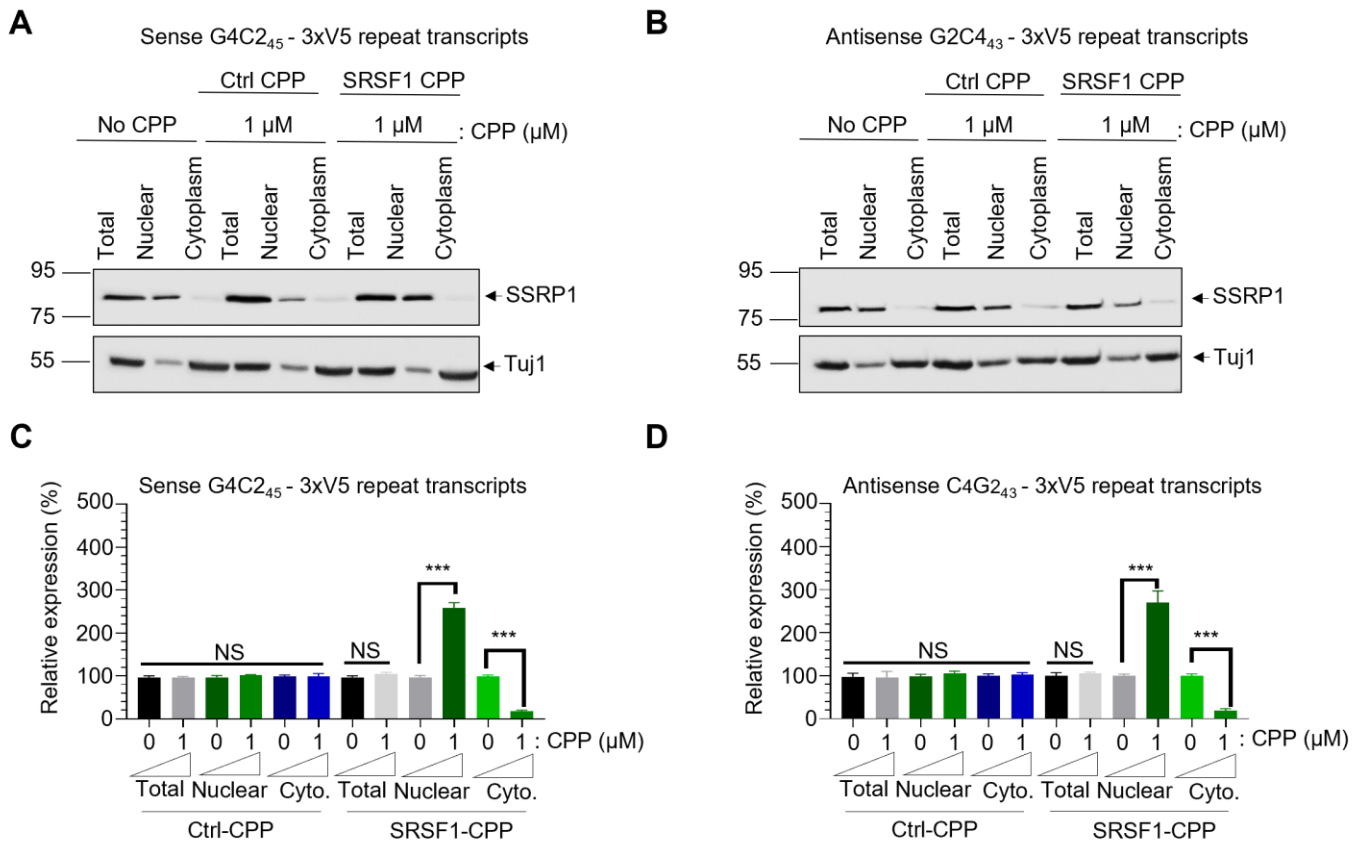


Figure 4.5. The recombinant SRSF1-CPP inhibits nuclear export of sense and antisense C9orf72-repeat transcripts. HEK293T cells transfected with either no DPR control or sense G4C₂₄₅-3xV5 or antisense G2C₄₄₃-3xV5 plasmids. After 6 hours post transfection, the media was changed with media containing of Ctrl or SRSF1-CPPs. The fractionation experiment was performed at 72-hour post transfection to isolate the total, nuclear and cytoplasmic compartments of the cell. **(A)** Representative Western blot images of total, nuclear and cytoplasmic fractions extracted from HEK293T cells transfected with either no DPR control or sense (G_{4C})₄₃-3xV5 plasmids and treated with Ctrl and SRSF1-CPPs. Blots were probed with nuclear modelling factor SSRP1 and cytoplasmic TUJ1 antibodies. **(B)** Representative western blot images of total, nuclear and cytoplasmic fractions extracted from HEK293T cells transfected with either no DPR control or antisense (G_{2C})₄₃-3xV5 plasmids and treated with Ctrl and SRSF1-CPPs. Blots were probed with nuclear modelling factor SSRP1 and cytoplasmic TUJ1 antibodies. **(C)** The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. Total, nuclear and cytoplasmic mRNA levels of sense G4C₂₄₅-3xV5 repeat transcripts were quantified with qRT-PCR followed by normalization to U1snRNA levels and to 100% in untreated cells (n=3, mean ± SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001). **(D)** The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. Total, nuclear and cytoplasmic mRNA levels of sense G2C₄₄₃-3xV5 repeat transcripts were quantified with qRT-PCR followed by normalization to U1snRNA levels and to 100 % in untreated cells (n=3, mean ± SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons, NS = Non-significant, p ≥ 0.05; * = p<0.05; ** = p<0.01; *** = p<0.001; **** = p<0.0001. Molecular weight marker bands are indicated in kDa.

4.3.6. Recombinant SRSF1-CPP Inhibits Dose-dependent Sense (poly-GA, GP, GR) and Antisense (poly-GP, PR, PA) DPR Expression

HEK 293T cells transfected with either V5-tagged sense or antisense RAN translation reporter plasmids, were treated with increasing concentrations (0.0625, 0.125, 0.25, 0.5 and 1 μ M) of either recombinant SRSF1, synthetic SRSF1 or control CPP (2 μ M) for 72 hours to compare their effect on DPR production. A V5 antibody was used to detect the expression of DPRs in all three reading frames. (**Figure 4.6.A and Figure 4.6.B**). No DPR product was detected in the no DPR control extracts, confirming V5 antibody specificity. As expected, DPRs from the sense and antisense repeat transcripts were detected in the lanes where cells were treated with Ctrl peptide at 0.5 and 1 μ M, to an equal level as with the DPR control cells, indicating no effect on DPR production (**Figure 4.6.A and Figure 4.6.B**). However, an increasing concentration of the SRSF1 peptide, whether recombinant or synthetic, gradually reduced sense DPR production, with significant reduction observed at 0.5 and 1 μ M concentrations in comparison to the control peptide and DPR control cell extract (**Figure 4.6.C**). This indicates that the SRSF1 peptide is able to enter the cells and inhibit the NXF1:SRSF1 interaction which leads to an inhibition of RAN translation of sense DPRs in dose-dependent manner. The progressive reduction in DPR expression was also observed when cells were transfected with antisense repeat transcripts and treated with recombinant or synthetic SRSF1 peptide at increasing concentrations (**Figure 4.6.B**). As observed in the case of sense DPRs, the Ctrl peptide has no inhibitory effect on antisense DPR production as they are still detectable with the V5 antibody. The sense and antisense DPR expression levels from HEK cells were repeated in triplicate and quantified using GeneSYS software. The quantification data show that SRSF1-CPPs lead to similar inhibitory dose-responses for sense and antisense DPRs with IC₅₀ concentrations of approximately 0.5 μ M (**Figure 4.6.C and Figure 4.6.D**).

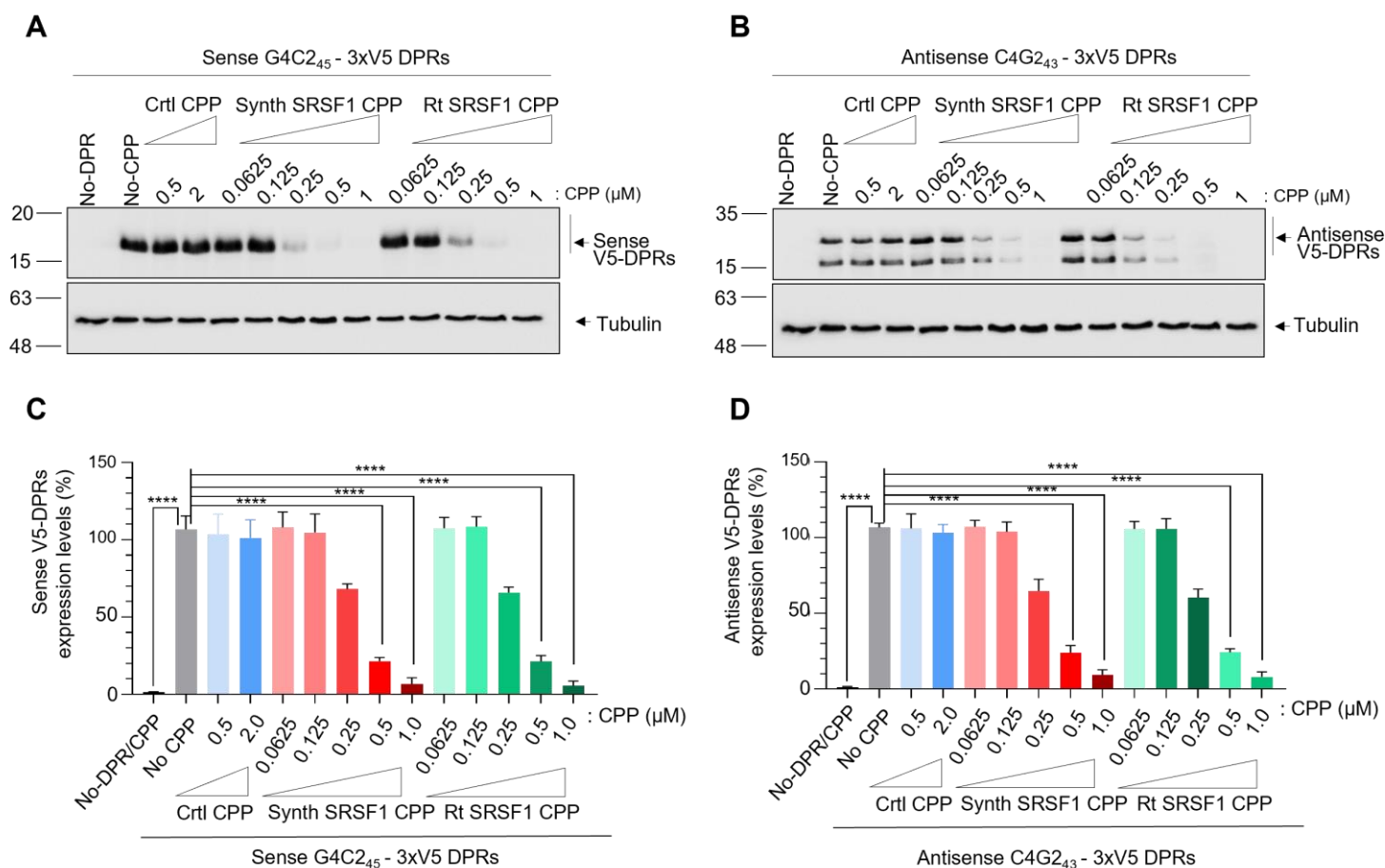


Figure 4.6. The recombinant and synthetic SRSF1-CPP inhibits expression of sense antisense DPRs in dose dependent manner. (A) Representative western blot images from HEK cells transfected with no DPR control or G4C2_{x38}-V5 for 6 hours and then media changed with media containing increasing concentrations of synthetic and recombinant control and SRSF1 cell-permeable peptides. Blots are probed for V5 and tubulin to show peptide uptake by the cells. (B) Representative western blot images from HEK cells transfected with no DPR control or C4G2_{x39}-V5 for 6 hours and then media changed with media containing increasing concentrations of synthetic and recombinant control and SRSF1 cell-permeable peptides. Blots are probed for V5 and tubulin to show peptide uptake by the cells. (C) The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. The quantification of synthetic and recombinant SRSF1/control peptide transfected with G4C2_{x38}-V5 into HEK cells were analysed relative to α -tubulin loading (n=3, mean $\pm$ SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001). (D) The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. The quantification of synthetic and recombinant SRSF1/control peptide transfected with C4G2_{x39}-V5 into HEK cells were analysed relative to α -tubulin loading (n=3, mean $\pm$ SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001). Molecular weight marker bands are indicated in kDa.

4.3.7. Recombinant SRSF1-CPP Inhibits Dose-dependent Reduction of Sense (poly-GA, GP, GR) and Antisense (poly-GP, PR, PA) DPR-associated Cytotoxicity

Nuclear export and DPR assays indicated that dose-dependent recombinant SRSF1 peptide is able to inhibit SRSF1:NXF1 dependent nuclear export of sense and antisense pathological repeat transcripts and reduce DPR expression. Therefore, I wanted to evaluate whether the recombinant SRSF1 peptide would rescue the cellular toxicity mediated by the RAN translation of DPRs using MTT assays. MTT assays are a rapid colorimetric method of measuring cell proliferation based on the enzymatic activity of mitochondria. The same number of cells were seeded in plates at the beginning of the assay prior to transfection with either the sense or antisense repeat transcript constructs and addition of either recombinant or synthetic SRSF1 at increasing concentrations (0.0625, 0.125, 0.25, 0.5, and 1 μ M) and Ctrl-CPPs (2 μ M) after 5 hours post-transfection. The effect of both CPPs on cell proliferation was evaluated by dividing averaged results of No CPP and both CPPs at each concentration by averaged results from No DPR/CPP.

Both the sense and antisense transcripts induced cellular toxicity in comparison to un-transfected cells which showed a significant decrease in cell proliferation by $\sim 80\%$, though no cell death was observed (**Figures 4.7.A and 4.7.B**). Treatment with the synthetic or recombinant Ctrl-CPPs at increasing concentrations for 72 hours resulted in no observable effect on the cellular toxicity compared to the No CPP lane where DPRs were expressed from either the G4C2₄₅-3xV5 or G2C4₄₃-3xV5 repeat transcript constructs (**Figures 4.7.A and 4.7.B**). Similar to the Ctrl peptide, synthetic and recombinant SRSF1-CPPs at 0.0625 and 0.125 μ M showed no significant reduction in cell viability, aligning well with the observed western blot data (**Figures 4.6.A and Figures 4.6.B**). However, the cell proliferation level of transfected cells (sense or antisense repeat transcripts) resulted in a $\sim 3 - 4$ fold increase upon addition of synthetic and recombinant SRSF1-CPPs at 0.25, 0.5 and 1 μ M, again supporting the western blot data (**Figures 4.6.A and Figures 4.6.B**) and suggesting a dose-dependent rescue of cytotoxicity.

4.3.8. Therapeutic Efficiency of SRSF1-CPP in an Animal Model of C9ORF72-ALS/FTD

4.3.8.1. Investigation of Oral Delivery of Synthetic SRSF1-CPP into C9ORF72-ALS Drosophila Model on Poly (GA), (GP), (GR) DPRs Expression and Locomotor Deficits

The therapeutic efficacy of a non-viral gene therapy approach using synthetic SRSF1-CPP has been investigated in a C9ORF72-ALS *Drosophila* model by our collaborators (Dr. Alvaro Sanchez-Martinez in Dr. Alexander Whitworth's group at the University of Cambridge). 10 µl of synthetic SRSF1-CPP at 1 mM was administered orally to larvae and adult *Drosophila*. This resulted in a significant decrease in the cytoplasmic level of G₄C₂x36 repeat transcripts and a corresponding increase in nuclear repeat transcripts upon addition of SRSF1 peptide into the larvae's food. Administration of the ctrl peptide showed no reduction or increase in total, nuclear and cytoplasmic fractions (**Figure 4.8.A**).

Crawling and climbing assays were used to assess the effect of SRSF1-CPP treatment on motor function in C9ORF72-ALS larvae and adult (2 days old) *Drosophila*, respectively. Treatment with SRSF1-CPP for 5 days (larvae) and 10-days (for adult *Drosophila*) resulted in a reduction in motor deficits compared to non-transgenic groups (**Figure 4.8.B** and **Figure 4.8.C**). Subsequently, the effect of SRSF1-CPP treatment on DPR inhibition was investigated using dot-blot assays. The assay showed that 10 µM of SRSF1-CPP treatment for 10 days reduces poly GA, GP, and GR DPR production in adult *Drosophila* compared to non-transgenic groups (**Figure 4.8.D**). We were unable to validate intracellular delivery of the CPPs in *Drosophila* brains using V5 immunofluorescence microscopy or dot blot analysis. This was likely due to a very small proportion of CPP uptake via the oral administration route. However, using dot blots, we were able to confirm that Ctrl- and SRSF1-CPPs are delivered to the heads of adult flies treated with 60 µM CPPs (**Figure 4.8.E**). Overall, these results align with *in vitro* data generated with both synthetic and recombinant SRSF1-CPP. The SRSF1-CPP showed promise as a therapeutic agent in a *Drosophila* model of C9ORF72-ALS/FTD; therefore, I wanted to test the therapeutic proof of principle efficacy of synthetic and recombinant SRSF1-CPP on C9ORF72-ALS mice (Castelli et al., 2023).

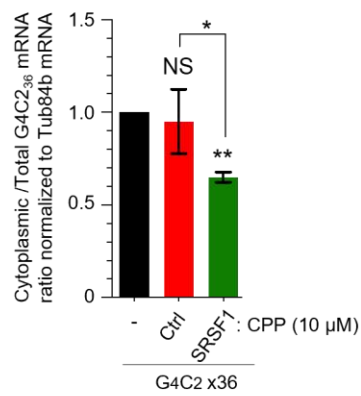
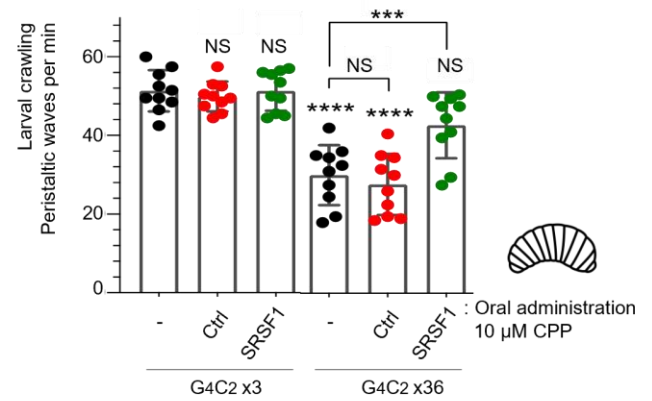
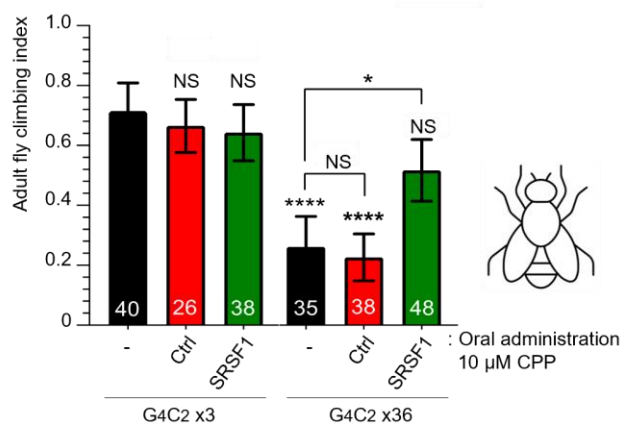
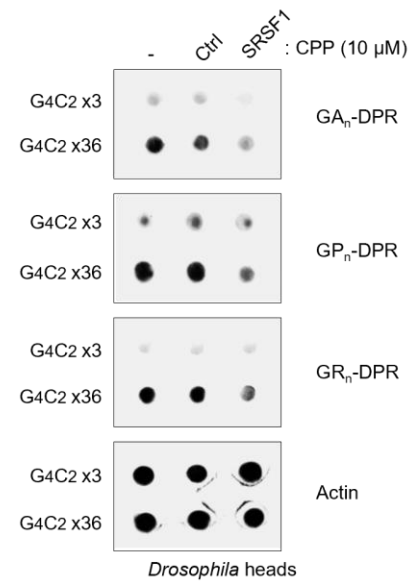
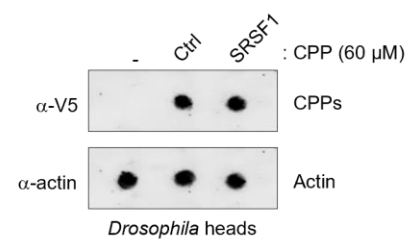
A**B****C****D****E**

Figure 4.8. The synthetic SRSF1-CPP inhibits RAN-translation of DPRs and reduce motor deficits in vivo in C9ORF72-ALS/FTD *Drosophila*. 10 µl of Ctrl-CPP and SRSF1-CPP at 1 mM were orally treated to *Drosophila* model of C9ORF72-ALS/FTD **(A)** The quantification of total and cytoplasmic expression level of sense pathogenic repeat transcripts from non-transgenic and transgenic CPP-treated *larvae* was conducted by qRT-PCR. The experiment was repeated biologically 3 times and normalized against Tub84b mRNA expression level. Bar chart shows mRNA levels (n=3, mean ± SEM; one-way ANOVA with Tukey post hoc's correction for multiple comparisons; NS: non-significant; **: p<0.01, ***: p<0.001;). **(B)** Crawling assays of non-transgenic and CPP-treated *larvae* expressing G4C2₃₆ pathogenic repeat expansions. Numbers of animals used in the assays are indicated at the bottom of each bar. (n=3, mean ± SEM; one-way ANOVA with Tukey post hoc's correction for multiple comparisons; NS: non-significant; **: p<0.01, ***: p<0.001;). **(C)** Climbing assay of healthy and transgenic CPP-treated adult *Drosophila*. Numbers of animals used in the assays are indicated at the bottom of each bar. (n=3, mean ± SEM; one-way ANOVA with Tukey post hoc's correction for multiple comparisons; NS: non-significant; **: p<0.01, ***: p<0.001). **(D)** Dot blot analysis of poly-GA, GP and GR DPRs in non-transgenic and transgenic C9ORF72-ALS/FTD adult *Drosophila*. Actin was used as loading control. **(G)** Dot blot assays from adult heads of *Drosophila* fed with food supplemented or not with 60 µM of Ctrl- or SRSF1-CPP for 10 days. Anti-V5 and actin antibodies were respectively used to probe for V5-tagged CPPs and as a loading control. **This figure, which shows the results of experiment on Panel A, B, C, D and E, is published in the Journal of Science Translational Medicine, where I am co-author in that publication (Castelli et al., 2023). The experiments were performed by my collaborators at the University of Cambridge.**

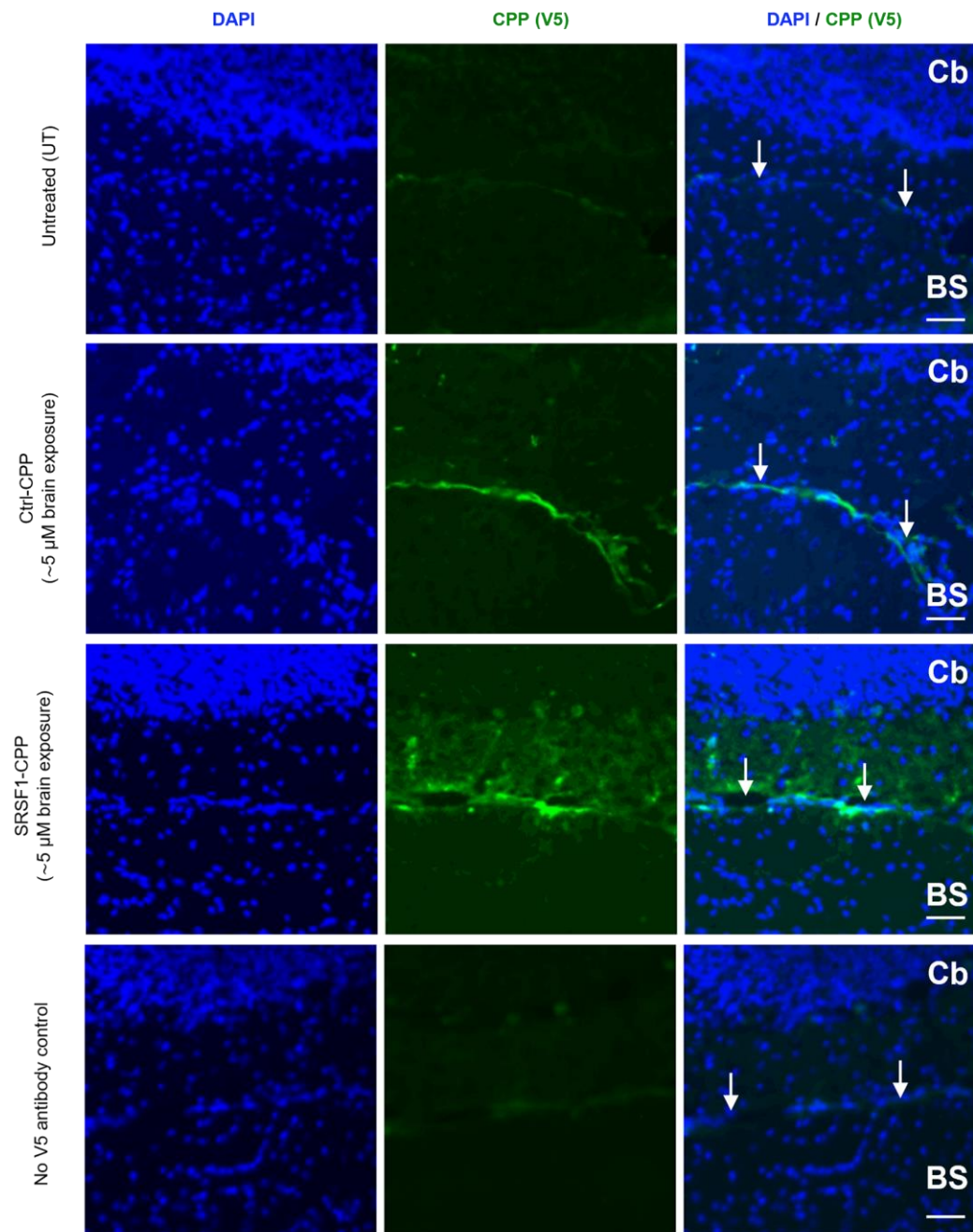
4.3.8.2. Investigation of Cisterna Magna Delivery of Recombinant and Synthetic SRSF1-CPPs Into The *C9ORF72*-ALS Mouse Model Upon Levels of Poly (GP) DPR Expression

The *C9ORF72*-ALS/FTD BAC 500 transgenic mouse model was used for the efficacy study. 5 µl of synthetic (custom chemically synthesised) or recombinant SRSF1 (produced by *E.coli*) and Ctrl-CPPs at 1 mM, which corresponds to 5 µM brain exposure, was administered via the cisterna magna injection at a rate of 1 µl per minute to female transgenic mice at postnatal day P30. Transgenic and non-transgenic mice in the control group did not receive anything. The mice were monitored daily to ensure their overall conditions were stable until the end of the study. 3 out of 5 transgenic mice receiving recombinant SRSF1-CPP died either during recovery or after recovery; therefore, the recombinant SRSF1-CPP injection was terminated. Detailed information about the deaths is presented in **Figure 4.9.D**.

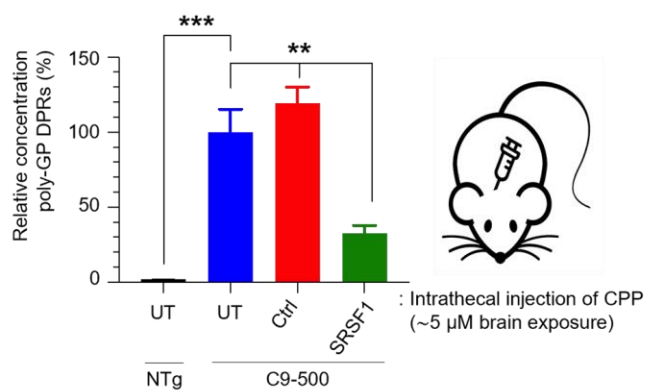
Immunohistochemistry and MSD-ELISA results presented in this chapter were performed using tissue obtained from mice injected with synthetic SRSF1-CPP. Tissue from control and treatment groups was harvested 4 days post injection. 1 out of 5 mice in each group was used to assess the localization peptides across the brain by immunohistochemistry (**Figure 4.9.A**); the remaining mice were used in the MSD-ELISA assay (**Figure 4.9.B**) to quantify DPR levels. Immunofluorescence results show a gradient uptake of SRSF1 and Ctrl-CPPs in the cerebellum from cells neighbouring the cerebrospinal fluid (**Figure 4.9.A**). It appears that both CPPs were taken up by Purkinje neuron cells. Cerebellum/brain stem homogenates were also used to generate the data showing reduced DPR expression upon injection of SRSF1 -CPP 4 days post administration.

To provide a proof-of-concept of whether the SRSF1-CPP is functional in a *C9ORF72*-ALS/FTD mammalian model, single intrathecal injection of CPPs into the cisterna magna was performed. The MSD assay showed that the synthetic SRSF1-CPP was functional and resulted in a 70 % reduction of poly GP DPRs through a single direct CSF delivery in comparison to untreated or Ctrl-CPP-treated C9-500 transgenic mice (**Figure 4.9.B**). The weights of the brain and peripheral organs of the CPP-treated and untreated animals were taken to determine whether the CPP injection resulted in immediate weight loss or any other major side effects. No adverse effects were observed upon injection of synthetic and ctrl-CPPs (**Figure 4.9.C**).

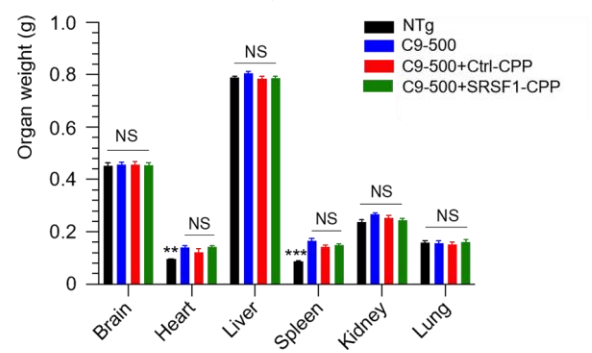
A



B



C



D

SEX	Genotype	Pup's Number	Treatment	Injection Record
Female	Transgenic	307123	Rt SRSF1 CPP	Died during recovery, around 30 minutes
Female	Transgenic	307123	Rt SRSF1 CPP	Good recovery
Female	Transgenic	307461	Rt SRSF1 CPP	Recovered but died after 24 hours post injection
Female	Transgenic	307461	Rt SRSF1 CPP	Good recovery
Female	Transgenic	307461	Rt SRSF1 CPP	Died during recovery, around 30 minutes

Figure 4.9. The synthetic SRSF1-CPP inhibits RAN-translation of DPRs in vivo in C9ORF72-ALS/FTD mouse brains. 5 µl of Ctrl-CPP and SRSF1-CPP at 1 mM (~5 µM brain exposure) were injected into the cerebrospinal fluid (CSF) via cisterna magna injections in nontransgenic (NTg) and transgenic C9ORF72-ALS/FTD (C9-500) mice at post-natal day 30. Animals were sacrificed 4 days post-delivery. **(A)** Representative immunofluorescence microscopy images of untreated (UT) and CPP-injected C9-500 cerebellum/brainstem stained with DAPI and V5 antibody to detect V5-tagged CPPs. The no V5 antibody control shows that the staining signal is specific of the V5 antibody. Ctrl-CPP exhibits a more diffuse pattern as expected with an inactive peptide which does not interact with any molecular target and is prone to diffusion and quick turnover. Arrows delineate the CSF space between the cerebellum (Cb) and the brain stem (BS). Scale bar: 50 µm. In total, 3 images from brain stem and 3 images from Cerebellum from 1 mouse. **(B)** Meso Scale Discovery (MSD)-ELISA quantification of poly-GP DPRs in non-transgenic (NTg) and transgenic C9ORF72-ALS/FTD (C9-500) mice which received intrathecal injections or not (-) with Ctrl-CPP and SRSF1-CPP at post-natal day 30. Animals were sacrificed 4 days post-delivery. Bar chart indicates poly-GP DPRs in brain tissue homogenates (mean ± SEM; one-way ANOVA with Tukey post hoc's correction for multiple comparisons; **: p<0.01, ***: p<0.001; N=3 mice/group). Poly-GP DPRs were quantified against a standard curve established with a GPx7 peptide and levels normalized to 100% for the untreated C9-500 mice. **(C)** Weights of CNS and peripheral organs from NTg and C9-500 Tg mice treated or not with Ctrl-CPP and SRSF1-CPP. Bar chart shows weight in grams (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; NS: non-significant; **: p<0.01, ***: p<0.001; N=3 mouse organs/group). **(D)** Detail information about injection of recombinant (Rt) SRSF1-CPP in C9ORF72-ALS/FTD mice. **This figure is published in the Journal of Science Translational Medicine where I am co-author in that publication (Castelli et al., 2023). I performed the experiments on Panel A, B, C and D.**

4.4. Discussion

In the work described in this chapter, it was aimed to validate the functionality of recombinant SRSF1-CPP in comparison with synthetic SRSF1-CPP as a potential new therapeutic strategy for inhibiting the SRSF1: NXF1 dependent nuclear export of pathological SRSF1 bound sense and antisense *C9orf72* repeat transcripts *in vitro* and *in vivo* *C9ORF72*-ALS/FTD models. CPP-based therapeutic agents or drugs provided promising pre-clinical and clinical results in several relevant fields such as cancer and neuroscience (Guidotti et al., 2017); (Habault and Poyet, 2019). However, investigation of potential effect of CPP-dependent therapeutics in ALS and FTD diseases were notably absent. In relation to this concept, this chapter provides a view towards use of CPP as a potential therapeutic agent in *C9ORF72*/ALS/FTD disease as well as molecular tool for manipulating nuclear export mechanism *in vivo* studies.

In this chapter, all *in vitro* characterization studies of recombinant and synthetic SRSF1 peptide were conducted in HEK293T cells, which are generated by transformation of human embryonic kidney cells with sheared adenovirus 5 DNA (Graham et al., 1977). The HEK293T cell line provides several advantages due to the rapid growth rate, low maintenance required, high transfectability with calcium phosphate or polyethylenimine (PEI) and the ability to achieve efficient and stable protein expression (Dumont et al., 2016; Tan et al., 2021). HEK293T cells are also widely used in the field of neuroscience studies. This was based on the rationale that embryonic kidney cells and neurons originate from same stem cells which indicates that key biological process such as transcription, translation and protein folding are quite similar with nerve cells thereby providing neuron-like insights during characterization of therapeutic approaches relevant to neurological disorders (Lin et al., 2014). For example, Shaw *et al* revealed neuronal progenitor properties of HEK cells in their study. The HEK293T cells which were used express neurofilament subunits including NF-L, NF-M NF-H. In addition, gene expression profiling revealed that strong expression of several genes, which are already highly upregulated in hematopoietic and neuronal stem cells, was detected in HEK293T cells (Shaw et al., 2002).

Other studies also pointed out that HEK293T cells exhibit relevant neural phenotypes. For example, microarray analysis of several studies found that expression of several receptors which are highly expressed in brain such corticotrophin releasing factor type 1 (Dautzenberg et al., 2000), somatostatin receptor subtype SSTR2 (Daaka et al., 1997), spingosine-1 phosphate (van Koppen et al., 1996) was detected in HEK293T cells. Overall, these are the reasons why HEK293T cells line was used in this thesis. Although the HEK cell model is commonly used as a valuable tool for studying certain aspects of cell biology, they have some limitations that should be considered when using them as a model. For example, they do not replicate the complexity of primary neurons nor do they replicate the genetic background of human patients with *C9ORF72*-ALS/FTD.

Functional validation studies were initially started by performing high content immunofluorescence imaging and quantification of SRSF1 and ctrl peptides treatment into HEK293T cells. Fluorescence images indicated that SRSF1 and Ctrl-CPPs were delivered into HEK cells (**Figure 4.1.A and 4.1.B**). The accumulation of SRSF1 peptide in the nucleus was slightly higher than in the cytoplasm (**Figure 4.1.C**). This might be explained by interaction of the peptide with histone proteins which have binding affinity for peptides, but the mechanism of this interaction is still not identified. Increased peptide in the nucleus may result from overexpression of histones. The lower level of peptide in the cytoplasm could result from the peptide shuttling between the nucleus and cytoplasm, due to a lack of a rigid surface for preventing diffusion of peptide-like small molecules. Indeed, it is likely that the SRSF1 and control peptide can be exposed to cytoplasmic peptidases causing degradation, but it is still uncertain whether small peptides are faced with degradation by peptidases (Reits et al., 2003). The efficiency of dose-dependent cytoplasmic and nuclear delivery with recombinant SRSF1 and Ctrl peptides were compared by performing the same experiment with synthetic SRSF1 and Ctrl-CPP (**Figure 4.2**). Immunofluorescence images of the synthetic peptides showed that recombinant SRSF1 and Ctrl peptide exhibited similar functionally.

The next functional validation study was aimed at characterizing the ability of recombinant SRSF1-CPP to bind NXF1 in HEK cells, thereby allowing observations into whether an increasing concentration of SRSF1-CPP disrupts the interaction between endogenous SRSF1 and FLAG-tagged NXF1 (**Figure 4.3**). Co-immunoprecipitation assay showed that the Ctrl peptide had no effect on the interaction; however, the increasing concentration of recombinant SRSF1 that binds to NXF1, this prevents the binding of endogenous SRSF1 to NXF1.

To precisely characterize the molecular mechanism of my hypothesis, a GST pull-down assay was performed (**Figure 4.4.A**). The pull-down assay revealed that recombinant SRSF1-CPP is able to bind TAP/NXF1 protein by directly in a dose-dependent manner. This confirms that the SRSF1 (89-120) region of SRSF1-CPP is involved in the direct binding with TAP/NXF1. However, this result raised another question of whether the peptide would interact with other cellular proteins belonging the family of NXF factors. Therefore, specificity of the interaction between SRSF1-CPP with NXF2 and NXF3 which are other cellular proteins that show sequence and structural homology to NXF1 (**Figure 4.4.B and 4.4.C**) were investigated. The pull-down assay showed that the SRSF1-CPP exhibited weak binding to the NXF2 and NXF3 proteins at an increasing concentration in comparison to NXF1. This may be because increasing the concentration of SRSF1-CPP did not correlate with a proportional increase in pull-down signal obtained from either the Coomassie blue stained gel or western blot. In parallel to the pull-down assay, expression of NXF2 and NXF3 proteins within HEK cells, patient derived-neurons, and mouse brain and testis tissues was quantified with qRT-PCR (**Figure 4.4.E**). NXF2 and NXF3 were specifically expressed in mouse testis, which aligns with reports in the literature (Pan et al., 2009); however, neither protein is expressed in mouse brain, HEK cells or patient-derived neurons.

The direct and specific interaction of recombinant SRSF1-CPP to NXF1 was confirmed in two different assays as mentioned above but this raised again another question whether the recombinant SRSF-CPP inhibits the nuclear export of pathological *C9orf72* repeat transcripts in HEK cells. To test this, nuclear export assay was performed and sense (G₄C₂)₃₈ and (G₂C₄)₃₉ antisense transcripts in the presence of recombinant control and SRSF1-CPPs were quantified with qRT-PCR (**Figure 4.5.C and 4.5.D**). The nuclear export assay indicated an inhibitory role of SRSF1-CPP in SRSF1:NXF1 dependent nuclear export of pathological *C9orf72* repeat transcripts, which correlates with DPR production.

The effect of SRSF1 and Ctrl-CPPs on DPR expression was examined (**Figure 4.6**). V5 Western blotting showed that there is a dose-dependent inhibitory effect of synthetic and recombinant SRSF1-CPPs on sense and antisense DPRs expression. This is due to SRSF1-CPP binding specifically to the NXF1 receptor, thus preventing pathogenic repeat transcripts bound to endogenous SRSF1 from being exported into the cytoplasm. This allows for translation of less of the abnormal pre-mRNA with a corresponding reduction in toxic DPR production, as the SRSF1-CPP concentration is increased. It is well known that expression of toxic sense and antisense DPRs has an impact on cell proliferation (Hautbergue et al., 2017). Since recombinant SRSF1-CPP has been shown to reduce the expression of DPRs in HEK cells, the next step was to assess the dose-dependent effect of SRSF1-CPP on cell proliferation using an MTT assay (**Figure 4.7**). The MTT assay result correlates with the western blot data indicating that the SRSF1-CPP, whether synthetic or recombinant, resulted in an increase in cell survival with increasing concentration of SRSF1-CPP.

The promising results obtained from *in vitro* studies prompted investigations into the neuroprotective effect of synthetic SRSF1-CPP in a *C9ORF72*-ALS/FTD *Drosophila* model. It's worth mentioning that the recombinant SRSF1-CPP was tested in *Drosophila* and exhibited biological activity similar to the synthetic SRSF1-CPP. However, this study was conducted by our collaborator at the University of Cambridge, which is why I did not present the data. Furthermore, while no toxicity was observed in *Drosophila* following the recombinant SRSF1-CPP treatment, it was evident in mouse models. This discrepancy is attributed to the differences in immune responses between mice and *Drosophila*. *Drosophila* does not possess acquired/adaptive immunity and it relies on humoral and cell-mediated innate immunity for its defense against pathogens, such as bacteria, viruses, fungi, and parasites. Therefore, *Drosophila* has an immune system that is distinct from that of mammals (Salminen and Vale, 2020; Sheehan et al., 2018). Synthetic SRSF1-CPP was administered to wild type and transgenic *Drosophila* expressing sense DPRs (**Figure 4.8**). SRSF1-CPP treatment resulted in disrupting SRSF1:NXF1 dependent nuclear export (**Figure 4.8.A**), thereby reducing expression of sense DPRs (**Figure 4.8.D**) as well as reduction in motor deficits in *C9ORF72*-ALS larvae and adult flies (**Figure 4.8.B and 4.8.C**).

Our *in vivo* results align with data generated using TAT-based CPP for the treatment of other repeat expansion-associated neurological diseases. For example, polyglutamine (polyQ) disease is one of the microsatellite repeat expansion neurodegenerative diseases resulting from pathological expansion of cytosine-adenine-guanine (CAG) repeats. This repeat expansion causes aggregation of poly-Q proteins which in turn induce cellular toxicity and neuronal death. To prevent these effects, Popiel and colleagues designed a polyQ binding peptide fused with PTD (QBP1). When the peptide was administered orally to a *Drosophila* model of polyQ disease, it attenuated premature death by inhibiting aggregation of polyQ proteins (Popiel et al., 2007). The use of *Drosophila melanogaster*, or fruit fly, models in the study of ALS and FTD has contributed a lot to reveal pathogenic mechanisms of disease and to testing therapeutic approaches due to their rapid ability to expand mutants, ease of genetic manipulation and performance of large scale genome screening (Layalle et al., 2021); (Liguori et al., 2021). However, there are several limitations to using this model that should be taken into account. One of the main limitations is the timing of the phenotypes observed in the fly models, specifically that they are typically seen at early developmental stages such as the larval, pupal, or early adult stages, whereas in human ALS/FTD the symptoms typically manifest in the later stages of life (50 to 69 years). This timing discrepancy between the fly models and human disease means that the cognitive and behavioural impairments of the disease may not be accurately recapitulated in the fly models, which could affect the applicability of the results to human ALS/FTD (Snow et al., 2022; Yuva-Aydemir et al., 2018). Another limitation of using *Drosophila* models for ALS/FTD research is that some of the models utilize overexpression of disease-causing genes in the eyes of the *D.melanogaster* which is used as a measure of the neuropathology of the disease. While this method can provide valuable information about the toxic effects of certain genes, it does not replicate the complex neural circuitry and pathology found in the human brain. Additionally, the simplicity of the immune system in *Drosophila* compared to mammalian system constitutes another issue because this feature does not allow investigation of the important relationship between neuroinflammation and neurodegenerative disease (Pandey and Nichols, 2011). Furthermore, disease pathogenicity of ALS is not limited to the affected cells alone, it also involves the interaction between neurons and non-neuronal cells. Therefore, it is crucial to investigate the neuron-glia cell crosstalk to fully comprehend the mechanisms of the disease. However, *Drosophila* models do not accurately reflect the human disease pathway physiologically (Pandey and Nichols, 2011; Van Harten et al., 2021). Indeed, *Drosophila* does not have a blood-brain barrier which is an important consideration in the treatment of human neurological disorders (Cauchi and van den Heuvel, 2006).

Having confirmed the safety and efficacy of SRSF1-CPP in a *Drosophila* model of *C9ORF72*-ALS, I tested the therapeutic proof of principle efficacy of synthetic and recombinant SRSF1-CPP on *C9ORF72*-ALS mice (**Figure 4.9**). C9-500 bacterial artificial chromosome (BAC) mice line were utilised. This model is a widely used and well-established animal model for studying the genetic and cellular mechanisms underlying *C9ORF72*-ALS/FTD disease. The model is based on the insertion of a human *C9orf72* repeat expansion into the mouse genome, mimicking the genetic mutation that is found in many cases of human ALS and FTD (Liu et al., 2016). Researchers have found that this model exhibits a number of key features that are similar to those seen in human patients, including accumulation of sense and antisense DPRs and RNA foci formation. In terms of behavioural and functional features, this model has been shown to exhibit progressive motor dysfunction, including muscle weakness, tremors, paralysis, and reduced lifespan (Liu et al., 2016), although several research groups recently reported conflicting results relating to disease phenotypes observed on in this mouse line (see full discussion in Chapter 5) (Mordes et al., 2020).

A single intrathecal injection of recombinant or synthetic SRSF1-CPP to female C9-500 BAC mice via the cisterna magna was performed (see detailed discussion in next chapter). No adverse effects were observed following injection of synthetic SRSF1 and Ctrl-CPPs. However, 3 out of 5 transgenic mice injected with recombinant CPP died. This may have been due to the presence of excess urea following the purification of recombinant SRSF1-CPP. The presence of urea at high concentration can lead to an increase in the production of reactive oxygen species (ROS) by mitochondrial and cytosolic mechanisms in various cells such as beta and adipocytes cells (Magder, 2006). Excessive ROS damage to cell membranes causing cell death and impairing important chemical pathways which can cause tissue injury or death, septic shock and multi organ system failure (Giardino et al., 2017). It has been shown that urea at 10 nmol/L, which is relatively low in comparison to the urea concentration in human blood (1.8 to 7.1 mmol/L), accelerated the production of ROS observed in a patient with chronic kidney disease (D'Apolito et al., 2017).

There are also several factors, including poor injection techniques or mishandling, might contribute to mice mortality after recombinant SRSF1-CPP injection. Improper injection can lead to direct trauma, such as puncturing vital organs or introducing air emboli into the bloodstream. Significant blood loss can arise if a major blood vessel is punctured, and improper post-operative care can introduce a slew of complications (Lehnhardt et al., 2005). Furthermore, the use of non-sterile equipment or techniques may expose the animals to pathogens, causing systemic infections. Administering the wrong dosage, especially of substances with a narrow therapeutic margin, can result in overdosing and potentially be fatal. Improper handling can subject mice to undue stress, which, if chronic, can weaken their immune systems and raise susceptibility to infections. An unexpected yet vital consideration is the potential for substances to induce allergic reactions, leading to anaphylactic shock. Mice also face risks from poor restraint, which can lead to injuries, and thermal stress due to their narrow thermoneutral zone. Lastly, the behavior of mice post-procedure can be unpredictable; at times, cage mates may become aggressive after handling, leading to injuries or fatalities. Proper training in handling and injection techniques, paired with close post-procedure observation, is essential to mitigate these risks. However, while the potential reasons outlined above could account for the death of the mice injected with recombinant CPP, it's worth noting that I only used three mice in this study. Given this small sample size, drawing definitive conclusions about the causes becomes challenging.

Due to concerns regarding the safety of the recombinant SRSF1-CPP injections, prompting the veterinary staff in the animal house to advise the discontinuation of the experiment. Beyond the immediate safety issues, broader ethical considerations and animal welfare concerns further necessitated the cessation of our work. As a result, the experiment was terminated entirely. The findings from the recombinant CPP research suggest that the recombinant SRSF1-CPP methodology may not be suitable for mammalian systems. This led to my decision to explore gene therapy in the subsequent chapter. Given the challenges tied to the recombinant approach, it seems improbable to secure approval for *in vivo* or clinical studies. Therefore, I will not be pursuing recombinant peptide for *in vivo* experiments. However, the recombinant CPP approach still holds potential for *in vitro* and *Drosophila* studies.

The synthetic peptide approach presents itself as a promising alternative for in vivo studies. However, should mice treated with peptides experience mortality, it's essential to conduct systematic analyses to determine the underlying causes. The first step usually involves a necropsy, which offers a comprehensive visual examination of every organ and tissue, flagging any deviations in properties such as hue, consistency, dimensions, and contour, along with noting external abnormalities like lesions or excretions. Following this, a histopathological examination is crucial. During this process, tissue samples from essential organs, including the heart, liver, kidneys, lungs, and brain, are collected. These samples are then prepared and stained, enabling detailed microscopic evaluation. Concurrently, post-mortem blood analyses can offer a complete blood count (CBC) and a biochemical profile, reflecting the state of organ function and overall health. Furthermore, organs or potential infection sites can undergo microbiological tests to detect possible bacterial, fungal, or viral infections. If toxins are believed to be a factor in death, toxicological tests on blood, urine, and tissues can be performed to detect the presence of toxic substances or metabolites. Advanced methods like immunohistochemistry can pinpoint specific disease-related proteins or markers. Advanced molecular methods such as PCR or RNA sequencing can be used to detect specific pathogens, especially if there's a suspicion of a viral infection. Complementing these tests, cytological studies examine cell abnormalities, and parasitological examinations could help discern if parasitic infections played a role in the death.

Immunofluorescence and poly GP DPR expression analysis by MSD-Elisa were performed. The V5 immunofluorescence images confirmed uptake of both CPPs into the cerebellum. Interestingly, the ctrl-CPP did not show good distribution in comparison to SRSF1-CPP (**Figure 4.9.A**). This may be due to the fact that the ctrl-CPP does not target anything specifically in brain, and therefore might be cleared quickly. Alternatively, the size of ctrl-CPP peptide is smaller than SRSF1-CPP making it more difficult to detect its distribution across the tissue. A poly (GP) DPR assay indicated that the SRSF1-CPP is functional in a *C9ORF72*-ALS/FTD mice, reducing DPRs expression (**Figure 4.9.B**). While I have demonstrated encouraging results using synthetic SRSF1-linker CPPs as potential methods for oral delivery or gene therapy in animal models, several considerations must be addressed before advancing these SRSF1-CPPs on the translational pathway for potential treatments in patients. Firstly, the current study was limited by a small sample size, with only three animals per treatment group. This number should be increased in future long-term cohort studies.

Secondly, the pharmacodynamic and pharmacokinetic parameters must be ascertained through nasal or oral delivery methods to determine the optimal dosage and regimen for repeated dosing. Additionally, there is a need for proof-of-concept efficacy and safety studies, which should encompass repeated peptide injections and comprehensive toxicology assessments.

Similar *in vivo* results were observed with a study investigating a neuroprotective role of a peptide on the SOD1^{G93A} animal model of ALS. Bcl-XL is one of the anti-apoptotic proteins which belongs to Bcl-2 protein family that has been found to have a rescue effect on cell death caused by ALS-linked mutant SOD1 (Azzouz et al., 2000). In this respect, the peptide was designed by creating the BH4 domain of Bcl-XL and fusing with TAT₄₈₋₅₇ to assess the therapeutic efficacy *in vivo* in SOD1^{G93A} transgenic mice. The BH4-TAT peptide was administered daily via intraperitoneal injection. The administration of the peptide delayed disease onset by 107.5 ± 1.8 days, reduced the death of motor neurons and astrocytes, and improved motor function (Martorana et al., 2011). This study supports my work and suggests that CPP based innovative drug-like components are of potential therapeutic benefit for ALS. No further studies have since been published on ALS/FTD disease based on using CPP as a therapeutic tool. Research into CPPs has revealed many potential applications and yielded promising results. However, these peptides have not yet been approved for use as a therapy for patients. This is because of limitations that stem from their chemical and functional properties such as instability, difficulty in crossing cellular membranes, and the risk of inducing an immune response (Xie et al., 2020); (Reissmann, 2014). The stability is one of main issues for *in vivo* application of CPPs because CPPs have been found to degrade rapidly inside cells and living organisms caused due to enzymatic degradation caused by the presence of proteases and peptidases in biological fluids. This instability is a significant drawback for their therapeutic use, as it limits their circulation half-life and bioavailability (Fominaya et al., 2015; Pujals and Giralt, 2008). Additionally, CPPs have a lack of specificity in targeting specific cells or tissues for therapeutic applications. This issue can cause cellular toxicity (Govindarajan et al., 2012). This lack of specificity can lead to widespread distribution of CPP and CPP-conjugated therapeutics thereby reducing drug efficiency due to lower local concentrations. To address this, CPPs are often applied directly to the target tissue (Walrant et al., 2017).

Additionally, CPPs can cause detrimental immune responses in patients due to their polypeptide nature, leading to reduction in treatment efficacy and causing stress on the immune system (Järver et al., 2010). Furthermore, the pharmacokinetics and pharmacodynamics of CPP-based treatments have not been extensively studied, presenting a limitation in this field (Langel, 2019).

4.5. Experimental Conclusion

In summary, this chapter carried out a series of validation experiments to determine the effectiveness of recombinant and synthetic SRSF1-CPPs *in vitro* and in a *C9ORF72*-ALS/FTD mouse model. The results demonstrate that both recombinant and synthetic SRSF1-CPPs can effectively enter HEK cells, interact specifically with NXF1, inhibit the nuclear export of abnormal transcripts and decrease the expression of harmful DPRs in both *in vitro* and *in vivo* *C9ORF72*-ALS/FTD models. This highlights the potential of SRSF1-CPPs as a promising therapeutic option for treating *C9ORF72*-ALS/FTD.

5. Chapter 5 - Engineering a Viral Vector Expressing SRSF1 Cell Permeable Peptides for The Development of Novel Gene Therapy Approaches

5.1. Introduction

In the previous chapters, I started initially with the development of an expression and purification protocol for generating recombinant SRSF1-CPP, as detailed in Chapter 3. In the subsequent work documented in Chapter 4, I investigated the potential functionality of the recombinant SRSF1-CPP as a therapeutic agent for *C9ORF72* ALS/FTD, employing a combination of both in vitro and in vivo models. Encouragingly, results demonstrated that the recombinant SRSF1-CPP exhibited functional capabilities in the *C9ORF72*-ALS/FTD reporter cell model. This not only underscores its therapeutic promise but also suggests its utility as a molecular tool for modulating nuclear export processes within cells. Yet, in the mammalian models of *C9ORF72*-ALS/FTD, its efficacy was unfortunately lacking. In contrast, when tested in vivo, the synthetic variant of SRSF1-CPP exhibited remarkable efficacy. As delineated by (Castelli et al., 2023), there was a significant reduction in poly-GP expression by around 70%. Furthermore, the synthetic SRSF1-CPP showcased its potential as a neuroprotective agent, evident in its efficacy within *C9ORF72*-ALS/FTD patient-derived motor neurons and *Drosophila* models.

However, the journey with SRSF1-CPP was not straightforward. Two primary challenges emerged (Lipinski, 2000; Nugrahadhi et al., 2023) its poor bioavailability and inherent chemical-physical limitations prompted us to recalibrate our approach towards gene therapy vector design. Reinforcing this decision, data depicted in **Figure 5.1** revealed that the SRSF1-linker CPP possessed a half-life of approximately 24 hours in HEK cell culture medium. Such a transient presence highlights the impracticality of frequent intrathecal injections. Shifting our focus to a gene therapy vector strategy, which allows for the expression of the SRSF1-linker CPP, offers a compelling solution. This method holds the potential for continuous peptide production in the CNS with just a single intrathecal injection. It's noteworthy to mention that, to date, the expression of peptides from a viral gene therapy vector has not been documented in scientific literature. This places our viral vector design in a unique position, adding a novel contribution to the field. Our venture into this unexplored domain is not only an innovative step but also a point of significant academic intrigue, given the absence of prior work expressing peptides from such AAV viral vectors. To facilitate the expression of CPP, we plan to investigate the functions of both RNA polymerases II and III, particularly their roles in transcribing long coding and short non-coding transcripts (see **section 5.3.2** for a detailed description of the viral vector design).

5.2. Aims and Objectives

- Generate scAAV9 expressing SRSF1-linker CPP comprising amino-acids 89-120 which inhibits the interaction of *C9orf72*-repeat RNA-bound SRSF1 with the nuclear export machinery, under either an RNA polymerase II (Chicken- β -actin hybrid CBh) or III (histone H1) promoter.
- Generate scAAV9 expressing SRSF1-RRM2 α 1 CPP encompassing amino-acids 132-144, a peptide recently identified in the lab as an inhibitor of the interaction between SRSF1 and *C9orf72*-repeat RNAs, under either H1 or CBh promoter.
- Evaluate the functionality of the constructs in a transfected HEK293T cell model of *C9ORF72*-ALS/FTD.
- Test proof-of-concept efficacy of scAAV9 expressing SRSF1-CPPs in a *C9ORF72*-ALS/FTD mouse model while comparing their potency to the established scAAV9 mediating the partial depletion of SRSF1.

5.3. Results

5.3.1. Stability of SRSF1-CPP

The stability of CPPs remains a prevalent concern in the realm of biopharmaceutical research, frequently emerging as a bottleneck in various stages of biological drug development. Given the ubiquity of this challenge, I was compelled to scrutinize the *in vitro* stability of my SRSF1-CPP within the medium of HEK cells. To this end, a concentration of 1 μ M SRSF1 peptide was introduced to the HEK cell medium, followed by incubation in a cell culture environment over a span of 12 days. Throughout this period, aliquots of the medium were systematically extracted daily to undergo V5 western blot analysis. Intriguingly, the outcomes of the western blot delineated that the peptide commenced degradation merely 24 hours post-incubation.

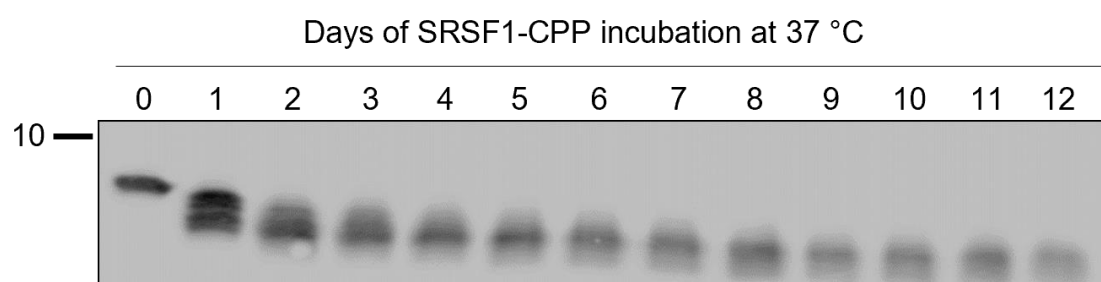


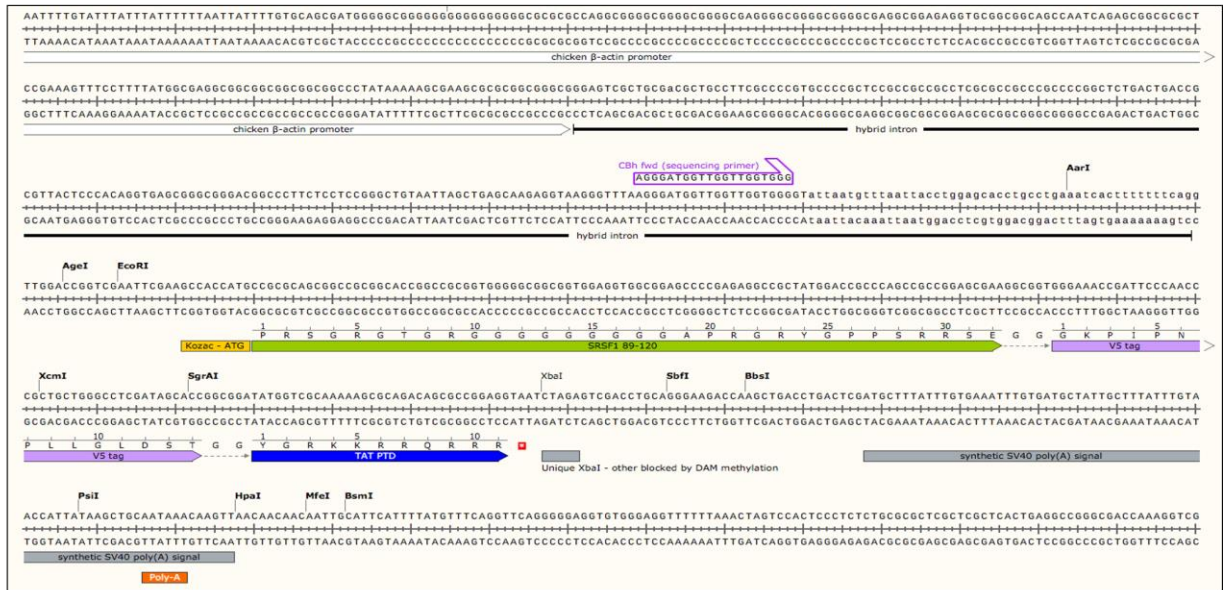
Figure 5.1. Stability assessment of SRSF1-CPP in HEK cell culture medium over a 12-day period. The SRSF1-CPP, at a concentration of 1 μ M, was introduced to the HEK cell culture medium and subsequently incubated. On each day of incubation, 50 μ l of the medium was sampled for V5 western blot analysis. The western blot image presented in this Figure is from one biological repeat.

5.3.2. Cloning of SRSF1-linker and RRM2 α 1 CPPs Into scAAV Constructs

I have previously shown in chapter 4 that both the recombinant or synthetic SRSF1-linker CPPs are able to specifically bind to NXF1 and inhibit the nuclear export of pathogenic sense and antisense *C9orf72* repeat transcripts in *in vivo* and *in vitro* *C9ORF72*-ALS/FTD disease models. In parallel to the SRSF1-linker CPP, Prof. Guillaume Hautbergue identified another cell permeable peptide comprising the alpha helix 1 of the RRM2 domain of SRSF1 (RRM2 α 1) which binds to pathogenic *C9orf72* repeat transcripts rather than NXF1. I therefore wanted to express both peptides in scAAV9 viral vectors and compare their efficacy head-to-head in a proof-of-principle gene therapy study in *C9ORF72*-ALS/FTD mice. However, there is no report of viral-mediated expression of peptides in the scientific literature. One of the potential challenges is that the peptides are encoded by short RNA sequences (typical on non-small non-coding interfering RNAs) which have to be translated like much larger protein-encoding mRNAs. In most case, RNA polymerase II is responsible for transcribing both protein-coding genes and long non-coding RNAs (LNC-RNAs). Protein-coding genes transcribed by this polymerase typically have transcripts that are longer than 300 nucleotides, corresponding to proteins that are more than 100 amino acids long. This length is considered the minimum for a protein based on bioinformatics methods. Long non-coding RNAs transcribed by RNA polymerase II often possess micro-open reading frames that are more than 200 nucleotides in length. Regardless of whether the transcripts are protein-coding or LNC-RNAs, those transcribed by RNA polymerase II generally exceed 300 nucleotides.

On the other hand, RNA polymerase III typically transcribes small non-coding RNAs that are less than 200 nucleotides, such as tRNAs. For instance, the peptide in my study is 60 nucleotides long, and such peptides are often transcribed by RNA polymerase III and are non-coding. Given these distinctions, I have focused on both RNA polymerase II and RNA polymerase III for my work and therefore, I decided to build four different viral scAAV plasmids in which expression of the peptides is driven by either RNA polymerase II (CBh) or RNA polymerase III (H1) promoters. RNA polymerase III synthesises small transcripts, RNA polymerase II produces large transcripts, so we did not know which one would work the best. The plasmid maps of linker and RRM2 α 1 CPP constructs under H1 (**Figure 5.2.A and 5.2.B**) and CBh promoters (**Figure 5.2.C and 5.2.D**) are provided below.

C



D

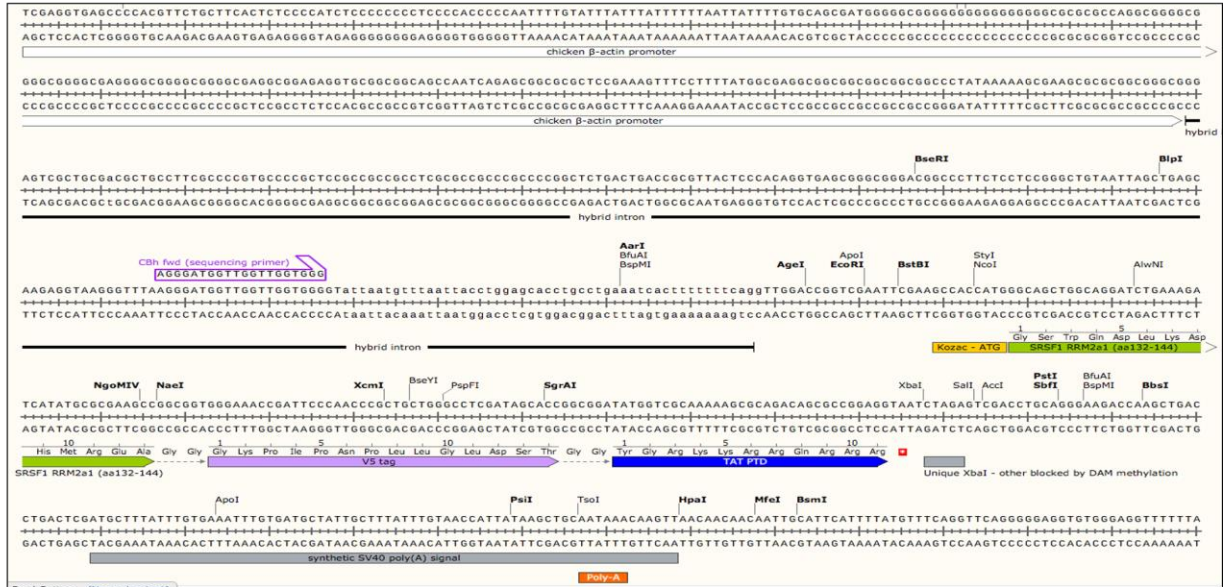


Figure 5.2. Plasmid maps of engineered scAAV SRSF1(89-120)-V5-TAT PTD and scAAV SRSF1-RRM2a1(132-144) V5-TAT PTD. (A) Plasmid map showing integration of SRSF1(89-120)-V5-TAT PTD with the H1 promoter between the BamHI and HindIII restriction sites. **(B)** Plasmid map showing integration of the SRSF1-RRM2a1 (132-144) V5-TAT PTD with the H1 promoter between the BamHI and HindIII restriction sites. **(C)** Plasmid map showing integration of SRSF1(89-120)-V5-TAT PTD with CBh promoter between the EcoRI and XbaI restriction sites. **(D)** Plasmid map showing integration of SRSF1-RRM2a1 (132-144) V5-TAT PTD with the CBh promoter between the EcoRI and XbaI restriction sites. Both DNA sequences and encoded amino-acids in the one-letter code are highlighted on the map.

5.3.2.1. Cloning of SRSF1-linker and SRSF1-RRM2 α 1 CPPs Into scAAV_Efs1 α GFP

To generate a viral construct capable of driving expression of either SRSF1-linker 89-120 or SRSF1-RRM2 α 1 (132-144) V5 TAT-PTD CPPs under the control of the RNA polymerase III H1 promoter, Taq PCR were first utilized to amplify the constructs by annealing two respective oligonucleotides of each peptide. PCR products underwent phenol precipitation and pellets were resuspended prior to digesting with BamHI and HindIII. A previously engineered scAAV_H1_SRSF1 shRNA10 vector was digested with BamHI and HindIII and treated with CIAP. The digested top band corresponding to scAAV_H1-SRSF shRNA10+CIAP was extracted from the agarose gel (**Figure 5.3.A**). The BamHI/HindIII digested linker and RRM2 α 1 CPP constructs were ligated into pre-digested and CIAP treated scAAV_H1-SRSF1shRNA10 vector to obtain scAAV_H1-SRSF1-linker CPP_EF1 α -eGFP and scAAV_H1-SRSF1-RRM2 α 1 CPP_EF1 α -eGFP.

The resulting constructs were transformed into DH5 α competent cells and purified via a miniprep. All clones were digested using BamHI and HindIII restriction enzymes to determine successful integration. Bands corresponding to the inserts were expected at 210 bp for SRSF1-linker CPP and at 169 bp for SRSF1-RRM2 α 1 CPP. All clones exhibited the presence of an insert (**Figure 5.3.B and 5.3.C**). Subsequent digestion with SmaI was performed to confirm the presence of inverted terminal repeats (ITRs) in the plasmids which are replicated the regions in scAAV, and which are usually prone to shrinkage in *E. coli*. The absence of ITRs would prevent the manufacturing of the peptide by scAAV. The expected band sizes of scAAV_H1 SRSF1-linker CPP/SmaI were 3689 bp, 11981 bp, 11 bp and 11 bp, and for scAAV_H1 SRSF1-RRM2 α 1 CPP, the band sizes were 3689 bp, 1940 bp, 11 bp and 11 bp. ITRs were present in all clones following SmaI digestion (**Figure 5.3.D and 5.3.E**). Clone #2 for linker and clone #1 for RRM2 α 1 CPPs were selected for mega-prep to generate a high plasmid yield; the integrity of these plasmids was assessed via restriction digest using BamHI, HindIII, and subsequently SmaI to ensure that the final plasmids would be able to express CPPs in HEK cells post transfection.

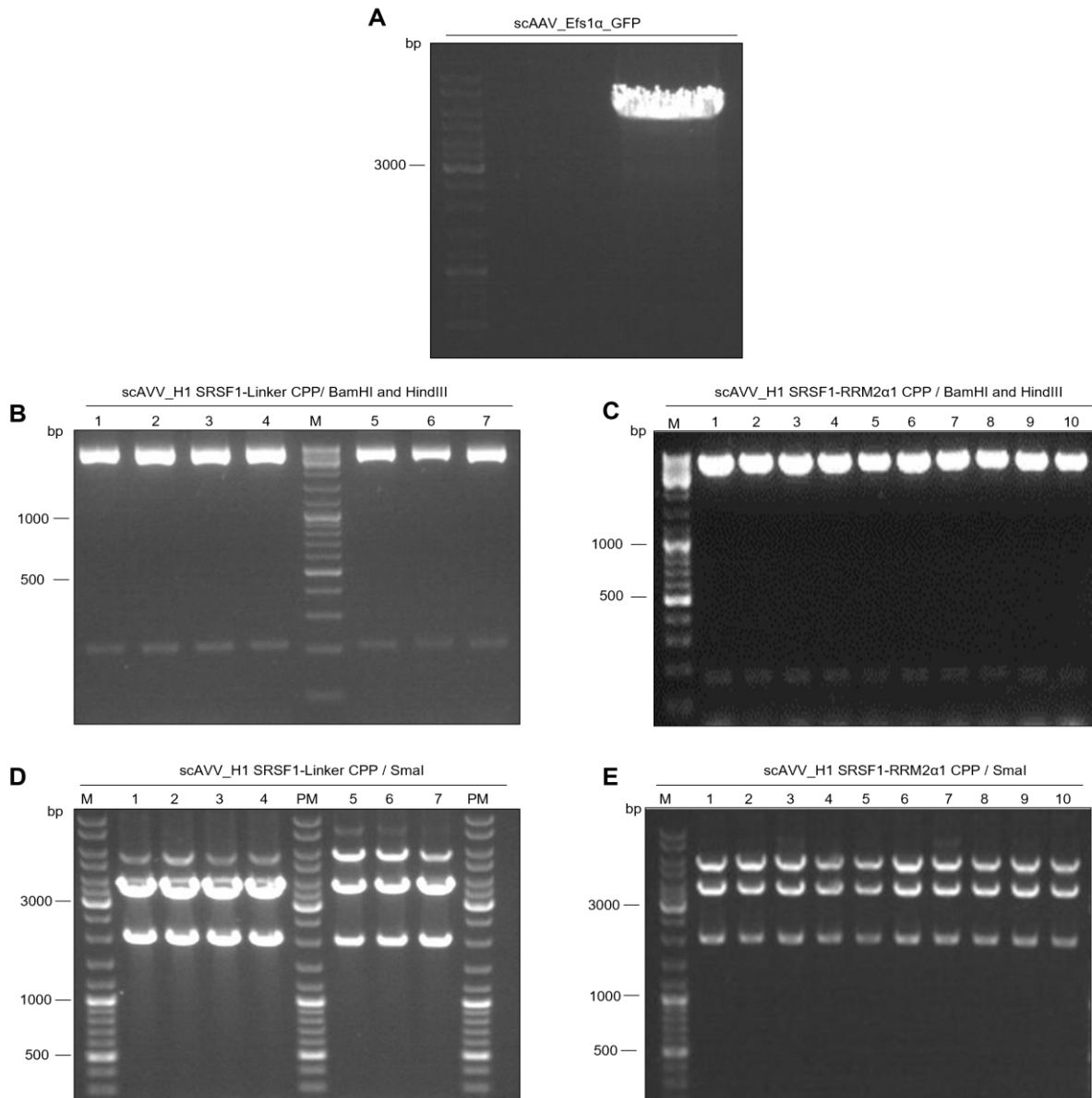


Figure 5.3. Cloning of SRSF1-linker and RRM2 α 1 CPPs into scAAV_Efs1 α GFP. (A) Preparation cut of the vector: the plasmid scAAV_H1-SRSF1-shR10_GFP was digested with BamHI and HindIII and treated with CIP and the restricted vector band corresponding to scAAV_H1-SRSF1-shR10_GFP + CIP was extracted from a 1 % agarose gel. (B) Digest of the scAAV_H1 SRSF1-linker CPP mini prep using BamHI and HindIII. The bottom bands of 210 bp indicate that the scAAV_H1 SRSF1-linker CPP insert has been integrated successfully into the excised vector. (C) Digest of the scAAV SRSF1_H1 RRM2 α 1 CPP mini prep using BamHI and HindIII. The bottom bands of 169 bp indicate that the scAAV_H1 SRSF1-RRM2 α 1 CPP insert has been integrated successfully into the excised vector. (D) SmaI digest run on a 3 % agarose gel to check the integrity of the ITRs in the mini prep of selected scAAV_H1 SRSF1-linker CPP clones. (E) SmaI digest run on a 3 % agarose gel to check the integrity of the ITRs in the mini prep of selected scAAV_H1 SRSF1-RRM2 α 1 CPP clones. Molecular weight marker bands are indicated in bp. N=1; no statistical analysis was performed.

5.3.2.2. Cloning of SRSF1-linker and RRM2 α 1 CPPs Into scAAV_CBh-SRSF1-m4

I generated an alternative method of expressing linker or RRM2 α 1 SRSF1-CPPs under the control of RNA polymerase II CBh promoter. Plasmid constructs were engineered using the same PCR protocol as used for plasmids expressing peptides under an H1 promoter. The resulting PCR products were obtained after digesting with EcoRI and XbaI. A scAAV_CBh-SRSF1-m4 vector was digested with EcoRI and XbaI and treated with CIAP. The digested products were run on an agarose gel and the upper band corresponding to EcoRI/XbaI-cut scAAV_CBh-SRSF1-m4 +CIAP was extracted (**Figure 5.4.A**). The EcoRI/XbaI digested linker and RRM2 α 1 plasmids were then ligated into EcoRI/XbaI-cut scAAV_CBh-SRSF1-m4 +CIAP vector to generate scAAV_CBh SRSF1 89-120 V5 CPP and scAAV_CBh SRSF1 RRM2 α 1 (132-144) V5 CPP.

EcoRI and XbaI digestion of all clones following DH5 α transformation were performed to infer successful ligation of inserts into the vector. Bands were expected at 198 bp for SRSF1_CBh linker CPP and at 147 bp for SRSF1_CBh RRM2 α 1 CPPs (**Figure 5.4.B and 5.3.C**). The agarose image shows that four clones are at the correct size for the linker and two clones exhibit the correct size for the RRM2 α 1 CPPs (**Figure 5.4.B and 5.4.C**). Following that, SmaI digestion was undertaken to check the integrity of ITRs. The expected band sizes of scAAV_CBh SRSF1-linker CPP/SmaI were 2724 bp, 1430 bp, 77 bp, 11 bp and 11 bp. For scAAV_CBh SRSF1-RRM2 α 1 CPP, the expected band sizes were 2724 bp, 1373 bp, 77 bp, 11 bp and 11 bp (**Figure 5.4.D and 5.4.E**). ITRs were present in all clones following SmaI digestion. I chose clone #3 for the SRSF1_CBh linker and clone #3 for the SRSF1_CBh RRM2 α 1 CPP for further large-scale plasmid purification. The final plasmids were re-examined using EcoRI and XbaI digestion, followed by SmaI digestion.

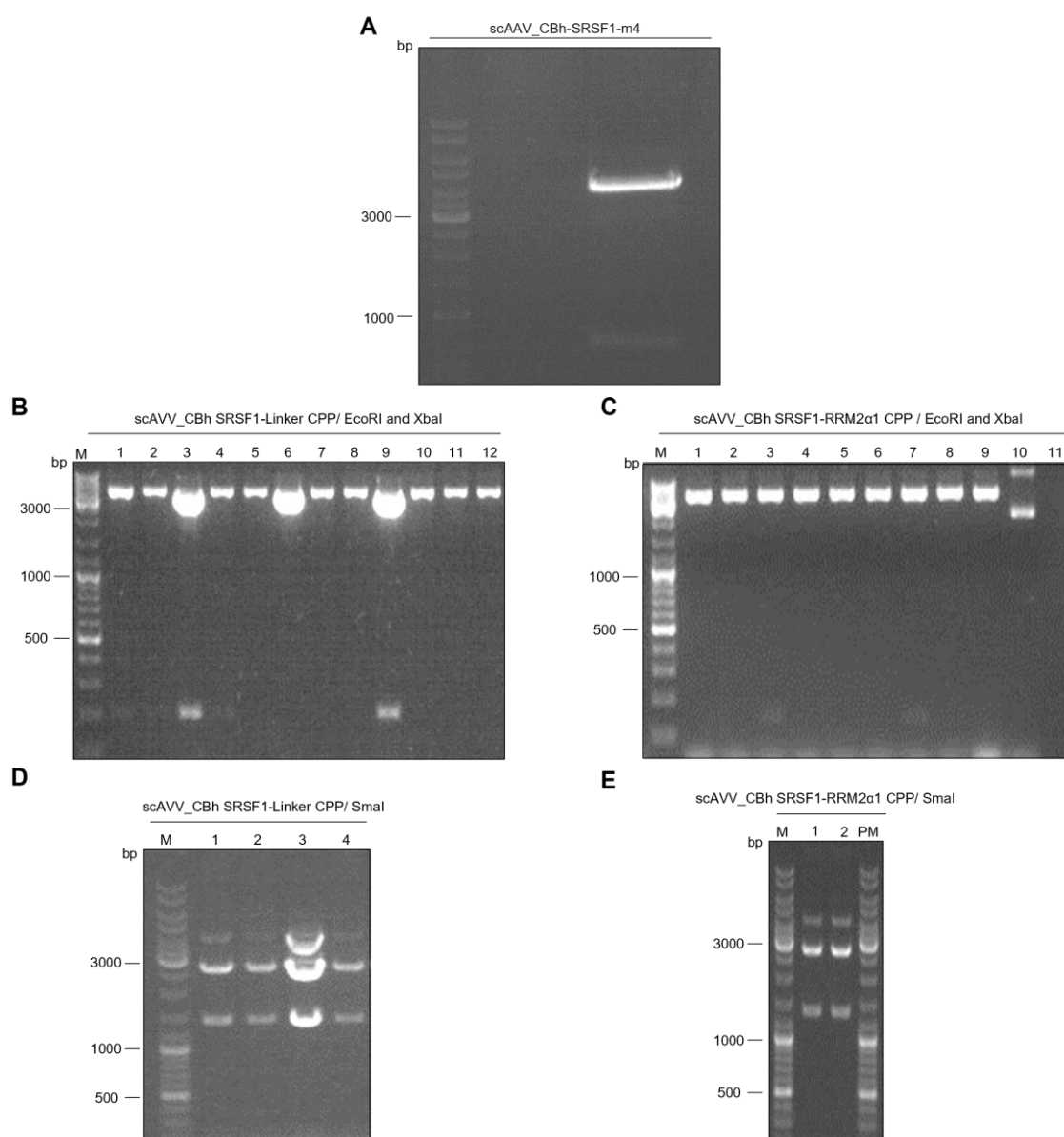


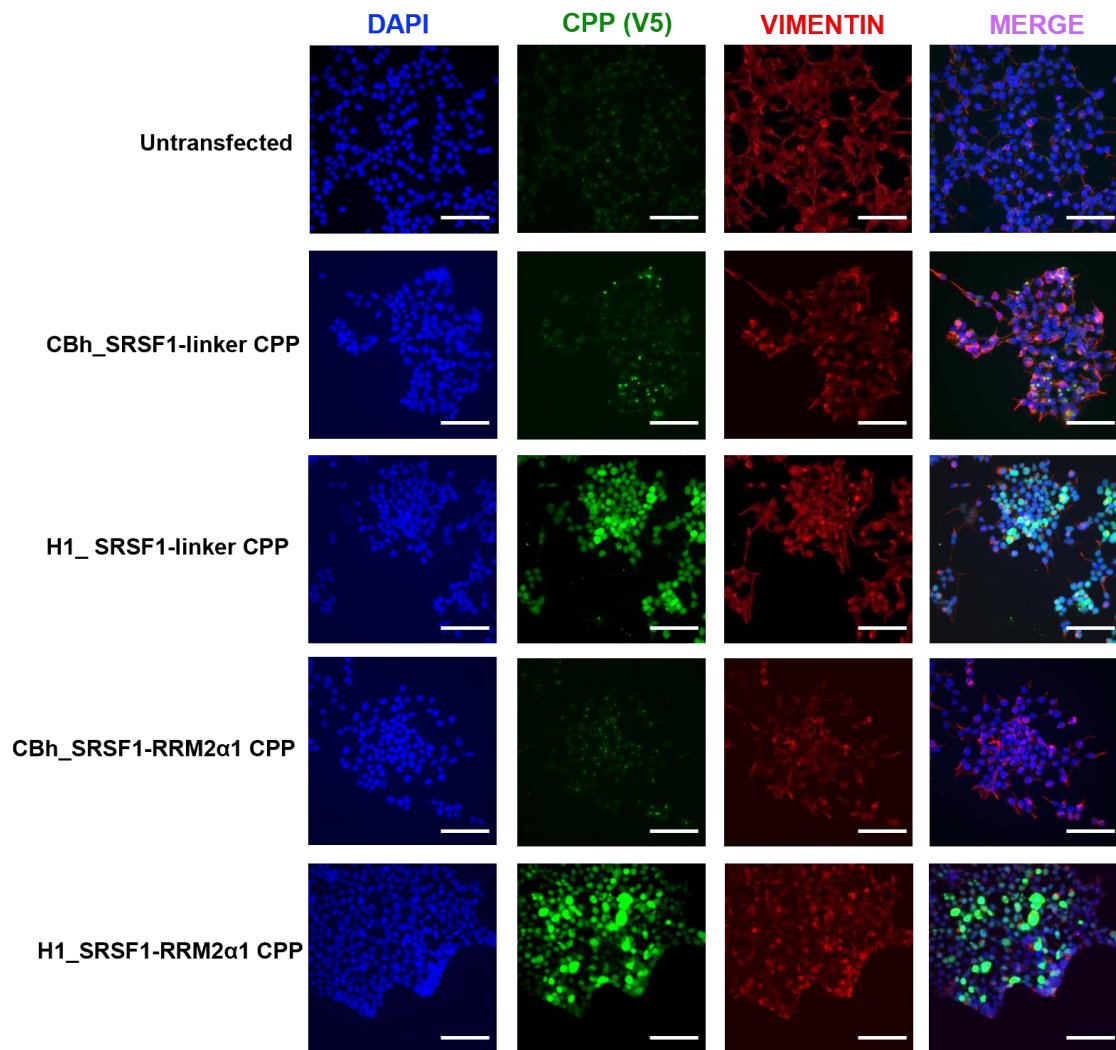
Figure 5.4. Cloning of SRSF1-linker and RRM2 α 1-CPPs into scAAV_CBh-SRSF1-m4. **(A)** Preparation of the cut vector: the plasmid scAAV_CBh-SRSF1-m4 was digested with EcoRI and XbaI and treated with CIP and the restricted vector top band corresponding to scAAV_CBh-SRSF1-m4 + CIP was extracted from a 1 % agarose gel. **(B)** Digest of the scAAV_CBh SRSF1-linker CPP mini prep using EcoRI and XbaI. The bottom bands of 198 bp indicate that the scAAV_CBh SRSF1-linker CPP insert has been integrated successfully into the excised vector. **(C)** Digest of the scAAV_CBh SRSF1-RRM2 α 1 CPP mini prep using EcoRI and XbaI. The bottom bands of 147 bp indicate that the scAAV SRSF1-RRM2 α 1 CPP insert has been integrated successfully into the excised vector. **(D)** SmaI digest run on a 3 % agarose gel to check the integrity of the ITRs in the mini prep of selected scAAV_CBh SRSF1-linker CPP clones. **(E)** SmaI digest run on a 3 % agarose gel to check the integrity of the ITRs in the mini prep of selected scAAV_CBh SRSF1-RRM2 α 1 CPP clones. Molecular weight marker bands are indicated in bp. N=1; no statistical analysis was performed.

5.3.3. In vitro Validation of The Functionality of The Constructs

5.3.3.1. Transfection of SRSF1-linker and RRM2 α 1 CPP Plasmids Under H1 Promoters Successfully Expressed V5 tagged CPPs in HEK293T Cells

After successful cloning of the SRSF1-linker and RRM2 α 1 CPPs into scAAV_H1-SRSF1-shRNA10 and scAAV_CBh-SRSF1-m4 vectors, the functionality of all constructs was validated in a HEK cell model of *C9ORF72*-ALS/FTD. Expression of linker or RRM2 α 1 CPPs under H1 and CBh promoters was tested following transient transfection with the recently generated four different plasmids (H1_SRSF1-linker CPP, CBh_SRSF1-linker CPP, H1_RRM2 α 1-CPP and CBh_SRSF1-RRM2 α 1 CPP) for 48 hours followed by fixation and immunostaining. As a negative control, untransfected cells were used for the immunofluorescence assay. The visualisation of SRSF1-linker and RRM2 α 1 CPPs expression was undertaken by using confocal microscope. The immunofluorescence images (**Figure 5.5.A**) showed a modest increase in CPP expression in HEK cells under the control of a CBh promoter compared to untransfected cells. However, when the cells were transfected with plasmids controlling peptide expression under an H1 promoter, there was a significant increase in the amount of V5 tagged SRSF1-CPPs compared to cells transfected with plasmids expressing CPP under CBh promoters or untransfected cells. The quantification data showed that expression levels of linker and RRM2 α 1 CPP were more pronounced under the control of RNA polymerase III H1 promoter and approximately more than 10 times higher than when the RNA polymerase II CBh promoter was utilised (**Figure 5.5.B**).

A



B

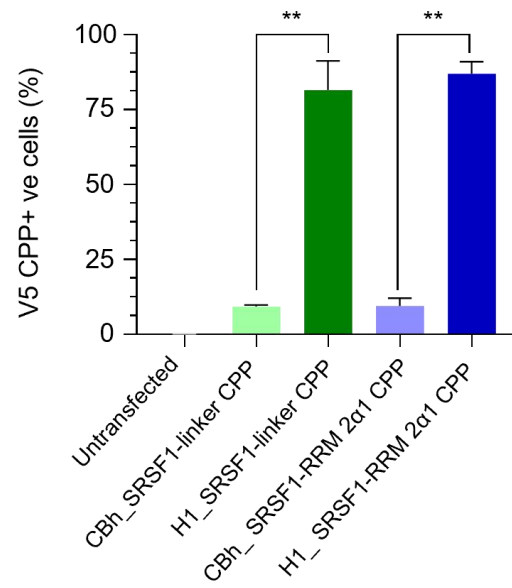


Figure 5.5. scAAV_H1 transfection facilitates expression of V5-tagged SRSF1-linker and SRSF1-RRM2 α 1 CPPs in HEK293T cells. (A) HEK293T cells were transfected with scAAV SRSF1-linker CPP plasmid with CBh or H1 promoters and scAAV SRSF1-RRM2 α 1 CPP plasmid with CBh or H1. Cells were stained with anti-V5 (green) to visualize V5-tagged SRSF1-CPP expression, Vimentin (red) to delineate cytoplasm and Hoechst (blue) to delineate nuclei. Confocal microscope images were acquired at 100X magnification. Images were acquired with a Confocal microscope (Zeiss) at 100X and the same exposure at 72 hours post transfection. Scale bar = 20 μ m. (B) Quantification of fluorescent intensity of the scAAV SRSF1-linker CPP plasmids with CBh or H1 promoters and scAAV SRSF1-RRM2 α 1 CPP plasmid with CBh or H1. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat and statistical assessment was performed from approximately 3000 cells cultured per biological repeat with each plasmid. (n=3, Bar charts are means of cell numbers $\pm$ SD or SEM; one-way ANOVA with Tukey post hoc's correction for multiple comparisons).

5.3.3.2. SRSF1-linker and RRM2 α 1 CPPs Expressed Under H1 Promoters Inhibit Sense (poly-GA, GP, GR) and Antisense (poly-GP, PR, PA) DPRs Expression

Following the identification of which promoter (RNA polymerase III H1 and RNA polymerase II CBh) is able to drive efficient expression of linker and RRM2 α 1 CPPs, my next aim was to examine whether the engineered SRSF1-CPP plasmids are able to suppress toxic sense and antisense DPR expression in a *C9ORF72*-ALS/FTD HEK cells model. This will help us to decide which SRSF1-CPP constructs will be applied to the C9-500 BAC mouse model as a therapeutic treatment strategy. HEK cells were co-transfected with either scAAV_H1-SRSF1-linker, scAAV_CBh-SRSF1-linker, scAAV_H1-SRSF1-RRM2 α 1 or scAAV_CBh-SRSF1-RRM2 α 1 V5-TAT PTD CPPs along with (G₄C₂)₄₅-3xV5 or (C₄G₂)₄₃-3xV5 repeat transcript plasmids expressing either sense or antisense DPRs. An empty vector carrying only the scAAV_H1 and scAAV-H1-Ctrl-shRNA were co-transfected with either sense or antisense repeat plasmids serving as negative controls. scAAV-SRSF1 shRNA 10 was also co-transfected along with sense and antisense repeat constructs as a positive control to assess impact of SRSF1-CPP expression alone in the cell line on DPR expression without a confounding effect of expressed peptide. A V5 antibody was used to detect the expression of DPR proteins in all three reading frames. Tubulin was used as a loading control. SRSF1 antibody was used to show that the cells treated with sAAV-SRSF1-RNAi reduce expression of endogenous SRSF1 in cells thereby reducing sense and antisense DPR expression. A GFP antibody was used to determine the transfection efficiency of CPP plasmids.

Western blotting shows successful expression of sense and antisense DPRs following co-transfection with either sense and antisense DPR plasmids alongside lines expressing an empty vector or ctrl-shRNA (**Figure 5.6.A and 5.6.B**). This indicates that the empty vector and ctrl-shRNA did not interfere with the production of DPRs. Sense and antisense DPR expression were not detected in untransfected cells, which also confirms the specificity of the V5 antibody. The transfection of linker and RRM2 α 1 CPP plasmids under the control of CBh promoters did not inhibit sense and antisense DPR production. However, transfection of cells with SRSF1-linker and RRM2 α 1 CPP plasmids under the H1 promoter reduced the expression of DPRs compared to cells transfected with either ctrl-shRNA or sense and antisense DPR plasmids.

It was also observed that scAAV-SRSF1-RNAi transfected cells show a reduction in sense and antisense DPR expression indicating that the shRNA was effective as a positive control (**Figure 5.6.A and 5.6.B**). GFP was detected in cells transfected with either ctrl-shRNA, SRSF1-RNA or H1_ SRSF1-linker and RRM2 α 1 CPPs indicating a high transfection efficiency. The quantification of the DPR assay aligns with western blot data indicating a significant reduction in the expression of sense or antisense DPRs in cells treated with plasmids expressing the CPPs under RNA polymerase III H1 promoter or treated with SRSF1-RNAi compared to cells transfected with the plasmids expressing CPPs under CBh promoter and ctrl-shRNA (**Figure 5.6.C and 5.6.D**). This highlights the overall indication given from the images that the CPPs expressed under the control of the H1 promoter were functional and were able to inhibit RAN translation sense and antisense repeat transcripts in HEK cells.

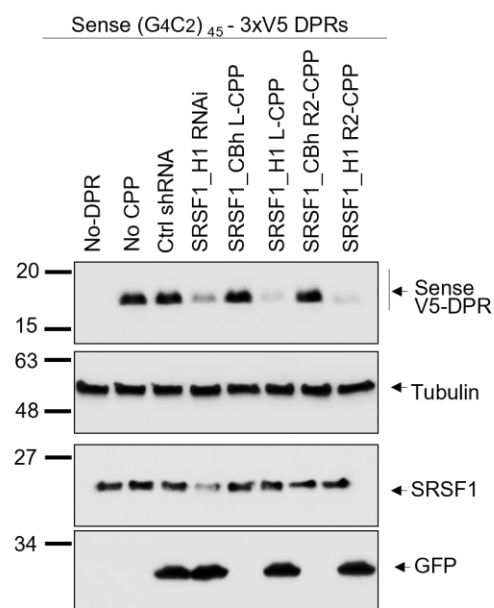
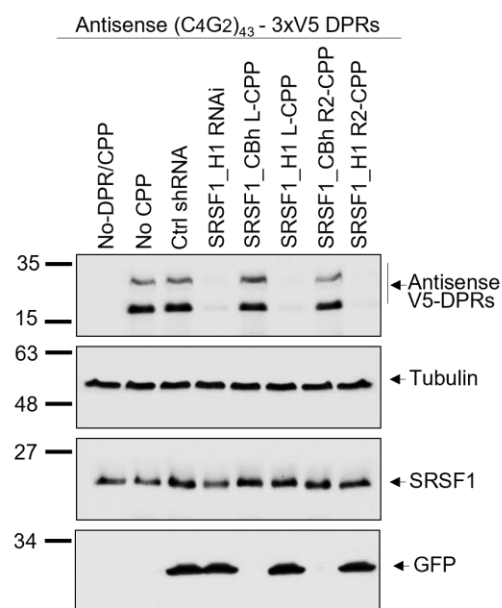
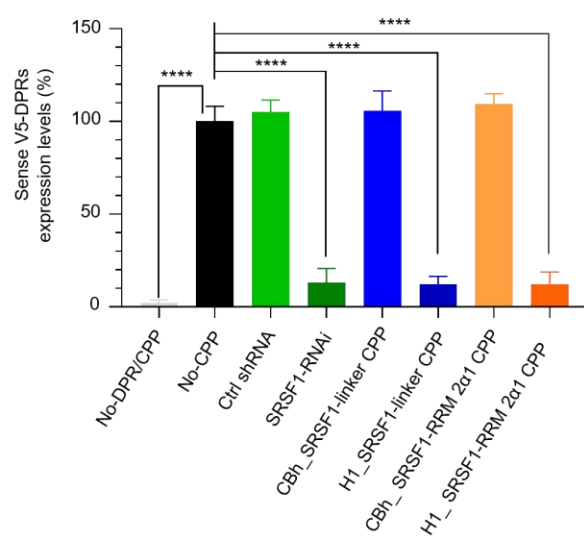
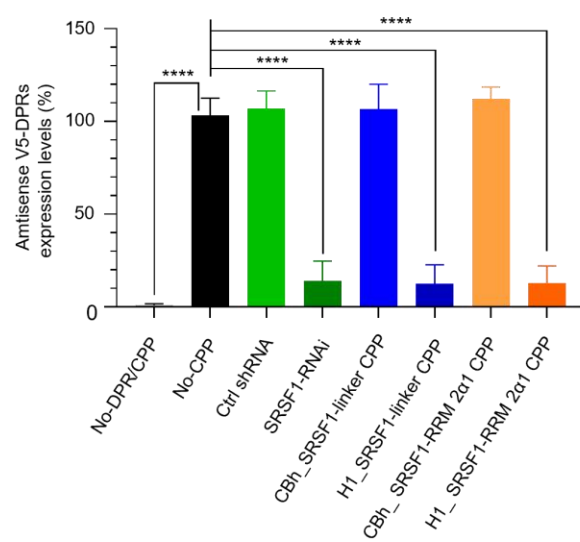
A**B****C****D**

Figure 5.6. scAAV_H1 SRSF1-linker and scAAV_H1 SRSF1-RRM2 α 1 CPPs inhibits expression of sense and antisense DPRs. (A) Representative western blot images from HEK cells co-transfected with either scAAV_CBh SRSF1-linker CPP or scAAV_H1 SRSF1-linker CPP or scAAV_CBh SRSF1-RRM2 α 1 CPP or scAAV_H1 SRSF1-RRM2 α 1 CPP along with (G₄C₂)₄₅-V5 for 72 hours. Blots were probed for V5 and tubulin, SRSF1 and GFP antibodies. **(B)** Representative western blot images from HEK cells co-transfected with either scAAV_CBh SRSF1-linker CPP or scAAV_H1 SRSF1-linker CPP or scAAV_CBh SRSF1-RRM2 α 1 CPP or scAAV_H1 SRSF1-RRM2 α 1 CPP along with (C₄G₂)₄₃-V5 for 72 hours. Blots were probed for V5 and tubulin, SRSF1 and GFP antibodies. **(C)** The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. The quantification of scAAV SRSF1-linker and RRM2 α 1 plasmids with CBh and H1 promoters transfected with (G₄C₂)₄₅-V5 into HEK cells analysed relative to α -tubulin loading (n=3, mean $\pm$ SEM; one-way ANOVA with Tukey post-hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001). **(D)** The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. The quantification of scAAV SRSF1-linker and RRM2 α 1 plasmids with CBh and H1 promoters transfected with (C₄G₂)₄₃-V5 into HEK cells analysed relative to α -tubulin (n=3, mean $\pm$ SEM; one-way ANOVA with Tukey post-hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001).

5.3.3.3. SRSF1-linker and RRM2 α 1 CPPs Expressed Under H1 Promoters Reduce Sense (poly-GA, GP, GR) and Antisense (poly-GP, PR, PA) DPR-associated Toxicity

The previous results (section 5.3.2.2) suggests that the CPP plasmids containing the H1 promoter reduce toxic sense and antisense DPR expression. Therefore, I wanted to further investigate whether CPPs expressed under an H1 promoter are able to rescue DPR-associated toxicity. HEK cells were co-transfected with all designed plasmids expressing CPPs either under H1 or CBh promoters along with either G4C2₄₅-3xV5 or G2C4₄₃-3xV5 repeat transcript constructs. It is important to note that same number of cells (25,000) were seeded into 24-well plates prior to performing the transfection. The ctrl-shRNA and SRSF1-RNAi were used as negative and positive controls respectively.

MTT assays were performed at 72-hours post transfection. Transfection of (G₄C₂)₄₅-3xV5 or (C₄G₂)₄₃-3xV5 repeat transcript plasmids alone significantly decreased cell proliferation by ~ 80% but no cell death was observed (**Figure 5.7.A and 5.7.B**). The expression of ctrl-shRNA and linker and RRM2 α 1 CPPs under CBh promoters resulted in no rescue effect on cell proliferation in comparison to the absence of DPR treatment. However, cell proliferation increased by ~ 3 to 4-fold upon co-transfection with SRSF1-RNAi or plasmids expressing linker and RRM2 α 1 CPPs under the H1 promoter which aligns with the western blot data (see **section 5.3.2.2**). The rescue effect was evaluated by dividing the averaged results of all transfected plasmids by the averaged results from the untransfected lane.

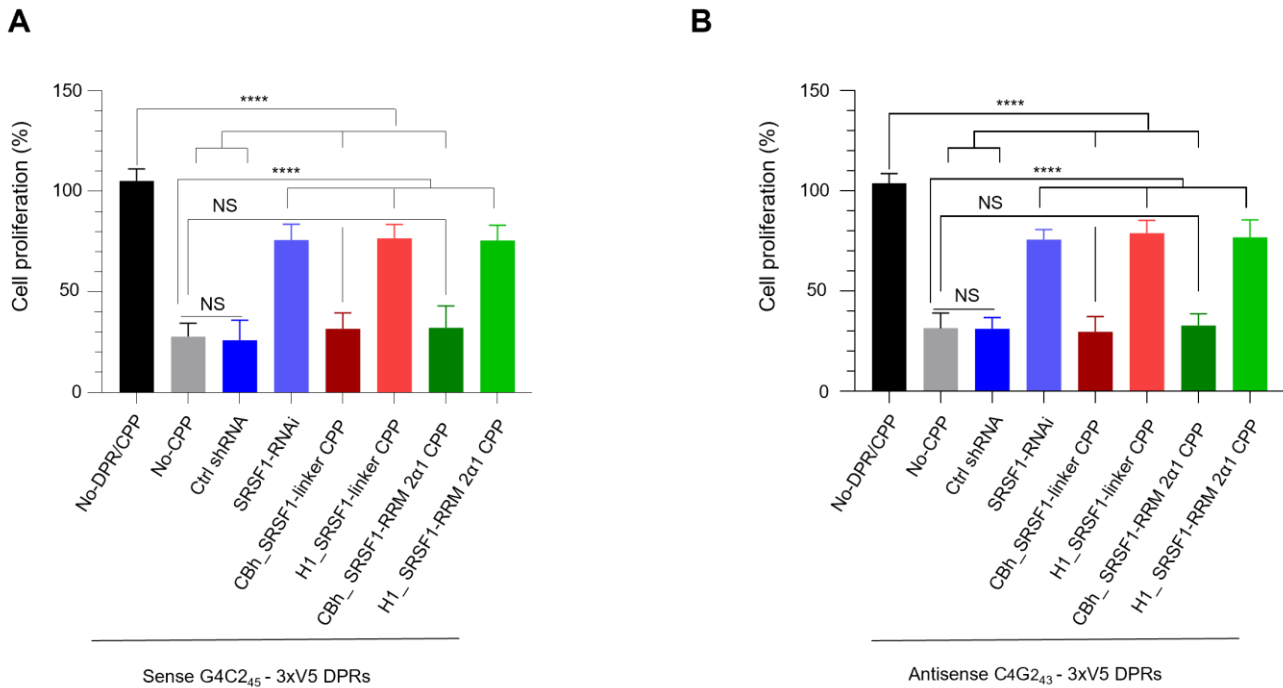


Figure 5.7. scAAV_H1 SRSF1-linker and scAAV_H1 SRSF1-RRM2α1 CPPs rescue sense and antisense DPRs-associated toxicity. (A) MTT cell proliferation assay from HEK cells co- transfected with either scAAV_CBh SRSF1-linker CPP or scAAV_H1 SRSF1-linker CPP or scAAV_CBh SRSF1- RRM2α1 CPP or scAAV_H1 SRSF1-RRM2α1 CPP along with (G₄C₂)₄₅ -V5 for 72 hours. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. The quantification of scAAV SRSF1-linker and RRM2α1 plasmids with CBh and H1 promoters transfected with (G₄C₂)₄₅ -V5 into HEK cells were analysed relative to No-DPR/CPP lane (n=3, mean ± SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001). **(B)** MTT cell proliferation assay from HEK cells co- transfected with either scAAV_CBh SRSF1-linker CPP or scAAV_H1 SRSF1-linker CPP or scAAV_CBh SRSF1-RRM2α1 CPP or scAAV_H1 SRSF1-RRM2α1 CPP along with (C₄G₂)₄₃ -V5 for 72 hours. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. The quantification of scAAV SRSF1-linker and RRM2α1 plasmids with CBh and H1 promoters transfected with (C₄G₂)₄₃ -V5 into HEK cells were analysed relative to No-DPR/CPP lane (n=3, mean ± SEM; one-way ANOVA with Tukey post- hoc's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001).

5.3.4. In Vivo Validation of The Functionality of scAAV9 Viruses

5.3.4.1. Sex Determination of C9-500 BAC mice

Sex differences were considered during the design of the study and therefore only female mice were included for the efficacy study. However, so as not to waste male C9-500 mice, they were used for a short-term safety study to show that injection of our therapeutic treatments was safe in both sexes. To identify sex differences of neonatal C9-500 mice, DNA samples were extracted from tail clippings and a PCR was performed. This was followed by running of the PCR products agarose gel (**Figure 5.8.A and 5.8.B**) An expected band at ~200 bp was observed on the gel indicating that these pups were male, and no bands were expected in females.

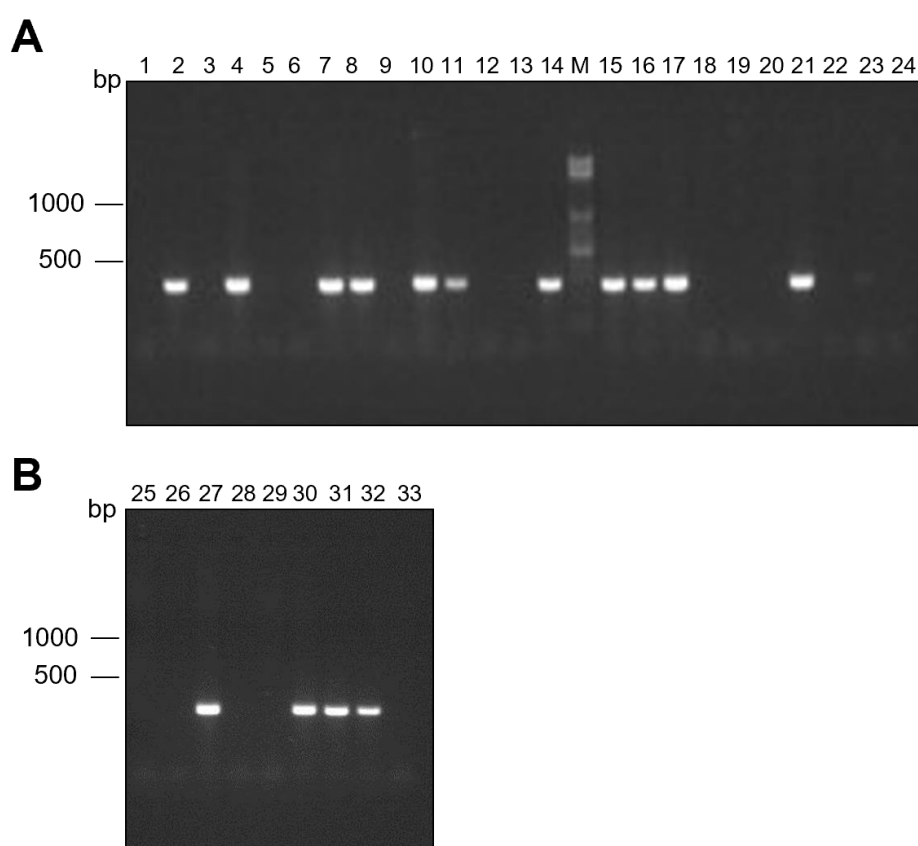


Figure 5.8. PCR-based sexing analysis using sex 1 sry_Fw and sex 1 sry_Rev primers for gender identification of neonatal mice at P1 using DNA extracted from the tail. Lanes 1-33 (A-B) are the genomic DNA samples from P1 mice. A single PCR product was expected at 200bp indicating males. Lanes 2 ,4 ,7 ,8 ,10 ,11 ,14 ,15 ,16 ,17 ,21 ,23 ,27 ,30 ,31 ,32 are from males. Lanes 1, 3, 5, 6, 9, 12, 13, 18, 19, 20, 22, 24, 25, 26, 28, 29, 33 are from females. Molecular weight marker bands are indicated in bp. N=1; no statistical analysis was performed.

5.3.4.2. Genotyping of C9-500 BAC mice

Following gender determination, C9-500 BAC mice were genotyped to determine non-transgenic and transgenic groups prior to administration of the viruses. DNA extracted from tail clippings from each litter was used to perform a genotyping PCR. The PCR products were resolved on agarose gels (**Figure 5.9.A and 5.9.B**). The gels show a single band at ~314 bp indicating transgenic mice while no band is expected for non-transgenic mice.

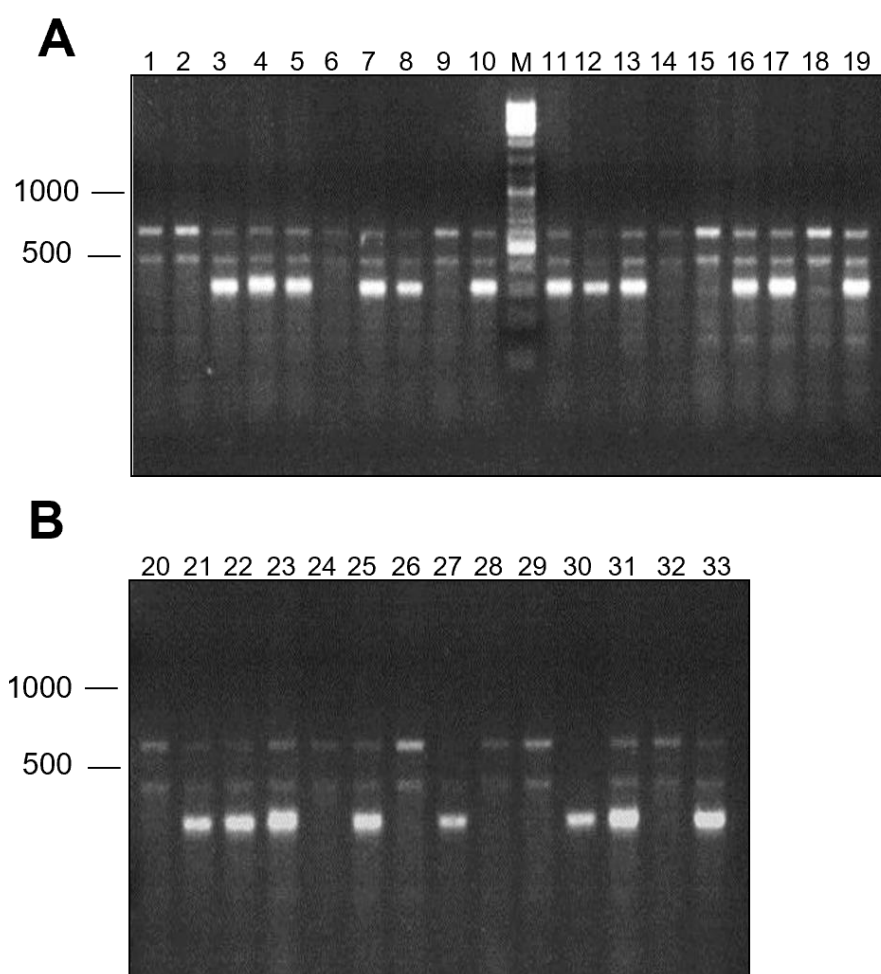


Figure 5.9. PCR-based genotyping analysis using C9-GT-F, C9-GT-R, Vgll4-F, Vgll4-R primers for genotyping of neonatal mice at P1 using DNA extracted from tail. Lanes 1-33 (A-B) are the genomic DNA samples from P1 mice. PCR product expected at 314bp indicating transgenic mice. Lanes 1, 2, 6, 9, 14, 15, 18, 20, 24, 26, 28, 29, 32 are from non-transgenic mice. Lanes 3, 4, 5, 7, 8, 10, 11, 12, 13, 16, 17, 19, 21, 22, 23, 25, 27, 30, 31, 33 are from transgenic mice. Molecular weight marker bands are indicated in bp. N=1; no statistical analysis was performed.

Having completed identification of gender (see **section 5.3.3.1**) and genotyping (see **section 5.3.3.2**) analysis of mice litters, two viruses that inhibit DPRs *in vitro* (see **section 5.3.2.2**) were injected intrathecally into the cisterna magna of post-natal day 1 or day 2 C9-500 BAC transgenic mice. One virus co-expressed SRSF1-linker CPP + GFP under the control of an H1 promoter, and the other co-expressed SRSF1-RRM2 α 1 CPP + GFP under control of an H1 promoter. In parallel to these approaches, viruses expressing SRSF1-RNAi or GFP were injected as positive controls. Non-transgenic and transgenic untreated mice did not receive any form of injection or exposure to the anaesthetic procedure. After injection, the mice were monitored daily for the first week and then every 2 days until 4-weeks post injection was completed. At the end of the study, all tissues were harvested and stored for analysis of poly (GP) DPRs, CPP biodistribution, and histopathology.

5.3.4.3. Determination of Repeat Size in C9ORF72-ALS/FTD Mice

Southern blotting is a common and reliable technique that can be used to identify the size of *C9orf72* repeat expansions and determine whether pathological repeat transcripts are present in DNA samples extracted from brain tissues from mice utilised in this study for performing biodistribution, histopathological and MSD ELISA. Although we identified genetic identity of new-born mice by genotyping PCR prior to injection, the fundamental question still arose as to whether each transgenic animal in this study is still carrying the *C9orf72* repeat expansion. To address this, DNA samples were first extracted from brain samples and subjected to digestion with AluI and DdeI enzymes to obtain DNA fragments containing the repeat expansions. The cut DNA fragments were resolved on an agarose gel and transferred onto a membrane. DNA fragments on the membrane were covalently immobilized via UV-crosslinking. The membrane was hybridised with a (G₄C₂)₅ rich probe labelled with digoxigenin (DIG) prior to incubation with an anti-DIG secondary antibody. The membrane was exposed on an X-ray film and then placed in development solution until the bands of interest were detected.

The southern blot analysis (**Figure 5.10**) shows that non-transgenic animals do not exhibit any band corresponding to a repeat expansion; however, DNA of ten mice including the positive control exhibits one single distinct expansion fragment in comparison to non-transgenic littermates indicating that all of the transgenic mice contained a pathological repeat expansion. The mice show a variety of repeat expansion lengths ranging from approximately 550 to 820 repeats. The number of G_4C_2 repeats was achieved by subtracting the size of the flanking regions on the 5' and 3' end of the repeats which are 72 bp and 182 bp from the size of the corresponding fragments in base pairs and then dividing by 6 base pairs (**Figure 5.10**).

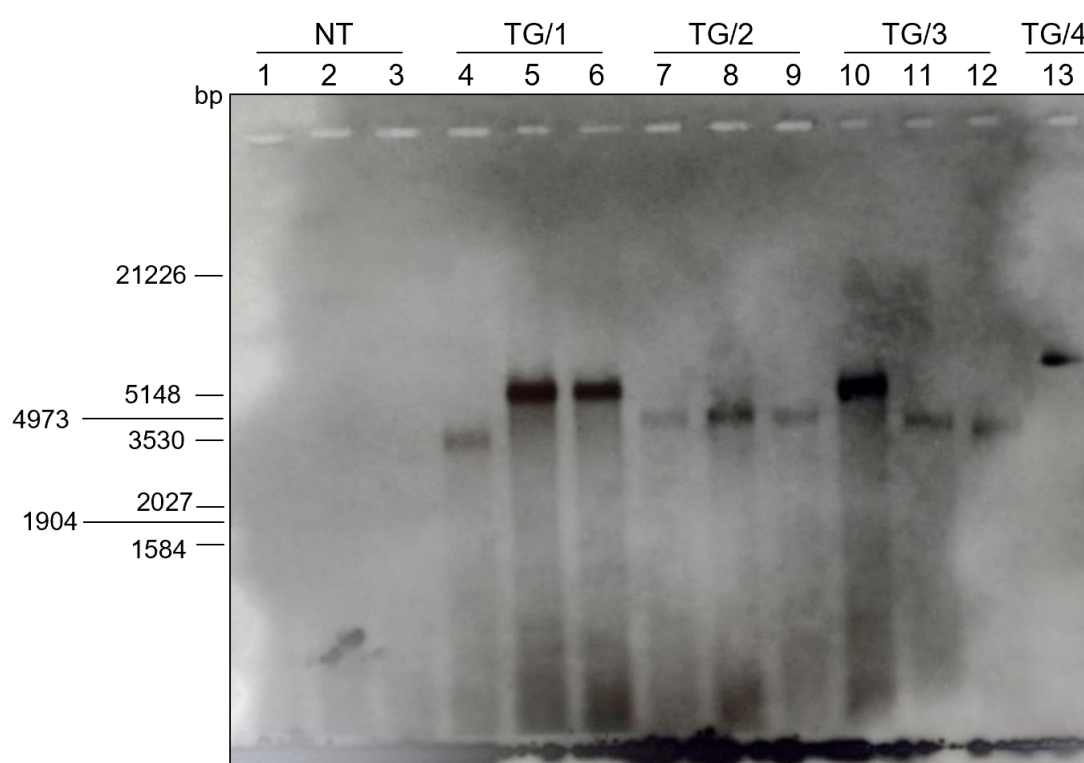


Figure 5.10. Southern blot analysis for identification of *C9ORF72* repeat expansions length in C9-500 BAC mice. Genomic DNA was extracted from whole brain and digested with AluI/DdeI enzymes then subjected to hybridization to detect the presence of GGGGCC repeat expansions using a hexanucleotide southern blot probe. Lanes show different repeat sizes ranging from approximately 350 to 500 and repeats are maintained in the animal. **Lanes (1-3)** are from non-transgenic mice. **Lanes 4-6** are from transgenic untreated animals. **Lanes 7-9** are from transgenic animals treated with scAAV9_H1 SRSF1-linker CPP. **Lanes 10-12** are from transgenic animals treated with scAAV9_H1 SRSF1-RRM2 α 1 CPP. **Lane 13** is from a transgenic animal used as a positive control. Molecular weight marker bands are indicated in bp. N=1; no statistical analysis was performed.

5.3.4.4. Biodistribution of Gene Therapy CPP Delivery in Neonatal *C9ORF72*-ALS/FTD Mice

To find out whether our therapeutic viral constructs could transduce and show higher tropism in disease related tissues, DNA was extracted from different tissue types of each experimental group and the biodistribution of the pre-clinical scAAV9 SRSF1-linker and scAAV9 SRSF1-RRM2 α 1 CPP viral vectors was assessed by qPCR. In the lab, we previously developed a qPCR assay to quantify the biodistribution of the preclinical scAAV9_SRSF1-shRNA_GFP DNA in mouse tissues using primers annealing in the GFP cassette. However, the GFP sequence required replacing with a stuffer of human origin in order to test our viral vectors in a future clinical study. Therefore, new primers were required to quantify both the preclinical and clinical vector genomes in mouse tissues. Primer pairs annealing to the H1 promoter-driven SRSF1-shRNA cassette were designed to develop a further validated qPCR assay. Primer pairs annealing to the GFP coding frame (GFP), the BGH terminator sequence of the GFP expression cassette (BGH), or the H1 promoter driven SRSF1-shRNA cassette (H1cas2) were investigated using standard curves generated from 9 serial dilutions of the original plasmids. The R^2 values of the standard curves show that assay linearity is achieved over a large range of vector genome copy numbers (1×10^1 to 1×10^9 molecules) (**Figure 5.11.A**).

The next step in the study was to conduct a pilot assay to determine the distribution of scAAV9_SRSF1 RNAi_GFP in brain tissue using all 3 primer pairs. The assay involved injecting low and high doses of the scAAV9 SRSF1 RNAi_GFP into the brain. Mice with low and high biodistribution were chosen here to highlight whether all 3 primer pairs provide the same quantification. The results indicate that the H1cas2 and BGH primers yield similar vector genome copy numbers, which are higher than those obtained using GFP primers (**Figure 5.11.B**). This confirms the validity of the assay and the new primer pairs, which bind to the H1 promoter-driven SRSF1-shRNA cassette (H1cas2). The viral distribution of scAAV9_SRSF1-shRNA_GFP administered at P120 in *C9ORF72*-ALS/FTD (C9-Tg) mice was further quantified with H1cas2 in the cerebellum/brainstem (Cb/Bs), cortex (Cx) and lumbar spinal cord (sc) one-month post-injection (**Figure 5.11.C**). Although untreated mice (red, salmon) showed non-specific background signals in the corresponding tissue, the low and high dose treatment groups showed higher tropism in the CNS regions. Therefore, for future clinical and subsequent biodistribution experiments, it was decided to use the new primer pair annealing to the H1 promoter-driven SRSF1 RNAi, SRSF1-linker or SRSF1-RRM2 α 1 CPP (H1cas2).

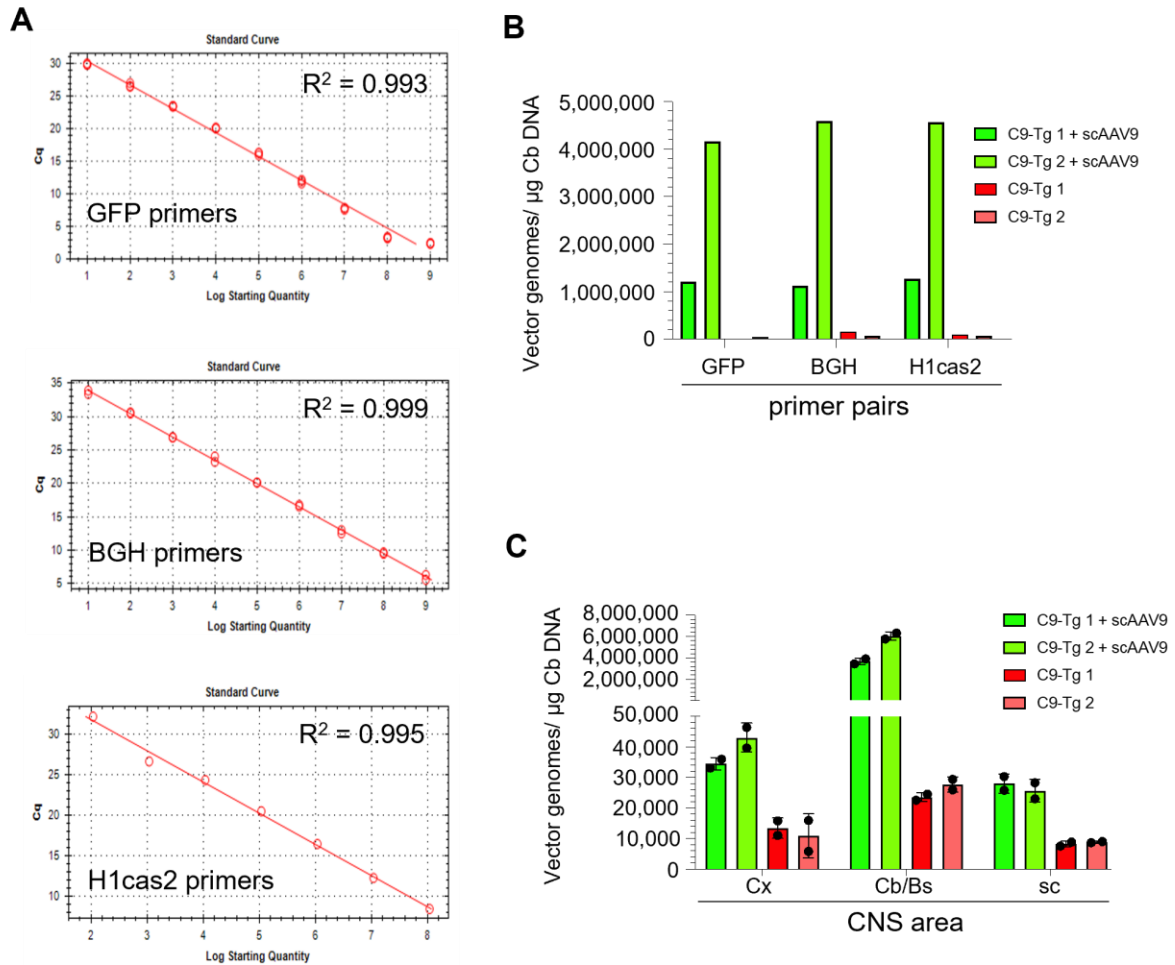


Figure 5.11. Viral genome quantification development assay and validation in *C9ORF72*-ALS/FTD mouse brains. (A) Standard curves using linearised scAAV9_SRSF1-shRNA_GFP plasmid and 3 different optimised pairs of qPCR primers. R^2 values indicate high linearity of the standard curves over a large range of copy numbers (1×10^1 to 1×10^9 molecules or vector genomes). **(B)** qPCR quantification of vector genomes (vg) in untreated C9-Tg and C9-Tg mice injected at P120 via the cisterna magna with 2.5×10^{11} vg of scAAV9. The H1cas2 primer pair was used for quantification. Mice with low and high biodistribution were chosen to demonstrate that all 3 primer pairs provide the same quantification. Untreated mice (red, salmon) show non-specific background qPCR signal in the corresponding tissues. **(C)** qPCR quantification of vector genomes (vg) in untreated C9-Tg and C9-Tg mice injected via the cisterna magna with 2.5×10^{11} vg of scAAV9 (C9-Tg 1 and 2 respectively) at P120. Mice were sacrificed one-month post-injection. Quantification was performed in the cortex (Cx), cerebellum/brain stem (Cb/Bs) and in the lumbar spinal cord (sc). The H1cas2 primer pair was used for quantification. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. No statistical analysis was performed.

Having confirmed the primer pairs, the qPCR quantification of the vector genome of scAAV9 SRSF1-linker and RRM2 α 1 CPPs was undertaken initially in the cortex and spinal cord. The H1cas2 primer pair was used for quantification. The results show that high transduction of scAAV9 SRSF1-linker CPP, scAAV9 SRSF1-RRM2 α 1 CPP was detected in cerebellum and spinal cord at 4-weeks post injection in contrast to the non-transgenic and transgenic untreated group (**Figure 5.12.A**). A low transduction level of scAAV9 SRSF1-linker and RRM2 α 1 CPPs was observed in peripheral tissues including heart, spleen, liver, muscle, lung, and kidney (**Figure 5.12.B**). Among these organs, off-target distribution of virus in the heart was higher than other organs.

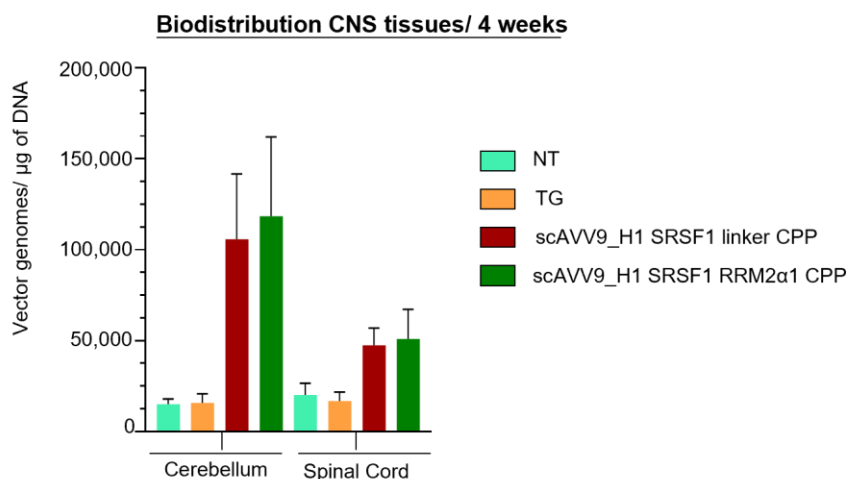
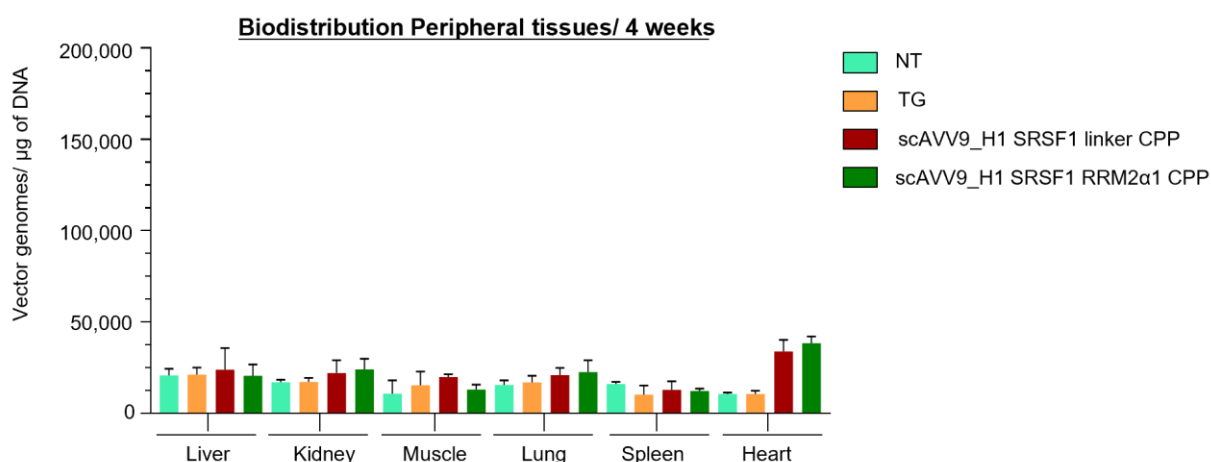
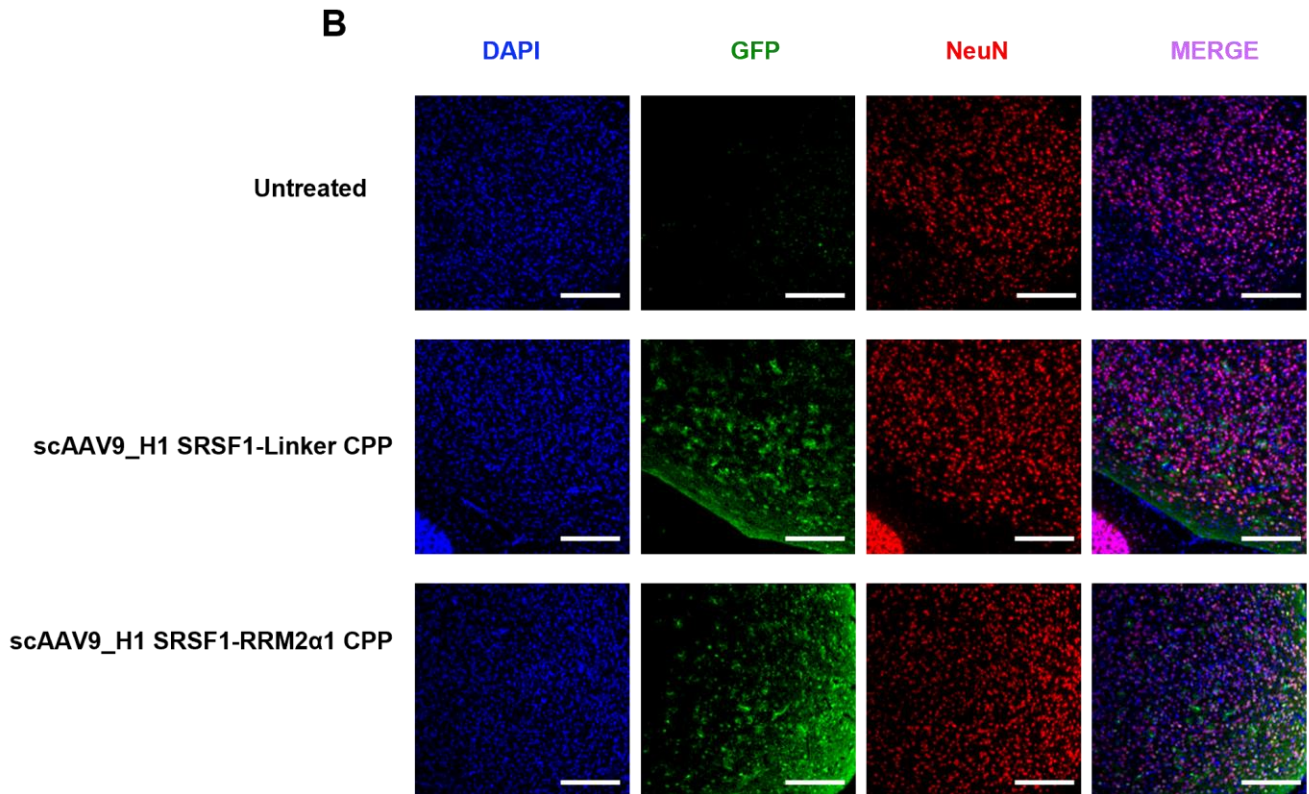
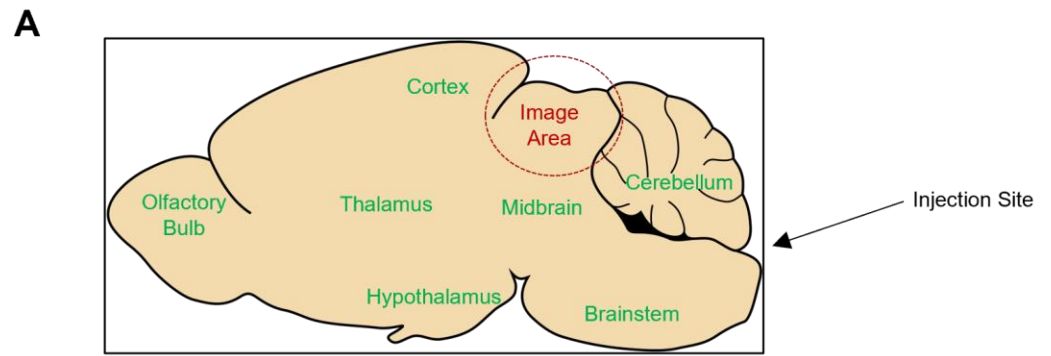
A**B**

Figure 5.12. Biodistribution analysis of C9-500 BAC mice injected with scAAV9_H1 SRSF1-linker and scAAV9_H1 SRSF1-RRM2α1 CPP. (A) qPCR quantification of DNA extracted from the CNS tissues (n=3) of 4 weeks old C9-500-BAC mice from the 2 treatment groups revealed that cisterna magna injection of therapeutic viruses was able to cross the Blood-Brain-Barrier and accumulate in the cerebellum and spinal cord. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. (B) The same qPCR quantification on DNA extracted from the peripheral tissues (n=3) of 4 weeks old C9-500-BAC mice from the 2 treatment groups. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat. No statistical analysis was performed.

5.3.4.5. Histopathological Analysis of Gene Therapy CPP Delivery in Neonatal *C9ORF72*-ALS/FTD Mice

Having shown the biodistribution of virus across different tissues by qPCR, I next performed histopathological analysis of the brains following postnatal day-one cisterna magna injection of scAAV9 SRSF1-linker and RRM2 α 1 CPP viruses to evaluate the widespread transduction within the brain tissue. In order to evaluate transduction, I investigated GFP expression as a marker of transduction, in animals 30 days-post injection. Post-fixed brain from C9-500 BAC mice was embedded with sagittal orientation and sectioned at 10 μ m thickness using a microtome into a 6-slide series containing 3 sections per slide and then incubated with anti-GFP and anti-NeuN antibodies and nuclei were counterstained with DAPI.

The results revealed by histopathological analysis showed that all scAAV9 SRSF1 viruses injected into mice were able to express GFP in the midbrain region of the brain compared to untreated controls. The location of the GFP signal was detected in regions close to the injection site such as cerebellum and cortex however it was predominantly localized in the peripheral area of the midbrain, as illustrated in **Figure 5.13.A**. GFP expression within the midbrain confirmed the successful scAAV_H1 SRSF1 linker and scAAV_H1 SRSF1 RRM2 α 1 viral transductions (**Figure 5.13.B**). To further characterise the cell population transduced by the viruses, tissue was co-stained with NeuN. The double-stained immunofluorescence images revealed that some of the GFP-positive brain cells were identified as neurons, but there was also a subset of non-neuronal cells that expressed GFP, as indicated in the enlarged images (**Figure 5.13.C**). The morphology of these cells was described as resembling microglial cells with oval-shaped nuclei, minimal cytoplasmic stain, and slender elongated processes.



C

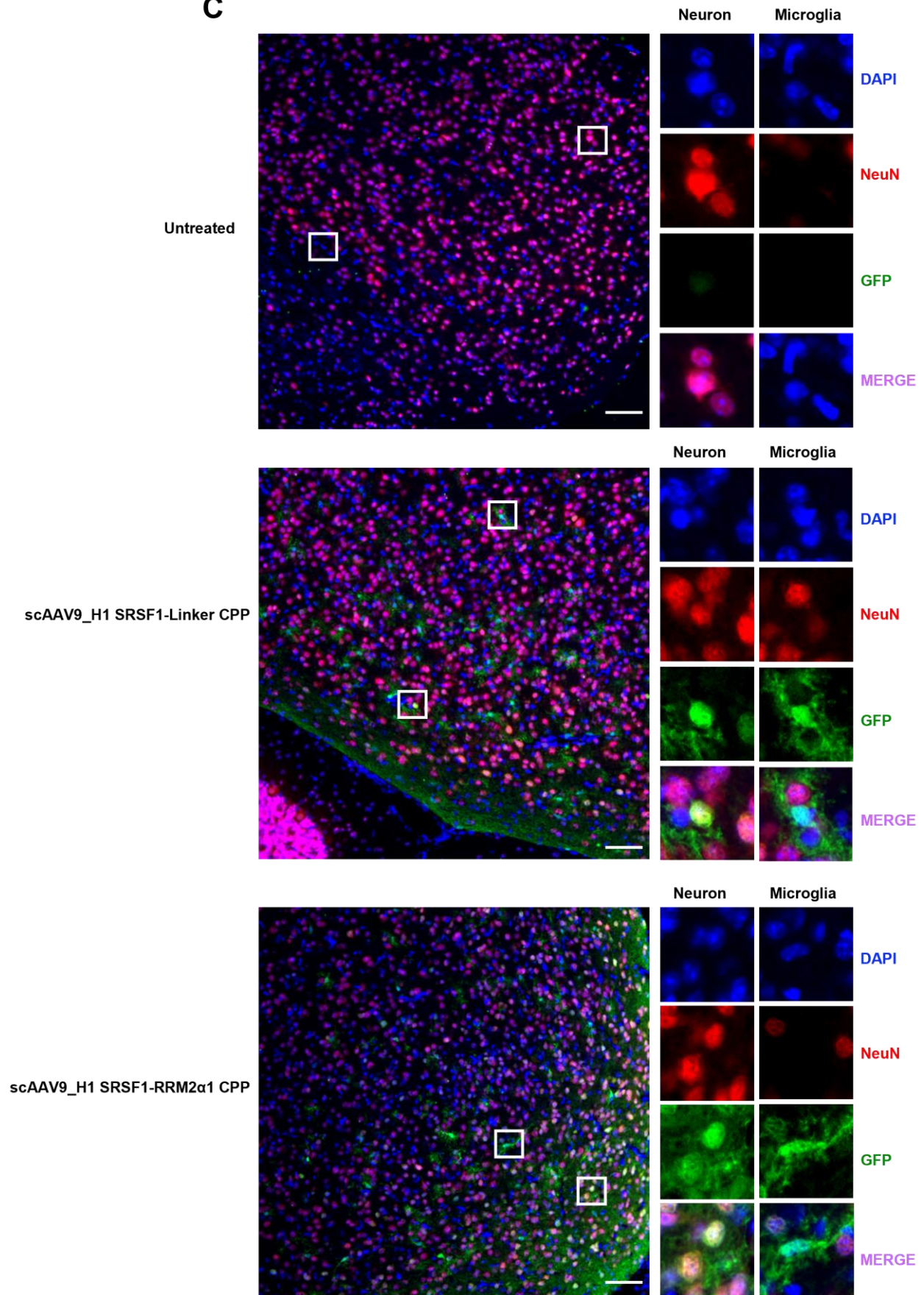


Figure 5.13. C9ORF72-ALS/FTD mouse brain biodistribution of scAAV9_H1 SRSF1-linker and scAAV9_H1 SRSF1-RRM2 α 1 CPP one month post injection in neonatal mice. (A) Schematic illustration of Mouse brain anatomy showing brain regions, immunofluorescence image area and injection site of scAAV9_H1 SRSF1-linker CPP or scAAV9_H1 SRSF1-RRM2 α 1 CPP. **(B)** Immunohistochemical analysis of GFP expression in P1 neonatal mice injected intrathecally with either scAAV9_H1 SRSF1-linker CPP or scAAV9_H1 SRSF1-RRM2 α 1 CPP at a dose of 5×10^{10} vector genome (vg). Animals were sacrificed one month post injection prior to anti-GFP immunofluorescence microscopy in brain. Representative images were captured on the sections of midbrain. GFP expression is displayed as green staining. DAPI (blue channel) and NeuN (red channel) are used to stain nuclei and neurons respectively. Scale bars represent 500 μ m. In total, 12 images from midbrain taken from 4 mice. **(C)** Enlarged fluorescence images of GFP expression in neuronal and microglial cells. All images in this panel are generated from **Figure 5.12.B**. GFP expression is displayed as the green channel. DAPI (blue channel) and NeuN (red channel) are used to stain nuclei and neurons respectively. Boxes indicate examples of magnified cells with neuronal and microglial morphology. Scale bars represent 500 μ m. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat.

5.3.4.6. Therapeutic Efficiency of Gene Therapy CPP Intervention in *C9ORF72*-ALS/FTD Mice

Having confirmed that the scAAV9 SRSF1-linker and RRM2 α 1 CPPs exhibited higher tropism specifically within the CNS, the next aim was to investigate the therapeutic efficacy of SRSF1-CPP in C9-500 BAC mice. The viral vectors were injected into the cisterna magna between post-natal day 1 and day 2 at a dosage of 5×10^{10} vg/animal (scAAV9 SRSF1-linker, scAAV9 SRSF1-RRM2 α 1), 4×10^{10} vg/animal and 8×10^{10} vg/animal (scAAV9 SRSF1 RNAi, positive control, 6×10^{10} vg/animal (scAAV9 Ctrl-shRNA, negative control). The cortex region of the brain was extracted from each experimental group and the level of poly (GP) was quantified by MSD ELISA. Poly-GP DPRs in cortices from P1 *C9ORF72*-ALS/FTD mice were injected with 4×10^{10} and 8×10^{10} vg of scAAV9_SRSF1-RNAi.

No reduction was observed 1 month post injection (**Figure 5.14.A**) and 2 months post injection are required to detect a statistically significant reduction of DPRs by approximately 50 % (data not shown). However, total poly (GP) DPR level detected in the scAAV9 SRSF1-linker CPP treatment group were reduced by ~33 % compared to non-treated C9-500 BAC non-transgenic and transgenic mice. Similar to the linker CPP data, treatment of mice with scAAV9 SRSF1-RRM2 α 1 CPP resulted in ~ 30 % reduction in poly (GP) protein level at 4 weeks post injection compared to the control groups (**Figure 5.14.B**). It should be noted that protein samples analysed by MSD ELISA were taken from the same regions of the cortex for all the mice.

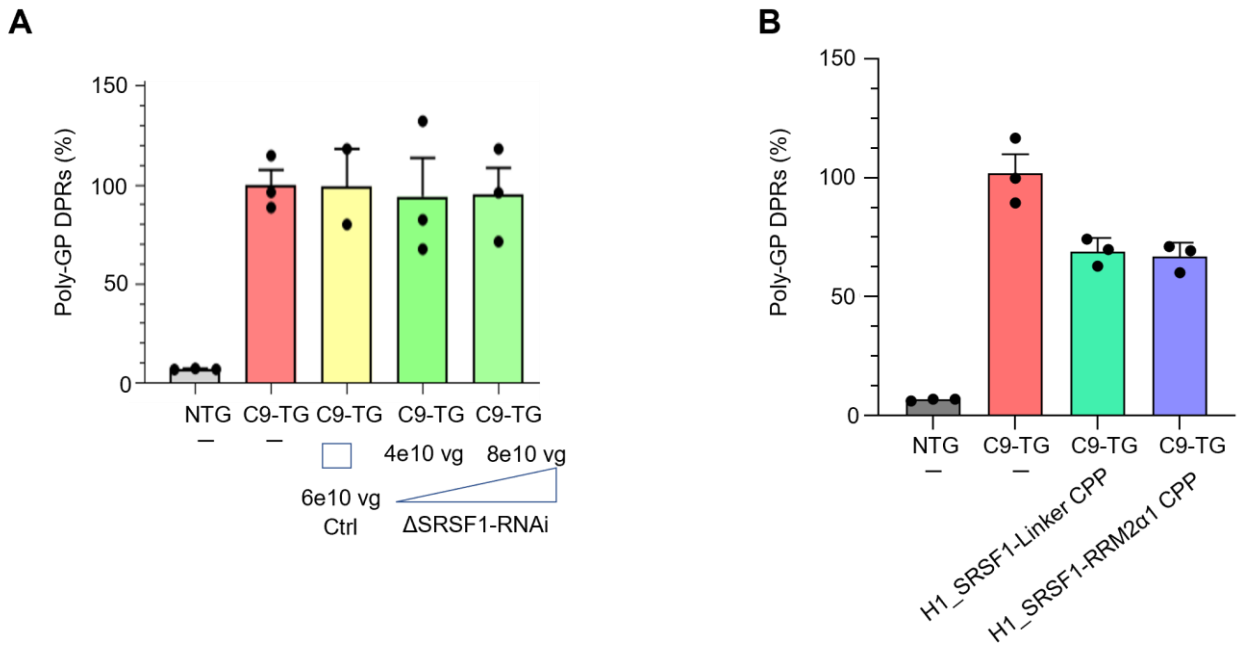


Figure 5.14. Quantification of poly-GP DPRs in mouse brains at 1-month post-injection at P1 C9-500 BAC Mice. **(A)** Quantification of Poly-GP DPRs using Meso Scale Discovery (MSD)-ELISA assays in the cerebellum/brainstem of non-transgenic (NTG), untreated transgenic *C9ORF72*-ALS/FTD (C9-500) and transgenic *C9ORF72*-ALS/FTD mice administered with scAAV9_Ctrl-shRNA_GFP (Ctrl) or 2 doses of scAAV9_SRSF1-shRNA_GFP (Δ SRSF1) at post-natal day 1 injection. **(B)** MSD-ELISA quantification of poly-GP DPRs in the cerebellum/brainstem of NTG, untreated transgenic *C9ORF72*-ALS/FTD (C9-500) and transgenic *C9ORF72*-ALS/FTD (C9-500) mice received intrathecal injections with scAAV9_H1 SRSF1-linker and scAAV9_H1 RRM2 α 1 CPP at post-natal day 1. Bar chart indicates poly-GP DPRs in brain tissue homogenates. The experiment was repeated biologically 3 times with 3 technical replicates for every biological repeat, (n=3, mean $\pm$ SEM; one-way ANOVA with Tukey post-hoc's correction for multiple comparisons, NS: non-significant). Poly-GP DPRs were quantified against a standard curve established with a GPx7 peptide and levels normalized to 100 % for the untreated C9-500 mice.

5.4. Discussion

I previously described the promising therapeutic potential of the recombinant or synthetic SRSF1-linker CPP in *C9ORF72* HEK cell model, *Drosophila* and *C9ORF72*-ALS/FTD mice (see **Chapter 4**). However, for their delivery as therapeutic molecules, they suffer from two serious shortcomings: their poor metabolic stability and short half-life. These cause a loss of their secondary structure, and thus reduced activity versus time (Al Musaimi et al., 2022). In light of this, I explored the half-life of a recombinant or synthetic peptide within HEK cell culture medium and observed peptide degradation occurring approximately 24 hours post-introduction (**Figure 5.1**). We do not know the exact reason of degradation, but degradation of peptides can proceed via a multitude of routes, broadly categorized into chemical and physical pathways. The chemical degradation of peptides arises from modifications that either form or break covalent bonds, culminating in the generation of distinct chemical entities (Manning et al., 2010). Such alterations can be realized through mechanisms like oxidation, hydrolysis, β -elimination, deamidation, racemization, isomerization, and disulfide shuffling. Conversely, physical degradation concerns the perturbations in the non-covalent associations within peptides, which can induce changes in their secondary structure, drive adsorption, promote aggregation, or lead to precipitation (Nugrahi et al., 2023). Furthermore, therapeutic peptides encounter several obstacles when it comes to their administration. Notably, peptides are typically vulnerable to digestive enzymes and possess a restricted capacity to traverse intestinal membranes, resulting in diminished bioavailability following oral intake (Nugrahi et al., 2023).

Concerns over the rate of peptide turnover and feasibility of the injection route were considered. To address these concerns, I designed an scAAV9 viral vector continuously expressing the SRSF1-linker CPP following a single intrathecal injection into the mice, thereby eliminating potential issues and providing long-term expression. We have previously generated another SRSF1-CPP construct called SRSF1-RRM2 α 1 in the lab, which is able to bind directly to the pathological repeat RNA. In the present chapter, my aim was to evaluate the therapeutic potential of an scAAV9 viral vector co-expressing the linker and RRM2 α 1 peptides in *in vitro* and *in vivo* *C9ORF72*-ALS/FTD models. To achieve these aims, plasmids expressing SRSF1-linker and RRM2 α 1 CPPs were initially designed under RNA polymerase III H1 (**Figure 5.2**) and RNA polymerase II CBh (**Figure 5.3**) and cloned into a backbone of an scAAV vector.

The design of a single vector construct expressing a small molecule such as the SRSF1-linker CPP (6.5 kDa) is technically challenging as short RNAs are not easily translated into protein. Also, there have been no previous reports describing this method. Therefore, several steps were considered to enhance the transduction efficiency of the virus both *in vitro* and *in vivo*. The first step was to use an EF1 α (Elongation Factor 1 alpha) core promoter, also known as the EFS promoter. The EFS promoter is commonly used in the design of adeno-associated viral (AAV) vectors for gene therapy. It provides several advantages to the design of AAV9 viral vectors. The EFS promoter has been shown to drive consistent transgene expression in different animal models and across different cell types, which is important for achieving consistent therapeutic effects (Heckl et al., 2014). One study used retroviral vectors for stable transgene expression in human induced pluripotent stem cells (iPSCs). However, the efficiency of retroviral transduction and transgene expression can vary depending on the specific promoter used. Therefore, the authors investigated transgene expression of a CBX3 vector with and without EFS and SFFV promoters. They found that, in comparison to vectors containing the SFFV promoters, cells that were stably transduced with vectors containing the EFS promoter showed a significant increase in EGFP expression, with a minimum of 4-fold higher levels measured by mean fluorescence intensity. Additionally, the RNA expression levels were at least 2-fold higher in cells transduced with the EFS promoter. They also examined the silencing effect of the promoter during differentiation. To do so, they differentiated stably transduced iPSCs into endothelial cells. The results showed that the CBX3.EFS configuration maintained stable transgene expression, while EGFP could not be detected in cells with the CBX3.SFFV promoter configuration (Hoffmann et al., 2017). The EFS promoter provides low toxicity, which is important for minimizing side effects associated with gene therapy, and long-term robust transgene expression in the target cells (Booth et al., 2016; Poletti and Biffi, 2019). Indeed, the EFS promoter has been effectively used in human clinical trials for the genetic disorder X-linked severe combined immune deficiency. These trials were registered under the NCT03601286 and NCT03311503 (Poletti and Biffi, 2019). Also, according to the study by Ackermann et al., the use of the full-length A2UCOE in combination with the EFS promoter resulted in stabilization of transgene expression (Ackermann et al., 2014). Taken this evidence together, I decided to use the EFS promoter during the design of our viral vectors for providing stable SRSF1-linker and RRM2 α 1 CPP expression.

Conventional single stranded or self-complementary AAV9 virus (scAAV) vectors were considered for the present study. It is known that scAAV9 vectors have several advantages over traditional AAV vectors in terms of their ability to transduce cells and their potential for use in gene therapy. scAAV9 provides high transduction efficiency, broad tropism and can transduce both dividing and non-dividing cells, including neurons and glia, which makes it an attractive vector for use in the nervous system (McCarty et al., 2001). Additionally, being self-complementary, it does not require a helper virus for replication or packaging, which reduces the risk of insertional mutagenesis and host immune responses (McCarty et al., 2001). One report investigated the effectiveness of Anc80L65 in delivering genetic material to the CNS compared to traditional AAV9 and scAAV9 using intravenous, intracerebroventricular, and intraparenchymal injection methods. Immunostaining results showed that after administering Anc80L65 intravenously, the number of cells that displayed the EGFP signal was notably higher compared to AAV9 (which had the lowest EGFP signal) but lower than scAAV9 (which had the highest EGFP signal) seen throughout the whole brain tissue. They also quantified the ability of Anc80L65, AAV9, and scAAV9 to deliver genetic material after intravenous injection. The percentages of transduced astrocytes and neurons in the cortex were analysed and the number of double-positive EGFP/GS or EGFP/NeuN cells among all astrocytes or neurons were calculated. The results showed indicate that a single dose of Anc80L65 was able to transduce 6.8 % of neurons and 26.7 % of astrocytes in the cortical mantle, a significantly greater transduction compared to AAV9 (2.3 % and 6.2 % respectively) although not as high as scAAV9 (11.9 % and 45.9 % respectively) (Hudry et al., 2018). Another study examined factors affecting the transduction efficiency of AAV9 in the CNS of adult mice such as the dose, the component of the delivery vehicle, the co-administration of mannitol, and the use of either single-stranded or self-complementary AAV9. They initially investigated the CNS transduction ability of AAV9 in comparison to other AAV serotypes (AAV1, AAV5, AAV6 and AAV8). The results showed that AAV9 has superior transduction efficiency in the CNS compared to other AAV serotypes. In order to evaluate the ability of single-stranded and self-complementary AAV vectors to deliver larger transgenes to the brain, ssAAV9 and scAAV9 viral vectors expressing GFP were injected into the tail vein of adult mice, and the GFP expression was analysed in the brain and spinal cord using IHC. Transduction of cells with neuronal and glial morphology was observed in the brain and spinal cord after ssAAV9 but it was substantially lower than the levels observed with scAAV9. The number of GFP-positive cells was similar to that of a significantly lower dose of 2.5×10^{10} vg of scAAV9, illustrating the superior efficiency of scAAV9 over ssAAV9 (Gray et al., 2011).

Similar observations were also made in previous studies demonstrating significantly higher transgene expression obtained via scAAV9 compared to ssAAV9 (Wang et al., 2003) (McCarty, 2008). The scAAV vector system, despite its advantages, also has certain drawbacks. The main disadvantage is the restricted packaging capacity, caused by the limited size of the AAV genome. The genome capacity of scAAV (2.4 kb) is lower than ssAAV which is around 4.7 kb. The genome capacity that scAAV can deliver can be raised to 3.3 kb, but this results in a substantial number of genomes being packaged in a single-stranded form. Therefore, the utilization of this vector system is limited to only a select number of diseases. The scAAV system has another limitation (Raj et al., 2011). The high gene expression efficiency of scAAV may trigger a stronger immune response *in vivo*. This could potentially have negative effects on the treatment outcome and require additional measures to mitigate the immune response (McCarty et al., 2001).

After successfully cloning the SRSF1-linker and RRM2 α 1 CPP sequences into the vector backbone of scAAV, the constructs' efficacy were validated first in HEK293T cells in order to move subsequent animal study. High content immunofluorescence imaging showed that significant amount of linker and RRM2 α 1 peptide expression was observed in plasmids containing an H1 promoter (**Figure 5.4**). Next, I assessed whether the designed CPP plasmids/constructs were able to reduce sense and antisense DPR production. Constructs expressing the linker or RRM2 α 1 peptides under the control of an H1 promoter reduced the expression of sense and antisense DPRs compared to ctrl-shRNA transfected and untransfected cells, as expected. It was also reported that plasmids with CBh promoters did not reduced sense and antisense DPR production. This indicates that the amount of expressed peptide under control of a CBh promoter is insufficient to inhibit SRSF1/NXF1 dependent nuclear export of pathogenic repeat transcripts into the cytoplasm and subsequent DPR expression by either binding repeat RNA for RRM2 α 1 CPP or binding the endogenous SRSF1 protein for linker CPP (**Figure 5.5**). The final *in vitro* functional validation study was MTT cell proliferation assay showing that constructs driving SRSF1-linker and RRM2 α 1 peptide expression under the control of an H1 promoter increased in cell proliferation by ~80 % compared to untransfected and ctrl-shRNA-transfected cells. Also, in alignment with western blot data, plasmids expressing CPP under the control of a CBh promoter had no effect on cell proliferation(**Figure 5.6**).

Having validated the constructs for driving peptide expression under an H1 promoter as a potential and robust *in vitro* gene therapy, I took the study one step further by investigating their therapeutic efficacy in an *in vivo* model of *C9ORF72*-ALS/FTD. In the literature, several animal models of ALS have been generated to study the disease mechanisms and develop new treatments for *C9ORF72*-ALS/FTD. One of the commonly used models is the zebrafish, which provides several advantages such as ease of genetic manipulation, cost-effective maintenance, and rapid anatomical development (Chia et al., 2022; Shaw et al., 2018). The optical transparency of zebrafish in their embryonic or early adult stage allows real-time visualization of axon pathology and motor neuron death (Cong et al., 2017). In recent years, this model has also been used for high-throughput rapid drug screening in the CNS (Patten et al., 2017). While drug screening is often performed using the *Drosophila* model, its CNS complexity is not the same as the human CNS (Boeynaems et al., 2016; Jovičić et al., 2015), making the zebrafish model a superior choice for assessing the efficacy of therapeutic tools compared to *Drosophila*. Furthermore, the fully-sequenced zebrafish genome shows high levels of gene conservation (over 80 %) with the human genome, making it a suitable model for studying human genetic disease (Howe et al., 2013). The zebrafish model system has been used to simulate the major pathological mechanisms of *C9orf72* mutations related to ALS/FTD. For instance, Shaw and colleagues generated a novel transgenic zebrafish model that expresses RNA foci and DPRs through RAN translation of sense and antisense repeat transcripts, which are pathological hallmarks of the *C9orf72* mutation (Shaw et al., 2018). They also reported that this model is capable of demonstrating motor and cognitive dysfunction.

Moreover, a zebrafish model was generated to support the haploinsufficiency mechanism, where a knockdown of *C9orf72* resulted in motor and behavioural impairment (Ciura et al., 2013). Overall, the zebrafish model plays a valuable role in discovering and testing potential treatments for ALS and FTD. However, it has several limitations as a model for ALS. One of the main disadvantages is the absence of the corticospinal and rubrospinal tracts in its central nervous system, limiting its usefulness in studying upper and motor neuron disorders, which are characteristic of ALS (Babin et al., 2014).

Another disadvantage is the difficulty in achieving long-term transgene expression in zebrafish, which is crucial for gene therapy applications. For example, one study evaluating the use of zebrafish as a model organism to assess the transduction efficacy of adenoviral vectors expressing GFP found that while fish injected with the virus maintained GFP expression up to 32 days (Gulías et al., 2019), the rodent model showed long-term expression of over one year (Brunetti-Pierri et al., 2013). For this reason, I chose the C9-500 BAC mice model as an in vivo model for investigating the efficiency of SRSF1-CPP mediated non-viral and viral gene therapy approaches.

The C9-500 bacterial artificial chromosome (BAC) mouse line is considered to be an excellent model for studying the underlying pathogenic mechanisms of *C9ORF72*-ALS as it recapitulates many of the key features observed in human patients including paralysis, neurodegeneration, decreased survival, and expression of RNA foci and DPRs. This C9-500 BAC mouse model developed on an FVB background contains a whole human *C9orf72* gene sequence with ~ 500 hexanucleotide repeat expansion and important upstream and downstream regulatory regions of the *C9orf72* gene which were isolated from a patient-derived lymphoblastoid cell line to drive expression of sense and antisense repeat transcripts. We used this rodent model to assess the effects of our designed therapeutic gene therapy approach on poly (GP) DPR expression (Liu et al., 2016). Researchers have used this animal model to study the effects of different genetic and chemical agents such as small molecule drug inhibitors to identify potential targets for therapeutic intervention. Additionally, it has also been used to study the behaviour and pathology of the disease and to understand the correlation between the genetic mutation and the pathology.

For example, one recent example used the C9-500 BAC mouse model to investigate the efficacy of an antibody based therapeutic approach for *C9ORF72*-ALS/FTD. In this study, human recombinant antibodies against C9 RAN proteins and their chimeric variants were generated by utilizing libraries of memory B cells from healthy elderly individuals (Nguyen et al., 2020). Three of the generated antibodies (α -GA1, α -GP1, and α -GA2), that have a high binding specificity and low nanomolar affinity towards GA or GP, were able to identify RAN targets in cells and in the brains of C9 mice and humans. They initially focused on investigating the effects of the antibodies on GA protein expression in HEK293T cells. To address this, the authors treated the cells with α -GA1, α -GP1, α -GA2, or control antibodies of the IgG isotype class.

Western blot analysis showed that treatment with α -GA1 resulted in a statistically significant 50 % reduction in GA protein levels, whereas no significant changes in protein expression were observed in cells treated with α -GP1, α -GA2, or control antibodies of the IgG isotype class. After they completed all the *in vitro* functionality experiments for confirming their efficacy by several biochemical methods such as MTT, LDH and proximity ligation assays, they treated the C9-BAC ALS/FTD mouse model delivering α -GA1, α -GP1, α -GA2 antibodies via a peripheral route to examine their effects on RAN protein expression, behavioural deficits and survival *in vivo*. The result showed that following administration of the α -GA1 antibody to C9-BAC mice, a decrease in GA, GP and GR protein expression, an improvement in behavioural symptoms, a reduction in neuroinflammation and neurodegeneration and an increase in the survival rate were observed (Nguyen et al., 2020).

The C9-500 BAC mouse model has also been used to investigate the effect of metformin; a drug commonly used to manage type 2 diabetes. It was found to have a positive effect on aging and neurodegenerative disorders (Gantois et al., 2019) (Barzilai et al., 2016). It has been also observed that metformin regulates protein translation by suppressing the MAPK/Erk pathway in fragile X syndrome (Gantois et al., 2017). Mice were treated with metformin and molecular analysis showed that metformin is able to reduce expression of poly GA, GP in the cortex. Additionally, when the C9-500 mice were treated with metformin, parameters including gait, gait/stance, and gait/stride were found to be better compared to the control group. Indeed, metformin treatment was found to decrease neuroinflammation and loss of motor neurons in this model (Zu et al., 2020). Based on these studies, the C9-500 BAC mouse line was considered an appropriate model for investigating the efficacy of our potential therapeutic viral vectors as a *in vivo* proof-of-concept experiment.

C9-500 BAC transgenic mice are an exciting preclinical animal model for testing therapeutic hypotheses for *C9ORF72*-ALS/FTD. However, this model has limitations that have been reported by several research groups in terms of detecting disease-specific phenotypes. A recent study validated the presence of the repeat expansion in this model via southern blot and found that the length of the repeat expansion is stable. However, the researchers were not able to observe any changes in behavioural phenotypes relevant to ALS or FTD (Mordes et al., 2020). Concerns regarding inconsistencies in the behavioural phenotype of C9-500 BAC mice was not an issue in this study as this model recapitulates the molecular features of ALS/FTD such as the expression of DPRs.

During the experimental design, sex differences were considered. Mice gender and genotype were identified with sexing (**Figure 5.7**) and genotyping PCRs (**Figure 5.8**) and then female mice were used for injection of our gene therapy viral vectors. The reason for choosing female mice was that approximately 30-35 % of female transgenic C9-500 BAC mice have been reported to develop severe ALS/FTD disease phenotypes, such as sudden weight loss and rapid paralysis, more frequently than male mice (Liu et al., 2016). This makes female mice a more appropriate model to study the effects of the repeat expansion on the disease. Another reason is that the *C9orf72* repeat expansion is known to have a stronger effect on females, as the repeat expansion is often associated with a greater risk of developing ALS/FTD in females compared to males. This is supported by the 16 % higher incidence rate of ALS in female patients with *C9orf72* repeat expansion compared to males (Curtis et al., 2017). The same study also reported that females were found to have a 33 % higher incidence of GRN-related FTD, similar to the higher prevalence seen in *C9ORF72*-related ALS (Curtis et al., 2017). In agreement with this, it was also reported that bulbar-onset ALS with women patients has a higher frequency of rapid disease progression compared to male patients (Palese et al., 2019; Richards et al., 2020). Another study reported in the literature used C9-500 BAC mice to evaluate the effect of reduced *C9orf72* protein level on motor coordination and balance through the utilization of the rotarod test. The results of the rotarod assay indicated that there was a greater degree of motor dysfunction in the female mice compared to the male mice (Shao et al., 2019). Similar observations were made in another study that generated the C9-500 mouse line. These findings were consistent with the results of a separate study that generated the C9-500 mice line. In that study, mice with long repeats of C9-500 and C9-500/32 have decreased survival and a high incidence of death or abnormalities by the age of 12 months, particularly in females (76-82 %) compared to males (50%). In addition, female mice from the C9-36/29 line with high gene expression also showed decreased survival, but disease onset is later and less severe compared to those from long expansion lines. However, males from the C9-36/29 line do not differ significantly from non-transgenic males (Liu et al., 2016). Due to these reasons, we focused on analysing the efficacy of the gene therapy vectors in female transgenic mice, specifically by measuring their levels of poly(GP) DPRs, looking at the distribution of the gene therapy vectors and performing histological analysis. Additionally, in order not to waste resources and to make the best use of the available transgenic mice, male transgenic mice were used to demonstrate that the injection of our gene therapy vectors was safe and did not cause any adverse effects in both genders.

In this project, I also considered route of administration, which is highly linked with the efficiency and safety of AAV delivery into the central nervous system (CNS). Various administration routes have been used for delivery of gene therapy vectors for the treatment of neurological diseases. Several studies also compared the targeting efficiency of virus delivered into mice. In this regard, one study attempted to compare the transduction efficiency of five different routes for targeting dorsal root ganglia (DRG) neurons via delivering recombinant AAV (rAAV) serotype 6 into C57BL/6 mice. The strongest transduction in lumbar DRG cells was attained with intrathecal injection in comparison to subcutaneous, intramuscular, intravenous, or sciatic nerve injections. The presence of virus in the spinal cord and most brain regions was observed via intrathecal injection, suggesting an efficient route for delivering rAAV2/6 into the brain (Towne et al., 2009). A similar observation was reported in another study where intrathecal injection of scAAV9 CBh GFP into adult pigs resulted in successful and broad transduction in motor neurons within the ventral horn (Federici et al., 2012).

Due to these positive outcomes, I focused on the intrathecal delivery route for delivering our gene therapy vectors into the CNS. However, intrathecal injection of virus can be achieved with three different routes: (I) intracerebroventricular (ICV), (II) the cisterna magna (CM), and (III) lumbar puncture (LP) (Piguet et al., 2021). These three methods have been tested with several AAV serotypes in animal models to define the best route of injection by comparing their CNS targeting efficacy and biodistribution (Hocquemiller et al., 2016). For example, a single injection of haemagglutinin (HA) tagged AAV9 with co-injection of gadolinium, an MRI contrast agent, via CM, LP, ICV and CM+LP delivery routes was used to monitor the pattern of virus distribution between delivery routes with real time MRI. MRI analysis showed a higher gadoteridol signal distribution in broad regions of the brain and spinal cord via CM injection. They also reveal that the highest virus concentration was seen in CSF with CM administration followed by ICV and CM+LP (Ohno et al., 2019).

Similar observations were seen another study performed in our department (SITraN). These investigators reported that the CM injection route provided broad AAV transduction across the brains of the SOD1-ALS mice model (Iannitti et al., 2018). It is worth to mentioned that although cisterna magna injection produces widespread CNS biodistribution, it can lead to severe brain damage if the needle is misplaced, particularly as mice at this age are not fully developed anatomically (Pellot and De Jesus, 2022). This also could result in obtaining a low level transduction of virus into the target tissues as the viral vector is absorbed by other tissues surrounding the cisterna magna (Pellot and De Jesus, 2022). Overall, I chose the CM injection route for delivering scAAV9_H1 SRSF1-linker CPP and scAAV9_H1 RRM2 α 1 CPP or scAAV9_H1 SRSF1 RNAi_GFP (positive control).

After collecting all tissues from untreated and treated mice, southern blot hybridisation was conducted to determine the length of the *C9orf72* G₄C₂ repeat expansions. Although transgenic animals showed that each DNA sample contained a hexanucleotide repeat expansion, there is an ambiguous smear was observed, indicating possible somatic instability of the G₄C₂ repeat expansion (**Figure 5.9**). A similar observation has been described previously in the literature. For example, Buchman *et al* detected the presence of a blurry pattern on a Southern blot membrane. This was attributed to somatic heterogeneity in DNA from blood samples and poor-quality DNA from lymphoblastoid cells and post mortem tissue from ALS patients. However, they managed to identify the size of the pathological repeat expansion despite the southern blot result presenting in a difficult-to-interpret pattern (Buchman et al., 2013).

A similar smear pattern was also observed in another study that aimed to examine whether there is a relationship between the size of G₄C₂ repeat expansion in peripheral DNA and clinical variations of *C9ORF72*-ALS/FTD by developing semi-automated quantification method (Suh et al., 2015). One of the issues related to identifying the size of the repeat expansion is somatic instability which is a common feature of microsatellite expansion disorders (Pearson et al., 2005). Repeat expansion instability also accounts for observed differences in repeat length across various tissue types (Goula et al., 2012). For example, it has been reported that repeat expansions in cerebellum tend to be shorter than in peripheral tissue and blood. The reason underlying the instability of *C9orf72* repeat expansions is still not clear but has been attributed to problems with DNA metabolism such as during DNA replication (Liu et al., 2010). Alternatively, mismatches within repair pathways might result in instability (López Castel et al., 2010).

Similar observations were also seen in other repeat expansion diseases caused by trinucleotide repeat expansions. It was found that the length of CAG repeat expansion was shorter in cerebellum than other region of brain and peripheral tissues in the transgenic C9-500 BAC animal model and patients with Huntington's disease. The difference of repeat length in the cerebellum may be due to a higher neuronal density compared to other brain regions due to encompassing post mitotic neurons (Zhao et al., 2021). Despite the issues mentioned above in the southern blot, we were able to determine the number of G₄C₂ repeat expansions in C9-500 BAC animals used in this study.

The next step was to investigate the biodistribution of viral vectors in several mouse tissues to determine where the virus is preferentially accumulating and how it is being transported throughout the body. This information is crucial for identifying any potential off-target effects. Initially, three different qPCR primers were validated and qPCR data showed that although three primer pairs achieved high linearity and provided similar readouts of vector genome copy in brain tissue, I decided to utilise the H1cas2 primer pair as this would be more likely to apply to a future clinical vector which will not contain GFP gene and BGH terminators therefore the H1cas2 primer was chosen for viral vector quantification (**Figure 5.10**). The biodistribution of scAAV9 SRSF1-linker and RRM2 α 1 CPP viruses was assessed across different tissues from each experimental group. The qPCR quantification data showed that there is a higher distribution of viral vector was observed in the brain and spinal cord, indicating that AAV9 vectors passed the blood brain barrier and delivered efficiently into the CNS of *C9orf72* repeat expansion mice (**Figure 5.11**).

In the literature, a study utilizing quantitative PCR to analyze viral distribution following intracisternal magna and intracerebroventricular injections yielded analogous results. This research observed the presence of vector genomes across both the brain and spinal cord. In alignment with these findings, prominent transgene expression was identified in multiple regions of the cerebral cortex and throughout the spinal cord (Hinderer et al., 2018). In a recent investigation, researchers delivered intrathecal injections of AAV9 and AAV-F to nonhuman primates. They then assessed the distribution of transduced virus in the spinal cord and specific peripheral tissues, including the liver, spleen, and heart by qPCR. Observations highlighted transduced cells within regions of the cerebellum, brain stem, and hippocampus, and also across the lumbar, thoracic, and cervical segments of the spinal cord (Beharry et al., 2022).

As for viral distribution across peripheral tissues, low levels of viral genomes were detected across peripheral tissues. This could be due to redistribution of cisterna magna injected virus from the subarachnoid space into the systemic circulation; therefore, limited distribution of viral particles is seen unevenly in other tissues. The presence of viral vector in peripheral tissues was not considered significant because the viral vector predominantly accumulated in CNS tissues at a level 7.5 to 15-fold higher compared to peripheral tissues. One study reported that the distribution of the virus in the brain was stable over time; however, the distribution of viral vector in the liver, spleen, kidney, and lung was found to decrease over time, suggesting that off-target effects of our viral vectors are likely to decrease when we conduct a long-term efficacy study (Bobo et al., 2020). However, we detected a higher biodistribution of scAAV9 SRSF1-linker and RRM2 α 1 CPP viruses in heart tissue (**Figure 5.11**). Several studies reported that SRSF1 is required as a splicing factor and plays a key role in the embryonic development of the heart because knockdown of SRSF1 at the embryonic stage in tadpole, causes delay in the development of the retina and oedema in the heart and head were observed (Lee et al., 2016b). A similar observation was reported in other studies that SRSF1 plays an essential role in alternative splicing in the heart and deletion of SRp20(SRSF3), SC35 (SRSF2) and ASF/SF2(SRSF1) resulted in embryonic lethality (Xu et al., 2005). Therefore, I expected to see the distribution of our SRSF1 gene therapy viral vector in the heart. However, the presence of scAAV9 SRSF1-linker CPP and RRM2 α 1 CPPs in the heart did not result in any death or behavioural or developmental side effects.

In term of presence of viral vectors in other peripheral tissues, there are several examples in the literature of studies that have investigated the off-target effects of AAV9 in peripheral tissues. One study investigated the biodistribution of scAAV9-SMN1 at three different doses (2.5×10^{10} vg/mouse, 4×10^{10} vg/mouse, and 10^{11} vg/mouse) delivered via either ICV injection or a combination of ICV and IV injection in a mouse model of spinal muscular atrophy (SMA). They evaluated the distribution of the viral vector across different tissues, including the thoracic region of the spinal cord, liver, heart, tibialis anterior and gastrocnemius muscles, for both routes of administration.

Administration of the scAAV9-SMN1 vector via the ICV route at a dose of 4×10^{10} vg/mouse led to a high number of vector copies in the spinal cord and liver, whereas a lower dose of 2.5×10^{10} vg/mouse resulted in fewer copies in the spinal cord. Off-target effects seen in liver have been documented in several studies (Dirren et al., 2014; Inagaki et al., 2006); these studies align with our results. Both doses resulted in a low number of copies in the heart and skeletal muscles. The highest dose of 10^{11} vg/mouse produced a similar number of copies in the spinal cord and liver as the intermediate dose group, but a significantly higher accumulation of the vector in the skeletal muscles and specifically in the heart, which is likely due to the significant migration of the vector from the CSF to peripheral organs. When they assessed vector distribution in the ICV+IV route it was found that VCN (vector copy number) in the spinal cord remained constant with increasing IV doses, while VCN in the liver and heart increased with the IV vector dose up to >15 and >9 vg/dg respectively (Armbruster et al., 2016).

Another study investigated the distribution of an intrathecal administered single dose of scAAV9 expressing GFP in Cynomolgus Macaques (13 to 17 months of age), which are a widely used non-human primate model. Biodistribution of the viral DNA and GFP protein expression was quantified by digital PCR and MSD ECLIA across brain, spinal cord, and peripheral tissues. Intrathecal administration led to a high expression GFP DNA, mRNA, and protein in tissues such as liver, skeletal muscle, heart, spinal cord, and sensory ganglia neurons. However, it had minimal presence in various parts of the brain, indicating that the virus primarily spreads to systemic tissues, spinal cord, and sensory ganglia neurons. Limited presence of virus into the brain parenchyma suggests that viral vector in the CSF was primarily distributed into systemic circulation (Meseck et al., 2022).

Histopathological characterisation of scAAV9 SRSF1-linker and RRM2 α 1 CPPs injection revealed that GFP protein expression was detected in several regions of brain and neuros as well as glial cells. (**Figure 5.12**). A comparable outcome was observed in another study. In this research, they demonstrated that intracisternal injection of the scAAV vector resulted in widespread GFP expression throughout the CNS. This expression covered a broad region, extending from the olfactory area all the way down to the spinal cord. Notably, more than 50 % of the Purkinje cells in the cerebellum exhibited GFP expression. Beyond this, the study also reported robust and intense transgene expression in a variety of neurons and glial cells across the brain (Fu et al., 2003).

A similar observation was noted in another study, which reported robust GFP expression in both neuronal and glial cells following intrathecal injection of AAV9 into non-human primates (Samaranch et al., 2014). A recent study analyzed the amount of viral DNA from AAV serotypes AAV9, AAV6, and a combined AAV1&2, all of which expressed green fluorescent protein (GFP). The findings revealed that the viral vectors AAV1&2-syn-GFP and AAV6-syn-GFP were effectively transported from the bloodstream to targeted brain regions, including the hippocampus, striatum, thalamus, and cortex. This led to successful GFP transgene expression in these areas. Notably, a pronounced GFP expression was observed in the hippocampus for both vector types. A detailed analysis of the brain tissue showed GFP expression in cell bodies and their extensions in the specified regions. Importantly, GFP-positive cell bodies consistently co-existed with NeuN-positive cells (Weber-Adrian et al., 2021).

Finally, the efficacy of our SRSF1 based gene therapy vector on the expression of poly (GP) DPRs was assessed in C9-500 BAC mice. Research on poly-(GP) is of significant importance as it is translated from both the sense and antisense repeat transcripts and can be identified in the CSF of individuals with a *C9orf72* expansion via MSD ELISA, indicating that poly (GP) levels in CSF could be used as potential biomarker of disease and therapeutic efficacy (Gendron et al., 2017; Lehmer et al., 2017). Additionally, the levels of poly-(GP) DPRs are consistent and remain unchanged with the progression of the disease or over 6 to 18 months, which reinforces poly-(GP) as a biomarker for measuring the effects of therapeutic treatment in *C9ORF72*-ALS/FTD. This notion is supported by a study investigating inhibition of DPRs using an antisense oligonucleotide (ASO) targeting *C9orf72* repeat expansions in patient-derived lymphoblastoid cell lines, patient derived neurons and mice expressing (G₄C₂)₆₆ repeat expansions.

This study showed a significant decrease in the number of cells with G₄C₂ repeat RNA foci and also intra and extracellular levels of poly(GP) was attained by the ASO administration. *In vivo* assessment of ASO treated mice resulted in significant reduction in repeat containing mRNA and subsequent reductions in the number of RNA foci and the levels of poly (GP) DPR proteins in the CSF (Gendron et al., 2017). There are also other reasons that researchers choose to focus on poly (GP) for their studies. Firstly, poly (GP) is one of the most frequently found DPRs in the brain tissue of individuals with C9ALS/FTD. Secondly, it is more soluble than the most commonly found DPR, poly(GA).

Thirdly, it was found that poly (GP) DPRs levels are higher in the mouse line that we used in this study (C9-500 BAC FVB) than other C9 BAC mouse lines (Liu et al., 2016). For these reasons, detection of poly (GP) DPRs by MSD ELISA was decided as the protein of choice in the present study. An optimized MSD ELISA was utilized to review how cisterna magna injection of scAAV9 SRSF1-linker CPP or RRM2 α 1 CPP or SRSF1-RNAi reduced the level of poly (GP) DPR expression in the cortex region of brain in comparison to the untreated group. The assay shows that scAAV9 SRSF1-linker CPP and RRM2 α 1 CPP treatments reduced poly (GP) expression by ~ 33 % and 30 % (**Figure 5.13**).

While I observed the efficacy of our gene therapy vector in a one-month study, it is important to acknowledge several limitations within this investigation. One notable limitation is the retention of approximately 70% of poly (GP) proteins within the brain over this relatively short timeframe. Consequently, a more extended study period may be warranted to establish a more robust reduction in DPR levels. Further, I strongly hypothesize that extended exposure to our gene therapy methods will yield a more substantial reduction effect. This assertion is supported by results from the cisterna magna injection of SRSF1-RNAi, which exhibited a notable dose-dependent decrease, with up to 80% reduction in poly-(GP) DPRs levels at 3 months post-injection, relative to both untreated and control groups (data not shown). Additionally, it should be noted that the current study employed a small cohort of only three animals per treatment group for the one-month duration. Future studies should consider increasing the sample size to enhance statistical power. Another aspect of limitation in this study pertains to the lack of comprehensive biodistribution analysis and the absence of toxicology assessments. Addressing these aspects in future research would provide a more comprehensive understanding of the gene therapy's safety and distribution profile.

On a positive note, a notable strength of this thesis chapter is the pioneering expression of CPP from the viral vector. This achievement holds the potential to offer a sustained expression of therapeutic peptides, mitigating concerns related to peptide turnover that often arise during human clinical trials. Another notable advantage of this therapeutic intervention is its single-injection approach, eliminating the need for multiple invasive injections in both animal models and potentially in human patients.

For the subsequent phases of this research, a longitudinal study on mice is planned, spanning durations of 6 and 12 months. A minimum of 12 animals per treatment group will be included to evaluate safety and efficacy at various time points and using different doses of viral vectors. Following this, a Good Laboratory Practice (GLP) toxicology study will be conducted on non-human primates. It is anticipated that this phase will be facilitated by private sector entities. Upon the successful completion of these stages, we will initiate clinical trials.

5.5. Experimental Conclusion

The results presented in this chapter highlight the therapeutic potential of scAAV9_H1 SRSF1-linker CPP_GFP and scAAV9_H1 SRSF1-RRM2 α 1 CPP_GFP as gene therapy strategies in ALS/FTD. Both gene therapy approaches provide an efficient means of inhibiting the expression of toxic DPRs in cells or animal models of *C9ORF72*-ALS/FTD. The efficacy of the viral vectors was confirmed based on several readouts containing quantification of sense and antisense DPR levels in HEK293T cells with western blot, MTT cell proliferation assay, immunochemistry, biodistribution, histopathological analysis, and poly (GP) DPR quantification with an MSD ELISA system. These promising approaches provide a strong basis for continuing development of an SRSF1-mediated gene therapy approach for *C9ORF72*-ALS/FTD as a neuroprotective strategy for future clinical trial. Although we present exciting and highly innovative data here, further studies are required to determine the long-term safety and efficacy in pre-clinical disease models.

6. Chapter 6 - General Conclusions and Discussion

6.1. Project Aims and Main Findings

My lab previously reported that the nuclear export adaptor protein SRSF1 (Serine/arginine- rich splicing factor 1) triggers the nuclear export of pathological *C9orf72* repeat transcripts and the subsequent RAN translation in the cytoplasm (Hautbergue et al., 2017). This finding supported the development of novel gene therapy approaches aimed at decreasing the expression levels of SRSF1 or expressing a dominant negative mutant. My project was set up to use an alternative approach based on the use of an SRSF1:NXF1-inhibitory CPP (SRSF1-linker CPP) which does not change the level of expression of SRSF1. This constitute an advantageous of our CPP based therapeutic approach because one of the recent study was found that full depletion of SRSF1 expression was resulted in cell death in liver of mice and induce NASH (Non-Alcoholic Steatohepatitis) pathology (Arif et al., 2023). We hypothesised that the SRSF1-linker CPP would inhibit the interaction between *C9orf72*-repeat RNA-bound SRSF1 with NXF1 in the nuclear export machinery, conferring neuroprotection in *C9ORF72*-ALS/FTD cell and animal models by inhibiting the nuclear export and RAN translation of sense and antisense *C9orf72*-repeat transcripts.

My first chapter focused on developing a protocol for expression and purification of recombinant SRSF1-linker CPP by testing several expression and purification conditions which resulted in obtaining a successful and robust recombinant peptide production method. My second chapter describes the validation of the functionality of the recombinant SRSF1-linker CPP in HEK293T cell models of *C9ORF72*-ALS/FTD. Briefly, my data demonstrated that the SRSF1-linker CPP is intracellularly delivered into the nucleus and cytoplasm, inhibits the nuclear export of *C9orf72*-repeat transcripts, the RAN translation of DPRs and rescues the DPR associated cytotoxicity. I then tested the potential therapeutic effect of SRSF1-linker CPP in the C9-500 BAC mouse models of *C9ORF72*-ALS/FTD. However, injection of the recombinant SRSF1-linker CPP , but not a chemically-synthesised one (Castelli et al., 2023), cause mice to die within a few days. Also, western blot data showed that half-life of SRSF1-CPP was very short. Therefore, in the third chapter I designed and developed for the first time a novel scAAV9 based gene therapy approach expressing continuously SRSF1-linker CPP which was successful in inhibiting DPR expression in mice in a pilot proof-of-principle study.

6.2. Developing a Robust Protocol to Produce Recombinant SRSF1-linker CPP

The first chapter of thesis presented expression and purification of recombinant SRSF1-linker CPP. However, achieving high purity, sufficient quantity and biologically active recombinant peptide is challenging and I encountered several practical problems in performing recombinant upstream and downstream processes.

One of the first and main problems I faced in this section was to encounter insufficient expression of SRSF1-linker CPP in *E.coli* cells which is generally due to the toxic effect of heterologous protein in the cell (Burgess-Brown, 2017). This leads to low growth rate and inadequate proliferation of cells and subsequent low production efficiency of the peptide. To tackle this challenge, I, along with several other researchers referenced in the literature, have found benefits in fusing our proteins of interest with a variety of fusion tags, including histidine and GST tags. For example, chimeric lactoferrin and Lasioglossin LL-III antimicrobial peptides were fused with 6xhistidine tag to increase the expression and solubility of peptide. (Tanhaeian et al., 2020; Tanhaeian et al., 2018). As an additional strategy to enhance peptide expression, several studies in the literature have employed varying culture media. Different media compositions contain distinct levels of nutrients, carbon sources, nitrogen sources, vitamins, and growth factors, all of which can influence cell growth and metabolism (Combe and Sokolenko, 2021). Accumulation of inclusion bodies, which are insoluble peptide precipitates, is another major common significant issue when expressing peptide in a recombinant cell system, as it hampers the ability to obtain soluble expressed peptide (Bhatwa et al., 2021). This situation is also the second main challenge I encountered after large scale purification of the SRSF1-linker CPP. To address the precipitation issue, both I and several others in the literature have utilized denaturing agents such as urea and guanidine hydrochloride have been used to eliminate protein/peptide precipitation or to solubilize and denature proteins (Chen et al., 2017; Herbert, 1999; Rabilloud, 1996; Rodríguez et al., 2014; Schlager et al., 2012; Yang et al., 2011).

My third challenge was non-specific retention of SRSF1-linker CPP onto the chromatography matrix which resulted in obtaining only a low quantity of peptide during elution. This is also a very common problem in the literature and has been reported by several studies as inhibiting the purification yield (Al Musaimi et al., 2023; Birrenbach et al., 2021; Thomassen et al., 2005). Several strategies were employed in this thesis and also in various studies related to experimenting with different chromatographic parameters.

For example, Rekker et al. reported a correlation between the hydrophobicity of the mobile phase and retention of peptide. Therefore, he and others suggested that optimization of pH, ionic strength, or composition in the mobile phase could help reduce the retention time of peptide (Al Musaimi et al., 2023; Badgett et al., 2018; Islam et al., 2014; Meek, 1980; Rekker, 1977). In addition to these strategies, the use of different chromatography media was proposed in other studies (Moldoveanu and David, 2013; Robards and Ryan, 2022). Based on similar strategies applied in the literature, the recombinant SRSF1-CPP was produced in substantial quantities and with high purity. In addition to the challenges addressed in this project, there are other common difficulties that can affect peptide structure and functionality, such as improper formation of disulphide bonds (Remans et al., 2022) and proteolytic degradation (Roca-Pinilla et al., 2020) that have been reported in the literature during recombinant production systems.

6.3. Functionality of SRSF1-linker CPP in Human Cell Models of *C9ORF72*-ALS/FTD

The next stage of the thesis concentrated on investigating whether recombinant SRSF1-linker CPP is a potential drug-like therapeutic approach for HEK models of *C9ORF72*-ALS/FTD. *In vitro* studies showed over 90% uptake of SRSF1 and ctrl-CPPs by HEK293T cells. Co-immunoprecipitation assays indicated that SRSF1-linker CPP can interact with NXF1, disrupting endogenous SRSF1 and FLAG-tagged NXF1 interactions in a dose-dependent way. Direct interaction of the SRSF1-linker CPP with Tap/NXF1 was confirmed through the GST pull-down assay, where it directly competed with SRSF1:NXF1 interactions. The peptide specifically interacts with NXF1, showing weak or no binding to other NXF proteins. Nuclear export assays revealed that SRSF1-linker CPP treatment prevents nuclear escape of pathological *C9orf72* repeat transcripts, impacting both sense and antisense mRNA levels in the nucleus and cytoplasm. This peptide also reduced sense and antisense DPR expression dose-dependently, as verified by western blotting. Lastly, SRSF1-linker CPP treatment increased cell survival against DPR-associated toxicity.

Overall, I have shown that recombinant SRSF1-linker CPP confers a cryoprotection effect in HEK cells model of *C9ORF72*-ALS/FTD. Also, similar findings in the HEK cells were also obtained by use of synthetic SRSF1-linker CPP. This was carried out in the context of a published paper where I am listed as one of the co-authors (Castelli et al., 2023). In this paper, we also investigated the potential neuroprotective effect of synthetic SRSF1-linker CPP in primary and patient-derived neuron cells.

A V5 immunofluorescence assay showed that the synthetic CPPs were effectively delivered to cells and remained in over 90 % of primary rat neurons. Sense and antisense DPR production was inhibited in rat cortical neurons where they were transfected with (G₄C₂)₄₅-3xV5 or (C₄G₂)₄₃-3xV5 repeat transcripts. The qRT-PCR quantification revealed that the addition of synthetic SRSF1-linker CPP into iNeurons selectively inhibited nuclear export of pathological *C9orf72* sense and antisense repeat transcripts leading to significant reduction in the cytoplasmic content and corresponding increase in nuclear content of sense and antisense pathological repeat transcripts (Castelli et al., 2023).

Similarly, CPP-based therapeutic applications have been also applied to several neurological diseases. For example, a prion protein (PrP)-derived CPP namely, NCAM1- A β CPP and NCAM1-PrP CPP was used to target amyloid- β formation and reduce associated toxicity in N2a mouse neuroblastoma cells. They found that exposure to both peptides rescued the neuronal toxicity of A β ₄₂ fibres. Additionally, NCAM1-PrP and NCAM1-(Zhang et al., 2018b) A β peptides were found to promote an exchange of the toxic structure of A β ₄₂ into a non- toxic formation resulted in inhibition of toxic A β ₄₂ oligomers (Henning-Knechtel et al., 2020). In a more recent study, a novel therapeutic D- amino acid peptide called ISAD1 and its reversed form ISAD1rev were designed to inhibit Tau protein aggregation in an N2a cell model of Alzheimer disease. They found that increasing concentrations of ISAD1 and ISAD1rev inhibited the formation of Tau aggregates. They also revealed that the peptides were able to inhibit Tau and Tau^{RDAK} fibril formation caused from the mutations *Tau*^{AK280}, *Tau*^{A152T} and *Tau*^{P301L}. Additionally, an MTT assay showed that increased N2a cell survival was attained by increasing the concentration of the peptides (Aillaud et al., 2022). Overall, the literature indicates that successful treatment with peptide-based therapeutic substances into an *in vitro* model of neurological disease exhibits promise as a therapeutic tool by reducing toxic protein expression or aggregation, rescuing cellular toxicity and reaching the target site which are all in agreement with our *in vitro* results for SRSF1-linker CPPs.

6.4. In vivo Functionality of SRSF1-linker CPP

6.4.1. Drosophila Model of C9ORF72-ALS/FTD

As promising *in vitro* results had been obtained, the study moved its translational path to a transgenic C9ORF72-ALS/FTD *in vivo* model to investigate the potential neuroprotective effect of synthetic or recombinant SRSF1-linker CPP. Collaborators from the University of Cambridge explored the therapeutic potential of synthetic SRSF1-linker CPP in a C9ORF72-ALS Drosophila model. The oral administration of this compound reduced cytoplasmic repeat transcripts and increased nuclear ones, leading to improved locomotion in C9ORF72-ALS larvae and adult flies and inhibition of sense DPRs translation. (Castelli et al., 2023). The CPP strategy has also been used for the treatment of polyQ disease. Zhang *et al* designed a novel CPP called BIND (beta-structured inhibitor for neurodegenerative diseases) to investigate the inhibitor effect of the peptide on an *in vivo Drosophila* model of polyQ disease. The BIND peptide was administered orally to *Drosophila* bearing retinal degeneration induced by the expression of a gmr-GAL4 driver.

The peptide treatment rescued the retinal degeneration in a dose dependent manner in adult flies without causing any severe side effects. Administration of the peptide suppressed nucleolar stress and rescued rRNA transcription (Zhang et al., 2018b). The same research group also assessed whether TAT-BIND ameliorated neurodegeneration in a *Drosophila* model of C9ORF72-ALS/FTD. Oral administration of the BIND peptide resulted in suppression of the formation G₄C₂ RNA foci and rescue of the retinal degeneration, improving climbing ability, and extending the lifespan of flies expressing (G₄C₂)₃₆ repeat expansions (Zhang et al., 2019).

6.4.2. Mouse Model of C9ORF72-ALS/FTD

I next evaluated the safety and efficacy of recombinant and synthetic SRSF1-linker CPP in a transgenic C9-500 BAC mouse line which expresses (G₄C₂)₅₀₀ repeat transcripts as well as sense and antisense DPRs. In order to inject the SRSF1-linker CPP, the optimal injection route first had to be determined. There is a large body of literature investigating the efficient and safe way for delivery of gene therapy approaches into the central nervous system (CNS) (Fischell and Fishman, 2021; Hocquemiller et al., 2016).

In this project, I chose the cisterna magna injection route for P1 mice because direct CM delivery of therapeutics into the pups provides the following advantages: **(i)** broad brain transduction in small and large animals (Bey et al., 2017; Hinderer et al., 2018); **(ii)** provides less peripheral distribution of the therapeutic virus and also prevents hepatotoxicity (Gray et al., 2013); **(iii)** requires a small volume of injection to achieve a high dose ;**(iv)** spinal cord transduction was higher than achieved using the ICV or IV route (Lukashchuk et al., 2016). Having decided the route of administration, I directly injected recombinant and synthetic SRSF1-linker CPP via the cisterna magna. A single intrathecal injection of the recombinant SRSF1-linker CPP showed that cisterna magna injection of the recombinant peptide was not safe, resulting in the death of 3 animal out of 5. We do not know the exactly the cause of adverse effect but it is possible that it is involved presence of urea and or endotoxin. For example, a similar issue was also reported in the literature due to endotoxin contamination which are toxic and have the ability to disrupt the functionality of the target protein/peptide (Ongkudon et al., 2012). It has been reported that endotoxins can cause inflammation and secretion of cytokines (Lu et al., 2008). Indeed, it has been reported that the presence of LPS at concentrations above 0.6 units/ml in the bloodstream can cause severe morbidity and mortality (Yaroustovsky et al., 2013). However, no-off target behavioural or biological side effects were seen with synthetic SRSR1 peptide treatment.

Mice treated with the synthetic peptide showed gradient uptake of SRSF1 and Ctrl CPPs in the cerebellum, especially near the cerebrospinal fluid. Both CPPs were absorbed by Purkinje neurons. Further tests revealed the SRSF1-CPP reduced DPR production by 70% in a *C9ORF72*-ALS/FTD model when delivered intrathecally. CPP based therapeutics have also been tested in other animal disease models (Dietz and Bähr, 2007). For example, TAT PTD₄₇₋₅₇-haemagglutinin-Bcl-XL was engineered by Cao and colleagues to investigate its therapeutic effect in cerebral ischaemia. Intraperitoneal administration of TAT PTD₄₇₋₅₇-haemagglutinin-Bcl-XL to mice following middle cerebral artery occlusion (MCAO) resulted in improved neurological function and reduced necrosis (Cao et al., 2002). In another study, the researcher designed a peptide which has the C-terminal regulatory domain of p53 and TAT₄₇₋₅₇ to repair p53 activity in cancer cells. The TAT-p53 peptide was introduced to the TA3/St tumour models intraperitoneally. The peptide treatment prolonged the lifespan of the mice (Snyder et al., 2004).

It is worth to mention that additional studies will be required to determine its stability within the CSF and its bioavailability within the CNS. To address this, repeat brain exposure would need to be performed on a daily or weekly basis. However, this was not feasible via the cisterna magna route which is an invasive method for adult mice. Therefore, oral delivery of the SRSF1-linker CPP would be another route requiring investigation in order to determine pharmacodynamic and pharmacokinetic parameters especially as we had already demonstrated that oral delivery of the peptide conferred neuroprotection in a *Drosophila* model of *C9ORF72*-ALS/FTD (Castelli et al., 2023). Additionally, intranasal delivery of the peptide would constitute an attractive route of administration in repeat-dosing future studies which will help us to determine the kinetic parameters of the CPP. Indeed, this delivery route has also been used for the treatment of other neurological diseases. For example, a peptide which consisted of NF- κ B binding domain and TAT47–56 was administered intranasally into rats exposed to hypoxia–ischaemia (HI) brain injury. The result showed that microglial activation and brain damage were alleviated (Yang et al., 2013). Although intranasal and oral delivery may have potential benefits for drug delivery, it would still require repeat dosing due to the stability half-life of the peptide (Langel, 2019). Relevant to this, it was reported that daily intranasal co-injections of L-penetratin, a cell-permeable peptide and insulin or extendin-4, a glucagon-like peptide-1 (GLP-1) receptor agonist, ameliorated the cognitive dysfunction phenotype of a mouse model. (Kamei et al., 2016); (Kamei et al., 2018). Also, an intranasal may not be an ideal method of treatment for *C9ORF72*-ALS and FTD because it would be unlikely to achieve exposure to the spinal cord (Kamei et al., 2018; Kamei et al., 2016).

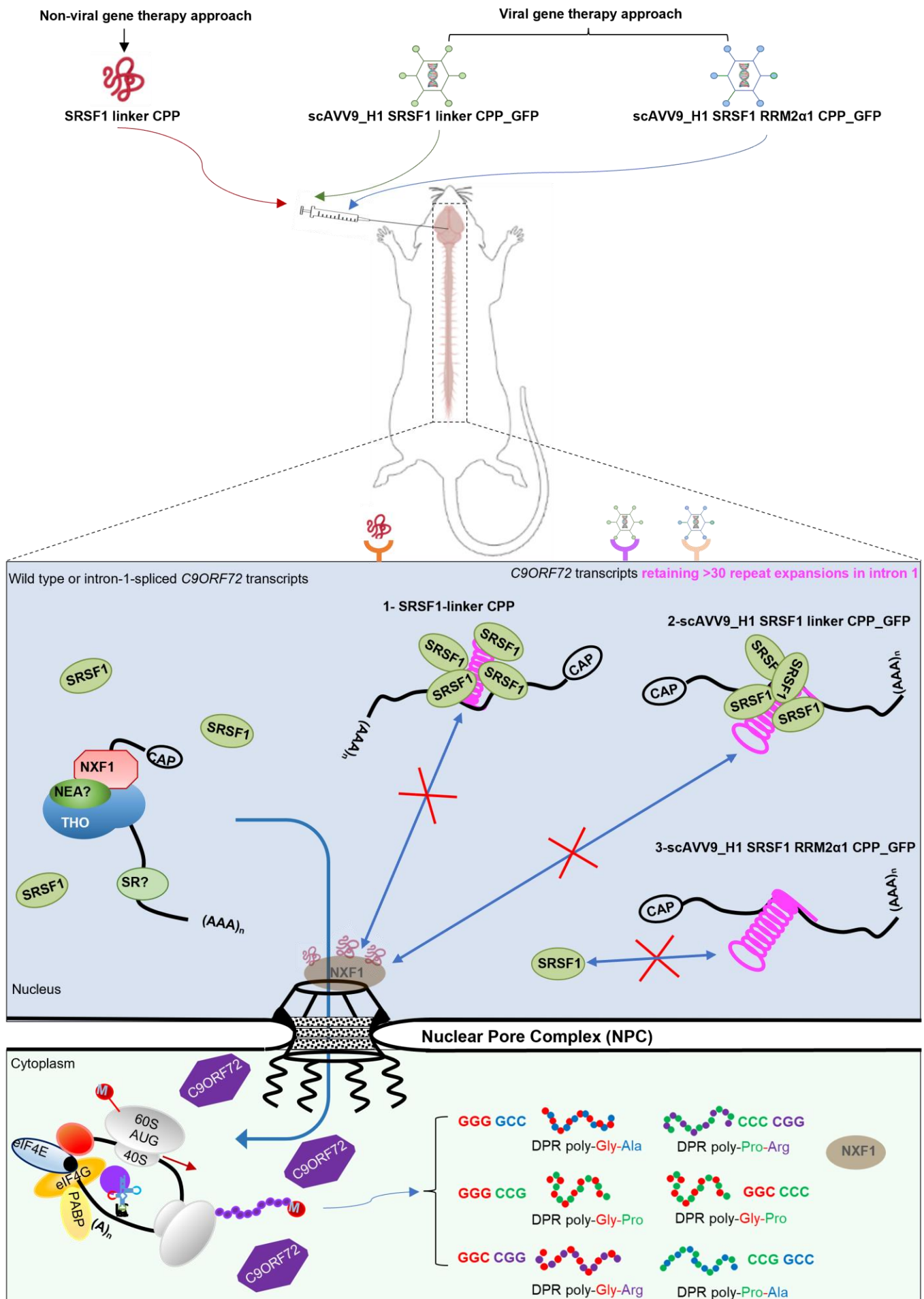


Figure 6.1. Project outcomes and highlights: the promise of therapeutic *in vivo* non-viral and viral gene therapy approaches for the treatment of C9ORF72-ALS/FTD. The schematic representation summarises the primary results of the project with emphasis on development of therapeutic tools for their *in vivo* application. The proposed therapeutic approaches rely on using either drug-like SRSF1 linker CPP or gene therapy approaches respectively expressing SRSF1 linker CPP or SRSF1 RRM2 α 1 CPP. These approaches can decrease toxic sense and antisense DPR production *in vitro* but can also reduce toxic poly (GP) expression by inhibiting SRSF1/NXF1 dependent nuclear export of sense and antisense repeat transcripts to the cytoplasm when administered via the cisterna magna injection route.

To overcome these potential concerns relating to stability and/or biodistribution after directly delivering the SRSF1-CPP which would need repeat dosing, a novel gene therapy approach involving continuous scAAV9-driven expression of SRSF1-linker and RRM2 α 1 CPPs were developed, as another therapeutic strategy (**Figure 6.1**) because this approach will provide long term and continuous expression of the SRSF1-linker CPP in mouse brain and we hypothesise that this would inhibit SRSF1 dependent nuclear export of pathological sense and antisense repeat transcripts thereby consequently providing long-term inhibitory and rescue effects on sense and antisense DPR production and their associated toxicity. To test this idea, we designed four therapeutic plasmids, two of which express variations of the SRSF1-linked CPP. Initial tests in HEK cells indicated that the peptide under the H1 promoter was more effective in reducing DPRs and their associated toxicities compared to the CBH one. Based on these promising results, I administered scAAV9_H1 SRSF1-linked CPPs to P1 C9-500 BAC mice for further studies. The results showed efficient viral distribution in the brain and spinal cord, with histopathology highlighting significant GFP expression in the cerebral cortex. Also, within a month, treatments lowered poly (GP) DPR brain levels.

Overall, these results suggest that our novel gene therapy viruses expressing SRSF1-linker and RRM2 α 1 CPPs offer advantages over present gene therapy approaches. Typically, when therapeutic viruses are injected into the brain, they do not distribute evenly across all neurons. However, our viruses express peptides with a cell-permeable domain, enabling the peptide to diffuse from cell to cell. This feature enabled us to observe a more rapid inhibitory effect on DPR production. In contrast, another scAAV9 gene therapy approach currently used in my lab, which depletes SRSF1 expression to inhibit DPR translation, only showed effects after 3 months of injection into C9-500 BAC mice (data not shown) indicating an earlier effect of the gene therapy viruses expressing SRSF1-linker and RRM2 α 1 CPPs.

6.5. Clinical Applications of CPP

To our understanding, this research marks an innovative approach in gene therapy, utilizing scAAV9 to express a cell-permeable peptide for therapeutic applications. To date, there are no known clinical trials or studies that directly parallel this approach, making direct comparisons with other research. Moving forward, it will be crucial to ascertain the most effective route of administration, systematically evaluate biodistribution profiles, and undertake in-depth toxicological analyses to fortify the robustness and potential therapeutic implications of our strategy.

It's worth noting that, while no one has previously endeavoured to express cell-permeable peptides in this context, peptides have been utilized for other purposes in scientific research. There are several *in vitro* and *in vivo* studies that try to inject AAV virus in conjugation with some CPP (Cho et al., 2017; Stalmans et al., 2015) or pre-incubating with CPP (Meng et al., 2021; Zhang et al., 2018c) in order to enhance the ability of the virus to cross the blood-brain barrier (BBB). For example, one recent study aimed to insert individual CPPs into the capsid of AAV9 variants. Researchers identified two variants, AAV.CPP.16 and AAV.CPP.21, which demonstrated notably enhanced transduction capabilities within the CNS in contrast to the control AAV9 vector. Additionally, AAV.CPP.16 has been shown to maintain its affinity for nerve cells in primates, effectively cross the BBB, and successfully transport anti-cancer agents in a mouse glioblastoma model (Yao et al., 2022). In another study, the LAH4 CPP was incubated with AAV9 virus expressing GFP. They showed that CPP conjugation enhanced transduction efficiency of the virus into HEK293T and human astrocyte cells by observing increased target gene (GFP) expression in comparison to the standard AAV9 vector. The *in vivo* results also showed that increased GFP expression in the brain and a decrease in the immune response were observed after systemically injecting the AAV-CPP complex into mice. This suggests that the CPP also improved the viral vector's internalization through the blood-brain barrier (Meng et al., 2021). Similar *in vitro* and *in vivo* findings were also observed in another study that used LAH4 CPP to improve the gene therapy potential of AAV2 viral vectors (Liu et al., 2014).

The injection of CPPs with viral vectors at the same time has garnered attention in the scientific community. However, the multifaceted challenges associated with this method raise concerns about its broad feasibility. Before introducing the viral vector, it's important to evaluate the peptide's impact on mice. Determining the optimal duration for peptide treatment is a significant hurdle, especially considering its limited half-life. This often leads to frequent and potentially invasive injections, particularly when targeting the cisterna magna in *C9ORF72* ALS/FTD mouse models. Additionally, a rigorous safety assessment is necessary: first for the CPP injection, then for the viral vector injection, and finally, when both are combined. This comprehensive safety validation further extends to primate studies, making the endeavor both time-intensive and expensive. Balancing these challenges with the expected benefits, especially in the context of the 3Rs principles, suggests that the potential advantages of this approach might be limited.

CPP approach has been directly employed in clinical trials for various diseases, indicating its broader potential in therapeutic applications yet no successful CPP therapy has been reported thus far. The following section provides examples of ongoing clinical trials. Xigen SA developed a CPP-derived therapeutic agent called XG-102, which inhibits c-Jun N-terminal kinase (JNK) activity. This underwent a Phase I clinical trial (NCT01570205) to evaluate dose-related adverse effects of the CPP; this trial was completed in 2012 (Deloche et al., 2014). The adverse effects of the group receiving XG-102 was similar to those receiving placebo. XG-102 entered Phase III clinical trials to evaluate its functional effects on intraocular inflammation and pain in patients after cataract surgery, but the results have not yet been published (NCT02235272). Auris Medical designed AM-111 CPP consisting of TAT₄₈₋₅₇ (GRKKRRQRRR) (D-JNKI-1), which inhibits c-Jun N-terminal kinase (JNK) activity, to treat inner ear disorders caused by cochlear injury (Omotehara et al., 2011). A Phase II clinical trial of AM-111 CPP was completed in 2014 (NCT00802425) (Suckfuell et al., 2014) and the treatment group showed permanent improvement in hearing and speech discrimination relative to the placebo group. Phase III clinical trials were also started and are currently ongoing (NCT02561091, NCT02809118).

Aileron Therapeutics has developed a peptide based anticancer drug called ALRN-6924 which is able to disrupt the interaction between the p53 tumour suppressor protein and its endogenous inhibitors. This, in turn, results in restoration of p53-dependent cell cycle arrest, apoptosis, and subsequently exhibiting antitumor effects. A phase I clinical trial showed promising results that ALRN-6924 exhibits anti-tumour activity and good tolerability. Therefore, they planned further clinical trials of this agent (NCT02264613) (Saleh et al., 2021). Pep-Therapy developed a drug called PEP-010 composed of a 30 amino-acid peptide that inhibits the Caspase-9/PP2A protein-protein interaction and induces apoptosis in cancer cells while not affecting healthy cells and signalling pathways. Additionally, the peptide demonstrated the ability to inhibit tumour growth in patient-derived xenograft (PDX) models of human breast, ovary, and lung primary tumours. Therefore, the company is planning to start Phase I clinical trial in the near future (Langel, 2019).

DrugCendR developed an anti-cancer drug called CEND-1. The ability of tumour specific targeting of CEND-1 was tested in various tumour models prior to its approval. Phase I clinical trials (NCT03517176) were completed in 2017 (Liu et al., 2017) with plans to move to phase III in the near future. Sarepta Therapeutics developed a novel compound, AVI-4658, consisting of an arginine-rich CPP which is conjugated with PMOs (phosphorodiamidate morpholino oligomers). In phase I and II clinical trials (NCT00159250, NCT00844597), intravenous administration of AVI-4658 met safety and tolerability criteria in Duchenne muscular dystrophy (DMD) patients. Treatment of AVI-4658 markedly increased the expression of dystrophin (Cirak et al., 2011). A Phase III clinical trial was initiated, and patients with (DMD) are currently being recruited (NCT02255552). DTS-108 is a novel peptidic prodrug of SN38 that was developed by Dia-tos. It combines a cationic peptide and SN38 to enhance the cytotoxicity of the active drug. The efficacy of DTS-108 was evaluated across various preclinical models, and it was found to exhibit significant antitumor efficacy in all of them. Indeed, it was found that when DTS-108 combined with other therapeutic agents that are relevant to clinical treatment, it demonstrated beneficial synergistic effects. A Phase I clinical trial was conducted and showed promising antitumor activity and safety therefore it is planning to move phase II clinical trials (Coriat et al., 2016).

6.6. Limitation and Future Work

The project has several limitations that could be addressed in future studies. For example, cisterna magna injection of recombinant SRSF1-linker CPP led to adverse effects in C9-500 BAC mice, whereby 3 out of 5 transgenic mice injected with recombinant CPP died. The precise cause of the mortality remains undetermined; however, it is possible that the presence of residual urea or endotoxins in the purified recombinant peptide was potentially hazardous to health, including the possibility of eliciting an immune response or adjuvant activity in mice. One of the first limitations is that this study did not incorporate additional steps to quantify the purity of the final recombinant peptide, despite achieving equivalent levels of purity to chemically synthesised peptides through an optimized production method. Therefore, additional studies are required to determine the presence of any residual contaminants or degradation products in the therapeutic SRSF1-linker CPP. Future work could involve the use of mass spectrometry (MS), a highly sensitive technique which is able to quantify individual components that may not be detectable by Coomassie gel staining.

Relevant to this, one study employed liquid chromatography-mass spectrometry (LC-MS) to identify and quantify host cell protein impurities in the therapeutic peptide Fc fusion protein. The results indicated that the MS approach was capable of accurately quantifying unknown protein impurities (Schenauer et al., 2012). It would therefore be interesting to utilize MS in our future work in order to design more effective purification techniques and improve the quality of the recombinant peptide. Additional future studies will be necessary to fully evaluate the therapeutic potential of the SRSF1-linker CPP and gene therapy approaches in *C9ORF72*-ALS/FTD mice. However, I chose to focus on the gene therapy strategy because there is no issue with peptide stability and no need for repeat injection. Also, I have shown that this strategy reduces DPR expression. Therefore, during my postdoc, I will first compare head-to-head the gene therapeutic efficacy and safety of SRSF1 depletion and expression of CPPs in scAAV9 as in a proof-of-concept study in *C9ORF72*-ALS/FTD mice. In another direction for future studies, I could also test direct delivery of the CPP but this presents more challenges because I need to first determine pharmacokinetics and pharmacodynamic parameters which are highly dependent on the stability of the CPP due to its susceptibility to rapid degradation or limited half-life in the plasma, and bioavailability within the target tissue (Lipinski, 2000). Therefore, I would need to improve the stability of SRSF1-linker CPP.

There are several methods in the literature for this purpose such as inverso and retro-inverso modifications and PEGylation (Langel, 2019). One prevalent strategy that I can use to stop degradation is the substitution of one or multiple L-amino acids with their D- amino acids (inverso), while preserving the functional properties of the peptide. There are several successful example in the literature that have benefited from this strategy. For example, sweet arrow Peptide (SAP), an amphipathic peptide rich in proline, showed protease resistance and lack of toxicity after replacing L-amino acids with their respective D- amino acids (Pujals and Giralt, 2008; Sánchez-Navarro et al., 2017). Additionally, recent literature provides evidence of multiple endeavours in the design of cyclic CPPs with the aim of enhancing their pharmacokinetic and cellular uptake properties. For example, cyclic arginine rich CPPs with higher peptide structural rigidity showed enhanced cellular uptake kinetics.

In the end for any translational therapeutic approaches, either injection of SRSF1-linker CPP or scAAV9 gene therapy approaches expressing CPPs, it will be required to evaluate the therapeutic efficacy on the motor and cognitive functions of *C9ORF72*-ALS/FTD mice over a 12-month study. Additionally, conducting a long-term toxicology and potential immune response studies will be performed in future experiments. Actually, these planned studies were my initial aims. Unfortunately, due to Covid interruptions in 2020, there was a delay of 8 months in obtaining my animal license. For this reason, I was only able to perform a safety and efficacy study in wild type mice and in C9 mice for a post-injection time of 1 month. Histopathological assessment showed that therapeutic viral vectors are able express GFP within the mouse CNS. Although several optimization studies were performed in the pathology section, SRSF1-linker or RRM2 α 1 CPPs expression were not detected within the brain by immunohistochemistry at 1 month post injection. Therefore, additional experiments to optimize the immunohistochemistry need to be conducted to confirm peptide expression and investigate any potential effects on motor neuron survival or degeneration in the area where the peptide is expressed. Biodistribution data indicated that there is higher tropism of viral vectors for the brain and spinal cord; however, they were also detected in the heart.

To eliminate this off-target effect, a future study including injection of mice at different ages would be interesting. Although P1 injecting is relatively easy and beneficial, ALS is an adult-onset disease and it is not likely to diagnose patients with *C9ORF72*-ALS/FTD at birth, therefore this age of mice does not reflect the disease conditions. On the other hand, it was important to test our approaches at P1 in relation to safety since SRSF1 is required for the developing heart at the embryonic stage. Also, several studies advised that distribution of viral vectors across the peripheral organs could be eliminated by delaying injection to later ages. Due to the rapid growth of young mice, with a rapid expansion in the size of organs and blood vessels, the expression of the viral vector may be diluted in target regions with distribution to peripheral organs (Cunningham et al., 2008; Wang et al., 2012)

6.7. Conclusions

Overall, this project supports the hypothesis that treatment with the SRSF1-linker CPP as a drug-like compound or expression of SRSF1-linker and RRM2 α 1 CPPs within CNS tissues confers neuroprotection in models of *C9ORF72*-ALS/FTD by preventing SRSF1/NXF1 dependent nuclear export of pathogenic sense and antisense repeat transcripts into the cytoplasm. I have provided novel methodology for overcoming challenges pertaining to the expression and purification of the SRSF1-linker CPP. This will help to develop the groundwork for the production of a variety of peptides and proteins in a desired amount and quality. Indeed, this optimised production method contributes to the development of inexpensive therapeutic drugs, and this knowledge could open avenues for the treatment of other diseases. SRSF1-linker CPP based manipulation of the nuclear export of pathogenic repeat transcripts provides very encouraging and promising results, especially reducing both sense and antisense DPR expression in cells, as much of the literature focuses on inhibiting DPR expression of sense repeat transcripts. Lastly, the scAAV9 viral vector expressing the SRSF1 linker or RRM2 α 1 CPP as therapeutic gene therapy approaches were designed. This was the first time such an approach was explored and I have reported that scAAV9 mediated inhibition of SRSF1/NXF1 dependent nuclear export of pathological repeat transcripts can be achieved by expression of a small molecule like the SRSF1 Linker and RRM2 α 1 CPPs without significant off-target effects. This was the first study to show a reduction of poly (GP) DPRs in *C9ORF72*-ALS/FTD animal model of the target tissue resulting from the use of SRSF1 linker and RRM2 α 1 CPPs expressed by scAAV9 viral vector.

Within this thesis I have described the development of therapeutic recombinant SRSF1-linker CPP and scAAV9 viral vectors expressing SRSF1-linker and RRM2 α 1 CPPs targeting both sense and antisense repeat transcripts which constitute another novel aspect for this thesis. Conventional gene therapy vectors and small molecule inhibitors frequently target only sense transcripts and have poor blood-brain barrier penetrance. CPPs could potentially be used as therapeutics for disrupting the nuclear export of other microsatellite repeat transcripts, which are involved in over 40 incurable neurological disorders. They could also be used as promising molecular tools for dissecting the nature of the pathogenic mechanisms linked to the formation of intranuclear RNA foci and RAN translation of polymeric repeat proteins.

7. References

Ackermann, M., Lachmann, N., Hartung, S., Eggenschwiler, R., Pfaff, N., Happle, C., Mucci, A., Göhring, G., Niemann, H., Hansen, G., *et al.* (2014). Promoter and lineage independent anti-silencing activity of the A2 ubiquitous chromatin opening element for optimized human pluripotent stem cell-based gene therapy. *Biomaterials* 35, 1531-1542.

Adawy, A., and Groves, M.R. (2017). The Use of Size Exclusion Chromatography to Monitor Protein Self-Assembly. *Crystals* 7, 331.

Agrawal, P., Bhalla, S., Usmani, S.S., Singh, S., Chaudhary, K., Raghava, Gajendra P.S., and Gautam, A. (2015). CPPsite 2.0: a repository of experimentally validated cell-penetrating peptides. *Nucleic Acids Research* 44, D1098-D1103.

Aguilar-Montes de Oca, S., Montes-de-Oca-Jiménez, R., Carlos Vázquez-Chagoyán, J., Barbabosa-Pliego, A., Eliana Rivadeneira-Barreiro, P., and P, C.Z.-R. (2022). The Use of Peptides in Veterinary Serodiagnosis of Infectious Diseases: A Review. *Vet Sci* 9.

Ail, D., Malki, H., Zin, E.A., and Dalkara, D. (2023). Adeno-Associated Virus (AAV) - Based Gene Therapies for Retinal Diseases: Where are We? *Appl Clin Genet* 16, 111-130.

Aillaud, I., Kaniyappan, S., Chandupatla, R.R., Ramirez, L.M., Alkhashrom, S., Eichler, J., Horn, A.H.C., Zweckstetter, M., Mandelkow, E., Sticht, H., *et al.* (2022). A novel D-amino acid peptide with therapeutic potential (ISAD1) inhibits aggregation of neurotoxic disease-relevant mutant Tau and prevents Tau toxicity in vitro. *Alzheimer's Research & Therapy* 14, 15.

Al-Chalabi, A., Hardiman, O., Kiernan, M.C., Chiò, A., Rix-Brooks, B., and van den Berg, L.H. (2016a). Amyotrophic lateral sclerosis: moving towards a new classification system. *The Lancet Neurology* 15, 1182-1194.

Al-Chalabi, A., van den Berg, L.H., and Veldink, J. (2016b). Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nature Reviews Neurology* 13, 96.

Al-Saif, A., Al-Mohanna, F., and Bohlega, S. (2011). A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis. *Annals of Neurology* 70, 913-919.

Al Musaimi, O., Lombardi, L., Williams, D.R., and Albericio, F. (2022). Strategies for Improving Peptide Stability and Delivery. *Pharmaceuticals (Basel)* 15.

Al Musaimi, O., Valenzo, O.M.M., and Williams, D.R. (2023). Prediction of peptides retention behavior in reversed-phase liquid chromatography based on their hydrophobicity. *Journal of Separation Science* 46, 2200743.

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., *et al.* (2007). The molecular architecture of the nuclear pore complex. *Nature* 450, 695-701.

Alcázar-Román, A.R., Tran, E.J., Guo, S., and Wente, S.R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nature Cell Biology* 8, 711-716.

Alliegro, M.C. (2000). Effects of Dithiothreitol on Protein Activity Unrelated to Thiol–Disulfide Exchange: For Consideration in the Analysis of Protein Function with Cleland's Reagent. *Analytical Biochemistry* 282, 102-106.

Almarza, D., Bussadori, G., Navarro, M., Mavilio, F., Larcher, F., and Murillas, R. (2011). Risk assessment in skin gene therapy: viral-cellular fusion transcripts generated by proviral transcriptional read-through in keratinocytes transduced with self-inactivating lentiviral vectors. *Gene Ther* 18, 674-681.

Alves, I.D., Goasdoué, N., Correia, I., Aubry, S., Galanth, C., Sagan, S., Lavielle, S., and Chassaing, G. (2008). Membrane interaction and perturbation mechanisms induced by two cationic cell penetrating peptides with distinct charge distribution. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1780, 948-959.

Amado, D.A., and Davidson, B.L. (2021). Gene therapy for ALS: A review. *Mol Ther* 29, 3345-3358.

Arakawa, T., and Timasheff, S.N. (1982). Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry* 21, 6545-6552.

Arif, W., Mathur, B., Saikali, M.F., Chembazhi, U.V., Toohill, K., Song, Y.J., Hao, Q., Karimi, S., Blue, S.M., Yee, B.A., *et al.* (2023). Splicing factor SRSF1 deficiency in the liver triggers NASH-like pathology and cell death. *Nature Communications* 14, 551.

Armbruster, N., Lattanzi, A., Jeavons, M., Van Wittenberghe, L., Gjata, B., Marais, T., Martin, S., Vignaud, A., Voit, T., Mavilio, F., *et al.* (2016). Efficacy and biodistribution analysis of intracerebroventricular administration of an optimized scAAV9-SMN1 vector in a mouse model of spinal muscular atrophy. *Mol Ther Methods Clin Dev* 3, 16060.

Asoh, S., Ohsawa, I., Mori, T., Katsura, K.-i., Hiraide, T., Katayama, Y., Kimura, M., Ozaki, D., Yamagata, K., and Ohta, S. (2002). Protection against ischemic brain injury by protein therapeutics. *Proceedings of the National Academy of Sciences* 99, 17107-17112.

Au, H.K.E., Isalan, M., and Mielcarek, M. (2022). Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings. *Frontiers in Medicine* 8.

Azzouz, M., Hottinger, A., Paterna, J.-C., Zurn, A.D., Aebischer, P., and Büeler, H. (2000). Increased motoneuron survival and improved neuromuscular function in transgenic ALS mice after intraspinal injection of an adeno-associated virus encoding Bcl-2. *Human Molecular Genetics* 9, 803-811.

Azzouz, M., Ralph, G.S., Storkebaum, E., Walmsley, L.E., Mitrophanous, K.A., Kingsman, S.M., Carmeliet, P., and Mazarakis, N.D. (2004). VEGF delivery with retrogradely transported lentivector prolongs survival in a mouse ALS model. *Nature* 429, 413-417.

Babi Leko, M., Zupunski, V., Kirincich, J., Smilovi, D., Hortobagyi, T., Hof, P.R., and Simic, G. (2019). Molecular Mechanisms of Neurodegeneration Related to C9orf72 Hexanucleotide Repeat Expansion. *Behavioural Neurology* 2019, 18.

Babin, P.J., Goizet, C., and Raldúa, D. (2014). Zebrafish models of human motor neuron diseases: advantages and limitations. *Prog Neurobiol* 118, 36-58.

Badgett, M.J., Boyes, B., and Orlando, R. (2018). Peptide retention prediction using hydrophilic interaction liquid chromatography coupled to mass spectrometry. *J Chromatogr A* 1537, 58-65.

Bainbridge, J.W.B., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., *et al.* (2008). Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis. *New England Journal of Medicine* 358, 2231-2239.

Balbás, P. (2001). Understanding the art of producing protein and nonprotein molecules in *Escherichia coli*. *Molecular Biotechnology* 19, 251-267.

Balendra, R., and Isaacs, A.M. (2018). C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nature Reviews Neurology* 14, 544-558.

Bandaranayake, A.D., and Almo, S.C. (2014). Recent advances in mammalian protein production. *FEBS Lett* 588, 253-260.

Bañez-Coronel, M., Ayhan, F., Tarabochia, A.D., Zu, T., Perez, B.A., Tusi, S.K., Pletnikova, O., Borchelt, D.R., Ross, C.A., Margolis, R.L., *et al.* (2015). RAN Translation in Huntington Disease. *Neuron* 88, 667-677.

Bannwarth, S., Ait-El-Mkadem, S., Chaussonot, A., Genin, E.C., Lacas-Gervais, S., Fragaki, K., Berg-Alonso, L., Kageyama, Y., Serre, V., Moore, D.G., *et al.* (2014). A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain* 137, 2329-2345.

Barnes, B., Caws, T., Thomas, S., Shephard, A.P., Corteling, R., Hole, P., and Bracewell, D.G. (2022). Investigating heparin affinity chromatography for extracellular vesicle purification and fractionation. *Journal of Chromatography A* 1670, 462987.

Barzilai, N., Crandall, J.P., Kritchevsky, S.B., and Espeland, M.A. (2016). Metformin as a Tool to Target Aging. *Cell Metab* 23, 1060-1065.

Bauer, P.O. (2016). Methylation of C9orf72 expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. *Neuroscience Letters* 612, 204-209.

Becker, L.A., Huang, B., Bieri, G., Ma, R., Knowles, D.A., Jafar-Nejad, P., Messing, J., Kim, H.J., Soriano, A., Auburger, G., *et al.* (2017). Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature* 544, 367-371.

Beharry, A., Gong, Y., Kim, J.C., Hanlon, K.S., Nammour, J., Hieber, K., Eichler, F., Cheng, M., Stemmer-Rachamimov, A., Stankovic, K.M., *et al.* (2022). The AAV9 Variant Capsid AAV-F Mediates Widespread Transgene Expression in Nonhuman Primate Spinal Cord After Intrathecal Administration. *Hum Gene Ther* 33, 61-75.

Bell, M.R., Engleka, M.J., Malik, A., and Strickler, J.E. (2013). To fuse or not to fuse: What is your purpose? *Protein Science* 22, 1466-1477.

Belzil, V.V., Bauer, P.O., Prudencio, M., Gendron, T.F., Stetler, C.T., Yan, I.K., Pregent, L., Daugherty, L., Baker, M.C., Rademakers, R., *et al.* (2013). Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathologica* 126, 895-905.

Berry, J.D., Cudkowicz, M.E., Windebank, A.J., Staff, N.P., Owegi, M., Nicholson, K., McKenna-Yasek, D., Levy, Y.S., Abramov, N., Kaspi, H., *et al.* (2019). NurOwn, phase 2, randomized, clinical trial in patients with ALS: Safety, clinical, and biomarker results. *Neurology* 93, e2294-e2305.

Bey, K., Ciron, C., Dubreil, L., Deniaud, J., Ledevin, M., Cristini, J., Blouin, V., Aubourg, P., and Colle, M.A. (2017). Efficient CNS targeting in adult mice by intrathecal infusion of single-stranded AAV9-GFP for gene therapy of neurological disorders. *Gene Ther* 24, 325-332.

Bhatwa, A., Wang, W., Hassan, Y.I., Abraham, N., Li, X.-Z., and Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in *Escherichia coli* and Strategies to Address Them for Industrial Applications. *Frontiers in Bioengineering and Biotechnology* 9.

Biferi, M.G., Cohen-Tannoudji, M., Cappelletto, A., Giroux, B., Roda, M., Astord, S., Marais, T., Bos, C., Voit, T., Ferry, A., *et al.* (2017). A New AAV10-U7-Mediated Gene Therapy Prolongs Survival and Restores Function in an ALS Mouse Model. *Molecular Therapy : the journal of the American Society of Gene Therapy* 25, 2038-2052.

Birney, Y.A., and O'Connor, B.F. (2001). Purification and Characterization of a Z-Pro-prolinal-Insensitive Z-Gly-Pro-7-amino-4-methyl Coumarin-Hydrolyzing Peptidase from Bovine Serum—A New Proline-Specific Peptidase. *Protein Expression and Purification* 22, 286-298.

Birrenbach, O., Faust, F., Ebrahimi, M., Fan, R., and Czermak, P. (2021). Recovery and Purification of Protein Aggregates From Cell Lysates Using Ceramic Membranes: Fouling Analysis and Modeling of Ultrafiltration. *Frontiers in Chemical Engineering* 3.

Bobo, T.A., Samowitz, P.N., Robinson, M.I., and Fu, H. (2020). Targeting the Root Cause of Mucopolysaccharidosis IIIA with a New scAAV9 Gene Replacement Vector. *Molecular Therapy - Methods & Clinical Development* 19, 474-485.

Boehringer, A., and Bowser, R. (2018). RNA Nucleocytoplasmic Transport Defects in Neurodegenerative Diseases. In *RNA Metabolism in Neurodegenerative Diseases*, R. Sattler, and C.J. Donnelly, eds. (Cham: Springer International Publishing), pp. 85-101.

Boeynaems, S., Bogaert, E., Michiels, E., Gijssels, I., Sieben, A., Jovičić, A., De Baets, G., Scheveneels, W., Steyaert, J., Cuijt, I., *et al.* (2016). *Drosophila* screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Sci Rep* 6, 20877.

Bogomolovas, J., Simon, B., Sattler, M., and Stier, G. (2009). Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expression and Purification* 64, 16-23.

Boisguerin, P., Redt-Clouet, C., Franck-Miclo, A., Licheheb, S., Nargeot, J., Barrère-Lemaire, S., and Lebleu, B. (2011). Systemic delivery of BH4 anti-apoptotic peptide using CPPs prevents cardiac ischemia–reperfusion injuries in vivo. *Journal of Controlled Release* 156, 146-153.

Bolten, S.N., Rinas, U., and Scheper, T. (2018). Heparin: role in protein purification and substitution with animal-component free material. *Applied Microbiology and Biotechnology* 102, 8647-8660.

Booth, C., Gaspar, H.B., and Thrasher, A.J. (2016). Treating Immunodeficiency through HSC Gene Therapy. *Trends Mol Med* 22, 317-327.

Borges, E.M., and Collins, C.H. (2012). Effects of pH and temperature on the chromatographic performance and stability of immobilized poly(methyloctylsiloxane) stationary phases. *Journal of Chromatography A* 1227, 174-180.

Borghouts, C., and Weiss, A. (2009). Production and Purification of Monomeric Recombinant Peptide Aptamers: Requirements for Efficient Intracellular Uptake and Target Inhibition. In *Peptides as Drugs*, pp. 145-186.

Boros, B.D., Schoch, K.M., Kreple, C.J., and Miller, T.M. (2022). Antisense Oligonucleotides for the Study and Treatment of ALS. *Neurotherapeutics* 19, 1145-1158.

Brown, B.L., Hadley, M., and Page, R. (2008). Heterologous high-level E. coli expression, purification and biophysical characterization of the spine-associated RapGAP (SPAR) PDZ domain. *Protein Expression and Purification* 62, 9-14.

Brown, R.H., and Al-Chalabi, A. (2017). Amyotrophic Lateral Sclerosis. *New England Journal of Medicine* 377, 162-172.

Brudar, S., and Hribar-Lee, B. (2021). Effect of Buffer on Protein Stability in Aqueous Solutions: A Simple Protein Aggregation Model. *The Journal of Physical Chemistry B* 125, 2504-2512.

Brunetti-Pierri, N., Ng, T., Iannitti, D., Cioffi, W., Stapleton, G., Law, M., Breinholt, J., Palmer, D., Grove, N., Rice, K., *et al.* (2013). Transgene expression up to 7 years in nonhuman primates following hepatic transduction with helper-dependent adenoviral vectors. *Hum Gene Ther* 24, 761-765.

Buchman, V.L., Cooper-Knock, J., Connor-Robson, N., Higginbottom, A., Kirby, J., Razinskaya, O.D., Ninkina, N., and Shaw, P.J. (2013). Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. *Molecular Neurodegeneration* 8, 12.

Bulcha, J.T., Wang, Y., Ma, H., Tai, P.W.L., and Gao, G. (2021). Viral vector platforms within the gene therapy landscape. *Signal Transduction and Targeted Therapy* 6, 53.

Burgess-Brown, N.A. (2017). *Heterologous gene expression in E. coli methods and protocols* (New York, NY: Humana Press).

Butti, Z., and Patten, S.A. (2019). RNA Dysregulation in Amyotrophic Lateral Sclerosis. *Frontiers in genetics* 9, 712-712.

Cao, G., Pei, W., Ge, H., Liang, Q., Luo, Y., Sharp, F.R., Lu, A., Ran, R., Graham, S.H., and Chen, J. (2002). In Vivo Delivery of a Bcl-xL Fusion Protein Containing the TAT Protein Transduction Domain Protects against Ischemic Brain Injury and Neuronal Apoptosis. *The Journal of Neuroscience* 22, 5423-5431.

Carmody, S.R., and Wenthe, S.R. (2009). mRNA nuclear export at a glance. *Journal of Cell Science* 122, 1933-1937.

Castanotto, D., and Rossi, J.J. (2009). The promises and pitfalls of RNA-interference-based therapeutics. *Nature* 457, 426-433.

Castelli, L.M., Cutillo, L., Souza, C.D.S., Sanchez-Martinez, A., Granata, I., Lin, Y.-H., Myszczyńska, M.A., Heath, P.R., Livesey, M.R., Ning, K., *et al.* (2021). SRSF1-dependent inhibition of C9ORF72-repeat RNA nuclear export: genome-wide mechanisms for neuroprotection in amyotrophic lateral sclerosis. *Molecular Neurodegeneration* 16, 53.

Castelli, L.M., Lin, Y.-H., Sanchez-Martinez, A., Gül, A., Mohd Imran, K., Higginbottom, A., Upadhyay, S.K., Márkus, N.M., Rua Martins, R., Cooper-Knock, J., *et al.* (2023). A cell-penetrant peptide blocking C9ORF72-repeat RNA nuclear export reduces the neurotoxic effects of dipeptide repeat proteins. *Science Translational Medicine* 15, eabo3823.

Castro, J., and Kipps, T.J. (2018). Gene Therapy for Neoplastic Hematology in Transplant Setting. *Advances and Controversies in Hematopoietic Transplantation and Cell Therapy*.

Cauchi, R.J., and van den Heuvel, M. (2006). The fly as a model for neurodegenerative diseases: is it worth the jump? *Neurodegener Dis* 3, 338-356.

Cesana, D., Ranzani, M., Volpin, M., Bartholomae, C., Duros, C., Artus, A., Merella, S., Benedicenti, F., Sergi, L., Sanvito, F., *et al.* (2014). Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. *Mol Ther* 22, 774-785.

Chaitanya, K.V. (2022). *Diagnostics and Gene Therapy for Human Genetic Disorders* (1st ed.).

Chamberlain, K., Riyad, J.M., and Weber, T. (2016). Expressing Transgenes That Exceed the Packaging Capacity of Adeno-Associated Virus Capsids. *Hum Gene Ther Methods* 27, 1-12.

Chandran, J.S., Joseph M Shaw, Pamela J Azzouz, Mimoun (2017). Gene Therapy in the Nervous System: Failures and Successes. *Advances in experimental medicine and biology*, 241-257.

Chandrudu, S., Simerska, P., and Toth, I. (2013). Chemical methods for peptide and protein production. *Molecules* 18, 4373-4388.

Chao, Q., Ding, Y., Chen, Z.-H., Xiang, M.-H., Wang, N., and Gao, X.-D. (2020). Recent Progress in Chemo-Enzymatic Methods for the Synthesis of N-Glycans. *Frontiers in Chemistry* 8.

- Chen, H., Li, N., Xie, Y., Jiang, H., Yang, X., Cagliero, C., Shi, S., Zhu, C., Luo, H., Chen, J., *et al.* (2017). Purification of inclusion bodies using PEG precipitation under denaturing conditions to produce recombinant therapeutic proteins from *Escherichia coli*. *Appl Microbiol Biotechnol* 101, 5267-5278.
- Chen, Y.-Z., Bennett, C.L., Huynh, H.M., Blair, I.P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M.L., Rabin, B.A., *et al.* (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *American Journal of Human Genetics* 74, 1128-1135.
- Cheng, H., Dufu, K., Lee, C.-S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human mRNA Export Machinery Recruited to the 5' End of mRNA. *Cell* 127, 1389-1400.
- Chia, K., Klingseisen, A., Sieger, D., and Priller, J. (2022). Zebrafish as a model organism for neurodegenerative disease. *Frontiers in Molecular Neuroscience* 15.
- Chiò, A., Battistini, S., Calvo, A., Caponnetto, C., Conforti, F.L., Corbo, M., Giannini, F., Mandrioli, J., Mora, G., and Sabatelli, M. (2014). Genetic counselling in ALS: facts, uncertainties and clinical suggestions. *Journal of Neurology, Neurosurgery & Psychiatry* 85, 478-485.
- Chiò, A., Logroscino, G., Traynor, B.J., Collins, J., Simeone, J.C., Goldstein, L.A., and White, L.A. (2013). Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. *Neuroepidemiology* 41, 118-130.
- Chitiprolu, M., Jagow, C., Tremblay, V., Bondy-Chorney, E., Paris, G., Savard, A., Palidwor, G., Barry, F.A., Zinman, L., Keith, J., *et al.* (2018). A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. *Nature Communications* 9, 2794.
- Cho, C.F., Wolfe, J.M., Fadzen, C.M., Calligaris, D., Hornburg, K., Chiocca, E.A., Agar, N.Y.R., Pentelute, B.L., and Lawler, S.E. (2017). Blood-brain-barrier spheroids as an in vitro screening platform for brain-penetrating agents. *Nat Commun* 8, 15623.
- Chow, C.Y., Landers, J.E., Bergren, S.K., Sapp, P.C., Grant, A.E., Jones, J.M., Everett, L., Lenk, G.M., McKenna-Yasek, D.M., Weisman, L.S., *et al.* (2009). Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. *American Journal of Human Genetics* 84, 85-88.
- Ciervo, Y., Ning, K., Jun, X., Shaw, P.J., and Mead, R.J. (2017). Advances, challenges and future directions for stem cell therapy in amyotrophic lateral sclerosis. *Mol Neurodegener* 12, 85.
- Cirak, S., Arechavala-Gomez, V., Guglieri, M., Feng, L., Torelli, S., Anthony, K., Abbs, S., Garralda, M.E., Bourke, J., Wells, D.J., *et al.* (2011). Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *The Lancet* 378, 595-605.

Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A., and Kabashi, E. (2013). Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Annals of Neurology* 74, 180-187.

Cole, C.N., and Scarcelli, J.J. (2006). Transport of messenger RNA from the nucleus to the cytoplasm. *Current Opinion in Cell Biology* 18, 299-306.

Combe, M., and Sokolenko, S. (2021). Quantifying the impact of cell culture media on CHO cell growth and protein production. *Biotechnol Adv* 50, 107761.

Cong, L., Wang, Z., Chai, Y., Hang, W., Shang, C., Yang, W., Bai, L., Du, J., Wang, K., and Wen, Q. (2017). Rapid whole brain imaging of neural activity in freely behaving larval zebrafish (*Danio rerio*). *Elife* 6.

Cooper-Knock, J., Moll, T., Ramesh, T., Castelli, L., Beer, A., Robins, H., Fox, I., Niedermoser, I., Van Damme, P., Moisse, M., *et al.* (2019). Mutations in the Glycosyltransferase Domain of GLT8D1 Are Associated with Familial Amyotrophic Lateral Sclerosis. *Cell Rep* 26, 2298-2306.e2295.

Cooper-Knock, J., Shaw, P.J., and Kirby, J. (2014a). The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathologica* 127, 333-345.

Cooper-Knock, J., Walsh, M.J., Higginbottom, A., Robin Highley, J., Dickman, M.J., Edbauer, D., Ince, P.G., Wharton, S.B., Wilson, S.A., Kirby, J., *et al.* (2014b). Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* 137, 2040-2051.

Coriat, R., Faivre, S.J., Mir, O., Dreyer, C., Ropert, S., Bouattour, M., Desjardins, R., Goldwasser, F., and Raymond, E. (2016). Pharmacokinetics and safety of DTS-108, a human oligopeptide bound to SN-38 with an esterase-sensitive cross-linker in patients with advanced malignancies: a Phase I study. *Int J Nanomedicine* 11, 6207-6216.

Cristofani, R., Crippa, V., Vezzoli, G., Rusmini, P., Galbiati, M., Cicardi, M.E., Meroni, M., Ferrari, V., Tedesco, B., Piccolella, M., *et al.* (2018). The small heat shock protein B8 (HSPB8) efficiently removes aggregating species of dipeptides produced in C9ORF72-related neurodegenerative diseases. *Cell Stress Chaperones* 23, 1-12.

Crombez, L., Morris, M.C., Dufort, S., Aldrian-Herrada, G., Nguyen, Q., Mc Master, G., Coll, J.-L., Heitz, F., and Divita, G. (2009). Targeting cyclin B1 through peptide-based delivery of siRNA prevents tumour growth. *Nucleic Acids Research* 37, 4559-4569.

Cummins, P.M., Dowling, O., and O'Connor, B.F. (2011). Ion-Exchange Chromatography: Basic Principles and Application to the Partial Purification of Soluble Mammalian Prolyl Oligopeptidase. In *Protein Chromatography: Methods and Protocols*, D. Walls, and S.T. Loughran, eds. (Totowa, NJ: Humana Press), pp. 215-228.

Cummins, P.M., Rochfort, K.D., and O'Connor, B.F. (2017). Ion-Exchange Chromatography: Basic Principles and Application. *Methods Mol Biol* 1485, 209-223.

- Cunningham, S.C., Dane, A.P., Spinoulas, A., Logan, G.J., and Alexander, I.E. (2008). Gene delivery to the juvenile mouse liver using AAV2/8 vectors. *Mol Ther* 16, 1081-1088.
- Curtis, A.F., Masellis, M., Hsiung, G.-Y.R., Moinuddin, R., Zhang, K., Au, B., Millett, G., Mackenzie, I., Rogaeva, E., and Tierney, M.C. (2017). Sex differences in the prevalence of genetic mutations in FTD and ALS. A meta-analysis 89, 1633-1642.
- D'Apolito, M., Colia, A.L., Lasalvia, M., Capozzi, V., Falcone, M.P., Pettoello-Mantovani, M., Brownlee, M., Maffione, A.B., and Giardino, I. (2017). Urea-induced ROS accelerate senescence in endothelial progenitor cells. *Atherosclerosis* 263, 127-136.
- Daaka, Y., Luttrell, L.M., and Lefkowitz, R.J. (1997). Switching of the coupling of the beta2-adrenergic receptor to different G proteins by protein kinase A. *Nature* 390, 88-91.
- Dautzenberg, F.M., Higelin, J., and Teichert, U. (2000). Functional characterization of corticotropin-releasing factor type 1 receptor endogenously expressed in human embryonic kidney 293 cells. *Eur J Pharmacol* 390, 51-59.
- Deceglie, S., Lionetti, C., Roberti, M., Cantatore, P., and Loguercio Polosa, P. (2012). A modified method for the purification of active large enzymes using the glutathione S-transferase expression system. *Analytical Biochemistry* 421, 805-807.
- DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., *et al.* (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245-256.
- Deloche, C., Lopez-Lazaro, L., Mouz, S., Perino, J., Abadie, C., and Combette, J.-M. (2014). XG-102 administered to healthy male volunteers as a single intravenous infusion: a randomized, double-blind, placebo-controlled, dose-escalating study. *Pharmacology Research & Perspectives* 2, e00020.
- Deng, H.-X., Chen, W., Hong, S.-T., Boycott, K.M., Gorrie, G.H., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H., *et al.* (2011). Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 477, 211-215.
- Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G., and Prochiantz, A. (1996). Cell Internalization of the Third Helix of the Antennapedia Homeodomain Is Receptor-independent. *Journal of Biological Chemistry* 271, 18188-18193.
- Dhuri, K., Bechtold, C., Quijano, E., Pham, H., Gupta, A., Vikram, A., and Bahal, R. (2020). Antisense Oligonucleotides: An Emerging Area in Drug Discovery and Development. *J Clin Med* 9.
- Dietz, G.P.H., and Bähr, M. (2007). Synthesis of Cell-Penetrating Peptides and Their Application in Neurobiology. In *Neuroprotection Methods and Protocols*, T. Borsello, ed. (Totowa, NJ: Humana Press), pp. 181-198.

- Dirren, E., Towne, C.L., Setola, V., Redmond, D.E., Jr., Schneider, B.L., and Aebischer, P. (2014). Intracerebroventricular injection of adeno-associated virus 6 and 9 vectors for cell type-specific transgene expression in the spinal cord. *Hum Gene Ther* 25, 109-120.
- Donnelly, C.J., Zhang, P.-W., Pham, J.T., Haeusler, A.R., Mistry, N.A., Vidensky, S., Daley, E.L., Poth, E.M., Hoover, B., Fines, D.M., *et al.* (2013). RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* 80, 415-428.
- Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., Than, M.E., Mackenzie, I.R.A., Capell, A., Schmid, B., *et al.* (2010). ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *The EMBO Journal* 29, 2841-2857.
- Dumont, J., Euwart, D., Mei, B., Estes, S., and Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Critical Reviews in Biotechnology* 36, 1110-1122.
- Duong-Ly, K.C., and Gabelli, S.B. (2014). Using ion exchange chromatography to purify a recombinantly expressed protein. *Methods Enzymol* 541, 95-103.
- Easton, L.E., Shibata, Y., and Lukavsky, P.J. (2010). Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. *RNA* 16, 647-653.
- Edwards, A.M., Arrowsmith, C.H., Christendat, D., Dharamsi, A., Friesen, J.D., Greenblatt, J.F., and Vedadi, M. (2000). Protein production: feeding the crystallographers and NMR spectroscopists. *Nature Structural Biology* 7, 970-972.
- Eggers, M., Vannoy, C.H., Huang, J., Purushothaman, P., Brassard, J., Fonck, C., Meng, H., Prom, M.J., Lawlor, M.W., Cunningham, J., *et al.* (2022). Muscle-directed gene therapy corrects Pompe disease and uncovers species-specific GAA immunogenicity. *EMBO Molecular Medicine* 14, e13968.
- Ejima, D., Yumioka, R., Arakawa, T., and Tsumoto, K. (2005). Arginine as an effective additive in gel permeation chromatography. *Journal of Chromatography A* 1094, 49-55.
- Elden, A.C., Kim, H.-J., Hart, M.P., Chen-Plotkin, A.S., Johnson, B.S., Fang, X., Armakola, M., Geser, F., Greene, R., Lu, M.M., *et al.* (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466, 1069-1075.
- Elmqvist, A., Hansen, M., and Langel, Ü. (2006). Structure–activity relationship study of the cell-penetrating peptide pVEC. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1758, 721-729.
- Esposito, D., and Chatterjee, D.K. (2006). Enhancement of soluble protein expression through the use of fusion tags. *Current Opinion in Biotechnology* 17, 353-358.
- Fadzen, C.M., Holden, R.L., Wolfe, J.M., Choo, Z.-N., Schissel, C.K., Yao, M., Hanson, G.J., and Pentelute, B.L. (2019). Chimeras of Cell-Penetrating Peptides Demonstrate Synergistic Improvement in Antisense Efficacy. *Biochemistry* 58, 3980-3989.

Fanou-Ayi, L., and Vijayalakshmi, M. (1983). Metal-chelate affinity chromatography as a separation tool. *Ann N Y Acad Sci* 413, 300-306.

Farg, M.A., Sundaramoorthy, V., Sultana, J.M., Yang, S., Atkinson, R.A.K., Levina, V., Halloran, M.A., Gleeson, P.A., Blair, I.P., Soo, K.Y., *et al.* (2014). C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Human Molecular Genetics* 23, 3579-3595.

Fecto, F., Yan, J., Vemula, S.P., Liu, E., Yang, Y., Chen, W., Zheng, J.G., Shi, Y., Siddique, N., Arrat, H., *et al.* (2011). SQSTM1 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *JAMA Neurology* 68, 1440-1446.

Federici, T., Taub, J.S., Baum, G.R., Gray, S.J., Grieger, J.C., Matthews, K.A., Handy, C.R., Passini, M.A., Samulski, R.J., and Boulis, N.M. (2012). Robust spinal motor neuron transduction following intrathecal delivery of AAV9 in pigs. *Gene Therapy* 19, 852-859.

Feng, W., Jiang, X., and Geng, X. (1995). The Influence of Urea on the Retention of Protein in Hydrophobic Interaction Chromatography. *Journal of Liquid Chromatography* 18, 217-226.

Ferraiuolo, L., Kirby, J., Grierson, A.J., Sendtner, M., and Shaw, P.J. (2011). Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis. *Nature Reviews Neurology* 7, 616.

Ferreira, C.M.H., Pinto, I.S.S., Soares, E.V., and Soares, H.M.V.M. (2015). (Un)suitability of the use of pH buffers in biological, biochemical and environmental studies and their interaction with metal ions – a review. *RSC Advances* 5, 30989-31003.

Figlewicz, D.A., Krizus, A., Martinoli, M.G., Meininger, V., Dib, M., Rouleau, G.A., and Julien, J.P. (1994). Variants of the heavy neurofilament subunit are associated with the development of amyotrophic lateral sclerosis. *Hum Mol Genet* 3, 1757-1761.

Fischell, J.M., and Fishman, P.S. (2021). A Multifaceted Approach to Optimizing AAV Delivery to the Brain for the Treatment of Neurodegenerative Diseases. *Frontiers in Neuroscience* 15.

Fominaya, J., Bravo, J., and Rebollo, A. (2015). Strategies to stabilize cell penetrating peptides for in vivo applications. *Ther Deliv* 6, 1171-1194.

Foust, K.D., Salazar, D.L., Likhite, S., Ferraiuolo, L., Ditsworth, D., Ilieva, H., Meyer, K., Schmelzer, L., Braun, L., Cleveland, D.W., *et al.* (2013). Therapeutic AAV9-mediated Suppression of Mutant SOD1 Slows Disease Progression and Extends Survival in Models of Inherited ALS. *Molecular Therapy* 21, 2148-2159.

Francis, D.M., and Page, R. (2010). Strategies to Optimize Protein Expression in *E. coli*. *Current Protocols in Protein Science* 61, 5.24.21-25.24.29.

Frankel, A.D., and Pabo, C.O. (1988). Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* 55, 1189-1193.

Freibaum, B.D., Lu, Y., Lopez-Gonzalez, R., Kim, N.C., Almeida, S., Lee, K.-H., Badders, N., Valentine, M., Miller, B.L., Wong, P.C., *et al.* (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* 525, 129-133.

Freibaum, B.D., and Taylor, J.P. (2017). The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. *Frontiers in Molecular neuroscience* 10, 35-35.

French, R.L., Grese, Z.R., Aligireddy, H., Dhavale, D.D., Reeb, A.N., Kedia, N., Kotzbauer, P.T., Bieschke, J., and Ayala, Y.M. (2019). Detection of TAR DNA-binding protein 43 (TDP-43) oligomers as initial intermediate species during aggregate formation. *Journal of Biological Chemistry* 294, 6696-6709.

Fritz, J.S. (2004). Early milestones in the development of ion-exchange chromatography: a personal account. *J Chromatogr A* 1039, 3-12.

Fu, H., Muenzer, J., Samulski, R.J., Breese, G., Sifford, J., Zeng, X., and McCarty, D.M. (2003). Self-complementary adeno-associated virus serotype 2 vector: global distribution and broad dispersion of AAV-mediated transgene expression in mouse brain. *Mol Ther* 8, 911-917.

Futaki, S., Suzuki, T., Ohashi, W., Yagami, T., Tanaka, S., Ueda, K., and Sugiura, Y. (2001). Arginine-rich Peptides: An Abundant Source of Membrane-Permeable Peptides Having Potential As Carriers For Intracellular Protein Delivery. *Journal of Biological Chemistry* 276, 5836-5840.

Gaberc-Porekar, V., and Menart, V. (2001). Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods* 49, 335-360.

Gaglione, R., Pane, K., Dell'Olmo, E., Cafaro, V., Pizzo, E., Olivieri, G., Notomista, E., and Arciello, A. (2019). Cost-effective production of recombinant peptides in *Escherichia coli*. *New Biotechnology* 51, 39-48.

Gantois, I., Khoutorsky, A., Popic, J., Aguilar-Valles, A., Freemantle, E., Cao, R., Sharma, V., Pooters, T., Nagpal, A., Skalecka, A., *et al.* (2017). Metformin ameliorates core deficits in a mouse model of fragile X syndrome. *Nat Med* 23, 674-677.

Gantois, I., Popic, J., Khoutorsky, A., and Sonenberg, N. (2019). Metformin for Treatment of Fragile X Syndrome and Other Neurological Disorders. *Annu Rev Med* 70, 167-181.

Gao, C., Mao, S., Ditzel, H.J., Farnes, L., Wirsching, P., Lerner, R.A., and Janda, K.D. (2002). A cell-penetrating peptide from a novel pVII-pIX phage-displayed random peptide library. *Bioorganic & Medicinal Chemistry* 10, 4057-4065.

Gecchele, E., Merlin, M., Brozzetti, A., Falorni, A., Pezzotti, M., and Avesani, L. (2015). A comparative analysis of recombinant protein expression in different biofactories: bacteria, insect cells and plant systems. *J Vis Exp*, 52459.

Gendron, T.F., Bieniek, K.F., Zhang, Y.-J., Jansen-West, K., Ash, P.E.A., Caulfield, T., Daugherty, L., Dunmore, J.H., Castanedes-Casey, M., Chew, J., *et al.* (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathologica* 126, 829-844.

Gendron, T.F., Chew, J., Stankowski, J.N., Hayes, L.R., Zhang, Y.J., Prudencio, M., Carlomagno, Y., Daugherty, L.M., Jansen-West, K., Perkerson, E.A., *et al.* (2017). Poly(GP) proteins are a useful pharmacodynamic marker for C9ORF72-associated amyotrophic lateral sclerosis. *Sci Transl Med* 9.

Gendron, T.F., and Petrucelli, L. (2018). Disease Mechanisms of C9ORF72 Repeat Expansions. *Cold Spring Harbor Perspectives in Medicine* 8.

Ghosh, S., Brown, A.M., Jenkins, C., and Campbell, K. (2020). Viral Vector Systems for Gene Therapy: A Comprehensive Literature Review of Progress and Biosafety Challenges. *Appl Biosaf* 25, 7-18.

Giardino, I., D'Apolito, M., Brownlee, M., Maffione, A.B., Colia, A.L., Sacco, M., Ferrara, P., and Pettoello-Mantovani, M. (2017). Vascular toxicity of urea, a new "old player" in the pathogenesis of chronic renal failure induced cardiovascular diseases. *Turk Pediatri Ars* 52, 187-193.

Gijselinck, I., Van Mossevelde, S., van der Zee, J., Sieben, A., Engelborghs, S., De Bleecker, J., Ivanoiu, A., Deryck, O., Edbauer, D., Zhang, M., *et al.* (2016). The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Molecular Psychiatry* 21, 1112-1124.

Goodman, L.D., Prudencio, M., Srinivasan, A.R., Rifai, O.M., Lee, V.M.-Y., Petrucelli, L., and Bonini, N.M. (2019). eIF4B and eIF4H mediate GR production from expanded G4C2 in a *Drosophila* model for C9orf72-associated ALS. *Acta Neuropathologica Communications* 7, 62.

Gordon, P.H. (2013). Amyotrophic Lateral Sclerosis: An update for 2013 Clinical Features, Pathophysiology, Management and Therapeutic Trials. *Aging and Disease* 4, 295-310.

Goula, A.V., Stys, A., Chan, J.P., Trottier, Y., Festenstein, R., and Merienne, K. (2012). Transcription elongation and tissue-specific somatic CAG instability. *PLoS Genet* 8, e1003051.

Govindarajan, S., Sivakumar, J., Garimidi, P., Rangaraj, N., Kumar, J.M., Rao, N.M., and Gopal, V. (2012). Targeting human epidermal growth factor receptor 2 by a cell-penetrating peptide-affibody bioconjugate. *Biomaterials* 33, 2570-2582.

Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. (1977). Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology* 36, 59-72.

Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., *et al.* (2008). Protein production and purification. *Nature Methods* 5, 135-146.

Gray, S.J., Matagne, V., Bachaboina, L., Yadav, S., Ojeda, S.R., and Samulski, R.J. (2011). Preclinical differences of intravascular AAV9 delivery to neurons and glia: a comparative study of adult mice and nonhuman primates. *Mol Ther* 19, 1058-1069.

Gray, S.J., Nagabhushan Kalburgi, S., McCown, T.J., and Jude Samulski, R. (2013). Global CNS gene delivery and evasion of anti-AAV-neutralizing antibodies by intrathecal AAV administration in non-human primates. *Gene Ther* 20, 450-459.

Green, K.M., Glineburg, M.R., Kearse, M.G., Flores, B.N., Linsalata, A.E., Fedak, S.J., Goldstrohm, A.C., Barmada, S.J., and Todd, P.K. (2017). RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nature Communications* 8, 2005.

Greenway, M.J., Andersen, P.M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., Patterson, V., Swingler, R., Kieran, D., Prehn, J., *et al.* (2006). ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nature Genetics* 38, 411-413.

Grima, J.C., Daigle, J.G., Arbez, N., Cunningham, K.C., Zhang, K., Ochaba, J., Geater, C., Morozko, E., Stocksdales, J., Glatzer, J.C., *et al.* (2017). Mutant Huntingtin Disrupts the Nuclear Pore Complex. *Neuron* 94, 93-107.e106.

Gros-Louis, F., Larivière, R., Gowing, G., Laurent, S., Camu, W., Bouchard, J.P., Meininger, V., Rouleau, G.A., and Julien, J.P. (2004). A frameshift deletion in peripherin gene associated with amyotrophic lateral sclerosis. *J Biol Chem* 279, 45951-45956.

Guidotti, G., Brambilla, L., and Rossi, D. (2017). Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends in Pharmacological Sciences* 38, 406-424.

Gulías, P., Guerra-Varela, J., Gonzalez-Aparicio, M., Ricobaraza, A., Vales, A., Gonzalez-Aseguinolaza, G., Hernandez-Alcoceba, R., and Sánchez, L. (2019). Danio Rerio as Model Organism for Adenoviral Vector Evaluation. *Genes (Basel)* 10.

Guo, J., Wu, C., Lin, X., Zhou, J., Zhang, J., Zheng, W., Wang, T., and Cui, Y. (2021). Establishment of a simplified dichotomic size-exclusion chromatography for isolating extracellular vesicles toward clinical applications. *Journal of Extracellular Vesicles* 10, e12145.

Gurney, L.R.I., Taggart, J., Tong, W.-C., Jones, A.T., Robson, S.C., and Taggart, M.J. (2018). Inhibition of Inflammatory Changes in Human Myometrial Cells by Cell Penetrating Peptide and Small Molecule Inhibitors of NFκB. *Frontiers in Immunology* 9.

Gustafsson, C., Govindarajan, S., and Minshall, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology* 22, 346-353.

Habault, J., and Poyet, J.-L. (2019). Recent Advances in Cell Penetrating Peptide-Based Anticancer Therapies. *Molecules* 24, 927.

Hadano, S., Hand, C.K., Osuga, H., Yanagisawa, Y., Otomo, A., Devon, R.S., Miyamoto, N., Showguchi-Miyata, J., Okada, Y., Singaraja, R., *et al.* (2001). A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nature Genetics* 29, 166-173.

Haeusler, A.R., Donnelly, C.J., Periz, G., Simko, E.A.J., Shaw, P.G., Kim, M.-S., Maragakis, N.J., Troncoso, J.C., Pandey, A., Sattler, R., *et al.* (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507, 195.

Hage, D.S. (1999). Affinity chromatography: a review of clinical applications. *Clin Chem* 45, 593-615.

Hage, D.S., and Cazes, J. (2005). *Handbook of Affinity Chromatography* (CRC Press).

Hand, C.K., Khoris, J., Salachas, F., Gros-Louis, F., Lopes, A.A.S., Mayeux-Portas, V., Brewer, C.G., Brown, R.H., Jr., Meininger, V., Camu, W., *et al.* (2002). A novel locus for familial amyotrophic lateral sclerosis, on chromosome 18q. *American Journal of Human Genetics* 70, 251-256.

Haney, E.F., Straus, S.K., and Hancock, R.E.W. (2019). Reassessing the Host Defense Peptide Landscape. *Frontiers in Chemistry* 7.

Hao, Z., Liu, L., Tao, Z., Wang, R., Ren, H., Sun, H., Lin, Z., Zhang, Z., Mu, C., Zhou, J., *et al.* (2019). Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. *Nature Communications* 10, 2906.

Hardiman, O., Al-Chalabi, A., Chio, A., Corr, E.M., Logroscino, G., Robberecht, W., Shaw, P.J., Simmons, Z., and van den Berg, L.H. (2017). Amyotrophic lateral sclerosis. *Nature Reviews Disease Primers* 3, 17071.

Harper, S., and Speicher, D.W. (2001). Expression and purification of GST fusion proteins. *Curr Protoc Protein Sci Chapter 6*, Unit 6.6.

Harper, S., and Speicher, D.W. (2011). Purification of Proteins Fused to Glutathione S-Transferase. In *Protein Chromatography: Methods and Protocols*, D. Walls, and S.T. Loughran, eds. (Totowa, NJ: Humana Press), pp. 259-280.

Hautbergue, G.M. (2016). Widespread RNA Dysregulation in Neurodegeneration: Challenges and Opportunities. *Austin Neurology* 1, 1002.

Hautbergue, G.M., Castelli, L.M., Ferraiuolo, L., Sanchez-Martinez, A., Cooper-Knock, J., Higginbottom, A., Lin, Y.-H., Bauer, C.S., Dodd, J.E., Myszczyńska, M.A., *et al.* (2017). SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nature Communications* 8, 16063-16063.

Hautbergue, G.M., Cleary, J.D., Guo, S., and Ranum, L.P.W. (2021). Therapeutic strategies for C9orf72 amyotrophic lateral sclerosis and frontotemporal dementia. *Current Opinion in Neurology* 34, 748-755.

Hautbergue, G.M., Hung, M.-L., Golovanov, A.P., Lian, L.-Y., and Wilson, S.A. (2008). Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proceedings of the National Academy of Sciences* 105, 5154-5159.

Hayashi, Y., Homma, K., and Ichijo, H. (2016). SOD1 in neurotoxicity and its controversial roles in SOD1 mutation-negative ALS. *Advances in Biological Regulation* 60, 95-104.

Heckl, D., Kowalczyk, M.S., Yudovich, D., Belizaire, R., Puram, R.V., McConkey, M.E., Thielke, A., Aster, J.C., Regev, A., and Ebert, B.L. (2014). Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat Biotechnol* 32, 941-946.

Heidenreich, M., and Zhang, F. (2016). Applications of CRISPR-Cas systems in neuroscience. *Nat Rev Neurosci* 17, 36-44.

Heitz, F., Morris, M.C., and Divita, G. (2009). Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *British Journal of Pharmacology* 157, 195-206.

Held, D., and Kilz, P. (2021). Size-exclusion chromatography as a useful tool for the assessment of polymer quality and determination of macromolecular properties. *Chemistry Teacher International* 3, 77-103.

Henning-Knechtel, A., Kumar, S., Wallin, C., Król, S., Wärmländer, S.K.T.S., Jarvet, J., Esposito, G., Kirmizialtin, S., Gräslund, A., Hamilton, A.D., *et al.* (2020). Designed Cell-Penetrating Peptide Inhibitors of Amyloid-beta Aggregation and Cytotoxicity. *Cell Reports Physical Science* 1, 100014.

Hentati, A., Bejaoui, K., Pericak-Vance, M.A., Hentati, F., Speer, M.C., Hung, W.-Y., Figlewicz, D.A., Haines, J., Rimmler, J., Ben Hamida, C., *et al.* (1994). Linkage of recessive familial amyotrophic lateral sclerosis to chromosome 2q33–q35. *Nature Genetics* 7, 425-428.

Herbert, B. (1999). Advances in protein solubilisation for two-dimensional electrophoresis. *Electrophoresis* 20, 660-663.

Hinderer, C., Katz, N., Buza, E.L., Dyer, C., Goode, T., Bell, P., Richman, L.K., and Wilson, J.M. (2018). Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN. *Hum Gene Ther* 29, 285-298.

Hochuli, E., Döbeli, H., and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *Journal of Chromatography A* 411, 177-184.

Hocquemiller, M., Giersch, L., Audrain, M., Parker, S., and Cartier, N. (2016). Adeno-Associated Virus-Based Gene Therapy for CNS Diseases. *Hum Gene Ther* 27, 478-496.

Hoffmann, D., Schott, J.W., Geis, F.K., Lange, L., Müller, F.J., Lenz, D., Zychlinski, D., Steinemann, D., Morgan, M., Moritz, T., *et al.* (2017). Detailed comparison of retroviral vectors and promoter configurations for stable and high transgene expression in human induced pluripotent stem cells. *Gene Therapy* 24, 298-307.

Hong, P., Koza, S., and Bouvier, E.S. (2012). Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *J Liq Chromatogr Relat Technol* 35, 2923-2950.

Hou, Y., Lautrup, S., Cordonnier, S., Wang, Y., Croteau, D.L., Zavala, E., Zhang, Y., Moritoh, K., O'Connell, J.F., Baptiste, B.A., *et al.* (2018). NAD(+) supplementation normalizes key Alzheimer's features and DNA damage responses in a new AD mouse model with introduced DNA repair deficiency. *Proc Natl Acad Sci U S A* *115*, E1876-e1885.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., *et al.* (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* *496*, 498-503.

Hu, J., Rigo, F., Prakash, T.P., and Corey, D.R. (2017). Recognition of c9orf72 Mutant RNA by Single-Stranded Silencing RNAs. *Nucleic Acid Ther* *27*, 87-94.

Huang, Y., and Steitz, J.A. (2005). SRprises along a Messenger's Journey. *Molecular Cell* *17*, 613-615.

Huang, Y., Yario, T.A., and Steitz, J.A. (2004). A molecular link between SR protein dephosphorylation and mRNA export. *Proceedings of the National Academy of Sciences of the United States of America* *101*, 9666-9670.

Hudry, E., Andres-Mateos, E., Lerner, E.P., Volak, A., Cohen, O., Hyman, B.T., Maguire, C.A., and Vandenberghe, L.H. (2018). Efficient Gene Transfer to the Central Nervous System by Single-Stranded Anc80L65. *Molecular Therapy - Methods & Clinical Development* *10*, 197-209.

Hunter, J.E., Molony, C.M., Bagel, J.H., O'Donnell, P.A., Kaler, S.G., and Wolfe, J.H. (2022). Transduction characteristics of alternative adeno-associated virus serotypes in the cat brain by intracisternal delivery. *Molecular Therapy - Methods & Clinical Development* *26*, 384-393.

Iannitti, T., Scarrott, J.M., Likhite, S., Coldicott, I.R.P., Lewis, K.E., Heath, P.R., Higginbottom, A., Myszczyńska, M.A., Milo, M., Hautbergue, G.M., *et al.* (2018). Translating SOD1 Gene Silencing toward the Clinic: A Highly Efficacious, Off-Target-free, and Biomarker-Supported Strategy for fALS. *Molecular Therapy - Nucleic Acids* *12*, 75-88.

Inagaki, K., Fuess, S., Storm, T.A., Gibson, G.A., McTiernan, C.F., Kay, M.A., and Nakai, H. (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol Ther* *14*, 45-53.

Islam, T., Aguilar-Yañez, J.M., Simental-Martínez, J., Ortiz-Alcaraz, C.I., Rito-Palomares, M., and Fernandez-Lahore, M. (2014). A novel strategy for the purification of a recombinant protein using ceramic fluorapatite-binding peptides as affinity tags. *Journal of Chromatography A* *1339*, 26-33.

Ito, H., Wate, R., Zhang, J., Ohnishi, S., Kaneko, S., Ito, H., Nakano, S., and Kusaka, H. (2008). Treatment with edaravone, initiated at symptom onset, slows motor decline and decreases SOD1 deposition in ALS mice. *Experimental Neurology* *213*, 448-455.

Jackson, D.A. (2011). The Cell Nucleus: Biogenesis, Structure, and Function. In *Reviews in Cell Biology and Molecular Medicine*.

Järver, P., Mäger, I., and Langel, Ü. (2010). In vivo biodistribution and efficacy of peptide mediated delivery. *Trends Pharmacol Sci* 31, 528-535.

Jayanthi, S., Gundampati, R.K., and Kumar, T.K.S. (2017). Simple and Efficient Purification of Recombinant Proteins Using the Heparin-Binding Affinity Tag. *Current Protocols in Protein Science* 90, 6.16.11-16.16.13.

Jia, B., and Jeon, C.O. (2016). High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biology* 6, 160196.

Jia, L., Gorman, G.S., Coward, L.U., Noker, P.E., McCormick, D., Horn, T.L., Harder, J.B., Muzzio, M., Prabhakar, B., Ganesh, B., *et al.* (2011). Preclinical pharmacokinetics, metabolism, and toxicity of azurin-p28 (NSC745104) a peptide inhibitor of p53 ubiquitination. *Cancer Chemotherapy and Pharmacology* 68, 513-524.

Jiang, J., and Ravits, J. (2019). Pathogenic Mechanisms and Therapy Development for C9orf72 Amyotrophic Lateral Sclerosis/Frontotemporal Dementia. *Neurotherapeutics* 16, 1115-1132.

Jiang, J., Zhu, Q., Gendron, T.F., Saberi, S., McAlonis-Downes, M., Seelman, A., Stauffer, J.E., Jafar-Nejad, P., Drenner, K., Schulte, D., *et al.* (2016). Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron* 90, 535-550.

Johnson, J.O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V.M., Trojanowski, J.Q., Gibbs, J.R., Brunetti, M., Gronka, S., Wu, J., *et al.* (2010). Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* 68, 857-864.

Johnson, J.O., Pioro, E.P., Boehringer, A., Chia, R., Feit, H., Renton, A.E., Pliner, H.A., Abramzon, Y., Marangi, G., Winborn, B.J., *et al.* (2014). Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nature Neuroscience* 17, 664-666.

Joliot, A., Pernelle, C., Deagostini-Bazin, H., and Prochiantz, A. (1991). Antennapedia homeobox peptide regulates neural morphogenesis. *Proceedings of the National Academy of Sciences* 88, 1864-1868.

Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S.B., Paul, J.W., 3rd, Sun, S., Herdy, J.R., Bieri, G., *et al.* (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci* 18, 1226-1229.

Kabashi, E., Valdmanis, P.N., Dion, P., Spiegelman, D., McConkey, B.J., Velde, C.V., Bouchard, J.-P., Lacomblez, L., Pochigaeva, K., Salachas, F., *et al.* (2008). TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nature Genetics* 40, 572.

Kamei, N., Okada, N., Ikeda, T., Choi, H., Fujiwara, Y., Okumura, H., and Takeda-Morishita, M. (2018). Effective nose-to-brain delivery of exendin-4 via coadministration with cell-penetrating peptides for improving progressive cognitive dysfunction. *Sci Rep* 8, 17641.

Kamei, N., Shingaki, T., Kanayama, Y., Tanaka, M., Zochi, R., Hasegawa, K., Watanabe, Y., and Takeda-Morishita, M. (2016). Visualization and Quantitative Assessment of the Brain Distribution of Insulin through Nose-to-Brain Delivery Based on the Cell-Penetrating Peptide Noncovalent Strategy. *Mol Pharm* 13, 1004-1011.

Kaneb, H.M., Folkmann, A.W., Belzil, V.V., Jao, L.-E., Leblond, C.S., Girard, S.L., Daoud, H., Noreau, A., Rochefort, D., Hince, P., *et al.* (2014). Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. *Human Molecular Genetics* 24, 1363-1373.

Karlsson, E., and Hirsh, I. (2011). Ion exchange chromatography. *Methods Biochem Anal* 54, 93-133.

Karlsson, G., and Winge, S. (2004). Separation of latent, prelatent, and native forms of human antithrombin by heparin affinity high-performance liquid chromatography. *Protein Expr Purif* 33, 339-345.

Katahira, J. (2015). Nuclear export of messenger RNA. *Genes (Basel)* 6, 163-184.

Kearse, M.G., Green, K.M., Krans, A., Rodriguez, C.M., Linsalata, A.E., Goldstrohm, A.C., and Todd, P.K. (2016). CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Molecular Cell* 62, 314-322.

Kenna, K.P., van Doormaal, P.T., Dekker, A.M., Ticozzi, N., Kenna, B.J., Diekstra, F.P., van Rheenen, W., van Eijk, K.R., Jones, A.R., Keagle, P., *et al.* (2016). NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. *Nat Genet* 48, 1037-1042.

Khosravi, B., LaClair, K.D., Riemenschneider, H., Zhou, Q., Frottin, F., Mareljic, N., Czuppa, M., Farny, D., Hartmann, H., Michaelson, M., *et al.* (2020). Cell-to-cell transmission of C9orf72 poly-(Gly-Ala) triggers key features of ALS/FTD. *The EMBO Journal* 39, e102811.

Kiernan, M.C., Vucic, S., Cheah, B.C., Turner, M.R., Eisen, A., Hardiman, O., Burrell, J.R., and Zoing, M.C. (2011). Amyotrophic lateral sclerosis. *The Lancet* 377, 942-955.

Kihira, T., Yoshida, S., Kondo, T., Iwai, K., Wada, S., Morinaga, S., Kazimoto, Y., Kondo, T., Okamoto, K., Kokubo, Y., *et al.* (2012). An increase in ALS incidence on the Kii Peninsula, 1960–2009: A possible link to change in drinking water source. *Amyotrophic Lateral Sclerosis* 13, 347-350.

Kilz, P., and H.Pasch (2000). Coupled LC Techniques in Molecular Characterization. In, pp. 7495-7543.

Kim, H.J., Kim, N.C., Wang, Y.-D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A., *et al.* (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495, 467-473.

Kim, J.-E., Hong, Y.H., Kim, J.Y., Jeon, G.S., Jung, J.H., Yoon, B.-N., Son, S.-Y., Lee, K.-W., Kim, J.-I., and Sung, J.-J. (2017). Altered nucleocytoplasmic proteome and transcriptome distributions in an in vitro model of amyotrophic lateral sclerosis. *PLOS ONE* 12, e0176462.

Kimble, M.E., Brill, A.L., and Pasker, R.L. (2013). Overview of Affinity Tags for Protein Purification. *Current Protocols in Protein Science* 73, 9.9.1-9.9.23.

Kino, Y., Washizu, C., Kurosawa, M., Yamada, M., Miyazaki, H., Akagi, T., Hashikawa, T., Doi, H., Takumi, T., Hicks, G.G., *et al.* (2015). FUS/TLS deficiency causes behavioral and pathological abnormalities distinct from amyotrophic lateral sclerosis. *Acta Neuropathologica Communications* 3, 24.

Kinoshita, Y., Ito, H., Hirano, A., Fujita, K., Wate, R., Nakamura, M., Kaneko, S., Nakano, S., and Kusaka, H. (2009). Nuclear Contour Irregularity and Abnormal Transporter Protein Distribution in Anterior Horn Cells in Amyotrophic Lateral Sclerosis. *Journal of Neuropathology & Experimental Neurology* 68, 1184-1192.

Köhler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Molecular Cell Biology* 8, 761-773.

Kopp, J., Slouka, C., Ulonska, S., Kager, J., Fricke, J., Spadiut, O., and Herwig, C. (2018). Impact of Glycerol as Carbon Source onto Specific Sugar and Inducer Uptake Rates and Inclusion Body Productivity in *E. coli* BL21(DE3). *Bioengineering* 5, 1.

Koren, E., and Torchilin, V.P. (2012). Cell-penetrating peptides: breaking through to the other side. *Trends in Molecular Medicine* 18, 385-393.

Kotterman, M.A., and Schaffer, D.V. (2014). Engineering adeno-associated viruses for clinical gene therapy. *Nat Rev Genet* 15, 445-451.

Kovanda, A., Zalar, M., Šket, P., Plavec, J., and Rogelj, B. (2015). Anti-sense DNA d(GGCCCC)_n expansions in C9ORF72 form i-motifs and protonated hairpins. *Scientific Reports* 5, 17944.

Kram, K.E., and Finkel, S.E. (2015). Rich Medium Composition Affects Survival, Glycation, and Mutation Frequency *Escherichia coli* during Long-Term Batch Culture. *Applied and Environmental Microbiology* 81, 4442-4450.

Kreisig, T., Prasse, A.A., Zscharnack, K., Volke, D., and Zuchner, T. (2014). His-tag protein monitoring by a fast mix-and-measure immunoassay. *Scientific Reports* 4, 5613.

Krishnan, G., Zhang, Y., Gu, Y., Kankel, M.W., Gao, F.B., and Almeida, S. (2020). CRISPR deletion of the C9ORF72 promoter in ALS/FTD patient motor neurons abolishes production of dipeptide repeat proteins and rescues neurodegeneration. *Acta Neuropathol* 140, 81-84.

Kumar, A., Galaev, I.Y., and Mattiasson, B. (2000). Polymer displacement/shielding in protein chromatography. *J Chromatogr B Biomed Sci Appl* 741, 103-113.

Kumar, P., Kizhakkedathu, J.N., and Straus, S.K. (2018). Antimicrobial Peptides: Diversity, Mechanism of Action and Strategies to Improve the Activity and Biocompatibility In Vivo. *Biomolecules* 8, 4.

Kuzmin, D.A., Shutova, M.V., Johnston, N.R., Smith, O.P., Fedorin, V.V., Kukushkin, Y.S., van der Loo, J.C.M., and Johnstone, E.C. (2021). The clinical landscape for AAV gene therapies. *Nat Rev Drug Discov* 20, 173-174.

Kwiatkowski, T.J., Bosco, D.A., LeClerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., *et al.* (2009). Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science* 323, 1205-1208.

Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., Kim, J., Yun, J., Xie, Y., and McKnight, S.L. (2014). Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* 345, 1139-1145.

Laflamme, C., McKeever, P.M., Kumar, R., Schwartz, J., Kolahdouzan, M., Chen, C.X., You, Z., Benaliouad, F., Gileadi, O., McBride, H.M., *et al.* (2019). Implementation of an antibody characterization procedure and application to the major ALS/FTD disease gene C9ORF72. *eLife* 8, e48363.

Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H.-R., Jiang, J., Watt, A.T., Chun, S., Katz, M., *et al.* (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proceedings of the National Academy of Sciences of the United States of America* 110, E4530-E4539.

Langel, Ü. (2019). CPP, Cell-Penetrating Peptides (Springer Singapore).

Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., and Brenner, S.E. (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* 446, 926-929.

Layalle, S., They, L., Ourghani, S., Raoul, C., and Soustelle, L. (2021). Amyotrophic Lateral Sclerosis Genes in *Drosophila melanogaster*. *Int J Mol Sci* 22.

LeCher Julia, C., Nowak Scott, J., and McMurry Jonathan, L. (2017). Breaking in and busting out: cell-penetrating peptides and the endosomal escape problem. In *Biomolecular Concepts*, pp. 131.

Lee, A.C.-L., Harris, J.L., Khanna, K.K., and Hong, J.-H. (2019). A Comprehensive Review on Current Advances in Peptide Drug Development and Design. *International Journal of Molecular Sciences* 20, 2383.

Lee, K.-H., Zhang, P., Kim, H.J., Mitrea, D.M., Sarkar, M., Freibaum, B.D., Cika, J., Coughlin, M., Messing, J., Molliex, A., *et al.* (2016a). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* 167, 774-788.e717.

Lee, S.-H., Lee, H.-K., Kim, C., Kim, Y.-K., Ismail, T., Jeong, Y., Park, K., Park, J.-W., Kwon, O.-S., Kang, B.S., *et al.* (2016b). The splicing factor SRSF1 modulates pattern formation by inhibiting transcription of tissue specific genes during embryogenesis. *Biochemical and Biophysical Research Communications* 477, 1011-1016.

- Lee, Y.-B., Chen, H.-J., Peres, J.N., Gomez-Deza, J., Attig, J., Stalekar, M., Troakes, C., Nishimura, A.L., Scotter, E.L., Vance, C., *et al.* (2013). Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Reports* 5, 1178-1186.
- Legrain, P., and Rosbash, M. (1989). Some cis and trans acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell* 57, 573-583.
- Lehmer, C., Oeckl, P., Weishaupt, J.H., Volk, A.E., Diehl-Schmid, J., Schroeter, M.L., Lauer, M., Kornhuber, J., Levin, J., Fassbender, K., *et al.* (2017). Poly-GP in cerebrospinal fluid links C9orf72-associated dipeptide repeat expression to the asymptomatic phase of ALS/FTD. *EMBO Mol Med* 9, 859-868.
- Lehnhardt, M., Vu, P., Kuhnen, C., Steinstraesser, L., Muehlberger, T., Druecke, D., Steinau, H.U., and Homann, H.H. (2005). [Serious complications of injections--retrospective analysis of incidences, complication-management, prophylaxis and economic aspects]. *Zentralbl Chir* 130, 162-169.
- Lei, J., Sun, L., Huang, S., Zhu, C., Li, P., He, J., Mackey, V., Coy, D.H., and He, Q. (2019). The antimicrobial peptides and their potential clinical applications. *Am J Transl Res* 11, 3919-3931.
- Levine, T.P., Daniels, R.D., Gatta, A.T., Wong, L.H., and Hayes, M.J. (2013). The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics (Oxford, England)* 29, 499-503.
- Li, C., and Samulski, R.J. (2020). Engineering adeno-associated virus vectors for gene therapy. *Nat Rev Genet* 21, 255-272.
- Li, Y. (2011). Self-cleaving fusion tags for recombinant protein production. *Biotechnology Letters* 33, 869-881.
- Licata, N.V., Cristofani, R., Salomonsson, S., Wilson, K.M., Kempthorne, L., Vaizoglu, D., D'Agostino, V.G., Pollini, D., Loffredo, R., Pancher, M., *et al.* (2022). C9orf72 ALS/FTD dipeptide repeat protein levels are reduced by small molecules that inhibit PKA or enhance protein degradation. *Embo j* 41, e105026.
- Liguori, F., Amadio, S., and Volonté, C. (2021). Fly for ALS: *Drosophila* modeling on the route to amyotrophic lateral sclerosis modifiers. *Cell Mol Life Sci* 78, 6143-6160.
- Lim, K., Ho, J.X., Keeling, K., Gilliland, G.L., Ji, X., Rüker, F., and Carter, D.C. (1994). Three-Dimensional structure of schistosoma japonicum glutathione s-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from hiv. *Protein Science* 3, 2233-2244.
- Lim, K.R., Maruyama, R., and Yokota, T. (2017). Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des Devel Ther* 11, 533-545.
- Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., *et al.* (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications* 5, 4767.

Lipinski, C.A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *Journal of Pharmacological and Toxicological Methods* 44, 235-249.

Liu, G., Chen, X., Bissler, J.J., Sinden, R.R., and Leffak, M. (2010). Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells. *Nat Chem Biol* 6, 652-659.

Liu, S., Li, Z., Yu, B., Wang, S., Shen, Y., and Cong, H. (2020). Recent advances on protein separation and purification methods. *Advances in Colloid and Interface Science* 284, 102254.

Liu, X., Lin, P., Perrett, I., Lin, J., Liao, Y.-P., Chang, C.H., Jiang, J., Wu, N., Donahue, T., Wainberg, Z., *et al.* (2017). Tumor-penetrating peptide enhances transcytosis of silicasome-based chemotherapy for pancreatic cancer. *The Journal of Clinical Investigation* 127, 2007-2018.

Liu, X., Yang, H., Xiong, P., Li, W., Huang, H.-H., and Zheng, Y. (2019). Comparative studies of Tris-HCl, HEPES and NaHCO₃/CO₂ buffer systems on the biodegradation behaviour of pure Zn in NaCl and SBF solutions. *Corrosion Science* 157, 205-219.

Liu, Y., Andreucci, A., Iwamoto, N., Yin, Y., Yang, H., Liu, F., Bulychev, A., Hu, X.S., Lin, X., Lamore, S., *et al.* (2022). Preclinical evaluation of WVE-004, an investigational stereopure oligonucleotide for the treatment of C9orf72-associated ALS or FTD. *Molecular Therapy - Nucleic Acids* 28, 558-570.

Liu, Y., Kim, Y.J., Ji, M., Fang, J., Siriwon, N., Zhang, L.I., and Wang, P. (2014). Enhancing gene delivery of adeno-associated viruses by cell-permeable peptides. *Molecular Therapy - Methods & Clinical Development* 1, 12.

Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O., Borchelt, D.R., Yachnis, A.T., and Ranum, L.P. (2016). C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. *Neuron* 90, 521-534.

Liu, Y.P., and Berkhout, B. (2014). HIV-1-based lentiviral vectors. *Methods Mol Biol* 1087, 273-284.

López Castel, A., Cleary, J.D., and Pearson, C.E. (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol* 11, 165-170.

Loreto, S., Cuypers, B., Brokken, J., Van Doorslaer, S., De Wael, K., and Meynen, V. (2017). The effect of the buffer solution on the adsorption and stability of horse heart myoglobin on commercial mesoporous titanium dioxide: a matter of the right choice. *Physical Chemistry Chemical Physics* 19, 13503-13514.

Lowe, C.R. (2001). Combinatorial approaches to affinity chromatography. *Curr Opin Chem Biol* 5, 248-256.

Lu, Y.-C., Yeh, W.-C., and Ohashi, P.S. (2008). LPS/TLR4 signal transduction pathway. *Cytokine* 42, 145-151.

Lukashchuk, V., Lewis, K.E., Coldicott, I., Grierson, A.J., and Azzouz, M. (2016). AAV9-mediated central nervous system-targeted gene delivery via cisterna magna route in mice. *Mol Ther Methods Clin Dev* 3, 15055.

Lund, M.K., and Guthrie, C. (2005). The DEAD-Box Protein Dbp5p Is Required to Dissociate Mex67p from Exported mRNPs at the Nuclear Rim. *Molecular Cell* 20, 645-651.

Lundstrom, K. (2019). RNA Viruses as Tools in Gene Therapy and Vaccine Development. *Genes (Basel)* 10.

Luty, A.A., Kwok, J.B.J., Dobson-Stone, C., Loy, C.T., Coupland, K.G., Karlström, H., Sobow, T., Tchorzewska, J., Maruszak, A., Barcikowska, M., *et al.* (2010). Sigma nonopioid intracellular receptor 1 mutations cause frontotemporal lobar degeneration–motor neuron disease. *Annals of Neurology* 68, 639-649.

Lydia M. Castelli, Y.-H.L., Laura Ferraiuolo, Alvaro Sanchez-Martinez, Ke Ning, Mimoun Azzouz, Alexander J. Whitworth, Pamela J. Shaw, Guillaume M. Hautbergue (2018). SRSF1-dependent nuclear export of C9ORF72 repeat-transcripts: targeting toxic gain-of-functions induced by protein sequestration as a selective therapeutic strategy for neuroprotection. *Therapeutic Targets for Neurological Diseases* 4, 2376-0478.

Mackenzie, I.R.A., Bigio, E.H., Ince, P.G., Geser, F., Neumann, M., Cairns, N.J., Kwong, L.K., Forman, M.S., Ravits, J., Stewart, H., *et al.* (2007). Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. *Annals of Neurology* 61, 427-434.

Mackenzie, I.R.A., and Rademakers, R. (2008). The role of transactive response DNA-binding protein-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Current Opinion in Neurology* 21, 693-700.

Madani, F., Lindberg, S., Langel, Ü., Futaki, S., and Graslund, A. (2011). Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. *Journal of Biophysics* 2011.

Mäe, M., El Andaloussi, S., Lundin, P., Oskolkov, N., Johansson, H.J., Guterstam, P., and Langel, Ü. (2009). A stearylated CPP for delivery of splice correcting oligonucleotides using a non-covalent co-incubation strategy. *Journal of Controlled Release* 134, 221-227.

Magder, S. (2006). Reactive oxygen species: toxic molecules or spark of life? *Critical Care* 10, 208.

Maguire, A.M., Simonelli, F., Pierce, E.A., Pugh, E.N., Mingozzi, F., Bennicelli, J., Banfi, S., Marshall, K.A., Testa, F., Surace, E.M., *et al.* (2008). Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis. *New England Journal of Medicine* 358, 2240-2248.

Majounie, E., Renton, A.E., Mok, K., Dopper, E.G.P., Waite, A., Rollinson, S., Chiò, A., Restagno, G., Nicolaou, N., Simon-Sanchez, J., *et al.* (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *The Lancet Neurology* 11, 323-330.

Malhotra, A. (2009). Chapter 16 Tagging for Protein Expression. In *Methods in Enzymology*, R.R. Burgess, and M.P. Deutscher, eds. (Academic Press), pp. 239-258.

Malm, D., and Nilssen, Ø. (2008). Alpha-mannosidosis. *Orphanet J Rare Dis* 3, 21.

Manjaly, Z.R., Scott, K.M., Abhinav, K., Wijesekera, L., Ganesalingam, J., Goldstein, L.H., Janssen, A., Dougherty, A., Willey, E., Stanton, B.R., *et al.* (2010). The sex ratio in amyotrophic lateral sclerosis: A population based study. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases* 11, 439-442.

Mann, D.M.A., Rollinson, S., Robinson, A., Bennion Callister, J., Thompson, J.C., Snowden, J.S., Gendron, T., Petrucelli, L., Masuda-Suzukake, M., Hasegawa, M., *et al.* (2013). Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathologica Communications* 1, 68-68.

Manning, M.C., Chou, D.K., Murphy, B.M., Payne, R.W., and Katayama, D.S. (2010). Stability of Protein Pharmaceuticals: An Update. *Pharmaceutical Research* 27, 544-575.

Marin, B., Hamidou, B., Couratier, P., Nicol, M., Delzor, A., Raymondeau, M., Druet-Cabanac, M., Lautrette, G., Boumediene, F., Preux, P.M., *et al.* (2014). Population-based epidemiology of amyotrophic lateral sclerosis (ALS) in an ageing Europe – the French register of ALS in Limousin (FRALim register). *European Journal of Neurology* 21, 1292-e1279.

Marinescu, G.C., Popescu, R.-G., and Dinischiotu, A. (2018). Size Exclusion Chromatography Method for Purification of Nicotinamide Mononucleotide (NMN) from Bacterial Cells. *Scientific Reports* 8, 4433.

Marks, J.R., Placone, J., Hristova, K., and Wimley, W.C. (2011). Spontaneous Membrane-Translocating Peptides by Orthogonal High-Throughput Screening. *Journal of the American Chemical Society* 133, 8995-9004.

Martier, R., Liefhebber, J.M., García-Osta, A., Miniarikova, J., Cuadrado-Tejedor, M., Espelosin, M., Ursua, S., Petry, H., van Deventer, S.J., Evers, M.M., *et al.* (2019). Targeting RNA-Mediated Toxicity in C9orf72 ALS and/or FTD by RNAi-Based Gene Therapy. *Mol Ther Nucleic Acids* 16, 26-37.

Martinez, B., and Peplow, P.V. (2021). Altered microRNA expression in animal models of Huntington's disease and potential therapeutic strategies. *Neural Regen Res* 16, 2159-2169.

Martorana, F., Brambilla, L., Valori, C.F., Bergamaschi, C., Roncoroni, C., Aronica, E., Volterra, A., Bezzi, P., and Rossi, D. (2011). The BH4 domain of Bcl-XL rescues astrocyte degeneration in amyotrophic lateral sclerosis by modulating intracellular calcium signals. *Human Molecular Genetics* 21, 826-840.

Maru, Y., Afar, D.E., Witte, O.N., and Shibuya, M. (1996). The Dimerization Property of Glutathione S-Transferase Partially Reactivates Bcr-Abl Lacking the Oligomerization Domain. *Journal of Biological Chemistry* 271, 15353-15357.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., *et al.* (2010). Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 465, 223.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. *Genes & Development* 19, 1512-1517.

Matos, C.A., Carmona, V., Vijayakumar, U.G., Lopes, S., Albuquerque, P., Conceição, M., Nobre, R.J., Nóbrega, C., and de Almeida, L.P. (2018). Gene Therapies for Polyglutamine Diseases. *Adv Exp Med Biol* 1049, 395-438.

Matsumoto, G., Shimogori, T., Hattori, N., and Nukina, N. (2015). TBK1 controls autophagosomal engulfment of polyubiquitinated mitochondria through p62/SQSTM1 phosphorylation. *Human Molecular Genetics* 24, 4429-4442.

Matsuzaki, K., Yoneyama, S., Murase, O., and Miyajima, K. (1996). Transbilayer Transport of Ions and Lipids Coupled with Mastoparan X Translocation. *Biochemistry* 35, 8450-8456.

Mayl, K., Shaw, C.E., and Lee, Y.B. (2021). Disease Mechanisms and Therapeutic Approaches in C9orf72 ALS-FTD. *Biomedicines* 9.

Mayor, S., and Pagano, R.E. (2007). Pathways of clathrin-independent endocytosis. *Nature Reviews Molecular Cell Biology* 8, 603-612.

McCarty, D.M. (2008). Self-complementary AAV Vectors; Advances and Applications. *Molecular Therapy* 16, 1648-1656.

McCarty, D.M., Jr., S.M.Y., and Samulski, R.J. (2004). Integration of Adeno-Associated Virus (AAV) and Recombinant AAV Vectors. *Annual Review of Genetics* 38, 819-845.

McCarty, D.M., Monahan, P.E., and Samulski, R.J. (2001). Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Therapy* 8, 1248-1254.

Mead, R.J., Shan, N., Reiser, H.J., Marshall, F., and Shaw, P.J. (2023). Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation. *Nature Reviews Drug Discovery* 22, 185-212.

Meek, J.L. (1980). Prediction of peptide retention times in high-pressure liquid chromatography on the basis of amino acid composition. *Proc Natl Acad Sci U S A* 77, 1632-1636.

Meijboom, K.E., Abdallah, A., Fordham, N.P., Nagase, H., Rodriguez, T., Kraus, C., Gendron, T.F., Krishnan, G., Esanov, R., Andrade, N.S., *et al.* (2022). CRISPR/Cas9-mediated excision of ALS/FTD-causing hexanucleotide repeat expansion in C9ORF72 rescues major disease mechanisms in vivo and in vitro. *Nat Commun* 13, 6286.

Mejzini, R., Flynn, L.L., Pitout, I.L., Fletcher, S., Wilton, S.D., and Akkari, P.A. (2019). ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now? *Frontiers in Neuroscience* 13.

Melander, W., and Horváth, C. (1977). Salt effect on hydrophobic interactions in precipitation and chromatography of proteins: an interpretation of the lyotropic series. *Arch Biochem Biophys* 183, 200-215.

Mendell, J.R., Al-Zaidy, S., Shell, R., Arnold, W.D., Rodino-Klapac, L.R., Prior, T.W., Lowes, L., Alfano, L., Berry, K., Church, K., *et al.* (2017). Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *New England Journal of Medicine* 377, 1713-1722.

Meng, Y., Sun, D., Qin, Y., Dong, X., Luo, G., and Liu, Y. (2021). Cell-penetrating peptides enhance the transduction of adeno-associated virus serotype 9 in the central nervous system. *Mol Ther Methods Clin Dev* 21, 28-41.

Meseck, E.K., Guibinga, G., Wang, S., McElroy, C., Hudry, E., and Mansfield, K. (2022). Intrathecal sc-AAV9-CB-GFP: Systemic Distribution Predominates Following Single-Dose Administration in Cynomolgus Macaques. *Toxicol Pathol* 50, 415-431.

Messens, J., and Collet, J.-F. (2006). Pathways of disulfide bond formation in *Escherichia coli*. *The International Journal of Biochemistry & Cell Biology* 38, 1050-1062.

Meyer, K., Ferraiuolo, L., Schmelzer, L., Braun, L., McGovern, V., Likhite, S., Michels, O., Govoni, A., Fitzgerald, J., Morales, P., *et al.* (2015). Improving single injection CSF delivery of AAV9-mediated gene therapy for SMA: a dose-response study in mice and nonhuman primates. *Molecular Therapy : the journal of the American Society of Gene Therapy* 23, 477-487.

Meyer, N.L., Hu, G., Davulcu, O., Xie, Q., Noble, A.J., Yoshioka, C., Gingerich, D.S., Trzynka, A., David, L., Stagg, S.M., *et al.* (2019). Structure of the gene therapy vector, adeno-associated virus with its cell receptor, AAVR. *eLife* 8, e44707.

Mikšík, I. (2000). Protein liquid chromatography: M. Kastner (Ed.), *Journal of Chromatography Library* Vol. 61, Elsevier, Amsterdam, 2000. *Journal of Chromatography B: Biomedical Sciences and Applications* 749, 143-144.

Miller, T., Cudkowicz, M., Shaw, P.J., Andersen, P.M., Atassi, N., Bucelli, R.C., Genge, A., Glass, J., Ladha, S., Ludolph, A.L., *et al.* (2020). Phase 1–2 Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *New England Journal of Medicine* 383, 109-119.

Miller, T.M., Cudkowicz, M.E., Genge, A., Shaw, P.J., Sobue, G., Bucelli, R.C., Chiò, A., Van Damme, P., Ludolph, A.C., Glass, J.D., *et al.* (2022). Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *N Engl J Med* 387, 1099-1110.

Miller, T.M., Kaspar, B.K., Kops, G.J., Yamanaka, K., Christian, L.J., Gage, F.H., and Cleveland, D.W. (2005). Virus-delivered small RNA silencing sustains strength in amyotrophic lateral sclerosis. *Annals of Neurology* 57, 773-776.

Miller, T.M., Pestronk, A., David, W., Rothstein, J., Simpson, E., Appel, S.H., Andres, P.L., Mahoney, K., Allred, P., Alexander, K., *et al.* (2013). An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. *The Lancet Neurology* 12, 435-442.

Milletti, F. (2012). Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discovery Today* 17, 850-860.

Minamiyama, S., Sakai, M., Yamaguchi, Y., Kusui, M., Wada, H., Hikiami, R., Tamaki, Y., Asada-Utsugi, M., Shodai, A., Makino, A., *et al.* (2023). Efficacy of oligodendrocyte precursor cells as delivery vehicles for single-chain variable fragment to misfolded SOD1 in ALS rat model. *Mol Ther Methods Clin Dev* 28, 312-329.

Mis, M.S.C., Brajkovic, S., Tafuri, F., Bresolin, N., Comi, G.P., and Corti, S. (2017). Development of Therapeutics for C9ORF72 ALS/FTD-Related Disorders. *Molecular Neurobiology* 54, 4466-4476.

Mitchell, J.C., McGoldrick, P., Vance, C., Hortobagyi, T., Sreedharan, J., Rogelj, B., Tudor, E.L., Smith, B.N., Klasen, C., Miller, C.C.J., *et al.* (2013). Overexpression of human wild-type FUS causes progressive motor neuron degeneration in an age- and dose-dependent fashion. *Acta Neuropathologica* 125, 273-288.

Mitsumoto, H., Brooks, B.R., and Silani, V. (2014). Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *The Lancet Neurology* 13, 1127-1138.

Mizielinska, S., Grönke, S., Niccoli, T., Ridler, C.E., Clayton, E.L., Devoy, A., Moens, T., Norona, F.E., Woollacott, I.O.C., Pietrzyk, J., *et al.* (2014). C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* 345, 1192-1194.

Mizielinska, S., Lashley, T., Norona, F.E., Clayton, E.L., Ridler, C.E., Fratta, P., and Isaacs, A.M. (2013). C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathologica* 126, 845-857.

Moens, T.G., Partridge, L., and Isaacs, A.M. (2017). Genetic models of C9orf72: what is toxic? *Current Opinion in Genetics & Development* 44, 92-101.

Mohanty, A.K., and Wiener, M.C. (2004). Membrane protein expression and production: effects of polyhistidine tag length and position. *Protein Expression and Purification* 33, 311-325.

Moiani, A., Paleari, Y., Sartori, D., Mezzadra, R., Miccio, A., Cattoglio, C., Cocchiarella, F., Lidonnici, M.R., Ferrari, G., and Mavilio, F. (2012). Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. *J Clin Invest* 122, 1653-1666.

Moldoveanu, S., and David, V. (2013). Parameters that Characterize HPLC Analysis. In, pp. 53-83.

Mordes, D.A., Morrison, B.M., Ament, X.H., Cantrell, C., Mok, J., Eggan, P., Xue, C., Wang, J.-Y., Eggan, K., and Rothstein, J.D. (2020). Absence of Survival and Motor Deficits in 500 Repeat C9ORF72 BAC Mice. *Neuron* 108, 775-783.e774.

- Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C., *et al.* (2013). The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLN/ALS. *Science* 339, 1335-1338.
- Moussa, E.M., Panchal, J.P., Moorthy, B.S., Blum, J.S., Joubert, M.K., Narhi, L.O., and Topp, E.M. (2016). Immunogenicity of Therapeutic Protein Aggregates. *Journal of Pharmaceutical Sciences* 105, 417-430.
- Müller-McNicoll, M., Botti, V., de Jesus Domingues, A.M., Brandl, H., Schwich, O.D., Steiner, M.C., Curk, T., Poser, I., Zarnack, K., and Neugebauer, K.M. (2016). SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. *Genes Dev* 30, 553-566.
- Münch, C., Sedlmeier, R., Meyer, T., Homberg, V., Sperfeld, A.D., Kurt, A., Prudlo, J., Peraus, G., Hanemann, C.O., Stumm, G., *et al.* (2004). Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. *Neurology* 63, 724-726.
- Munis, A.M. (2020). Gene Therapy Applications of Non-Human Lentiviral Vectors. *Viruses* 12.
- Nadeeshani, H., Li, J., Ying, T., Zhang, B., and Lu, J. (2022). Nicotinamide mononucleotide (NMN) as an anti-aging health product – Promises and safety concerns. *Journal of Advanced Research* 37, 267-278.
- Naso, M.F., Tomkowicz, B., Perry, W.L., 3rd, and Strohl, W.R. (2017). Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy* 31, 317-334.
- Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., *et al.* (2006). Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science* 314, 130-133.
- Nguyen, L., Montrasio, F., Pattamatta, A., Tusi, S.K., Bardhi, O., Meyer, K.D., Hayes, L., Nakamura, K., Banez-Coronel, M., Coyne, A., *et al.* (2020). Antibody Therapy Targeting RAN Proteins Rescues C9 ALS/FTD Phenotypes in C9orf72 Mouse Model. *Neuron* 105, 645-662.e611.
- Nguyen, T.K.M., Ki, M.R., Son, R.G., and Pack, S.P. (2019). The NT11, a novel fusion tag for enhancing protein expression in Escherichia coli. *Appl Microbiol Biotechnol* 103, 2205-2216.
- Nicolas, A., Kenna, K.P., Renton, A.E., Ticozzi, N., Faghri, F., Chia, R., Dominov, J.A., Kenna, B.J., Nalls, M.A., Keagle, P., *et al.* (2018). Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* 97, 1268-1283.e1266.
- Nishimura, A.L., Mitne-Neto, M., Silva, H.C.A., Richieri-Costa, A., Middleton, S., Cascio, D., Kok, F., Oliveira, J.R.M., Gillingwater, T., Webb, J., *et al.* (2004). A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. *American Journal of Human Genetics* 75, 822-831.

Nugrahad, P.P., Hinrichs, W.L.J., Frijlink, H.W., Schöneich, C., and Avanti, C. (2023). Designing Formulation Strategies for Enhanced Stability of Therapeutic Peptides in Aqueous Solutions: A Review. *Pharmaceutics* 15, 935.

O'Fágáin, C., Cummins, P.M., and O'Connor, B.F. (2011). Gel-filtration chromatography. *Methods Mol Biol* 681, 25-33.

O'Rourke, J.G., Bogdanik, L., Muhammad, A., Gendron, T.F., Kim, K.J., Austin, A., Cady, J., Liu, E.Y., Zarrow, J., Grant, S., *et al.* (2015). C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* 88, 892-901.

O'Rourke, J.G., Bogdanik, L., Yáñez, A., Lall, D., Wolf, A.J., Muhammad, A.K.M.G., Ho, R., Carmona, S., Vit, J.P., Zarrow, J., *et al.* (2016). C9orf72 is required for proper macrophage and microglial function in mice. *Science* 351, 1324-1329.

Ohno, K., Samaranch, L., Hadaczek, P., Bringas, J.R., Allen, P.C., Sudhakar, V., Stockinger, D.E., Snieckus, C., Campagna, M.V., San Sebastian, W., *et al.* (2019). Kinetics and MR-Based Monitoring of AAV9 Vector Delivery into Cerebrospinal Fluid of Nonhuman Primates. *Molecular Therapy - Methods & Clinical Development* 13, 47-54.

Okado-Matsumoto, A., and Fridovich, I. (2001). Subcellular distribution of superoxide dismutases (SOD) in rat liver: Cu,Zn-SOD in mitochondria. *J Biol Chem* 276, 38388-38393.

Omotehara, Y., Hakuba, N., Hato, N., Okada, M., and Gyo, K. (2011). Protection Against Ischemic Cochlear Damage by Intratympanic Administration of AM-111. *Otology & Neurotology* 32, 1422-1427.

Ongkudon, C.M., Chew, J.H., Liu, B., and Danquah, M.K. (2012). Chromatographic Removal of Endotoxins: A Bioprocess Engineer's Perspective. *ISRN Chromatography* 2012, 649746.

Özlem Bahadır, A. (2013). Ion-Exchange Chromatography and Its Applications. In *Column Chromatography*, F.M. Dean, and B.M. Barbara, eds. (Rijeka: IntechOpen), p. Ch. 2.

Özlem Bahadır, A., Gülçin Saltan, C.i.l., Serkan, O.z., and Burçin, E. (2013). Affinity Chromatography and Importance in Drug Discovery. In *Column Chromatography*, F.M. Dean, and B.M. Barbara, eds. (Rijeka: IntechOpen), p. Ch. 3.

Palese, F., Sartori, A., Logroscino, G., and Pisa, F.E. (2019). Predictors of diagnostic delay in amyotrophic lateral sclerosis: a cohort study based on administrative and electronic medical records data. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 20, 176-185.

Palm-Apergi, C., Lönn, P., and Dowdy, S.F. (2012). Do Cell-Penetrating Peptides Actually "Penetrate" Cellular Membranes? *Molecular Therapy* 20, 695-697.

Palomares, L.A., Estrada-Mondaca, S., and Ramírez, O.T. (2004). Production of recombinant proteins: challenges and solutions. *Methods Mol Biol* 267, 15-52.

Pan, J., Eckardt, S., Leu, N.A., Buffone, M.G., Zhou, J., Gerton, G.L., McLaughlin, K.J., and Wang, P.J. (2009). Inactivation of Nxf2 causes defects in male meiosis and age-dependent depletion of spermatogonia. *Dev Biol* 330, 167-174.

- Pancani, S., Rowson, J., Tindale, W., Heron, N., Langley, J., McCarthy, A.D., Quinn, A., Reed, H., Stanton, A., Shaw, P.J., *et al.* (2016). Assessment of the Sheffield Support Snood, an innovative cervical orthosis designed for people affected by neck muscle weakness. *Clinical Biomechanics* 32, 201-206.
- Pandey, U.B., and Nichols, C.D. (2011). Human Disease Models in *Drosophila melanogaster* and the Role of the Fly in Therapeutic Drug Discovery. *Pharmacological Reviews* 63, 411 - 436.
- Parente, E.S., and Wetlaufer, D.B. (1984). Influence of urea on the high-performance cation-exchange chromatography of hen egg white lysozyme. *Journal of Chromatography A* 288, 389-398.
- Park, J., Ryu, J., Kim, K.-A., Lee, H.J., Bahn, J.H., Han, K., Choi, E.Y., Lee, K.S., Kwon, H.Y., and Choi, S.Y. (2002). Mutational analysis of a human immunodeficiency virus type 1 Tat protein transduction domain which is required for delivery of an exogenous protein into mammalian cells. *Journal of General Virology* 83, 1173-1181.
- Parkes, M., Myant, C., Cann, P.M., and Wong, J.S.S. (2014). The effect of buffer solution choice on protein adsorption and lubrication. *Tribology International* 72, 108-117.
- Parkinson, N., Ince, P.G., Smith, M.O., Highley, R., Skibinski, G., Andersen, P.M., Morrison, K.E., Pall, H.S., Hardiman, O., Collinge, J., *et al.* (2006). ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology* 67, 1074-1077.
- Patra, A.K., Mukhopadhyay, R., Mukhija, R., Krishnan, A., Garg, L.C., and Panda, A.K. (2000). Optimization of Inclusion Body Solubilization and Renaturation of Recombinant Human Growth Hormone from *Escherichia coli*. *Protein Expression and Purification* 18, 182-192.
- Patten, S.A., Aggad, D., Martinez, J., Tremblay, E., Petrillo, J., Armstrong, G.A., La Fontaine, A., Maios, C., Liao, M., Ciura, S., *et al.* (2017). Neuroleptics as therapeutic compounds stabilizing neuromuscular transmission in amyotrophic lateral sclerosis. *JCI Insight* 2.
- Pavani, P., Kumar, K., Rani, A., Venkatesu, P., and Lee, M.-J. (2021). The influence of sodium phosphate buffer on the stability of various proteins: Insights into protein-buffer interactions. *Journal of Molecular Liquids* 331, 115753.
- Pearson, C.E., Edamura, K.N., and Cleary, J.D. (2005). Repeat instability: mechanisms of dynamic mutations. *Nature Reviews Genetics* 6, 729-742.
- Pellot, J.E., and De Jesus, O. (2022). Suboccipital Puncture. In *StatPearls* (Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.).
- Peters, O.M., Cabrera, G.T., Tran, H., Gendron, T.F., McKeon, J.E., Metterville, J., Weiss, A., Wightman, N., Salameh, J., Kim, J., *et al.* (2015). Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron* 88, 902-909.

- Peterson, J.M., Kline, W., Canan, B.D., Ricca, D.J., Kaspar, B., Delfín, D.A., DiRienzo, K., Clemens, P.R., Robbins, P.D., Baldwin, A.S., *et al.* (2011). Peptide-Based Inhibition of NF- κ B Rescues Diaphragm Muscle Contractile Dysfunction in a Murine Model of Duchenne Muscular Dystrophy. *Molecular Medicine* 17, 508-515.
- Piguet, F., de Saint Denis, T., Audouard, E., Beccaria, K., André, A., Wurtz, G., Schatz, R., Alves, S., Sevin, C., Zerah, M., *et al.* (2021). The Challenge of Gene Therapy for Neurological Diseases: Strategies and Tools to Achieve Efficient Delivery to the Central Nervous System. *Hum Gene Ther* 32, 349-374.
- Piubelli, L., Campa, M., Temporini, C., Binda, E., Mangione, F., Amicosante, M., Terreni, M., Marinelli, F., and Pollegioni, L. (2013). Optimizing *Escherichia coli* as a protein expression platform to produce *Mycobacterium tuberculosis* immunogenic proteins. *Microbial Cell Factories* 12, 115.
- Poletti, V., and Biffi, A. (2019). Gene-Based Approaches to Inherited Neurometabolic Diseases. *Hum Gene Ther* 30, 1222-1235.
- Pooga, M., Soomets, U., Hällbrink, M., Valkna, A., Saar, K., Rezaei, K., Kahl, U., Hao, J.-X., Xu, X.-J., Wiesenfeld-Hallin, Z., *et al.* (1998). Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nature Biotechnology* 16, 857-861.
- Popiel, H.A., Nagai, Y., Fujikake, N., and Toda, T. (2007). Protein Transduction Domain-mediated Delivery of QBP1 Suppresses Polyglutamine-induced Neurodegeneration In Vivo. *Molecular Therapy* 15, 303-309.
- Pouny, Y., Rapaport, D., Mor, A., Nicolas, P., and Shai, Y. (1992). Interaction of antimicrobial dermaseptin and its fluorescently labeled analogs with phospholipid membranes. *Biochemistry* 31, 12416-12423.
- Prasad, A., Bharathi, V., Sivalingam, V., Girdhar, A., and Patel, B.K. (2019). Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. *Frontiers in Molecular Neuroscience* 12.
- Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Prgent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P., *et al.* (2015). Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nature Neuroscience* 18, 1175.
- Pujals, S., and Giralt, E. (2008). Proline-rich, amphipathic cell-penetrating peptides. *Advanced Drug Delivery Reviews* 60, 473-484.
- Puzzo, F., Colella, P., Biferi, M.G., Bali, D., Paulk, N.K., Vidal, P., Collaud, F., Simon-Sola, M., Charles, S., Hardet, R., *et al.* (2017). Rescue of Pompe disease in mice by AAV-mediated liver delivery of secreted acid α -glucosidase. *Sci Transl Med* 9.
- Rabilloud, T. (1996). Solubilization of proteins for electrophoretic analyses. *Electrophoresis* 17, 813-829.

Rabilloud, T., Adessi, C., Giraudel, A., and Lunardi, J. (1997). Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *ELECTROPHORESIS* 18, 307-316.

Raj, D., Davidoff, A.M., and Nathwani, A.C. (2011). Self-complementary adeno-associated viral vectors for gene therapy of hemophilia B: progress and challenges. *Expert Rev Hematol* 4, 539-549.

Rakhit, R., and Chakrabartty, A. (2006). Structure, folding, and misfolding of Cu,Zn superoxide dismutase in amyotrophic lateral sclerosis. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1762, 1025-1037.

Ralph, G.S., Radcliffe, P.A., Day, D.M., Carthy, J.M., Leroux, M.A., Lee, D.C.P., Wong, L.-F., Bilsland, L.G., Greensmith, L., Kingsman, S.M., *et al.* (2005). Silencing mutant SOD1 using RNAi protects against neurodegeneration and extends survival in an ALS model. *Nature Medicine* 11, 429-433.

Raoul, C., Abbas-Terki, T., Bensadoun, J.-C., Guillot, S., Haase, G., Szulc, J., Henderson, C.E., and Aebischer, P. (2005). Lentiviral-mediated silencing of SOD1 through RNA interference retards disease onset and progression in a mouse model of ALS. *Nature Medicine* 11, 423-428.

Ratti, A., and Buratti, E. (2016). Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *Journal of Neurochemistry* 138, 95-111.

Raucher, D., and Ryu, J.S. (2015). Cell-penetrating peptides: strategies for anticancer treatment. *Trends in Molecular Medicine* 21, 560-570.

Reaume, A.G., Elliott, J.L., Hoffman, E.K., Kowall, N.W., Ferrante, R.J., Siwek, D.R., Wilcox, H.M., Flood, D.G., Beal, M.F., Brown, R.H., *et al.* (1996). Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. *Nature Genetics* 13, 43-47.

Reddy, K., Schmidt, M.H.M., Geist, J.M., Thakkar, N.P., Panigrahi, G.B., Wang, Y.-H., and Pearson, C.E. (2014). Processing of double-R-loops in (CAG)·(CTG) and C9orf72 (GGGGCC)·(GGCCCC) repeats causes instability. *Nucleic Acids Research* 42, 10473-10487.

Reissmann, S. (2014). Cell penetration: scope and limitations by the application of cell-penetrating peptides. *J Pept Sci* 20, 760-784.

Reits, E., Griekspoor, A., Neijssen, J., Groothuis, T., Jalink, K., van Veelen, P., Janssen, H., Calafat, J., Drijfhout, J.W., and Neefjes, J. (2003). Peptide Diffusion, Protection, and Degradation in Nuclear and Cytoplasmic Compartments before Antigen Presentation by MHC Class I. *Immunity* 18, 97-108.

Rekker, R.F. (1977). The Hydrophobic Fragmental Constant, Its Derivation and Application: A Means of Characterizing Membrane Systems (Elsevier Scientific Publishing Company : distributors for the U.S. and Canada, Elsevier/North Holland).

Remans, K., Lebendiker, M., Abreu, C., Maffei, M., Sellathurai, S., May, M.M., Vaněk, O., and de Marco, A. (2022). Protein purification strategies must consider downstream applications and individual biological characteristics. *Microbial Cell Factories* 21, 52.

Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., *et al.* (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257-268.

Rhee, M., and Davis, P. (2006). Mechanism of Uptake of C105Y, a Novel Cell-penetrating Peptide. *Journal of Biological Chemistry* 281, 1233-1240.

Richards, D., Morren, J.A., and Pioro, E.P. (2020). Time to diagnosis and factors affecting diagnostic delay in amyotrophic lateral sclerosis. *Journal of the Neurological Sciences* 417.

Riley, B.E., Loughheed, J.C., Callaway, K., Velasquez, M., Brecht, E., Nguyen, L., Shaler, T., Walker, D., Yang, Y., Regnstrom, K., *et al.* (2013). Structure and function of Parkin E3 ubiquitin ligase reveals aspects of RING and HECT ligases. *Nature Communications* 4, 1982.

Rizzu, P., Blauwendraat, C., Heetveld, S., Lynes, E.M., Castillo-Lizardo, M., Dhingra, A., Pyz, E., Hobert, M., Synofzik, M., Simón-Sánchez, J., *et al.* (2016). C9orf72 is differentially expressed in the central nervous system and myeloid cells and consistently reduced in C9orf72, MAPT and GRN mutation carriers. *Acta Neuropathol Commun* 4, 37.

Robards, K., and Ryan, D. (2022). Chapter 9 - Analyses. In *Principles and Practice of Modern Chromatographic Methods (Second Edition)*, K. Robards, and D. Ryan, eds. (Academic Press), pp. 399-451.

Roca-Pinilla, R., López-Cano, A., Saubi, C., Garcia-Fruitós, E., and Arís, A. (2020). A new generation of recombinant polypeptides combines multiple protein domains for effective antimicrobial activity. *Microbial Cell Factories* 19, 122.

Rodrigues, F.B., and Wild, E.J. (2018). Huntington's Disease Clinical Trials Corner: February 2018. *J Huntingtons Dis* 7, 89-98.

Rodrigues, F.B., and Wild, E.J. (2020). Huntington's Disease Clinical Trials Corner: April 2020. *J Huntingtons Dis* 9, 185-197.

Rodríguez, V., Asenjo, J.A., and Andrews, B.A. (2014). Design and implementation of a high yield production system for recombinant expression of peptides. *Microbial Cell Factories* 13, 65.

Rosano, G.L., and Ceccarelli, E.A. (2009). Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted *Escherichia coli* strain. *Microbial Cell Factories* 8, 41.

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

Rosano, G.L., Morales, E.S., and Ceccarelli, E.A. (2019). New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Science* 28, 1412-1422.

Rosen, D.R., Siddique, T., Patterson, D., Figlewicz, D.A., Sapp, P., Hentati, A., Donaldson, D., Goto, J., O'Regan, J.P., Deng, H.-X., *et al.* (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 362, 59-62.

Rosenfeld, J., and Strong, M.J. (2015). Challenges in the Understanding and Treatment of Amyotrophic Lateral Sclerosis/Motor Neuron Disease. *Neurotherapeutics* 12, 317-325.

Rowland, L.P., and Shneider, N.A. (2001). Amyotrophic Lateral Sclerosis. *New England Journal of Medicine* 344, 1688-1700.

Saleh, M.N., Patel, M.R., Bauer, T.M., Goel, S., Falchook, G.S., Shapiro, G.I., Chung, K.Y., Infante, J.R., Conry, R.M., Rabinowits, G., *et al.* (2021). Phase 1 Trial of ALRN-6924, a Dual Inhibitor of MDMX and MDM2, in Patients with Solid Tumors and Lymphomas Bearing Wild-type TP53. *Clin Cancer Res* 27, 5236-5247.

Salminen, T.S., and Vale, P.F. (2020). *Drosophila* as a Model System to Investigate the Effects of Mitochondrial Variation on Innate Immunity. *Frontiers in Immunology* 11.

Sama, R.R.K., Ward, C.L., and Bosco, D.A. (2014). Functions of FUS/TLS from DNA repair to stress response: implications for ALS. *ASN Neuro* 6, 1759091414544472.

Samaranch, L., Sebastian, W.S., Kells, A.P., Salegio, E.A., Heller, G., Bringas, J.R., Pivrotto, P., DeArmond, S., Forsayeth, J., and Bankiewicz, K.S. (2014). AAV9-mediated expression of a non-self protein in nonhuman primate central nervous system triggers widespread neuroinflammation driven by antigen-presenting cell transduction. *Mol Ther* 22, 329-337.

Sánchez-Navarro, M., Teixidó, M., and Giralt, E. (2017). Jumping Hurdles: Peptides Able To Overcome Biological Barriers. *Accounts of Chemical Research* 50, 1847-1854.

Sapp, P.C., Hosler, B.A., McKenna-Yasek, D., Chin, W., Gann, A., Genise, H., Gorenstein, J., Huang, M., Sailer, W., Scheffler, M., *et al.* (2003). Identification of two novel loci for dominantly inherited familial amyotrophic lateral sclerosis. *American Journal of Human Genetics* 73, 397-403.

Sarah, W., and Sergio, S. (2014). Recombinant Peptide Production in Microbial Cells. *Current Organic Chemistry* 18, 1005-1019.

Sareen, D., O'Rourke, J.G., Meera, P., Muhammad, A.K.M.G., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A., *et al.* (2013). Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. *Science Translational Medicine* 5, 208ra149-208ra149.

Sawant, R., and Torchilin, V. (2010). Intracellular transduction using cell-penetrating peptides. *Molecular BioSystems* 6, 628-640.

Scarrott, J.M., Herranz-Martín, S., Alrafiah, A.R., Shaw, P.J., and Azzouz, M. (2015). Current developments in gene therapy for amyotrophic lateral sclerosis. *Expert Opinion on Biological Therapy* 15, 935-947.

Schäfer, F., Seip, N., Maertens, B., Block, H., and Kubicek, J. (2015). Chapter Nine - Purification of GST-Tagged Proteins. In *Methods in Enzymology*, J.R. Lorsch, ed. (Academic Press), pp. 127-139.

Schenauer, M.R., Flynn, G.C., and Goetze, A.M. (2012). Identification and quantification of host cell protein impurities in biotherapeutics using mass spectrometry. *Analytical Biochemistry* 428, 150-157.

Schlager, B., Straessle, A., and Hafen, E. (2012). Use of anionic denaturing detergents to purify insoluble proteins after overexpression. *BMC Biotechnology* 12, 95.

Schmid, B., Hruscha, A., Hogl, S., Banzhaf-Strathmann, J., Strecker, K., van der Zee, J., Teucke, M., Eimer, S., Hegermann, J., Kittelmann, M., *et al.* (2013). Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration, vascular dysfunction, and reduced motor neuron axon outgrowth. *Proceedings of the National Academy of Sciences* 110, 4986-4991.

Sclip, A., Tozzi, A., Abaza, A., Cardinetti, D., Colombo, I., Calabresi, P., Salmona, M., Welker, E., and Borsello, T. (2014). c-Jun N-terminal kinase has a key role in Alzheimer disease synaptic dysfunction in vivo. *Cell Death & Disease* 5, e1019-e1019.

Sellier, C., Campanari, M.-L., Julie Corbier, C., Gaucherot, A., Kolb-Cheynel, I., Oulad-Abdelghani, M., Ruffenach, F., Page, A., Ciura, S., Kabashi, E., *et al.* (2016). Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *The EMBO Journal* 35, 1276-1297.

Shahryari, A., Burtscher, I., Nazari, Z., and Lickert, H. (2021). Engineering Gene Therapy: Advances and Barriers. *Advanced Therapeutics* 4, 2100040.

Shang, J., Yamashita, T., Nakano, Y., Morihara, R., Li, X., Feng, T., Liu, X., Huang, Y., Fukui, Y., Hishikawa, N., *et al.* (2017). Aberrant distributions of nuclear pore complex proteins in ALS mice and ALS patients. *Neuroscience* 350, 158-168.

Shao, Q., Liang, C., Chang, Q., Zhang, W., Yang, M., and Chen, J.-F. (2019). C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. *Acta Neuropathologica Communications* 7, 32.

Sharma, A., Lyashchenko, A.K., Lu, L., Nasrabad, S.E., Elmaleh, M., Mendelsohn, M., Nemes, A., Tapia, J.C., Mentis, G.Z., and Shneider, N.A. (2016). ALS-associated mutant FUS induces selective motor neuron degeneration through toxic gain of function. *Nature Communications* 7, 10465.

Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: Capping and polyadenylation. *Nature Structural Biology* 7, 838-842.

Shav-Tal, Y., Darzacq, X., Shenoy, S.M., Fusco, D., Janicki, S.M., Spector, D.L., and Singer, R.H. (2004). Dynamics of Single mRNPs in Nuclei of Living Cells. *Science* 304, 1797-1800.

Shaw, G., Morse, S., Ararat, M., and Graham, F.L. (2002). Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *Faseb j* 16, 869-871.

Shaw, M.P., Higginbottom, A., McGown, A., Castelli, L.M., James, E., Hautbergue, G.M., Shaw, P.J., and Ramesh, T.M. (2018). Stable transgenic C9orf72 zebrafish model key aspects of the ALS/FTD phenotype and reveal novel pathological features. *Acta Neuropathol Commun* 6, 125.

Shaw, P.J. (2005). Molecular and cellular pathways of neurodegeneration in motor neurone disease. *Journal of Neurology, Neurosurgery, and Psychiatry* 76, 1046-1057.

Sheehan, G., Garvey, A., Croke, M., and Kavanagh, K. (2018). Innate humoral immune defences in mammals and insects: The same, with differences ? *Virulence* 9, 1625-1639.

Sheffield, L.G., Miskiewicz, H.B., Tannenbaum, L.B., and Mirra, S.S. (2006). Nuclear Pore Complex Proteins in Alzheimer Disease. *Journal of Neuropathology & Experimental Neurology* 65, 45-54.

Shi, K.Y., Mori, E., Nizami, Z.F., Lin, Y., Kato, M., Xiang, S., Wu, L.C., Ding, M., Yu, Y., Gall, J.G., *et al.* (2017). Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. *Proceedings of the National Academy of Sciences* 114, E1111-E1117.

Shi, Y., Hung, S.T., Rocha, G., Lin, S., Linares, G.R., Staats, K.A., Seah, C., Wang, Y., Chickering, M., Lai, J., *et al.* (2019). Identification and therapeutic rescue of autophagosome and glutamate receptor defects in C9ORF72 and sporadic ALS neurons. *JCI Insight* 5.

Shukla, J., Bhargava, R., Dash, P.K., Dash, P.K., Parida, M., Tripathi, N., and Rao, P.V.L. (2009). Cloning and expression of domain III of the envelope gene of Japanese encephalitis virus: evaluation for early clinical diagnosis by IgM ELISA. *J Virol Methods* 158, 165-170.

Simón-Sánchez, J., Dopper, E.G.P., Cohn-Hokke, P.E., Hukema, R.K., Nicolaou, N., Seelaar, H., de Graaf, J.R.A., de Koning, I., van Schoor, N.M., Deeg, D.J.H., *et al.* (2012). The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. *Brain* 135, 723-735.

Simone, R., Balendra, R., Moens, T.G., Preza, E., Wilson, K.M., Heslegrave, A., Woodling, N.S., Niccoli, T., Gilbert-Jaramillo, J., Abdelkarim, S., *et al.* (2018). G-quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *EMBO Mol Med* 10, 22-31.

Singh, T., Murthy, A.S.N., Yang, H.-J., and Im, J. (2018). Versatility of cell-penetrating peptides for intracellular delivery of siRNA. *Drug Delivery* 25, 1996-2006.

Šiurkus, J., and Neubauer, P. (2011). Reducing conditions are the key for efficient production of active ribonuclease inhibitor in *Escherichia coli*. *Microbial Cell Factories* 10, 31.

Sivadasan, R., Hornburg, D., Drepper, C., Frank, N., Jablonka, S., Hansel, A., Lojewski, X., Sternecker, J., Hermann, A., Shaw, P.J., *et al.* (2016). C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nature Neuroscience* 19, 1610-1618.

Smeyers, J., Banchi, E.-G., and Latouche, M. (2021). C9ORF72: What It Is, What It Does, and Why It Matters. *Frontiers in Cellular Neuroscience* 15.

Snow, B.A., Stevenson, C.C., Kaur, J., Lee, S.G., Wei, Y., Miao, H., and Kim, W.J. (2022). Surface glia predominantly contribute to the development of ALS/FTD in *Drosophila* model. *bioRxiv*, 2022.2010.2007.511246.

Snyder, E.L., Meade, B.R., Saenz, C.C., and Dowdy, S.F. (2004). Treatment of Terminal Peritoneal Carcinomatosis by a Transducible p53-Activating Peptide. *PLOS Biology* 2, e36.

Sousa, D.A., Mulder, K.C.L., Nobre, K.S., Parachin, N.S., and Franco, O.L. (2016). Production of a polar fish antimicrobial peptide in *Escherichia coli* using an ELP-intein tag. *Journal of Biotechnology* 234, 83-89.

Sreedharan, J., Blair, I.P., Tripathi, V.B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J.C., Williams, K.L., Buratti, E., *et al.* (2008). TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science* 319, 1668-1672.

Staby, A., Sand, M.-B., Hansen, R.G., Jacobsen, J.H., Andersen, L.A., Gerstenberg, M., Bruus, U.K., and Jensen, I.H. (2005). Comparison of chromatographic ion-exchange resins: IV. Strong and weak cation-exchange resins and heparin resins. *Journal of Chromatography A* 1069, 65-77.

Stalmans, S., Bracke, N., Wynendaele, E., Gevaert, B., Peremans, K., Burvenich, C., Polis, I., and De Spiegeleer, B. (2015). Cell-Penetrating Peptides Selectively Cross the Blood-Brain Barrier In Vivo. *PLoS One* 10, e0139652.

Stevens, R.C. (2000). Design of high-throughput methods of protein production for structural biology. *Structure* 8, R177-R185.

Stewart, M. (2007). Ratcheting mRNA out of the Nucleus. *Molecular Cell* 25, 327-330.

Sträßer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A.G., Aguilera, A., Struhl, K., Reed, R., *et al.* (2002). TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* 417, 304-308.

Striegel, A., Yau, W.W., Kirkland, J.J., and Bly, D.D. (2009). *Modern Size-Exclusion Liquid Chromatography: Practice of Gel Permeation and Gel Filtration Chromatography* (Wiley).

Su, Y., Doherty, T., Waring, A.J., Ruchala, P., and Hong, M. (2009). Roles of Arginine and Lysine Residues in the Translocation of a Cell-Penetrating Peptide from ¹³C, ³¹P, and ¹⁹F Solid-State NMR. *Biochemistry* 48, 4587-4595.

Suckfuell, M., Lisowska, G., Domka, W., Kabacinska, A., Morawski, K., Bodlaj, R., Klimak, P., Kostrica, R., and Meyer, T. (2014). Efficacy and Safety of AM-111 in the Treatment of Acute Sensorineural Hearing Loss: A Double-Blind, Randomized, Placebo-Controlled Phase II Study. *Otology & Neurotology* 35, 1317-1326.

Suh, E., Lee, E.B., Neal, D., Wood, E.M., Toledo, J.B., Rennert, L., Irwin, D.J., McMillan, C.T., Krock, B., Elman, L.B., *et al.* (2015). Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. *Acta Neuropathol* 130, 363-372.

Sun, T., Chance, R.R., Graessley, W.W., and Lohse, D.J. (2004). A Study of the Separation Principle in Size Exclusion Chromatography. *Macromolecules* 37, 4304-4312.

Svensson, J., Andersson, C., Reseland, J.E., Lyngstadaas, P., and Bülow, L. (2006). Histidine tag fusion increases expression levels of active recombinant amelogenin in *Escherichia coli*. *Protein Expr Purif* 48, 134-141.

Tabet, R., Schaeffer, L., Freyermuth, F., Jambeau, M., Workman, M., Lee, C.-Z., Lin, C.-C., Jiang, J., Jansen-West, K., Abou-Hamdan, H., *et al.* (2018). CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nature Communications* 9, 152.

Takahashi, Y., Fukuda, Y., Yoshimura, J., Toyoda, A., Kurppa, K., Moritoyo, H., Belzil, V.V., Dion, P.A., Higasa, K., Doi, K., *et al.* (2013). ERBB4 mutations that disrupt the neuregulin-ErbB4 pathway cause amyotrophic lateral sclerosis type 19. *American Journal of Human Genetics* 93, 900-905.

Takata, M., Tanaka, H., Kimura, M., Nagahara, Y., Tanaka, K., Kawasaki, K., Seto, M., Tsuruma, K., Shimazawa, M., and Hara, H. (2013). Fasudil, a rho kinase inhibitor, limits motor neuron loss in experimental models of amyotrophic lateral sclerosis. *British Journal of Pharmacology* 170, 341-351.

Talbott, E.O., Malek, A.M., and Lacomis, D. (2016). Chapter 13 - The epidemiology of amyotrophic lateral sclerosis. In *Handbook of Clinical Neurology*, M.J. Aminoff, F. Boller, and D.F. Swaab, eds. (Elsevier), pp. 225-238.

Tan, E., Chin, C.S.H., Lim, Z.F.S., and Ng, S.K. (2021). HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. *Front Bioeng Biotechnol* 9, 796991.

Tanhaeian, A., Habibi Najafi, M.B., Rahnema, P., and Azghandi, M. (2020). Production of a Recombinant Peptide (Lasioglossin LL III) and Assessment of Antibacterial and Antioxidant Activity. *International Journal of Peptide Research and Therapeutics* 26, 1021-1029.

Tanhaeian, A., Azghandi, M., Razmyar, J., Mohammadi, E., and Sekhavati, M.H. (2018). Recombinant production of a chimeric antimicrobial peptide in *E. coli* and assessment of its activity against some avian clinically isolated pathogens. *Microbial Pathogenesis* 122, 73-78.

Thomassen, J., Faraday, D., Underwood, B., and Cleaver, J. (2005). The effect of varying transmembrane pressure and crossflow velocity on the microfiltration fouling of a model beer. *Separation and Purification Technology - SEP PURIF TECHNOL* 41, 91-100.

Tintaru, A.M., Hautbergue, G.M., Hounslow, A.M., Hung, M.-L., Lian, L.-Y., Craven, C.J., and Wilson, S.A. (2007). Structural and functional analysis of RNA and TAP binding to SF2/ASF. *EMBO reports* 8, 756-762.

- Toma, J.S., Shettar, B.C., Chipman, P.H., Pinto, D.M., Borowska, J.P., Ichida, J.K., Fawcett, J.P., Zhang, Y., Eggan, K., and Rafuse, V.F. (2015). Motoneurons derived from induced pluripotent stem cells develop mature phenotypes typical of endogenous spinal motoneurons. *J Neurosci* 35, 1291-1306.
- Towne, C., Pertin, M., Beggah, A.T., Aebischer, P., and Decosterd, I. (2009). Recombinant Adeno-Associated Virus Serotype 6 (rAAV2/6)-Mediated Gene Transfer to Nociceptive Neurons through Different Routes of Delivery. *Molecular Pain* 5, 1744-8069-1745-1752.
- Tran, E.J., Zhou, Y., Corbett, A.H., and Wenthe, S.R. (2007). The DEAD-Box Protein Dbp5 Controls mRNA Export by Triggering Specific RNA:Protein Remodeling Events. *Molecular Cell* 28, 850-859.
- Tran, H., Almeida, S., Moore, J., Gendron, T.F., Chalasani, U., Lu, Y., Du, X., Nickerson, J.A., Petrucelli, L., Weng, Z., *et al.* (2015). Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a *Drosophila* Model of C9ORF72 FTD/ALS. *Neuron* 87, 1207-1214.
- Tran, H., Moazami, M.P., Yang, H., McKenna-Yasek, D., Douthwright, C.L., Pinto, C., Metterville, J., Shin, M., Sanil, N., Dooley, C., *et al.* (2022). Suppression of mutant C9orf72 expression by a potent mixed backbone antisense oligonucleotide. *Nat Med* 28, 117-124.
- Trias, E., Ibarburu, S., Barreto-Núñez, R., Babdor, J., Maciel, T.T., Guillo, M., Gros, L., Dubreuil, P., Díaz-Amarilla, P., Cassina, P., *et al.* (2016). Post-paralysis tyrosine kinase inhibition with masitinib abrogates neuroinflammation and slows disease progression in inherited amyotrophic lateral sclerosis. *Journal of Neuroinflammation* 13, 177.
- Tschernutter, M., Schlichtenbrede, F.C., Howe, S., Balaggan, K.S., Munro, P.M., Bainbridge, J.W.B., Thrasher, A.J., Smith, A.J., and Ali, R.R. (2005). Long-term preservation of retinal function in the RCS rat model of retinitis pigmentosa following lentivirus-mediated gene therapy. *Gene Therapy* 12, 694-701.
- Tsumoto, K., Ejima, D., Senczuk, A.M., Kita, Y., and Arakawa, T. (2007). Effects of salts on protein-surface interactions: applications for column chromatography. *Journal of Pharmaceutical Sciences* 96, 1677-1690.
- Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J.S., and Arakawa, T. (2004). Role of Arginine in Protein Refolding, Solubilization, and Purification. *Biotechnology Progress* 20, 1301-1308.
- Uccelli, A., Milanese, M., Principato, M.C., Morando, S., Bonifacino, T., Vergani, L., Giunti, D., Voci, A., Carminati, E., Giribaldi, F., *et al.* (2012). Intravenous mesenchymal stem cells improve survival and motor function in experimental amyotrophic lateral sclerosis. *Mol Med* 18, 794-804.
- Uddin, F., Rudin, C.M., and Sen, T. (2020). CRISPR Gene Therapy: Applications, Limitations, and Implications for the Future. *Front Oncol* 10, 1387.
- Ueda, E.K., Gout, P.W., and Morganti, L. (2003). Current and prospective applications of metal ion-protein binding. *J Chromatogr A* 988, 1-23.

Urh, M., Simpson, D., and Zhao, K. (2009). Affinity chromatography: general methods. *Methods Enzymol* 463, 417-438.

van den Berg, A., and Dowdy, S.F. (2011). Protein transduction domain delivery of therapeutic macromolecules. *Current Opinion in Biotechnology* 22, 888-893.

Van Harten, A.C.M., Phatnani, H., and Przedborski, S. (2021). Non-cell-autonomous pathogenic mechanisms in amyotrophic lateral sclerosis. *Trends Neurosci* 44, 658-668.

van Koppen, C., Meyer zu Heringdorf, M., Laser, K.T., Zhang, C., Jakobs, K.H., Bünemann, M., and Pott, L. (1996). Activation of a high affinity Gi protein-coupled plasma membrane receptor by sphingosine-1-phosphate. *J Biol Chem* 271, 2082-2087.

van Rheenen, W., Shatunov, A., Dekker, A.M., McLaughlin, R.L., Diekstra, F.P., Pulit, S.L., van der Spek, R.A.A., Vösa, U., de Jong, S., Robinson, M.R., *et al.* (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature Genetics* 48, 1043.

Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., *et al.* (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323, 1208-1211.

Vatovec, S., Kovanda, A., and Rogelj, B. (2014). Unconventional features of C9ORF72 expanded repeat in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiology of Aging* 35, 2421.e2421-2421.e2412.

Viphakone, N., Hautbergue, G.M., Walsh, M., Chang, C.-T., Holland, A., Folco, E.G., Reed, R., and Wilson, S.A. (2012). TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nature Communications* 3, 1006.

Vivès, E., Brodin, P., and Lebleu, B. (1997). A Truncated HIV-1 Tat Protein Basic Domain Rapidly Translocates through the Plasma Membrane and Accumulates in the Cell Nucleus. *Journal of Biological Chemistry* 272, 16010-16017.

Volontè, F., Piubelli, L., and Pollegioni, L. (2011). Optimizing HIV-1 protease production in *Escherichia coli* as fusion protein. *Microbial Cell Factories* 10, 53.

Volontè, F., Pollegioni, L., Molla, G., Frattini, L., Marinelli, F., and Piubelli, L. (2010). Production of recombinant cholesterol oxidase containing covalently bound FAD in *Escherichia coli*. In *BMC Biotechnol*, pp. 33.

Von Hippel, P.H., and Wong, K.Y. (1965). On the conformational stability of globular proteins. The effects of various electrolytes and nonelectrolytes on the thermal ribonuclease transition. *J Biol Chem* 240, 3909-3923.

Walker, C., Herranz-Martin, S., Karyka, E., Liao, C., Lewis, K., Elsayed, W., Lukashchuk, V., Chiang, S.-C., Ray, S., Mulcahy, P.J., *et al.* (2017). C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nature Neuroscience* 20, 1225.

Walrant, A., Cardon, S., Burlina, F., and Sagan, S. (2017). Membrane Crossing and Membranotropic Activity of Cell-Penetrating Peptides: Dangerous Liaisons? *Accounts of Chemical Research* 50, 2968-2975.

Walsh, M.J., Cooper-Knock, J., Dodd, J.E., Stopford, M.J., Mihaylov, S.R., Kirby, J., Shaw, P.J., and Hautbergue, G.M. (2015). Invited Review: Decoding the pathophysiological mechanisms that underlie RNA dysregulation in neurodegenerative disorders: a review of the current state of the art. *Neuropathology and Applied Neurobiology* 41, 109-134.

Walsh, Matthew J., Hautbergue, Guillaume M., and Wilson, Stuart A. (2010). Structure and function of mRNA export adaptors. *Biochemical Society Transactions* 38, 232-236.

Wang, H., Guan, L., and Deng, M. (2023). Recent progress of the genetics of amyotrophic lateral sclerosis and challenges of gene therapy. *Front Neurosci* 17, 1170996.

Wang, L., Wang, H., Bell, P., McMenamin, D., and Wilson, J.M. (2012). Hepatic gene transfer in neonatal mice by adeno-associated virus serotype 8 vector. *Hum Gene Ther* 23, 533-539.

Wang, L., Wang, N., Zhang, W., Cheng, X., Yan, Z., Shao, G., Wang, X., Wang, R., and Fu, C. (2022). Therapeutic peptides: current applications and future directions. *Signal Transduction and Targeted Therapy* 7, 48.

Wang, W., Li, L., Lin, W.-L., Dickson, D.W., Petrucelli, L., Zhang, T., and Wang, X. (2013). The ALS disease-associated mutant TDP-43 impairs mitochondrial dynamics and function in motor neurons. *Human Molecular Genetics* 22, 4706-4719.

Wang, Y., Chen, X., Wang, Y., and Li, Y. (2021). Impact of salt concentration in mobile phase on antibody retention in Protein A, Protein L and KappaSelect affinity chromatography. *Protein Expression and Purification* 178, 105786.

Wang, Z.-F., Ursu, A., Childs-Disney, J.L., Guertler, R., Yang, W.-Y., Bernat, V., Rzuczek, S.G., Fuerst, R., Zhang, Y.-J., Gendron, T.F., *et al.* (2019). The Hairpin Form of r(G4C2) in c9ALS/FTD Is Repeat-Associated Non-ATG Translated and a Target for Bioactive Small Molecules. *Cell Chemical Biology* 26, 179-190.e112.

Wang, Z., Ma, H.I., Li, J., Sun, L., Zhang, J., and Xiao, X. (2003). Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors in vitro and in vivo. *Gene Ther* 10, 2105-2111.

Ward, M.E., Taubes, A., Chen, R., Miller, B.L., Sephton, C.F., Gelfand, J.M., Minami, S., Boscardin, J., Martens, L.H., Seeley, W.W., *et al.* (2014). Early retinal neurodegeneration and impaired Ran-mediated nuclear import of TDP-43 in progranulin-deficient FTLD. *The Journal of Experimental Medicine* 211, 1937-1945.

Waring, S.C., Esteban-Santillan, C., Reed, D.M., Craig, U.K., Labarthe, D.R., Petersen, R.C., and Kurland, L.T. (2004). Incidence of Amyotrophic Lateral Sclerosis and of the Parkinsonism-Dementia Complex of Guam, 1950–1989. *Neuroepidemiology* 23, 192-200.

Waugh, D.S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Protein Expression and Purification* 80, 283-293.

Weber-Adrian, D., Kofoed, R.H., Silburt, J., Noroozian, Z., Shah, K., Burgess, A., Rideout, S., Kügler, S., Hynynen, K., and Aubert, I. (2021). Systemic AAV6-synapsin-GFP administration results in lower liver biodistribution, compared to AAV1&2 and AAV9, with neuronal expression following ultrasound-mediated brain delivery. *Sci Rep* 11, 1934.

Webster, C.P., Smith, E.F., Bauer, C.S., Moller, A., Hautbergue, G.M., Ferraiuolo, L., Myszczyńska, M.A., Higginbottom, A., Walsh, M.J., Whitworth, A.J., *et al.* (2016). The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *The EMBO Journal* 35, 1656-1676.

Wegmuller, S., and Schmid, S. (2014). Recombinant Peptide Production in Microbial Cells. *Current Organic Chemistry* 18, 1005-1019.

Weidberg, H., and Elazar, Z. (2011). TBK1 Mediates Crosstalk Between the Innate Immune Response and Autophagy. *Science Signaling* 4, pe39-pe39.

Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S.S., Shi, Y., Lin, S., Shneider, N.A., Monaghan, J., Pandey, U.B., *et al.* (2014). Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* 84, 1213-1225.

Wender, P.A., Mitchell, D.J., Pattabiraman, K., Pelkey, E.T., Steinman, L., and Rothbard, J.B. (2000). The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *Proceedings of the National Academy of Sciences* 97, 13003-13008.

Wijesekera, L.C., and Nigel Leigh, P. (2009). Amyotrophic lateral sclerosis. *Orphanet Journal of Rare Diseases* 4, 3.

Witzel, S., Mayer, K., and Oeckl, P. (2022). Biomarkers for amyotrophic lateral sclerosis. *Current Opinion in Neurology* 35, 699-704.

Woerner, A.C., Frottin, F., Hornburg, D., Feng, L.R., Meissner, F., Patra, M., Tatzelt, J., Mann, M., Winklhofer, K.F., Hartl, F.U., *et al.* (2016). Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. *Science* 351, 173-176.

Wojciechowska, M., and Krzyzosiak, W.J. (2011). Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Human Molecular Genetics* 20, 3811-3821.

Wu, C.-H., Fallini, C., Ticozzi, N., Keagle, P.J., Sapp, P.C., Piotrowska, K., Lowe, P., Koppers, M., McKenna-Yasek, D., Baron, D.M., *et al.* (2012). Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* 488, 499-503.

Wu, X., Wang, Z., Li, X., Fan, Y., He, G., Wan, Y., Yu, C., Tang, J., Li, M., Zhang, X., *et al.* (2014). In vitro and in vivo activities of antimicrobial peptides developed using an amino acid-based activity prediction method. *Antimicrob Agents Chemother* 58, 5342-5349.

Wu, Z., Asokan, A., and Samulski, R.J. (2006). Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy. *Molecular Therapy* 14, 316-327.

Wyatt, T.J., Rossi, S.L., Siegenthaler, M.M., Frame, J., Robles, R., Nistor, G., and Keirstead, H.S. (2011). Human motor neuron progenitor transplantation leads to endogenous neuronal sparing in 3 models of motor neuron loss. *Stem Cells Int* 2011, 207230.

Xiao, S., MacNair, L., McGoldrick, P., McKeever, P.M., McLean, J.R., Zhang, M., Keith, J., Zinman, L., Rogaeva, E., and Robertson, J. (2015). Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Annals of Neurology* 78, 568-583.

Xie, J., Bi, Y., Zhang, H., Dong, S., Teng, L., Lee, R.J., and Yang, Z. (2020). Cell-Penetrating Peptides in Diagnosis and Treatment of Human Diseases: From Preclinical Research to Clinical Application. *Frontiers in Pharmacology* 11.

Xu, X., Yang, D., Ding, J.-H., Wang, W., Chu, P.-H., Dalton, N.D., Wang, H.-Y., Bermingham, J.R., Jr., Ye, Z., Liu, F., *et al.* (2005). ASF/SF2-Regulated CaMKII δ ; Alternative Splicing Temporally Reprograms Excitation-Contraction Coupling in Cardiac Muscle. *Cell* 120, 59-72.

Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D.L., Li, H., Hales, C.M., Gearing, M., Wingo, T.S., *et al.* (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America* 110, 7778-7783.

Yamashita, S., and Ando, Y. (2015). Genotype-phenotype relationship in hereditary amyotrophic lateral sclerosis. *Translational Neurodegeneration* 4, 13.

Yan, J., Xu, L., Welsh, A.M., Chen, D., Hazel, T., Johe, K., and Koliatsos, V.E. (2006). Combined immunosuppressive agents or CD4 antibodies prolong survival of human neural stem cell grafts and improve disease outcomes in amyotrophic lateral sclerosis transgenic mice. *Stem Cells* 24, 1976-1985.

Yang, D., Sun, Y.-Y., Lin, X., Baumann, J.M., Dunn, R.S., Lindquist, D.M., and Kuan, C.-Y. (2013). Intranasal delivery of cell-penetrating anti-NF- κ B peptides (Tat-NBD) alleviates infection-sensitized hypoxic-ischemic brain injury. *Experimental Neurology* 247, 447-455.

Yang, X., Ji, Y., Wang, W., Zhang, L., Chen, Z., Yu, M., Shen, Y., Ding, F., Gu, X., and Sun, H. (2021a). Amyotrophic Lateral Sclerosis: Molecular Mechanisms, Biomarkers, and Therapeutic Strategies. *Antioxidants (Basel)* 10.

Yang, Y., Hentati, A., Deng, H.-X., Dabbagh, O., Sasaki, T., Hirano, M., Hung, W.-Y., Ouahchi, K., Yan, J., Azim, A.C., *et al.* (2001). The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. *Nature Genetics* 29, 160-165.

Yang, Y., Xu, J., Ge, S., and Lai, L. (2021b). CRISPR/Cas: Advances, Limitations, and Applications for Precision Cancer Research. *Frontiers in Medicine* 8.

Yang, Z., Zhang, L., Zhang, Y., Zhang, T., Feng, Y., Lu, X., Lan, W., Wang, J., Wu, H., Cao, C., *et al.* (2011). Highly Efficient Production of Soluble Proteins from Insoluble Inclusion Bodies by a Two-Step-Denaturing and Refolding Method. *PLOS ONE* 6, e22981.

Yao, Y., Wang, J., Liu, Y., Qu, Y., Wang, K., Zhang, Y., Chang, Y., Yang, Z., Wan, J., Liu, J., *et al.* (2022). Variants of the adeno-associated virus serotype 9 with enhanced penetration of the blood–brain barrier in rodents and primates. *Nature Biomedical Engineering* 6, 1257-1271.

Yaroustovsky, M., Plyushch, M., Popov, D., Samsonova, N., Abramyan, M., Popok, Z., and Krotenko, N. (2013). Prognostic value of endotoxin activity assay in patients with severe sepsis after cardiac surgery. *J Inflamm (Lond)* 10, 8.

Yiu, E.M., and Kornberg, A.J. (2015). Duchenne muscular dystrophy. *J Paediatr Child Health* 51, 759-764.

Yoon, S.Y., Hunter, J.E., Chawla, S., Clarke, D.L., Molony, C., O'Donnell, P.A., Bagel, J.H., Kumar, M., Poptani, H., Vite, C.H., *et al.* (2020). Global CNS correction in a large brain model of human alpha-mannosidosis by intravascular gene therapy. *Brain* 143, 2058-2072.

Yuan, H., Li, N., Xun, Y., Wang, L., Feng, X., Wang, Y., Tao, Y., and Wang, S. (2021). An Efficient Heparin Affinity Column Purification Method Coupled with Ultraperformance Liquid Chromatography for the Quantification of Native Lactoferrin in Breast Milk. *Journal of Food Quality* 2021, 4675343.

Yuva-Aydemir, Y., Almeida, S., and Gao, F.B. (2018). Insights into C9ORF72-Related ALS/FTD from *Drosophila* and iPSC Models. *Trends Neurosci* 41, 457-469.

Zamiri, B., Mirceta, M., Abu-Ghazalah, R., Wold, M.S., Pearson, C.E., and Macgregor, R.B. (2018). Stress-induced acidification may contribute to formation of unusual structures in C9orf72-repeats. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1862, 1482-1491.

Zapadka, K.L., Becher, F.J., Santos, A.L.G.d., and Jackson, S.E. (2017). Factors affecting the physical stability (aggregation) of peptide therapeutics. *Interface Focus* 7, 20170030.

Zaro, J.L., and Shen, W.-C. (2015). Cationic and amphipathic cell-penetrating peptides (CPPs): Their structures and in vivo studies in drug delivery. *Frontiers of Chemical Science and Engineering* 9, 407-427.

Zhang, K., Daigle, J.G., Cunningham, K.M., Coyne, A.N., Ruan, K., Grima, J.C., Bowen, K.E., Wadhwa, H., Yang, P., Rigo, F., *et al.* (2018a). Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. *Cell* 173, 958-971.e917.

Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P., Daley, E.L., Miller, S.J., Cunningham, K.M., Vidensky, S., *et al.* (2015a). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* 525, 56-61.

Zhang, Q., An, Y., Chen, Z.S., Koon, A.C., Lau, K.F., Ngo, J.C.K., and Chan, H.Y.E. (2019). A Peptidyl Inhibitor for Neutralizing (r)(GGGGCC)(exp)-Associated Neurodegeneration in C9ALS-FTD. *Mol Ther Nucleic Acids* 16, 172-185.

Zhang, Q., Chen, Z.S., An, Y., Liu, H., Hou, Y., Li, W., Lau, K.F., Koon, A.C., Ngo, J.C.K., and Chan, H.Y.E. (2018b). A peptidyl inhibitor for neutralizing expanded CAG RNA-induced nucleolar stress in polyglutamine diseases. *Rna* 24, 486-498.

Zhang, X., He, T., Chai, Z., Samulski, R.J., and Li, C. (2018c). Blood-brain barrier shuttle peptides enhance AAV transduction in the brain after systemic administration. *Biomaterials* 176, 71-83.

Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S., and Yang, S.H. (2015b). Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids* 4, e264.

Zhang, Z., Devarajan, P., Dorfman, A.L., and Morrow, J.S. (1998). Structure of the Ankyrin-binding Domain of Na,K-ATPase *. *Journal of Biological Chemistry* 273, 18681-18684.

Zhao, Q., Xu, W., Xing, L., and Lin, Z. (2016). Recombinant production of medium- to large-sized peptides in *Escherichia coli* using a cleavable self-aggregating tag. *Microbial Cell Factories* 15, 136.

Zhao, X., Kumari, D., Miller, C.J., Kim, G.Y., Hayward, B., Vitalo, A.G., Pinto, R.M., and Usdin, K. (2021). Modifiers of Somatic Repeat Instability in Mouse Models of Friedreich Ataxia and the Fragile X-Related Disorders: Implications for the Mechanism of Somatic Expansion in Huntington's Disease. *J Huntingtons Dis* 10, 149-163.

Zhao, X., Li, G., and Liang, S. (2013). Several Affinity Tags Commonly Used in Chromatographic Purification. *Journal of Analytical Methods in Chemistry* 2013, 581093.

Zhao, Z., Anselmo, A.C., and Mitragotri, S. (2022). Viral vector-based gene therapies in the clinic. *Bioeng Transl Med* 7, e10258.

Zhou, L., Liu, Z., Xu, G., Li, L., Xuan, K., Xu, Y., and Zhang, R. (2020a). Expression of Melittin in Fusion with GST in *Escherichia coli* and Its Purification as a Pure Peptide with Good Bacteriostatic Efficacy. *ACS Omega* 5, 9251-9258.

Zhou, Q., Lehmer, C., Michaelsen, M., Mori, K., Alterauge, D., Baumjohann, D., Schludi, M.H., Greiling, J., Farny, D., Flatley, A., *et al.* (2017). Antibodies inhibit transmission and aggregation of C9orf72 poly-GA dipeptide repeat proteins. *EMBO Mol Med* 9, 687-702.

Zhou, Q., Mareljic, N., Michaelsen, M., Parhizkar, S., Heindl, S., Nuscher, B., Farny, D., Czuppa, M., Schludi, C., Graf, A., *et al.* (2020b). Active poly-GA vaccination prevents microglia activation and motor deficits in a C9orf72 mouse model. *EMBO Mol Med* 12, e10919.

Zorko, M., and Jerala, R. (2010). Production of recombinant antimicrobial peptides in bacteria. *Methods Mol Biol* 618, 61-76.

Zu, T., Cleary, J.D., Liu, Y., Bañez-Coronel, M., Bubenik, J.L., Ayhan, F., Ashizawa, T., Xia, G., Clark, H.B., Yachnis, A.T., *et al.* (2017). RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. *Neuron* 95, 1292-1305.e1295.

Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A.C., *et al.* (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proceedings of the National Academy of Sciences* 108, 260-265.

Zu, T., Guo, S., Bardhi, O., Ryskamp, D.A., Li, J., Khoramian Tusi, S., Engelbrecht, A., Klippel, K., Chakrabarty, P., Nguyen, L., *et al.* (2020). Metformin inhibits RAN translation through PKR pathway and mitigates disease in C9orf72 ALS/FTD mice. *Proc Natl Acad Sci U S A* *117*, 18591-18599.

Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T.M., Harms, M.B., Falchook, A.E., Subramony, S.H., *et al.* (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proceedings of the National Academy of Sciences of the United States of America* *110*, E4968-E4977.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol* *72*, 9873-9880.